



**A BLOOM-TAXONÓMIÁRA ÉPÜLŐ TUDÁSMÉRŐ
E-TESZTEK SZERKESZTÉSE ÉS MOTIVÁCIÓS
SZEREPE A FELSŐFOKÚ
MATEMATIKAOKTATÁSBAN**

Egyetemi doktori (PhD) értekezés

Papp Gabriella

Témavezető: Dr. habil. Korenova Lilla

DEBRECENI EGYETEM

Természettudományi és Műszaki Tudományi Doktori Tanács

Matematika- és Számítástudományok Doktori Iskola

Debrecen, 2026

Ezen értekezést a Debreceni Egyetem Természettudományi és Műszaki Tudományi Doktori Tanács, Matematika- és Számítástudományok Doktori Iskola Didaktika (Szak módszertan) programja keretében készítettem a Debreceni Egyetem természettudományi doktori (PhD) fokozatának elnyerése céljából.

Nyilatkozom arról, hogy a tézisekben leírt eredmények nem képezik más PhD disszertáció részét.

Debrecen, 20.

.....
a jelölt aláírása

Tanúsítom, hogy Papp Gabriella doktorjelölt 2020 - 2024 között a fent megnevezett Doktori Iskola Didaktika (Szak módszertan) programjának keretében irányításommal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotó tevékenységével meghatározóan hozzájárult. Nyilatkozom továbbá arról, hogy a tézisekben leírt eredmények nem képezik más PhD disszertáció részét.

Az értekezés elfogadását javasolom.

Debrecen, 20..

.....
a témavezető aláírása

**A BLOOM-TAXONÓMIÁRA ÉPÜLŐ TUDÁSMÉRŐ
E-TESZTEK SZERKESZTÉSE ÉS MOTIVÁCIÓS
SZEREPE A FELSŐFOKÚ
MATEMATIKAOKTATÁSBAN**

Értekezés a doktori (Ph.D.) fokozat megszerzése érdekében
a matematika- és számítástudományok tudományágban

Írta: Papp Gabriella okleveles matematikus, matematika és informatika tanár

Készült a Debreceni Egyetem Matematika- és Számítástudományok doktori iskolája
(Didaktika (Szak módszertan) programja) keretében

Témavezető: Dr. habil. Korenova Lilla

Az értekezés bírálói:

Dr.
Dr.

A bírálóbizottság:

elnök:	Dr.
tagok:	Dr.
	Dr.
	Dr.
	Dr.

Az értekezés védésének időpontja: 20... ..

Köszönetnyilvánítás

Köszönöm témavezetőmnek, Dr. habil. Korenova Lilla tanárnőnek, hogy az életre, önálló munkára tanítva elősegítette szakmai fejlődésem.

Köszönettel tartozom a II. Rákóczi Ferenc Kárpátaljai Magyar Egyetem vezetőségének, akik engedélyezték és támogatták a kutatást a szükséges csoportokban. Valamint a Matematika és Informatika Tanszék korábbi tanszékvezetőjének, Dr. Kucsinka Katalinnak és további jelenlegi és volt munkatársaknak, akik munkám során támogattak és mellettem álltak.

Köszönöm RNDr. Fehér Zoltánnak, a komáromi Selye János Egyetem (Gazdaságtudományi és Informatikai Kar, Matematika Tanszék) adjunktusának a segítőkészségét, hogy első megkeresésre elvállalta az elkészült tesztek megíratását és lelkiismeretesen véghezvitte azt.

Köszönet illeti a Matematika- és Számítástudományok Doktori Iskola oktatóit a tanulmányaim alatt nyújtott szakmai támogatásért. Külön köszönet jár Herendiné Dr. Kónya Eszter egyetemi docensnek, akihez minden felmerülő kérdéssel fordulhattam.

Hálás vagyok családomnak és barátaimnak, akik végig mellettem álltak, bíztattak, és a legnehezebb időszakban sem hagyták, hogy feladjam.

Hála mindenért Istennek, aki nemcsak a tanulás vágyát ültette el bennem, hanem minden boldog percben és nehézségben velem volt, kirendelve a szükséges emberek mellett a kellő időt, erőt és energiát a munka elvégzéséhez.

Köszönöm!

Tartalomjegyzék

BEVEZETÉS.....	7
1. A TÉMAVÁLASZTÁS INDOKLÁSA.....	8
2. KUTATÁSI KÉRDÉSEK	9
1. IRODALMI ÁTTEKINTÉS.....	11
1.1. TESZTELMÉLET.....	11
1.1.1. <i>Tesztek előkészítése taxonómia segítségével.....</i>	<i>15</i>
1.1.2. <i>A tesztek jószágmutatói</i>	<i>19</i>
1.1.3. <i>Itemtípusok és a feladatírás fontosabb szabályai</i>	<i>24</i>
1.2. E-TESZTEKKEL TÁMOGATOTT OKTATÁS	29
1.2.1. <i>Előhívásos tanulás</i>	<i>32</i>
1.2.2. <i>Motiváció a matematikaoktatásban</i>	<i>35</i>
2. A KUTATÁS MÓDSZERTANA	39
2.1.1. <i>Az órák megtervezése</i>	<i>39</i>
2.1.2. <i>A minta.....</i>	<i>40</i>
3. ELŐZETES KUTATÁSOK	42
3.1. MUNKAFELÜLET MEGVÁLASZTÁSÁRA VONATKOZÓ FELMÉRÉSEK.....	42
3.2. E-TESZTEK SZERKESZTÉSE ÉS MUNKAFELÜLETEK ÖSSZEHASONLÍTÁSA.....	43
3.3. E-TESZT SZERKESZTÉSE ÉS ALKALMAZÁSA A FELSŐFOKÚ MATEMATIKAOKTATÁSBAN (1. ÉS 2. CIKLUS)	47
3.4. E-TESZT SZERKESZTÉSE ÉS ALKALMAZÁSA A FELSŐFOKÚ MATEMATIKAOKTATÁSBAN (3. ÉS 4. CIKLUS)	54
4. EMPIRIKUS KUTATÁS.....	60
4.1. E-TESZTEK LÉTREHOZÁSA	60
4.2. E-TESZT, MINT MÉRŐESZKÖZ	64
4.3. A PONTSZERZÉS MINT KÜLSŐ MOTIVÁCIÓS DOMINANCIA	74
KÖVETKEZTETÉSEK, TOVÁBBI KUTATÁSI LEHETŐSÉGEK.....	79
FELHASZNÁLT IRODALOM.....	82
ÁBRÁK ÉS TÁBLÁZATOK JEGYZÉKE	91
ÖSSZEFOGLALÁS.....	92
SUMMARY	97
FÜGGELÉK.....	102
A. BEMENETI E-TESZT	102
B. PERMUTÁCIÓ	107

C.	VARIÁCIÓ	108
D.	KOMBINÁCIÓ	109
E.	KIMENETI E-TESZT.....	110
F.	KÉRDŐÍV	115

BEVEZETÉS

A matematika tanulása számos diák számára jelent komoly kihívást, ami gyakran vezet ahhoz az elterjedt vélekedéshez, miszerint a matematikára születni kell. A pedagógus szemszögéből nézve a legfontosabb cél nem csupán az, hogy a diákok megtanulják a tananyagban szereplő tételeket és algoritmusokat, hanem hogy megértsék azok mélyebb összefüggéseit, felismerjék a matematika kreatív és logikai jellegét. Úgy vélem, a pedagógus feladata az, hogy bebizonyítsa, a matematika megértése és megszeretése sokkal inkább egy fejleszthető folyamat eredménye, amely mindenki számára nyitva áll.

Nem tudhatjuk mikor, vagy mitől válunk igazán jó tanárrá, de a fejlesztő folyamat részeként bizonyosan mindenki arra törekszik, hogy legjobb tudása szerint tudja oktatni és motiválni az általa tanított diákokat. Ennek érdekében választjuk meg a tanítási-tanulási technikákat, a motiváció és a számonkérés lehetőségeit.

A pedagógiai értékelés egyik legelterjedtebb formája, sajátos területe a tudásszintmérés (Csapó, 2004). A tanulmányi munka értékelését a kezdetektől fogva folyamatos kritika éri (Székely, 2014). Az, hogy mi a fontos és értékes számunkra, attól a környezettől (kultúrától) függ, amelyben nevelkedtünk (Hercz, 2007).

Csapó Benő (1988) szerint az értékelés a pedagógiai folyamatok egyik alapvető szakasza, a visszacsatolás funkcióját tölti be (Csapó, 1988). Az egyik nehéz feladat az oktatás folyamatában ennek a funkciónak minél pontosabb és megbízhatóbb vizsgálata, a megszerzett tudás mérése. Többségében ezt hagyományos, órai kisebb felmérőkkel és zárótesztekkel végeztük és tesszük ezt még mindig. Ellenben a diákok – minden korosztályban – egyre nehezebben motiválhatók az órákon, egyre kevesebb tudásra szeretnének szert tenni. Ezzel a munkámmal a pedagógusoknak szeretnék segítséget nyújtani az e-tesztel segítt tanítást és az ehhez szükséges tesztkérdések előkészítésének menetét kutatva.

Az 1. fejezetben háttérkutatás segítségével, a teljesség igénye nélkül mutatom be a fontosabb tudnivalókat a tesztelmélettel és motivációval kapcsolatban. A 2. fejezetben a kutatás módszertanát, a tervezési folyamatot és a kutatási mintát mutatom be. A 3. fejezetben ismertetem az előzetes kutatásokat, alkalmazott módszereket és a tesztek leírását. A 4. fejezetben az empirikus kutatás eredményeinek bemutatásával adok választ a kutatási kérdések segítségével a hipotézisekre.

1. A témaválasztás indoklása

A pedagógusok különféle módszertani eszközökkel, változatos és gondolkodtató feladatok segítségével próbálják befogadhatóvá tenni a tanulást a diákok számára. Az e-tesztek mindennapi alkalmazása az utóbbi idők nehézségei miatt váltak általánossá, de felmerül a kérdés, valóban megfelelően mérnek-e, kellő tesztelméleti tudással készítjük-e őket? A tesztekkel kapcsolatos kérdések már korábban kezdtek foglalkoztatni, de a nehézségek tették lehetővé a könnyebb adaptálódást.

Ukrajnában az általános iskolát (2021-től hivatalosan Gimnázium) befejezve 2002-től minden 9. osztályt végzett diák számára kötelező az Állami záróvizsga (Державна підсумкова атестація, utóbbi években felmerült problémák miatt eltörölték), a középiskolát (2021-től hivatalosan Líceum) befejezve 2004-től választható, majd 2008-tól kötelező a Külső független értékelés (Зовнішнє незалежне оцінювання). A diákok jelentős része feltehetően a felkészülés folyamán és az éles vizsgákon találkozott a tudásszint mérésével független, nyílt és zárt kérdésekből álló tesztekkel. Bár itt még mindkét esetben papír alapú tesztekről beszélünk, felépítésük is hasonló volt, első részben a felsorolt válaszlehetőség közül kellett megjelölni a helyeset, második részben a kapott eredményt kellett a megfelelő helyen feltüntetni, érettségi esetén megjelent a párosítás, a harmadik részben pedig a kapott feladatok részletes megoldását kellett magyarázatokkal együtt leírni. Nem állítom, hogy a tanárok által választott technikák és tudásmérő tesztek nem készítik fel diákjaikat a vizsgára, de felmerült a kérdés, a „vizsgadruk” mellett mi okozhatja sokszor a matematikában jártas diákok esetén is a rosszabb eredményt?

Az érthető okok miatt 2022-ben a külső független értékelés Nemzeti többtárgyas teszt (Національний мультипредметний тест) néven online térbe került, aminek alkalmazásához nem voltak és véleményem szerint ma sincsenek hozzá szokva az érettségizők. Könnyen jelölnek részmegoldásokat, ahogy a válaszlehetőségek között látnak egy kapott eredményt. Ezt figyelmetlenségnek is betudhatjuk, de véleményem szerint azért is fordul elő nagy számban, mert nincsenek hozzá szokva a jól kidolgozott tesztekhez, melyek rossz válaszai (disztraktorok) a hibás gondolatmenettel vagy részeredményként kapott értékkel vannak beállítva. Úgy vélem, ezt lehetne valamelyest kiküszöbölni olyan oktatási módszerrel, ahol több e-teszttel találkoznak a diákok.

Munkám célja tehát egy olyan módszertani útmutató összeállítása, mely összefoglalja és kiemelni a fontosabb elemeket és tudnivalókat, amelyek segítséget nyújtanak azon tanárok számára, akik fejleszteni szeretnék a tesztkészítési és értékelési gyakorlatukat. Ennek érdekében az alapos háttérkutatót követően egy olyan mérőeszközt

hoztam létre, mely alkalmas a célcsoport tudásszintjének mérésére. A munka során figyelmet fordítottam a tesztelmélet alapvető követelményeire, a fejlesztés során kiemelten törekedtem arra, hogy erős jóságmutatókkal rendelkező mérőeszközt alakítsak ki. A módszertani útmutatóban az e-teszt sikerességének kiértékelését mutatom be.

További célom volt kérdőíves vizsgálat segítségével feltérképezni a kutatásban részt vevő hallgatók véleményét arról, hogy mennyire tartják motiválónak a tesztelést az oktatásban. Úgy gondolom, hogy a tesztelés nem csupán az ismeretek számonkérésére alkalmas, hanem megfelelően kidolgozott formában a tanulási folyamat szerves részévé válhat, így segítve a tanulók önértékelését és fejlődését, kiküszöbölve a független megméréstéseken felmerülő problémákat.

Munkám során fontosnak tartottam, hogy saját magam próbáljam ki a tesztek és tesztekkel társított oktatás hatékonyságát. Így témavezetőmmel a felsőoktatásba helyeztük a kutatást annak ellenére, hogy a téma iránti érdeklődést az alacsonyabb szintű oktatás fejtette ki.

A tárgyválasztás esetén arra törekedtem, hogy több irányban készítsek megbízható teszteket, közben olyan témát dolgozzak fel, mellyel már találkoztak a felsőoktatás előtt. Így esett a választás a Matematikai analízis és Informatika tárgyak vizsgatesztjei mellett a fő kutatás szempontjából a Valószínűségi számítás előfutáraként a Kombinatorika témakörre, mely még elég általános ahhoz, hogy a nem matematika szakos hallgatóknak is szükségük legyen rá; elég könnyű ahhoz, hogy a tesztek szerkesztésekor megfelelően tudjunk eljárni a tudásmérés szempontjából; és elég nehéz ahhoz, hogy úgy tekintsenek rá a hallgatók, mint amit nehéz értelmezni és megtanulni.

2. Kutatási kérdések

A háttérinformációk felkutatása és megismerése hasznosnak bizonyult, mégis úgy tűnt, minél többet kutatok, annál több kérdés merül fel, melyek megválaszolásában csak az segíthet, ha saját magamat is teszteljem, hogyan és milyen e-tesztet tudok létrehozni, s azokkal milyen eredményeket érek el. A szakirodalomban (Anderson & Krathwohl, 2001; Csapó, 1988; Demkanin et al., 2015; Einhorn, 2012; Hercz, 2007) fellelhető információkra támaszkodva teszteket hoztam létre. Egyrészt tudásmérő e-tesztet, másrészt az egyes részek elmélyítésére alkalmazható rövidebb, mini előhívó e-teszteket, melyek motivációs hatását is figyelemmel akartam kísérni.

Mindezt összefoglalva fő kutatási kérdésemet a következőképpen fogalmaztam meg:

Mennyire alkalmas a Bloom-taxonómia kognitív szintjeire épülő e-teszt a hallgatók differenciált tudásszintjének objektív mérésére a felsőfokú matematikaoktatásban — a jelenlegi reliabilitási mutatók tükrében —, és e mérési folyamatban miként befolyásolja a pontszerzési lehetőség az e-tesztek motivációs hatékonyságát?

Minden tanár tudja, hogy a tesztelés messze nem az egyetlen módja annak, hogy megismerjük a tanulót, és képet alkossunk a személyiségéről. A tanár minden beszélgetés során, a tanórai és tanórán kívüli viselkedés megfigyelésével, valamint a tanuló által készített különböző produktumok vizsgálatával szerez ismereteket a tanulóról (Demkanin et al., 2015). Emiatt a fő kérdés megválaszolása érdekében kvalitatív és kvantitatív kutatásokat egyaránt alkalmaztam, először a következő hipotézisekre és hozzájuk tartozó kérdésekre keresem a választ:

- H1.** Az e-tesztekben a hallgatók teljesítménye szignifikánsan romlik a Bloom-taxonómia magasabb kognitív szintjeit (pl. alkalmazás, elemzés) igénylő feladatoknál, függetlenül az előzetesen meghatározott nehézségi szinttől, valamint a mini tesztek kitöltésétől.
- K1.** Mely kognitív szinthez tartozó feladatok bizonyultak a legnehezebbnek a kombinatorika tesztben?
- K2.** Összhangban van-e a tanári előzetes nehézségi becslés a hallgatói eredményekkel (nehézségi index)?
- K3.** Milyen hatással van az előhívó e-teszt egy tudásmérő tesztre?
- H2.** A hallgatók számára az e-tesztek motivációs erejét elsősorban nem az azonnali visszacsatolás (önértékelés), hanem a végső értékelésbe való beszámíthatóság (pontgyűjtés) adja.
- K1.** Mennyire tudatos a válaszok megjelölése a tesztek kitöltése során?
- K2.** Mely funkciót tartják a hallgatók az e-tesztek legfőbb előnyének: a tanulást segítő visszajelzést vagy a jegyszerzés megkönnyítését?
- K3.** Van-e összefüggés a hallgatók motivációs szintje és a pontgyűjtési lehetőség között?

Ezekre a kérdésekre az irodalomkutatás mellett az elkészült tesztek eredményeinek statisztikai kiértékelése során szeretnék választ találni. Továbbá a tesztekéről és azok motivációs hatásáról egy kérdőív segítségével kutattam a hallgatói véleményt, melyet megvizsgáltam és következtetéseket vontam le.

1. IRODALMI ÁTTEKINTÉS

Az oktatási folyamatban különböző tesztek alkalmaznak a differenciálás és tudásszint mérése szempontjából. A matematikában jellemzően a hagyományos tesztek (nyitott feladatok) játszanak nagyobb szerepet, ahol a megoldás teljes leírása segíti az oktatót az adott témában megszerzett tudás felmérésére. Nagyobb létszám vagy távoktatás esetén, továbbá a tanulók motiválása érdekében egyre több oktató tért át az e-tesztek alkalmazására, ami időigényes és nehéz feladat, amennyiben a szükséges háttérinformációt is firtessen szerzi az oktató.

Ha felkészületlenül kezdünk e-feladatlap létrehozásába, könnyen eshetünk olyan hibába, hogy nem éri el a mérés célját, többek között a szükséges differenciáló hatást, vagy nem lesz megbízható. A dolgozatban elsősorban arra szeretném felhívni a figyelmet a megbízható feladatlap készítés és tudásszint mérés érdekében milyen meghatározó szakirodalmi információk ismeretében kell dolgoznia az oktátónak.

A klasszikus tesztelmélettel szemben a modern (valószínűségi, probabilisztikus) tesztelmélet, amely a tesztelmélet új generációja, valószínűségi eszközöket használ az itemek (tesztek legkisebb részei, amelyek önállóan pontozhatók) tulajdonságainak jellemzésére (Molnár, 2003). A két elmélet általában két nagyon különböző mérési keretrendszer képvisel (Eluwa et al., 2011). Az egyes itemek tulajdonságait vizsgáló modellek különböző paraméterek szerint csoportosíthatók. Eltérhetnek egymástól abban, hogy milyen típusú összefüggést feltételeznek a helyes válasz valószínűsége és a válaszoló képessége között (Molnár, 2003).

A kutatás ezen szakaszában a klasszikus tesztelmélet fundamentális téziseit, valamint a strukturált mérőeszközök fejlesztésének módszertani kritériumait rendszerezem. Ezt követően a információkutatás az elektronikus környezetbe történő adaptáció kérdésköre felé orientálódik.

1.1. Tesztelmélet

A legtöbb oktatási jelenség úgynevezett konstrukció, azaz olyan elvont fogalmak, amelyek az érzékszervekkel nem megragadhatók, és amelyekről utólag alkotunk képet a fejünkben. Ezek olyan fogalmak, mint a motiváció, jellem, figyelem, ízlés, megfontoltság stb (Demkanin et al., 2015). A teszt a tudományos kutatás alapja lehet, és számos gyakorlati területen alkalmazható (Hidegkuti & Balázs, 2015). A legelső és sokáig a

leggyakrabban használt tesztek pszichológusok fejlesztették ki, elsősorban az intelligencia mérésére (Csíkos & B. Német, 2002). Közös eredetük és párhuzamos fejlődésük miatt nehéz lenne olyan definíciót alkotni, amely fogalmilag egyértelműen elkülöníti a pszichológiai tesztek és a tudásmérőket (Csapó, 2004). A pedagógiai értékelés folyamatában a tudás értelmezése és meghatározása nagy jelentőséggel bír (Székely, 2014).

A tesztelmélet egy olyan gondolatrendszer, amely kifejezetten a mérési problémák kiküszöbölésére és csökkentésére szolgál az oktatásban és a társadalomtudományokban (Hidegkuti & Balázs, 2015). Az osztályozás körülbelül a 17. század óta része a tanításnak (Demkanin et al., 2015). A századforduló után megjelentek az értékelés eltörlésére vonatkozó nézetek, valamint olyan irányzatok, amelyek az értékelés objektívebbé tétele érdekében a tantárgytesztek, tantárgyi felmérőlapok kidolgozását, alkalmazását szorgalmazták (Székely, 2014).

A teszteléselméletek matematikai, főként valószínűségi és statisztikai összefüggésrendszereket (Csapó, 2004) használnak a mérési eljárások megtervezésére, a mérési eredmények meghatározására és a teszteredmények modellezésére (Hidegkuti & Balázs, 2015). A mérési módszerek ismertek voltak, de a hivatalos oktatásirányítás nem tartotta szükségesnek a tanulói eredmények objektív mérését (Székely, 2014).

Székely (2014) a Pedagógiai Lexikon meghatározása alapján az értékelést egy jelenség, esemény vagy tárgy értékének szisztematikus vizsgálatára vagy megítélésére irányuló folyamatként határozza meg (Székely, 2014). A háttérkutatás során többféle megfogalmazásban található meg a teszt fogalma:

- A teszt (eredeti értelemben próba) tipikus mérőeszköz a pedagógiában. A pedagógiai és pszichológiai terminológiában egy vizsgálandó tulajdonság kipróbálással történő felmérését jelenti (Csapó et al, 2008).
- Teszt olyan szabványosított eljárások leírására használt kifejezés, amelyek segítségével bizonyos viselkedési formákat leírhatunk vagy mérhetünk, hogy előre meghatározott módon kategóriákat vagy pontszámokat rendeljünk az egyénekhez vagy eseményekhez (Hidegkuti & Balázs, 2015).
- A teszt olyan mérőeszköz, amely a mérendő pszichikus jelenséget megfelelő skálán méri (Székely, 2014).
- A tudásszintmérő tesztek azt vizsgálják, azt mérik, hogy a tanuló mire képes, ha összes tudását összegezzük (Csapó, 1988).

- A didaktikai teszt egy vizsgatípus, melyben a tanuló általában meghatározott módon válaszol (zárt feladatok strukturált válasszal), és csak az esetek kisebb részében alkotja meg a választ (nyitott feladatok) (Demkanin et al., 2015).

A definíciók közös hangsúlya a mérés, miszerint a tanulók teljesítménye és képességei mérhetővé, összehasonlíthatóvá és számszerűsíthetővé válik. Munkám során ehhez a méréshez szeretnék útmutatást adni a tesztek szerkesztésének leírása és sikerességének kiértékelése révén.

Az értékelés alapvető kérdéseire világszerte különböző válaszokat adnak. A legfontosabb alapkérdések a következők: Mi az értékelés célja? Miért (értékelünk)? Ezeket számos más gyakorlati kérdés követi: Ki? Kik? Hol? Mikor? Hogyan? Mivel (értékel)? (Hercz, 2007; Fűzné, 2016).

A pedagógiai értékelés során a tanuló megfigyelt viselkedését és produktumát egy mérőeszközzel előre meghatározott kritériumok és viszonyítási pontok alapján egy skálához hasonlítják, hogy számszerűsítsék az elért eredményt (Molnár & Vígh, 2013). A pedagógiai értékelés egyik leggyakoribb formája, sajátos területe a tudásszintmérés. A tudásszintmérés elmélete a tesztelés pszichológiai elméletéből fejlődött ki. A tudásszintmérő tesztek pszichológiai tulajdonságokat, a tudás egy sajátos formáját is mérik. Általában az iskolai oktatásban szerepet játszó tudást (Csapó, 2004).

Az oktatási jelenségek mérése a mért jelenségek számszerűsítéseként vagy kvalitatív kutatási eljárásokkal történik. Lehetővé teszi a hosszabb, szóbeli vagy írásbeli elemzést és az eredmények értékelését mindenféle korlátozás nélkül (Csapó et al, 2008). Az oktatási jelenségek mennyiségi (kvantitatív) jellemzőire alkalmazott mérési eljárások sokkal gyakoribbak és hatékonyabbak (Demkanin et al., 2015), amelyek során a vizsgált teljesítményhez valamilyen számszerű értéket rendelnek, amely kifejezi a teljesítmény színvonalát és értékét (Csapó et al, 2008). A kvantitatív módszerek célja a társadalmi jelenségek szisztematikus vizsgálata statisztikai modellek és matematikai eszközök segítségével. Ezek a módszerek bizonyos értelemben "objektív" méréseket, strukturált adatgyűjtést alkalmaznak (Hajdu, 2018), valamint a tesztelmélet területén jönnek létre és kerülnek értékelésre (Korenova, 2013).

A pedagógiai értékelés szűkebb értelemben a tanítás és a tanulás módszere (Hercz, 2007). Az értékelés a mindenkire kiterjedő megerősítési és visszacsatolási folyamata (Parázso, 2012), amely a cél és az elért hatás közötti kapcsolatra, a tanítás és tanulás különböző szintjeinek és elemeinek visszacsatolására és értékelésére épül (Hercz, 2007). Az oktatási módszerként alkalmazott pedagógiai értékelés, mint tanítási módszer

információt nyújt a kitűzött célok elérésének szintjéről (Székely, 2014), mely lehetővé teszi, hogy ne csak a tanulók tevékenységét, hanem a teljes tanítási-tanulási folyamatot, annak hatékonyságát értékeljük, beleértve ennek a folyamatnak minden tényezőjét (Parázsó, 2012).

A szakirodalmi kutatások alapján három értékelési módot különböztetünk meg: formatív, szummatív és diagnosztikus (Fűzné, 2016). Bár a formatív tesztelés eredeti rendeltetése szerint a tanulás támogatását, a folyamatos visszacsatolást és a hallgatói fejlődés elősegítését szolgálná, a gyakorlatban a szummatív tesztelés gyakran az elszámoltathatóságra és az eredmények pusztá összegzésére fókuszál (Usiskin, 2019).

A tesztek leggyakrabban szummatív értékelési eszközökként használják, amelyek célja annak felmérése, hogy a hallgatók elérték-e a kurzus tanulási céljait (Brame & Biel, 2015). Ekkor az elsődleges cél a teszten nyújtott teljesítmény értékelése, ezért az eredményeket általában az altesztek és tesztek szintjén elemzik, nem pedig az itemek és feladatok szintjén (Molnár & Vigh, 2013).

A dolgozatban bemutatott kutatás módszertani keretét a formatív és szummatív értékelési funkciók szisztematikus integrációja és ezek hatékonyságvizsgálata alkotja. A dolgozat kritikai elemzése arra világít rá, hogy a két értékelési mód együttes alkalmazása átfogó képet ad a felsőfokú matematikai ismeretek elsajátításának dinamikájáról.

Az oktatási diagnosztikát a (1) kontextus, (2) a fejlődés és (3) a tanulási eredmények felismerésének elméletének és gyakorlatának tekintjük. Ez a három cél egyformán fontos (Niemierko, 2012). A diagnosztikai folyamatok funkciói a következők: tesztelés, visszajelzés, osztályozás, ellenőrzés, motiválás és oktatás (Horovčák & Stehlíková, 2007). A diagnosztikus (helyzetfeltáró) mérések és tesztek a tanítási-tanulási folyamat bármely szakaszában alkalmazhatók (Molnár & Vigh, 2013; Csapó, 2004) a tudás tartalmának és szerkezetének tükrözésére annak érdekében, hogy:

- az eredmények és a hiányosságok okai azonosítható legyen (Fűzné, 2016);
- a vizsgált ismeretkör minél teljesebben lefedett legyen feladatokkal;
- lépésről lépésre meghatározni, hogy a tanulók mit tudnak már és mit nem tudnak még (Molnár & Vigh, 2013).

Ezt a követelményt a tényleges tanulási eredmények elemzésével lehet teljesíteni. Az értékelés egyik kulcskérdése az egyénre szabott értékelési formák bevezetése, amely a legmodernebb formában, az online értékelésben valósulhat meg (Parázsó, 2012). A lezáró jellegű méréseknél a digitális platformok az objektivitást növelik. A szoftveres javítás kiküszöböli az emberi szubjektivitást és a figyelmetlenségből eredő tévedéseket.

A szummatív eredményekből kinyerhető statisztikai mutatók (pl. átlag, szórás, varianciaanalízis) pontos képet adnak a hallgatók közötti differenciáról (Parázsó, 2020; Négyesi, 2021).

1.1.1. Tesztek előkészítése taxonómia segítségével

Egy megfontoltan előkészített teszt esetén érdemes egy osztályozást figyelembe venni a feladatokra tekintve. A taxonómia egy speciális keretrendszer, melyben a kategóriák egy kontinuum mentén helyezkednek el (Anderson & Krathwohl, 2001). Mindegyik taxonómia tulajdonképpen csak egy hierarchikusan elrendezett osztályozás (Demkanin et al., 2015). Az oktatási célok taxonómiái, amelyeket a 20. század közepén a kognitív területen kezdeményeztek, még mindig kihívást jelentenek a tanulási és tanítási gyakorlatok számára az oktatás minden szintjén (Niemierko, 2012).

A nevelési céloknak többféle taxonómiája létezik, és a történelemből ismerjük Herbart négy formális nevelési szakaszát, amely ihlette B. S. Bloomot:

- Az elmélyülés szakasza:

1. az előzetes tudás tisztázása
2. új ismeretek befogadása - összeolvadás

- Reflexiós szakasz:

3. az új ismeretek integrálása a meglévő ismeretek rendszerébe
4. az új ismeretek alkalmazása - gyakorló módszer (Demkanin et al., 2015).

A tesztelméletben széles körben használják a tanulási célok (a megfelelő teszt) alapján történő osztályozást is (Sokolova & Totkov, 2005). Anderson & Krathwohl (2001) kiemelik, hogy az általuk meghatározott taxonómiában a célokat osztályozzák: Egy célmeghatározás tartalmaz egy igét és egy főnevet. Az ige általában a kívánt kognitív folyamatot írja le. A főnév általában azt a tudást írja le, amelyet a diákoktól elvárnak, hogy elsajátítsanak vagy felépítsenek (Anderson & Krathwohl, 2001).

M. D. Gall táblázatos áttekintést ad az olyan teoretikusok taxonómiáiról, mint Adams, Aschner, Bloom, Bremer, Carner, Clements, Schreiber, Guszak, Pate stb., akiknek taxonómiái különböző módon átfedik és eltérnek egymástól (Demkanin et al., 2015). A kognitív területre vonatkozó taxonómia kapta a legnagyobb figyelmet a tesztkonstruktorok részéről, mivel ez áll a legrégebben rendelkezésre, és mivel ez írja le azokat a képességeket, amelyek mérése a tesztkonstruktorokat a leginkább érdekli (Ebel & Frisbie, 1991).

Jelenleg Niemierko és Bloom taxonómiája a legismertebb. Bloom taxonómiája a különböző célok és készségek egy osztályozása, amelyeket az oktatók tanítványaik számára kitűznek (tanulási célok). A rendszert 1956-ban javasolta Benjamin Bloom, a Chicagói Egyetem oktatási pszichológusa (Shabatura, 2022).

Bloom eredeti kategóriái olyan kognitív területekhez kapcsolódnak, mint: Az ismeret, a megértés, az alkalmazás, az elemzés, a szintézis és az értékelés (Armstrong, 2010; Demkanin et al., 2015; Sokolova & Totkov, 2005; Shabatura, 2022). Ekkor a kognitív területen a hatszintű Bloom-taxonómiát úgy sűrítették össze, hogy a három legmagasabb szintet ("elemzés", "szintézis" és "értékelés") a gondolkodási folyamatok egyetlen problémamegoldó kategóriájába foglalták (Niemierko, 2012). E taxonómia egyik fő hozzájárulása az volt, hogy tudatosította, milyen intellektuális szinten írják meg a tanítási célokat és a tesztelési feladatokat (Ebel & Frisbie, 1991). Niemierko taxonómiája az összes magasabb szintű logikai-kognitív folyamatot egyetlen kategóriába (nem specifikus transzfer) foglalja össze (Demkanin et al., 2015).

Kognitív pszichológusok, tantervelméleti és oktatói kutatók, valamint tesztelési és értékelési szakemberek csoportja 2001-ben Bloom taxonómiájának felülvizsgálatát publikálta „A Taxonomy for Teaching, Learning, and Assessment” címmel (Armstrong, 2010). Az időtálló (már átdolgozott) Bloom-taxonómia tűnik a legígéretesebbnek (Demkanin et al., 2015), ami állandó tudományos kritika alatt maradt, különösen viselkedéseméleti eredete miatt, és a szinte általános használat dicsőségére (Niemierko, 2012). A 1. táblázat a taxonómia felülvizsgált kétdimenziós táblázatát mutatja.

1. táblázat. A felülvizsgált Bloom-taxonómia összefoglaló táblázata (Anderson & Krathwohl, 2001; Demkanin et al., 2015).

			A kognitív folyamat dimenziója					
			1	2	3	4	5	6
			Emlékezni	Megérteni	Alkalmazni	Elemezni	Értékelni	Létrehozni
A tudás dimenziója	A	Tárgyi tudás						
	B	Fogalmi tudás						
	C	Eljárási tudás						
	D	Metakognitív tudás						

A fenti táblázatban szereplő elemeket a következőképpen részletezi Anderson & Krathwohl (2001), Armstrong (2010), Demkanin és társai (2015), valamint Shabatura (2022):

1. Emlékezni (Remember): A releváns ismeretek előhívása, felismerése és felidézése a hosszú távú memóriából (Shabatura, 2022).
2. Megérteni (Understand): A jelentés konstruálása szóbeli, írott és grafikus üzenetekből (Anderson & Krathwohl, 2001) értelmezés, példamutatás, osztályozás, összefoglalás, következtetés, összehasonlítás és magyarázat útján (Shabatura, 2022).
3. Alkalmazni (Apply): Utasítások követése vagy egy eljárás alkalmazása megszokott vagy ismeretlen feladattípusokban (Demkanin et al., 2015).
4. Elemezni (Analyze): Az anyagot alkotóelemeire bontani, és meghatározni, hogy a részek hogyan kapcsolódnak egymáshoz és egy átfogó szerkezethez vagy célhoz (Anderson & Krathwohl, 2001), az egészhez vagy az anyag fő funkciójához (Demkanin et al., 2015) differenciálás, rendszerezés és hozzárendelés révén (Shabatura, 2022).
5. Értékelni (Evaluate): Kritériumok és szabványok alapján ítéletek meghozása ellenőrzés és kritikák útján (Shabatura, 2022).
6. Létrehozni (Create): A részek összekapcsolásával új, érdekes, belsőleg összetartozó egységek létrehozása (Demkanin et al., 2015), egy koherens vagy funkcionális egész kialakítása; az elemek új mintává vagy struktúrává történő átszervezése generálás, tervezés vagy gyártás útján (Shabatura, 2022).

A felülvizsgált taxonómiában a tudás ennek a hat kognitív folyamatnak az alapja, de a fent említett szakirodalmak külön kiemelik a megismerésben használt tudás dimenzióinak jellemzőit:

- A. Tárgyi tudás: Az alapvető elemek, amelyeket a diákoknak tudniuk kell ahhoz, hogy megismerkedjenek egy tudománnyal vagy problémákat oldjanak meg benne (Anderson & Krathwohl, 2001).
 - a) A terminológia ismerete (fogalmak, meghatározások);
 - b) Speciális részletek és elemek ismerete (Anderson & Krathwohl, 2001; Armstrong, 2010; Demkanin et al., 2015).
- B. Fogalmi tudás: Az alapelemek közötti kapcsolatok egy nagyobb struktúrán belül, amelyek lehetővé teszik, hogy együttesen működjenek (Anderson & Krathwohl, 2001; Demkanin et al., 2015).

- a) Osztályozások és kategóriák ismerete (tények, fogalmak, eszmék, értékek csoportokba sorolása, osztályok, fajták stb.);
 - b) Alapelvek és általánosítások ismerete (kapcsolatok, minták, viszonyok, összefüggések, területek és kontextusok);
 - c) Elméletek, modellek és struktúrák ismerete (elvek, hipotézisek, sémák) (Anderson & Krathwohl, 2001; Armstrong, 2010; Demkanin et al., 2015).
- C. Eljárási tudás: Konkrét eljárások, módszerek, algoritmusok, technikák ismerete, készségek alkalmazása egy kérdés/feladat/probléma megoldása során (Demkanin et al., 2015).
- a) Tantárgy-specifikus készségek és algoritmusok ismerete
 - b) Tantárgy-specifikus technikák és módszerek ismerete
 - c) A megfelelő eljárások alkalmazásához szükséges kritériumok ismerete (Anderson & Krathwohl, 2001; Armstrong, 2010; Demkanin et al., 2015).
- D. Metakognitív tudás: A megismerés általános ismerete, valamint a saját megismerés tudatossága és ismerete (Anderson & Krathwohl, 2001).
- a) Stratégiai tudás (a tanulás általános módszerei, a tanulási módszerek gyors olvasás és információkeresés, munka adatbázisokkal és nyilvántartásokkal, heurisztikák, problémamegoldási módszerek és stratégiák);
 - b) A kognitív feladatok ismerete, beleértve a megfelelő kontextusbeli és feltételes ismereteket (a feladatok vagy problémák megoldásának kognitív nehézségeinek ismerete, azok megoldásának háttere és módszere, beleértve a kontextust és a feltételeket);
 - c) Önismeret (saját erősségeinek és gyengeségeinek ismerete) (Anderson & Krathwohl, 2001; Armstrong, 2010; Demkanin et al., 2015).

Az átdolgozott taxonómia szerzői többrétegű választ javasolnak a „Miért használjuk Bloom taxonómiáját?” kérdésre, amelyet Armstrong (2010) néhány pontosítással egészített ki:

1. Fontos, hogy a pedagógiai megbeszélésekben célkitűzéseket (tanulási célokat) fogalmazzunk meg, hogy mind a tanárok, mind a diákok megértsék az adott eszmecsere célját.
2. A célok rendszerezése segít tisztázni a célokat saját maguk és a tanulók számára.
3. A szervezett célok megléte segít a tanároknak „megfelelő tanítást tervezni és végrehajtani”; „érvényes értékelési feladatokat és stratégiákat tervezni”; és

„biztosítani, hogy az oktatás és az értékelés összhangban legyen a célkitűzésekkel” (Anderson & Krathwohl, 2001; Armstrong, 2010).

Egy ideális értékelőeszköz megtervezéséhez a kétdimenziós taxonómiai mátrix széles körű lefedettsége szükséges. Ez a megközelítés lehetővé teszi a tanulók tudásának és specifikus kompetenciáinak pontos és differenciált mérését. A változatos osztályozás biztosítja, hogy a teszt ne csak az alapvető felidézést, hanem a magasabb szintű matematikai absztrakcióhoz elengedhetetlen komplex analitikai készségeket is értékelje. A kutatás során a felülvizsgált Bloom-taxonómiát alkalmaztam az általam létrehozott itemek rendszerezésére a tudás és a kognitív folyamatok mérésének vizsgálatában, ezekre 1. táblázat cellái alapján A1, A2, ..., D6 jelzésekkel hivatkozok az elemzések során.

1.1.2. A tesztek jóságmutatói

Minden mérőeszköz alkalmazásának célja a vizsgált tulajdonság pontos és megbízható mérése és objektív értékelése (Molnár & Vigh, 2013).

A tesztekkel és teszteléssel foglalkozó szakirodalomban a tesztek különböző jellemzőivel vagy tulajdonságaival találkozhatunk. Két csoportra oszthatjuk őket, amelyeken belül bizonyos hierarchia érvényesül:

- A tesztek alapvető vagy fő jellemzői (érvényesség, megbízhatóság),
- a tesztek egyéb jellemzői (nehézség, érzékenység, célszerűség, objektivitás) (Demkanin et al., 2015).

A klasszikus tesztelmélet azon a feltételezésen alapul, hogy a vizsgált személynek van egy megfigyelt pontszáma és egy valódi pontszáma. A megfigyelt pontszám az a tényleges eredmény, amelyet a vizsgázó egy adott teszten elér, két összetevőből áll össze: a valódi pontszámból és a mérési hibából. A valódi pontszám egy elméleti érték. Ez azt a pontszámot képviseli, amelyet a vizsgázó akkor érne el, ha a mérési körülmények tökéletesek lennének, és semmilyen külső vagy véletlenszerű tényező (hiba) nem befolyásolná az eredményt (Awopeju & Afolabi, 2016).

A szakirodalmak szerint a mérési hiba teljesítményben betöltött szerepének ellenőrzése, kontrollálása érdekében a mérőeszközöknek meg kell felelniük az objektivitás, a reliabilitás és a validitás jóságmutatóinak (Molnár & Vigh, 2013; Tóth, 2015). E három tulajdonságra koncentrálni hoztam létre a tesztek és vizsgáltam azokat különös figyelmet fordítva az elemzés során a megbízhatóságra.

A validitáson a teszt érvényességét értjük (Csapó, 2004), melynek egyik alapvető előfeltétele a részletes teszt specifikáció kidolgozása, amelynek egyértelmű tesztcélokat kell tartalmaznia (Demkanin et al., 2015). Amikor a tesztpontszámokra alkalmazzák, arra a konzisztenciára (pontosságra) utal, amellyel a pontszámok egy adott kognitív képességet mérnek (Ebel & Frisbie, 1991). A szakirodalmi kutatások szerint egy jó teszt akkor érvényes, ha tudjuk, mit akarunk mérni, hogyan mérjük, és ha a teszt valóban azt méri, amit mérni akarunk (Demkanin et al., 2015). Ezek alapján két aspektusa van: mit mérnek, és hogyan mérik az állandóságot (Ebel & Frisbie, 1991).

A tesztkészítési folyamat alapvető feladata a mérési cél, a követelmények és a teszt tartalmának meghatározása (Molnár & Víg, 2013). Ezután megvizsgáljuk azt a tulajdonságot, hogy a teszt valóban azt méri-e, aminek mérésére tervezték (Csapó, 2004). Ehhez elsősorban a tesztpontszámok felhasználói számára igazolni kell (1) a pontszámok értelmezését és (2) a pontszámok használatának megfelelőségét (Ebel & Frisbie, 1991).

Demkanin és munkatársai (2015) összefoglalják, hogy a meghatározás célja és módja szerint az érvényességet a következő típusokra osztjuk:

- *Tartalmi érvényesség* - a teszt tartalmának reprezentatív mintát kell képviselnie, azt jelzi, hogy a teszt tartalma mennyire van összhangban a tantervnek a pedagógiai dokumentumokban való bemutatásával. Néha ezt az érvényességet tantervi érvényességnek is nevezik (néha tantervi normativitásnak is);
- *Konstrukciós érvényesség* - az érvényességnek ez a formája azt jelzi, hogy a teszt milyen mértékben méri a tanulók egy adott tulajdonságát (a tanulók logikus gondolkodását, kreativitását stb.), vagyis azt, hogy a tesztben szereplő feladatok milyen mértékben tükrözik a teszt alapjául szolgáló elmélet mögöttes szempontjait;
- *Egyidejű érvényesség* - azt jelzi, hogy a teszt eredményei mennyire egyeznek egy másik teszttel, amely ugyanazt a dolgot vizsgálja;
- *Prediktív érvényesség* - a teszt azon képessége, hogy előre jelezzen egy jövőbeli eredményt;
- *Perceptuális érvényesség* (face validitás) - azt mutatja meg, hogy a tanárok és a diákok mit gondolnak a tesztről (a tesztet túl könnyű/nehez stb.) (Demkanin et al., 2015).

A teszt egyik legfontosabb tulajdonsága a megbízhatóság, idegen szóval reliabilitás (Csapó, 2004). A vizsgálat reliabilitása a mérés pontosságának és

megbízhatóságának mutatója (Demkanin et al., 2015), amely a mérési hibákat becsli csoportszinten (Molnár & Vigh, 2013).

A megbízhatósági együttható a pontsorozathoz kapcsolódó hiba mértékének mutatója, és így fontos információ a pontszámok értelmességének és hasznosságának értékeléséhez (Ebel & Frisbie, 1991). Minden egyes érték, amelyet egy tanuló a tesztben elér, a pontosság szempontjából két részre osztható:

- a helyes érték - azt kell tükröznie, amit a tanuló valóban tud,
- a hibás érték - az eredmény részét képezi, de nem tükrözi a tanuló tényleges tudását. A hibaértékek segítségével megtudhatjuk, hogy mit nem tud a tanuló, vagy mi a tanuló tanulási problémája (Demkanin et al., 2015).

A megbízhatóságot különböző módszerekkel lehet becsülni (Molnár & Vigh, 2013). Többek között egy univerzális mutató, a Cronbach-alfa (Molnár & Vigh, 2013) segítségével, ahol a varianciák összehasonlítását is felhasználjuk (Csapó, 2004). Cronbach 1951-ben publikálta a formulát:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_i V_i}{V_t} \right) \quad (\text{Cronbach, 1951; Panayides, 2013}). \quad (1)$$

Itt n az itemek számát, V_t a teljes pontszámok varianciáját, V_i pedig az item pontszámainak súlyozott varianciáját jelöli (Cronbach, 1951). Az alfa reprezentatív és átfogó képletként pozícionálható, nem csupán egy a sok megbízhatósági együttható közül (Cho & Kim, 2015).

A megbízhatóság számszerűsíthető egy 0 és 1 közötti számmal (Csíkos & B. Német, 2002), ahogy az eredményül kapott α megbízhatósági együttható 0 és 1 között mozog. Ideális esetben egy teszt megbízhatósága 1, ami azt jelenti, hogy ha egy ilyen tesztet ismételtelen elvégeznének, akkor minden tanuló minden egyes feladatnál ugyanolyan eredményt érne el, mint az első mérésnél. A gyakorlatban azonban az 1-es megbízhatóság nem fordul elő (Demkanin et al., 2015). Cortina (1993) különböző hipotetikus példákra keresztül bizonyította, hogy az alacsony item-interkorrelációk és a multidimenzionalitás ellenére is lehet magas (0,7-nél nagyobb) az alfa, feltéve, hogy elegendő számú item áll rendelkezésre (Cortina, 1993; Panayides, 2013).

Bár a „jó” α együttható előállítására vonatkozó szabványok teljesen önkényesek, és függenek a kérdéses skála elméleti ismereteitől, sok módszertanos 0,65 és 0,85 között javasolja a minimális α együtthatót (vagy sok esetben magasabbat); A 0,5-nél kisebb α együtthatók általában elfogadhatatlanok (Ebel & Frisbie, 1991; Demkanin et al., 2015; Goforth, 2015). Siswaningsih és társai egy táblázatban (lásd 2. táblázat) foglalja össze, a

Cronbach-alfa mely értéknél milyen osztályozást mutat. Kutatásom során ez alapján osztályoztam a létrehozott és alkalmazott tesztek megbízhatóságát.

2. táblázat. Cronbach-alfa osztályozása (Siswaningsih et al., 2017).

Cronbach-alfa	Osztályozás
$\alpha \geq 0,9$	Nagyon jó
$0,8 \leq \alpha < 0,9$	Jó
$0,7 \leq \alpha < 0,8$	Elfogadható
$0,6 \leq \alpha < 0,7$	Kétséges
$0,5 \leq \alpha < 0,6$	Rossz
$\alpha < 0,5$	Elfogadhatatlan

A belső konzisztenciának, mint a tételek egy csoportjának egymáshoz való kapcsolódásának a meghatározását számos szakértő elfogadta (Cortina, 1993; Cho & Kim, 2015). A reliabilitás értéke függ az itemek számától és minőségétől (Molnár & Vígh, 2013). A rosszul megválasztott disztraktorokkal (zavaró tényezőkkel, helytelen válaszokkal) rendelkező feladatok csökkentik a megbízhatóságot (Demkanin et al., 2015). Az alacsony megbízhatóság egy egészségtelen tesztelési helyzet tünete, ahogyan a magas láz is egészségtelen állapotot jelez. Egyik esetben sem tudjuk megmondani, hogy mi a probléma, de a szimptóma megmutatja, hogy hol keressük (Ebel & Frisbie, 1991). Általánosságban elmondható, hogy minél nagyobb az itemek száma, annál megbízhatóbb a mérés (Molnár & Vígh, 2013). Ha az itemek száma kicsi, vagy az átlagos korreláció alacsony, akkor a Cronbach-alfa is alacsony lesz (Hidegkuti & Balázs, 2015). Minél több feladatot és a tanterv minél több szempontját vizsgálják, annál megbízhatóbbak a teszteredmények (Demkanin et al., 2015). Gyakorlati szempontból egy tesztnek legalább 20-30 itemből kell állnia, és az itemek számának növelése is növelheti a megbízhatóságot (Ebel & Frisbie, 1991; Molnár & Vígh, 2013). Például, ha egy 20 itemes teszthez további 60 itemet adnánk hozzá, a megbízhatóság 0,5-ről 0,8-ra emelkedne (Ebel & Frisbie, 1991). Ettől függetlenül a magas Cronbach-féle alfa sem jelenti azt, hogy a teszt itemjei egy dimenziót mérnek (Hidegkuti & Balázs, 2015).

Cronbach és Hartmann (1954) szerint negatív reliabilitások merülhetnek fel, ha a tesztek belső konzisztenciája nulla, azaz ha az itemek közötti átlagos kovariancia nulla. Negatív "megbízhatósági együtthatókat" kapunk, ha megsértjük a féltesztek közötti egyenlő kovariancia feltételezését. Ha a feltételezések nem teljesülnek, akkor ezek a

képletek nem használhatók a közöttük várható összefüggés becslésére (Cronbach & Hartmann, 1954).

Csapó utal a feladat reliabilitásának és validitásának összefüggéseire (Csapó, 2004). Demkanin és munkatársai (2015) szerint a megbízhatóság az érvényesség szükséges, de nem elégséges feltétele (Demkanin et al., 2015). Ahhoz, hogy a teszt érvényes (valid) legyen, megbízhatónak (reliabilisnek) is kell lennie (Csíkos & B. Német, 2002), tehát jó reliabilitási mutatóval kell rendelkeznie. Az összefüggés fordítva nem áll fenn, ha egy teszt érvényessége gyenge, akkor is lehet nagyon magas a megbízhatósága. A tudásszintmérő tesztek érvényességét egy olyan technika, egy tesztkonstrukciós eljárás alkalmazásával lehet biztosítani, amely lehetővé teszi, hogy a tananyagot pontosan leképezzük a tananyag mérésére és tesztelésére használt eszközre. (Csapó, 2004). A megbízhatóság és a pontosság tehát az érvényesség előfeltétele, de nem az egyetlen előfeltétele (Demkanin et al., 2015).

A teszt objektivitása azt jelenti, hogy a teszt tárgyyszerű, tárgyilagos, nem szubjektív (Csapó, 2004), mindig mindenkit egyformán mér (Magyar & Molnár, 2013). Egyes szerzők a megbízhatóságot befolyásoló lényeges jellemzőnek (tényezőnek) tekintik. Ez a véletlenszerű vagy szubjektív tényezők kizárását (vagy erős csökkentését) is jelenti a tesztelésben (Demkanin et al., 2015). Pontosabban fogalmazva a teszttel végzett mérés eredménye független attól, ki végzi a teszttel a mérést (Csapó, 2004).

Az iskolákban használt tudásszintmérő tesztek objektivitása néhány egyszerű szabály betartásával kielégítő szintre emelhető (Csíkos & B. Német, 2002). A tesztelés objektivitását a következők biztosítják:

- (a) a teszt objektivitása,
- (b) a pontozás objektivitása,
- (c) a tesztelési folyamat objektivitása (Demkanin et al., 2015).

A teszt objektivitása a tételek helyes kiválasztásával, a kérdések egyértelmű megfogalmazásával és a válaszlehetőségekkel, a tételek megfelelő anyaglefedettségével, valamint a tételekre adott helyes válaszokkal kapcsolatos minden kétség kiküszöbölésével érhető el (Demkanin et al., 2015), biztosítja azt, hogy a teszteredmény független az adatfelvevő személyétől (Csapó, 2004). Az értékelés objektivitását pontos és egyenlő pontozási szabályokkal és az azonos válaszok azonos módon történő pontozásával (jelölésével) biztosítjuk (Demkanin et al., 2015), így eredménye független attól, hogy az eredmények javítását, kódolását, azaz értékelését ki végzi (Csapó, 2004).

A tesztelési folyamat objektivitását úgy érjük el, hogy minden vizsgált tanuló számára egyenlő tesztelési feltételeket biztosítunk (Demkanin et al., 2015).

Minden feladat a megoldási algoritmusnak megfelelően lépésekre (alternatív egységek, itemekret) bontható. Ez a művelet képezi a pontozás (súlyozás) alapját, mivel az itemek alapján készülnek a pontozási táblázatok és történik a pontozás. A súlyozás az egyes alternatív egységek pontszámának meghatározása a kezdeti egységes pontértékből kiindulva, a súlyozás szabályai szerint (empirikus súly, fontossági súly, szintsúly). Ily módon az egység tulajdonságai figyelembe vehetők és pontokban kifejezhetőek az egyes itemek pontértékének meghatározásakor.

1. A fontossági súly kifejezi, hogy a tanterv egyes részei (alternatív egységek) fontosabbak a többinél, és ezért magasabb pontszámot kapnak.
2. A szintsúly jelentősége a pontozásban az, hogy kifejezze, hogy a feladat mely elemei mutatnak magasabb tudásszintet, és mely elemek azok, amelyekhez alacsonyabb elsajátítási szint elegendő.
3. Az empirikus súlyt arra használják, hogy kiemeljék, hogy valamely feladatelemet a vizsgált személyek mekkora hányada oldja/oldotta meg hibátlanul, és hányan nem. Az ilyen módon nehezebbnek bizonyuló itemek több pontot érnek (N. Tóth, 2015).

1.1.3. Itemtípusok és a feladatírás fontosabb szabályai

Eluwa szerint bár a klasszikus tesztelméletet a legtöbb ideig a mérési közösség használta, az utóbbi évtizedekben az új generáció – mely az itemek tulajdonságait valószínűségelméleti eszközökkel jellemzi (Molnár, 2003) – egyre nagyobb teret nyert, és így a mérési keretrendszerek egyik kedvencévé vált (Eluwa et al., 2011).

A tanulóközpontú, tanulói aktivitáson alapuló, készségek fejlesztésére irányuló tanítás elképzelhetetlen jó minőségű feladatok nélkül. A tanulási, tanítási folyamat sokrétősége miatt azonban nagyon sokféle cél állhat a feladat mögött: az új anyag elsajátítása, gyakorlás, az alkalmazási készség fejlesztése, problémamegoldás, de lehet ráhangolódás, esetleg összefoglalás is. (Einhorn, 2012). A legtöbb tanár által készített tesztek célja, hogy a teljesítményt mérjék, amely alapján visszajelzést adhatnak a tanulóknak, és beszámolhatnak a szülőknek a fejlődésükről (Ebel & Frisbie, 1991). A tudásszintmérő tesztek tervezésének első lépése annak a tudásnak a meghatározása, amelyet a tesztek mérni hivatottak (Csíkos & B. Német, 2002). A mérés céljától függően

sokféle teszt és tesztfeladat tervezhető (Csapó, 2004), amelyek valamilyen tanulási tevékenységre összpontosítanak (Einhorn, 2012).

A tesztek feladatokból állnak, amelyek legkisebb részeit itemeknek nevezzük, amelyek önállóan pontozhatók. Egy item differenciáló ereje vagy diszkriminatív képessége azt méri, hogy az item mennyire élesen tesz különbséget a különböző tudásszintű tanulók között (Csapó, 1988). Az itemalkotásnak vannak részletes technikái, amelyek segítségével az adott vizsgálati célhoz legjobban illeszkedő, változatos teszt itemek állíthatók elő (Csíkos & B. Német, 2002). A tartalmi specifikációval a kezében a tesztfejlesztő következő döntése a használandó feladattípusokra vonatkozik (Ebel & Frisbie, 1991). Ehhez változatos kérdéstípusokat használnak (Tóth, 2015).

A feladatok a válaszadás módja szerint két fő kategóriába sorolhatók: feleletválasztásos és konstruált válaszok (Csapó, 2004; Csíkos & B. Német, 2002). Sokolova és Totkov (2005) ezeket a kategóriákat másként nevezi meg: szabad formájú válaszok (nyitott típus); kérdések és feladatok előre megadott válasszal (zárt típus). Tanulmányukban rámutatnak, hogy a Ruter-féle osztályozás a tesztkérdések és feladatok három alapvető csoportját tartalmazza (az elvárt válasz szerint): nyitott, félig nyitott és zárt (Sokolova & Totkov, 2005). Ezen osztályozás szerint: zártak azok a feladatok, amelyekben a tanulók csak megadott (alternatív) lehetőségek közül választhatnak; félig nyíltak azok a feladatok, amelyekben a tanulók önállóan (szabadon), de csak nagyon korlátozott keretek között dolgozhatnak; míg nyíltak azok a feladatok, amelyekben (Einhorn, 2012) a válaszadók maguk fogalmazzák meg válaszaikat (Sokolova & Totkov, 2005), azaz a válaszok produktív tevékenységet igényelnek és szabadon konstruálhatók (Einhorn, 2012).

Az egyes tesztfeladatok és következésképpen az egész teszt nehézségét a teszt szerzője határozza meg a tesztfeladatok összeállításakor és a tesztbe való felvételükkor (Demkanin et al., 2015). Az itemek meghatározásával pedig egyértelművé tehető a javítás, és a tanuló számára is világossá tehető, hogy mit oldott meg jól, és mit nem (Székely, 2014).

Az 1. ábrán a feladatok rendszerezését találjuk a tanulói tevékenység alapján, melyet Einhorn (2012) gyűjtött össze. Felhívja figyelmünket arra, hogy az ábra nem tartalmazza az összes lehetőséget, mivel a témában rejlő számtalan variáció miatt nem lehet ilyen listát készíteni (Einhorn, 2012).

FELADATTÍPUSOK

Feleletválasztás, igaz-hamis, egymáshoz rendelések

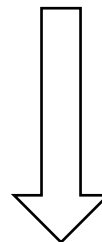
Kiegészítések

Rövid válasz

Hosszabb válasz

Szövegalkotás, rajz, problémamegoldó feladatok,
komplex feladatok

ZÁRT



NYITOTT

1. ábra. A feladatok rendszerezése a tanulói tevékenység alapján (Einhorn, 2012).

A feleletválasztós feladatokat számos értékelési környezetben használják, beleértve a képesség-, alkalmasság-, attitűd- és személyiségértékelési eszközöket (Moreno et al., 2015). Ellenben Csapó (2004) szerint kevésbé, vagy csak igen elmélyült tesztkészítés révén használhatók a készségek, a gondolkodás rugalmasságának, a tudás minőségének, a szövegértésnek, a szintézisnek a vizsgálatára. A feleletválasztós tesztek két fő részből állnak: egy problémából, amely lehet egy kérdés vagy egy hiányos állítás, és egy alternatívákból álló listából, amely tartalmazza az egyetlen helyes vagy legjobb választ és számos zavaró tényezőt (Varga, 2011). Az feleletválasztó-itelek írásának egyik legalapvető követelménye, hogy a feladatot kérdés- vagy kitöltési formában kell megfogalmazni (Crehan & Haladyna, 1991). Egy kérdésre általában több válaszlehetőséget adnak (Einhorn, 2012), jellemzően három disztraktor (zavaró tényező) és egy helyes válasz van (Demkanin et al., 2015), amelyekből a tanuló/vizsgáló kiválasztja a helyeset (Einhorn, 2012).

Emeljünk ki néhány fontosabb feleletválasztó itemet:

- *Alternatív választás:* csak két válaszlehetőség adott, amelyek közül az egyik mindig jó, a másik mindig rossz megoldás (Csapó, 2004). Használatának alapvető oka, hogy egyszerű és közvetlen eszközt biztosítanak az iskolai oktatás alapvető eredményeinek mérésére (Ebel & Frisbie, 1991).
- *Többszörös választás:* a feladatban több válasz is lehetséges, amelyek közül egy vagy több helyes lehet. A feladatnak tartalmaznia kell, hogy a megoldójának az adott válaszok közül egyet vagy többet kell megjelölnie (Csapó, 2004). A helyes válasz kitalálásának valószínűsége kisebb egy 5 válaszlehetőséget tartalmazó

feladat esetében, mint egy 2 vagy 3 választási lehetőséget tartalmazó feladat esetében, ha a rossz válaszok (disztaktorok) hihetőek (Haladyna & Downing, 1993).

- *Válaszok illesztése*: két halmaz elemei között kell kapcsolatot találni (Csapó, 2004). Valaminek a felismerése mindig egyszerűbb, mint a létrehozása. Ezért nagyon hasznos a tanulás azon fázisában, amikor a tanulók tudás- vagy készségszintje még nem teszi lehetővé, hogy produktív feladatokat oldjanak meg (Einhorn, 2012).

A feleletválasztó itemekről további részleteket és még több típusáról olvashatunk, valamint mintafeladatokat láthatunk a tanulmányokban (Crehan & Haladyna, 1991; Ebel & Frisbie, 1991; Einhorn, 2012; Haladyna & Downing, 1993; Moreno-Guerrero et al., 2020; Varga, 2011).

A feleletalkotó feladatok a tudás bonyolultabb összetevőinek mérésére is használhatók, amelyek gyakran speciális értékelési eljárást igényelnek (Csapó, 2004). Fontosabb típusai a következők:

- *Kiegészítés*: a hiányos szövegeket vagy mondatokat kell kiegészíteni önállóan vagy a megadott elemek segítségével. A feladattípus tudáselemek pontos mérését teszi lehetővé: definíciók, összefoglalások, folyamatok felidézésére alkalmas (Einhorn, 2012).
- *Rövid válasz*: a tesztfeladat célja a tudás ellenőrzése azáltal, hogy egy kérdésre válaszoló egy mondatot befejező szót, kifejezést vagy számot adjon meg (Ebel & Frisbie, 1991). Értékelése tökéletesen egyértelmű lehet, ha a kérdést úgy fogalmazzuk meg, hogy csak egy helyes válasz adható (Csapó, 2004). A legtöbb problémát az okozza, hogy mechanikus, ezért ismeretek alkalmazását nem lehet vele mérni vagy fejleszteni (Einhorn, 2012).
- *Hosszú válasz*: a feladat alkalmas a logikai összefüggések felismerésének mérésére, a megoldáshoz rövid érvelés, indoklás szükséges lehet szóban vagy írásban, írásbeli formáját nevezik nyílt végű kérdésnek is (Einhorn, 2012). A válasz értékelése nehezebb, mivel ugyanazt a tartalmat többféle mondatban is kifejezhetik, megfogalmazhatják (Csapó, 2004).
- *Esszé típusú válasz*: az ilyen típusú kérdések is a kifejtendő vagy hosszabb, részletes választ igénylő kérdések kategóriájába tartoznak (Einhorn, 2012). Viszonylag könnyű elkészíteni - a munka nehéz része általában a diákok

válaszainak osztályozása (Ebel & Frisbie, 1991). A legbonyolultabb értékelési feladatok, mivel egy mondatnál hosszabb, összefüggő válaszokat igényelnek (Csapó, 2004).

Székely (2014) a klasszikus valószínűségre alapozva vizsgálta az itemek nehézségi indexét, mely azt fejezi ki, hogy a tanuló milyen valószínűséggel oldja meg az itemet:

$$\frac{\text{jó megoldások száma}}{\text{itemet megoldó összes tanuló száma}} \quad (\text{Székely, 2014}) \quad (2)$$

A fenti törtből jól látható, hogy egy item akkor könnyű, ha a tört értéke 1-hez közelít, ha a tört értéke 0-hoz közelít, akkor az itemet kevesen oldották meg, ha 0, senki sem tudta a jó választ, azaz az item túlságosan nehéz (Székely, 2014).

Shabatura (2022) Bloom taxonómiájára hivatkozva (lásd 1. táblázat) jelenti ki, hogy a szerkesztési folyamat során hasznos a tudás és a kognitív folyamatok dimenzióit változtatatosan alkalmazni és úgy létrehozni az itemeket (Shabatura, 2022).

A feladatírás tartalmi alapelvei:

- A feladat szövege legyen életkornak és az elvárt felkészültségnek megfelelő;
- feladat szövege csak a megoldáshoz szükséges információkat tartalmazza. Javasolt a rövid mondatok alkalmazása;
- A feladat szövege ne legyen célzás vagy sugalmazó, és ne tartalmazza a megoldást vagy annak egy részét (Molnár & Víg, 2013).

Haladyna és Downing (1989) az általuk összegyűjtött 46 item-írási szabályból a következőket emelik ki, mint az általános (eljárési) szabály:

1. Használja a legjobb válasz vagy a helyes válasz formátumát;
2. Kerülje az összetett feleletválasztós formátumot (pl. A és D, A és C, Mindegyik, Egyik sem, A, B és C, stb.);
3. Formázza a feladatot függőlegesen, ne vízszintesen;
4. Hagyjon időt a szerkesztésre és más típusú itemek átdolgozására;
5. Használjon következetesen helyes nyelvtant, írásjeleket és helyesírást;
6. Az egyes tételek megfogalmazásakor minimalizálja a vizsgázó olvasási idejét;
7. Kerülje a trükkös tételeket, vagyis azokat, amelyek félrevezetik vagy megtévesztik a vizsgázókat, hogy helytelenül válaszoljanak. (Haladyna & Downing, 1989; Varga, 2011)

1.2. E-tesztekkel támogatott oktatás

Bár az 1.1. fejezet bevezetésében több meghatározást összegyűjtöttünk, egyet kell értenünk Csíkos és B. Német (2002) megállapításával, miszerint sokak számára - talán a hatvanas-hetvenes évek divatja miatt - a teszt még mindig egy írásbeli feleletválasztós teszt (Csíkos & B. Német, 2002). A közismert tesztek, amelyeket gyakran papír-ceruza (Paper and Pencil) teszteknek neveznek, nagyon fontos szerepet játszottak és játszanak ma is a tanítási-tanulási folyamatok irányításában, az oktatás eredményességének felmérésében. Ezek a tesztek többnyire rögzített formátumúak (Fixed Form) (Csapó et al, 2008). Az oktatási tesztelés technikái azonban az értékelés céljától függően igen változatosak lehetnek (Csíkos & B. Német, 2002). A tanulási eredmények mérésére didaktikai tesztek alkalmaznak. Ezeket mérhető, pedagógiaiilag releváns konstrukcióknak tekintjük (tudás, készségek, kompetenciák, kognitív műveletek stb.) (Demkanin et al., 2015).

A tesztkérdések és feladatok egy adott elektronikus tesztben különböző elvek és szabályok szerint választhatók ki, melyekre a korábbi fejezetek iránymutatóak. A pedagógiai tesztek alapvető célja az, hogy az oktatási, nevelési folyamatok irányításához információt szolgáltatassanak, és így a fejlődést, a fejlesztést, a változtatást segítsék (Csapó, 1988). Ezeknek a mérőeszközöknek

(1) egyértelműen kiértékelhető megoldásuk van.

(2) Főként a kognitív tudásterület, vagyis a tanuló pontosan elhatárolt, és a mérési folyamat elején előre meghatározott ismeret- és képességjellegű tudásának vizsgálatára alkalmasak.

(3) A kapott eredmények az egyéni és csoportszintű kiértékelést követően kvantifikálhatók (számszerűsíthetők), pontokban kifejezhetők és standardizálhatók (Molnár & Víg, 2013).

A legtöbb tesztelés során alkalmazott tesztek általában teljesítménytesztek, ahol az egyes tételekre adott válasz lehet helyes (korrekt) vagy helytelen (inkorrekt). Az ilyen itemeket és változókat dichotóm (kétértékű) változóknak nevezzük (Hidegkuti & Balázs, 2015). Ebben az esetben a teljesítmények két csoportba, elfogadható vagy elfogadhatatlan csoportba sorolhatók. Azt mondjuk, hogy a tanuló megfelelt, vagy nem felelt meg, a produktuma jó vagy rossz (Csapó et al, 2008).

A tömeges tesztelés során ugyanazon teszt több tucat különböző változata jelenik meg. Ezek a változatok abban az értelemben próbálnak hasonlóak lenni, hogy csak az itemek konkrét tartalmában és az itemek számában különböznek, a formátumuk azonos,

és az alapvető statisztikai jellemzőik is hasonlóak (Demkanin et al., 2015). Az összes feladat ezután egy feladatbankot alkot, amelyből a konkrét felmérések igényeinek megfelelően lehet kiválasztással vagy véletlen sorsolással a konkrét felmérések céljaira tesztek összeállítani (Csapó et al, 2008). Usiskin szerint az értékelés eredeti pedagógiai funkciója a tanulás támogatása lenne, azonban a nagyléptékű, elszámoltathatóság-központú szummatív tesztelés gyakran beszűkíti a tanulási célokat, és mechanikus tudáselemek mérésére redukálja az oktatást (Usiskin, 2019).

Az oktatás digitalizálásával különféle elektronikus eszközök segítik a méréseket (Korenova, 2013). Az online tanulási környezetek megjelenése alapvetően átalakította a formatív és szummatív értékelés szerepét és egymáshoz való viszonyát (Usiskin, 2019). Az online környezetben a formatív és szummatív értékelés határai elmosódhatnak (Négyesi, 2021). A Web, mint az online tesztelés platformja, új kihívást jelent. A világ tudományos és kulturális ismeretét egységbe szervezi, adatbankként is működik. A weboldalak kész tesztekkel kínálnak a tanároknak, akik személyre szabhatják azokat (Parázsó, 2012).

Korenova (2013) szerint kettősen definiálhatjuk az „e-teszt” kifejezést: 1. Szűkebb értelemben az e-teszt egy elektronikusan vezérelt didaktikai teszt, amelynek lehetősége van a multimédiás elemek gazdagítására. 2. Tágabb értelemben az e-teszt egy elektronikus interaktív anyag, amely kérdésrendszeren alapul és válaszokat keres, amelyek nemcsak mérésre, hanem az oktatási célok elérésére is létrejönnek (ezáltal az innovatív tanítási módszerek eszközeként szolgálhat).

Az e-teszt segítségével nemcsak a hallgatók tudását tudjuk meghatározni, hanem ezekkel az új digitális eszközökkel növelhetjük a hallgatók motivációját, felhasználhatjuk őket ismétlés, gyakorlás, ellenőrzött felfedezési módszerek során. Az e-teszt a hallgatók szempontjából nagyon vonzó, mivel a digitális világ nagyon közel áll hozzájuk (Korenova, 2013), könnyen kezelik az internetes felületeket (Parázsó, 2012).

Az e-tesztek felhasználásuk szerint az alábbiak szerint osztályozzuk:

- Meghatározzuk a hallgatók tudását;
- Javítani akarjuk a hallgatók motivációját, kihívás legyen a feladatok megoldása;
- Interaktív munkalapként használjuk (ez már nem hagyományos teszt, az elektronikus teszt itt csak eszköz) - ellenőrzött felfedezésben használható (helytelen válaszok után a hallgatót a hozzárendelt feladatokra irányítja) (Korenova, 2013).

Négyesi szerint a digitális platformokban rögzített tanulási adatok – például kvízeredmények– könnyen szummatív értékelési funkciót is betölthetnek, még akkor is, ha eredetileg fejlesztési céllal jöttek létre (Négyesi, 2021). Ez összhangban áll Usiskin kritikájával, miszerint a tesztelés túlsúlya torzíthatja a tanítási gyakorlatot és a tanulók tanuláshoz való viszonyát.

Négyesi (2021) rámutat arra, hogy a digitális környezetek lehetőséget teremtenek a formatív értékelés folyamatos, személyre szabott alkalmazására. Az azonnali visszajelzés, a tanulói hibákhoz kapcsolódó magyarázatok és a tanulási út dinamikus alakítása mind a formatív értékelés erősödését szolgálják (Négyesi, 2021). Az online teszt által kapott tanulmányi teljesítmény mérése mellett, az elektronikus értékelési kibertérben még számtalan paraméter létrehozható. Az önértékelés során megvalósul a tudás folyamatos ellenőrzése, lehetőség nyílik a tanulási stílusok, módszerek hatékonyságvizsgálatára, attitűdvizsgálat, szociometriai felmérésekre (Parázso, 2012).

Parázso (2020) kiemeli az online tesztek előnyeit és hátrányait:

Előnyök:

- A teszt eredményét azonnal szolgáltatva a diák érdeklődését fokozza, adrenalin szintjük megnő, a tananyag-elsajátítás hatékonyságát fokozza.
- A tanulónak lehetősége van arra, hogy bárholnan elérje a tesztfeladatot a megfelelő paraméterek ismeretében, akadályoztatása esetén (pl. fertőző betegség) is.
- Napjainkban megjelentek az online standardizált tesztek, feldolgozásuk a korszerű célszoftverek segítségével történik.
- Az online teszt beállítható, hogy a diák a feladatokat meghatározott sorrendben oldja meg. Kutatók megfigyelték, hogy az első válaszadás a jobb, mivel többségében a javítás során variál, ront az eredményen.
- A tanár időt takarít meg a tesztek javítására töltött idő lecsökkenésével.

Hátrányok, kihívások:

- Szem előtt kell tartani a diákok személyiségi jogainak sértetlenségét. Nevüket, adataikat kódolni kell.
- A tesztfelület nyitó felületének szerkesztésekor figyelembe kell venni a tanuló életkori sajátosságait. Gyakori eset, hogy a diák szövegértési kompetenciája nem megfelelő, amelynek következtében meg nem értések adódhatnak.

- Az online tesztek azon kérdéseinek megoldásában, ahol szavakkal, mondatokkal egyszerű választ adnak a diákok, előfordulhat, hogy a helyesírási vétségeket a szoftver rossz válaszként értékeli.
- A teszt felületére való lépést korlátozni kell jelszóval, így illetéktelen nem léphet be és védetté válik a feladatsor.
- A számítógép monitor előtt ülve a diákok nem érzékelik a feladat súlyát, játéknak tekinthetik. (A papíralapú dolgozat esetében a feladat inkább tudatosult). Előfordulhat, hogy chatelnek, előre, hátra tekingetnek.
- A teszt eredményeinek ismeretében a tanárnak rugalmasnak kell lennie, hisz előfordulhat, hogy ez a számonkérési forma az újdonság erejével hat. Abban az esetben, ha az osztály alacsony pontszámot ért el, újra kell írni.
- Biztonsági tervet kell kidolgozni arra az esetre, ha technikai hiba lép fel pl. áramkimaradás, szerver leállás (Parázsó, 2020).

A teszt alkalmazásának utolsó fázisa az eredmények feldolgozása, értelmezése. Itemszinten a helyes válasz 1, a helytelen vagy hiányzó válasz 0 pontot ér. Az adatokat Excel táblázatba gyűjtve néhány fontos következtetést már meg tudunk fogalmazni az itemek, a feladatok és a teszt szintjén, valamint a tanulócsoporthoz teljesítményét illetően. Az itemeket jellemző nehézségi indexet az átlag segítségével fejezhetjük ki, és azt várjuk, hogy az érték 0,25 és 0,80 között legyen; az ez alatti itemek túl nehezek, az előlöttiek pedig túl könnyűek az adott csoportban (Molnár & Vigh, 2013).

Végezetül meg kell jegyeznünk, hogy az itemek tulajdonságait valószínűségelméleti eszközökkel jellemezve nem feltétlenül lesz igaz a klasszikus tesztelmélet azon tétele, miszerint a teszt megbízhatósága a teszt hosszának (a itemek számának) növekedésével nő, azaz a mérési hiba csökken. Könnyen előfordulhat, hogy egy rövidebb teszt megbízhatóbb, mint egy lényegesen hosszabb tesztváltozat. Az itemek számának növelése csak abban az esetben növeli a mérés pontosságát, ha a hozzáadott itemek által hordozott információ növeli a tesztinformáció értékét egy adott képesség esetén (Hidegkuti & Balázs, 2015).

1.2.1. Előhívásos tanulás

Tanárként belátva, hogy az a technika, amivel engem is oktattak, már nem elegendő, új technikák és motiválási lehetőségek irányába kutattam. Megismerkedtem az előhívásos tanulási technikával, ami azt jelenti, hogy megpróbáljuk felidézni az

információt anélkül, hogy az előttünk lenne (Gonzalez, 2017). Ez a technika különbözik a tanórák elején gyakran alkalmazott kérdezési és visszakeresési technikáktól, amelyekkel ellenőrizni és mérni lehet a tanultakat. Az ötlet lényege, hogy a rövid távú memóriában tárolt információkat idézzük fel, így aktiválva a megtartást.

Az oktatási környezetben a tesztek általában az értékelés eszközeinek tekintik. A diákok tesztek írnak az órákon, hogy felmérjék, mit tanultak, és szabványosított tesztek írnak, hogy felmérjék tudásukat és alkalmasságukat (Roediger & Karpicke, 2006). A főként a teljesítmény pontos mérésére tervezett tesztek valószínűleg motiválják a tanulókat, és oktatási célúak is lehetnek (Ebel & Frisbie, 1991). Az értékelés (teszteredmények) bemutatása a vizsgázónak a teszt céljától függően többféleképpen is megvalósítható (Horovčák & Stehlíková, 2007), több dimenzióhoz kapcsolódó következtetésekkel és pedagógiai jellemzőkkel:

- Az önértékelésre és önképzésre szánt teszt tanulásra ösztönöz;
- A megszerzett pontszámok feltüntetésével átfogó képet ad a tanulási folyamatról;
- Flexibilitás (rugalmasság), azaz a követelmények és az eszközök a tanulók képességeihez és érdeklődéséhez igazodnak (Parázsó, 2012; Horovčák & Stehlíková, 2007).

Az előhívásos tanulás (Retrieval learning, Test-Enhanced Learning) azt jelenti, hogy megpróbáljuk felidézni az információt anélkül, hogy az előttünk lenne (Gonzalez, 2017).

Az elmúlt években a kognitív pszichológusok összehasonlították a felidézés előtti tanulást más tanulási módszerekkel - olyan stratégiákkal, mint az ismétlődő előadások, a tanulmányi útmutatók és a szövegek újraolvasása. Azt tapasztalták, hogy semmi sem ragadja meg olyan erősen a hosszú távú tanulást, mint az előhívásos tanulás (Gonzalez, 2017). A memóriában lévő tudáshoz való hozzáférés fenntartásának egyik kulcsa az, hogy időről időre használjuk ezt az információt - vagyis előhívjuk. Ellentétben más memóriarendszerekkel, például a magnó vagy a számítógép memóriájával, ahol a tárolt információ előhívása nem változtatja meg a memóriában lévő információ állapotát, az emberi memória jelentősen megváltozik az előhívás hatására (Bjork, 1988).

A tanulás és az előhívás javításának hagyományos megközelítése - mind a laboratóriumi vizsgálatokban, mind pedig az osztályteremre való kiterjesztéssel - a tanulási stratégiák megváltoztatását jelenti (Roediger at.al, 2011). A kognitív pszichológia egyik legkövetkezetesebb megállapítása, hogy a tesztelés jobban növeli az elmélyülést, mint a tanulás önmagában. Ez a hatás fokozható, ha a diákok visszajelzést

kapnak a sikertelen tesztekéről, és mind a rövid, mind a hosszú távú megtartás esetében megfigyelhető (Brame-Biel, 2015).

Egy évszázadnyi kutatás foglalkozott a tesztelési hatás vizsgálatával - azzal a ténnyel, hogy az aktív előhívás jobb elmélyüléshez vezet, mint a passzív újraolvasás (Roediger at.al, 2011). Bizonyított tény, hogy a tesztelés nemcsak a vizsgált információra, hanem a kapcsolódó információkra vonatkozóan is javítja a tanulók memóriáját (Brame & Biel, 2015). Bjork (1988) kiemeli, hogy az előhívás-gyakorlat gondolatának központi feltételezése, hogy az előhívás nem egyszerűen csak erősíti egy tárgynak a memóriában való reprezentációját, hanem magának az előhívási folyamatnak valamilyen aspektusát is erősíti. Az ezt a feltételezést alátámasztó megfontolások a következők:

(a) Ha a kezdeti előhívás egyszerűen csak megerősítené egy tárgy reprezentációját a memóriában, akkor azt várhatnánk, hogy a kezdeti előhívási teszt ugyanúgy megkönnyítené a későbbi visszatekintési tesztet, mint a későbbi előhívási tesztet. Jellemzően azonban a kezdeti előhívási teszt hatása egy későbbi visszatekintő tesztre sokkal kisebb, mint egy későbbi felidézésre.

(b) Ha viszont egy későbbi visszatekintő tesztet olyan módon nehezítünk meg, amely úgy tekinthető, hogy az előhívásos tesztet felidézésszerűbbé teszi, akkor a kezdeti előhívási teszt pozitív hatásai sokkal nagyobbak.

(c) Hasonlóképpen kimutatták azt, hogy amint egy kezdeti előhívó tesztet nehezebbé és felidézésszerűvé teszünk azáltal, hogy növeljük az alternatívák számát egy kezdeti feleletválasztásos teszten, a teszt teljesítménye csökken, de a későbbi visszahívás növekszik (Bjork, 1988).

Úgy tűnik, hogy a tesztelés megerősíti a további tanulást, így a diákok többet profitálhatnak a teszt utáni tanulási időszakokból (Brame & Biel, 2015). Létezik egy egyedülállóan optimális módja a tanulás sorrendjének az előzetes teszteléssel. A szinte azonnali első tesztet hosszabb késés esetén további teszteknek kell követniük egymás után. Az ilyen bővülő ütemezés egyfajta optimális alakítási eljárást jelent. Minden egyes egymást követő előretesztelés elősegíti a következő (hosszabb) intervallum utáni sikeres előhívást, és ahogy az intervallum hosszabbodik, úgy válik minden előhívás egyre erőteljesebbé, mint tanulási esemény (Bjork, 1988).

Ahogy Roediger és Karpicke (2006) fogalmazott: „Hogy egy nyilvánvaló dolgot mondjak, ha a diákok tudják, hogy rendszeresen (mondjuk hetente egyszer, vagy akár minden órán) tesztelik őket, akkor többet fognak tanulni, és a félév során elosztják a tanulást, ahelyett, hogy csak előtti időszakra torlódnának” (Roediger és Karpicke, 2006).

Az e-tesztelés nagyon vonzó a diákok számára, mert a digitális világ nagyon közel van hozzájuk (Korenova, 2013).

Korenova és munkatársai (2021) a kutatás következő módszereit jegyzik meg:

- az aktuális és a középszintű ellenőrzésre szolgáló értékelési eszközrendszer pedagógiai tervezése;
- a matematika tesztfeladatok bázisának gyakorlati értékelése az egyes e-tesztek létrehozásának részeként;
- kérdőíves felmérés - az év során tesztfeladatokat megoldó diákok körében végzett felmérés a visszajelzések megszerzése érdekében (Konnova et al., 2021).

A Muzsnay, Szabó és Szeibert (2024) által végzett kutatás tudományos konklúziója rávilágít a felsőfokú matematikaoktatás egyik legkritikusabb módszertani dilemmájára: a vizsgaközpontú, rövid távú ismeretszerzés és a valódi, tartós tudásmegőrzés közötti szakadékra. Kutatásuk bizonyította, hogy az előhívási gyakorlat (retrieval practice) nem csupán egy kiegészítő tanulási technika, hanem a hosszú távú tudásmegőrzés alapvető eszköze a komplex matematikai diszciplínák, például a számelmélet területén. Míg a hagyományos ismétlés (újraolvasás) gyakran csak a felismerés illúzióját kelti, az aktív felidézés strukturális változásokat idéz elő a memóriában, amelyek ellenállóbbá teszik a tudást a felejtéssel szemben (Muzsnay et al., 2024).

1.2.2. Motiváció a matematikaoktatásban

A mai felgyorsult és technikától függő világban a tanárok számára már nem elegendő megfelelő tudással rendelkezni, szükség van kreativitásra és technikai tudásra ahhoz, hogy felkeltse a diákok figyelmét a tananyag iránt, motiválja őket a tanulásra. A matematikát tekintve a nehézség már a régmúltban is jelen volt, sokan arra hivatkoztak, hogy a matematika megértésére születni kell. Ezzel a kifogással élve sokan elzárkóztak, ma is elzárkóznak a matematika tanulásától. Ekkor válik még fontosabbá a tanár szerepe az órákon, hogy fel tudja hívni a tanulók figyelmét, mennyi dologban veszi őt körbe a matematika, melyet ő maga is természetesnek tekint, és motiválni tudja őt a minél nagyobb tudás megszerzésére és logikus gondolkozásra.

Egyesek azt gondolják, hogy a motiváció olyan tulajdonság, mint a takarékoság, a kézügyesség vagy a zenei tehetség. Vannak, akik jó adaggal rendelkeznek belőle, míg mások egyszerűen alulmaradnak (Niermeyer & Seyffert, 2011). A kíváncsiság a

megismerésre, a tudás megszerzésére, a problémamegoldására irányuló belső késztetés legegyszerűbb formája (Ceglédy, 2011), míg a tevékenységek önmagukban örömet jelentenek. Ha ez kapcsolódik a tanulási tevékenységekhez, akkor a tanulás sikeres lesz, ha a tanuláshoz az öröm, a csoda és a tanulás iránti érdeklődés érzése társul (Fuqoha et al, 2018). Az emberek megfigyelnek, vizsgálnak dolgokat, folyamatokat, esetleg olyan tapasztalatokat szereznek, amelyeket pillanatnyilag nem tudnak hasznosítani, amelyeknek nem tulajdonítanak jelentőséget, de később ez a kíváncsiság nemes formákat ölthet, például a tudatosság formájában (Ceglédy, 2011).

A szakirodalmi kutatások alapján összegyűjthetjük a motiváció fogalmát. A szó latin eredetű, jelentése: cselekvés ösztönzői, kiváltói. A motívum szó pedig indítóokot, erkölcsi indítékokat jelent (Orosz, 1994). Az általános motivációt úgy határozzuk meg, mint a vágyat, hogy létrehozzunk valamit, hogy elérjünk valamit, és hogy változtassunk valamit. Ez a vágy alapvetően mindenkinben jelen van, bár különböző mértékben (Niermeyer & Seyffert, 2011). További kutatások szerint a motiváció definiálható, mint

- az emberi pszichológia és viselkedés összetett része, amely befolyásolja, hogy az egyének hogyan döntenek el, hogy befektetik-e az idejüket, mennyi energiát fordítanak egy adott feladatra, hogyan gondolkodnak és éreznek a feladattal kapcsolatban, és mennyi ideig maradnak a feladatban (Filgona et al., 2020).
- az a folyamat, amelynek során a céltudatos tevékenységek energiát kapnak, irányítottak és fenntartottak (Schunk & Usher, 2012).
- az aktívvá vált hajtóerő; energiaváltozás egy személyben, amelyet az „érzés” megjelenése jellemez, és amelyet megelőz a cél létezésére adott válasz; egy cselekvésből származó válasz, azaz a cél. (Fuqoha et al, 2018);
- az aktív tevékenység folyamatában rejlő, sajátos hierarchiában működő belső feszültség, amely minden emberi tevékenységben alapvető szerepet játszik (Orosz, 1994);
- az „önmagunkról alkotott meggyőződések” alkategóriája (Schukajlow et.al, 2023);

A kutatók számos különböző konstruktumot vizsgáltak, amelyek döntő fontosságúak a tanulók motivációja szempontjából. E konstruktumok tárgyalásának rendszerezése érdekében két nagy csoportra osztjuk őket. Az egyik csoportba tartozik az egyének kompetenciaérzése és a különböző eredmények elérésére irányuló

cselekvőképességük. A másik csoport a belső és külső motivációra, érdeklődésre, értékekre és célokra vonatkozik (Wigfield et al., 2012).

A konkrét motiváció az az ok, amiért egy személy elkötelezi magát egy bizonyos cél mellett. Ez abból a szubjektív jelentésből ered, amelyet a cél a személy számára jelent, és meghatározza a cél követéséhez szükséges kitartást és energiát (Niermeyer & Seyffert, 2011). Tágabb értelemben a motiváció magában foglalja az emberi viselkedés okait, és olyan pszichológiai erőket képvisel, amelyek az emberi viselkedés céljának irányát, intenzitását és tartósságát alakítják (Schukajlow et.al, 2023). A kortársak szintén fontos társadalmi befolyást gyakorolnak a motivációra. Ha a gyerekek szociális támogatást és elfogadást kapnak társaiktól, akkor erősebb a motivációjuk, jobbak a teljesítményük és elkötelezettebbek az iskolában (Wigfield et al., 2012).

A motiváció és az oktatás szempontjából az iskolai motivációt vizsgáló kutatók olyan dolgokat vizsgálnak, mint a diákok elkötelezettsége és érdeklődése a különböző tanulmányi tevékenységek iránt, a diákok választása a tanulmányi tevékenységek közül, kitartásuk a tevékenységek folytatásában, valamint a ráfordított erőfeszítések mértéke (Wigfield et al., 2012). A matematikaoktatásban az érzelmekkel és a motivációval kapcsolatos kutatásokban fordulópontot jelentett az affektusok taxonómiájának kidolgozása, az intenzitás és az időbeli stabilitás dimenziói mentén megkülönböztették a meggyőződéseket, attitűdöket és érzelmeket. A meggyőződéseket a legkevésbé intenzívnak és a legstabilabbnak javasolták. Ezzel szemben az érzelmeket tartották a legintenzívebbnek és legkevésbé stabilnak az affektív konstruktumok közül. Az attitűdök a meggyőződések és az érzelmek között horgonyoztak le az affektus e két kulcsfontosságú jellemzőjét illetően. Fontos, hogy az érzelmeket külön kategóriaként célozták meg, de azok szituációs változatára korlátozódtak, míg az érzelmek diszpozíciós részét nagyrészt figyelmen kívül hagyták. A motiváció az „önmagunkról alkotott meggyőződések” alkategóriájaként szerepelt (Schukajlow et.al, 2023).

A szakértők véleménye alapján azt a következtetést vonták le, hogy a tanulás motivációja az az erőfeszítés, amelyet a személy a tanulási célok miatt tesz, amelyek a szükségesség kérdésére vonatkoznak; nevezetesen annak szükségessége, hogy tegyen valamit a tanulás érdekében, annak szükségessége, hogy a tanulás után eredményeket érjen el, és annak szükségessége, hogy leküzdje a tanulási nehézségeket (Fuqoha et al, 2018). Közülük sokan a tanulóknak a teljesítményfeladatok elvégzésére vonatkozó képességükről és hatékonyságukról alkotott meggyőződésére összpontosítanak, mint a teljesítménymagatartás döntő motivációs közvetítőire (Wigfield et al., 2012). Ehhez

szükséges az előfeltételek biztosítása. Ez magában foglalja a tanulók pozitív hangulatának megteremtését, a tanulási célok tisztázását, az érdeklődést felkeltő problémahelyzetek biztosítását és a tanulók kíváncsiságának felkeltését (Cegléd, 2011). A számológépek egy döntő feltétel mellett egyszerűsíthetik a tanárok szerepét a matematikaórákon: az aktív tanulás elveinek megfelelően alkalmazzák őket. Ez azt jelenti, hogy funkciójuk támogató jellegű marad, és nem válik a tanulás céljává (Guncaga et al., 2019).

Minden tevékenységnek vannak olyan nehézségei, amelyek gyakran kétségbeesést keltenek. Egy tevékenység során felmerülő akadályoknak vagy nehézségeknek azonban arra kell ösztönözniük, hogy még hozzáértőbbek legyünk. Megjelenik a bátorítás, mint ami szükséges a probléma megoldásához. Ezért, ha valaki bajba kerül, meg kell próbálnia javítani magát (Fuqoha et al, 2018).

Azokat a tényezőket, melyek növelik a matematikaoktatás hatékonyságát, kialakítják a tantárgyhoz fűződő pozitív viszonyt és érdeklődést, motiváló tényezőknek nevezzük. Ezek lehetnek: a tananyag tartalmából adódóak; az alkalmazott módszerek, eszközök és munkaformák; az értékelés, valamint a tanár személyiségtulajdonságai (Orosz, 1994). Ezek ellenére megjelenhet a demotiváltság problémája. Cegléd (2011) kiemeli, hogy a demotiváltság oka nagyon gyakran az, hogy a diákok nem tudnak tanulni. A tanulás gyakran abból áll, hogy a tankönyv fölött ülnek, látszattanulást folytatnak (hogy megnyugtassák a szülőket vagy a tanárokat), de nem végeznek valódi munkát, vagy akár értelmetlen verbális „magolást” folytatnak (hogy megnyugtassák a lelkiismeretüket). Egyik gyerek sem lesz képes önállóan tanulni (Cegléd, 2011).

Az oktatás folyamán a tanárt terheli a motiváció kialakítása és fenntartása a tanórák alatt, valamint az egyéni házi feladatok megoldásakor. Ahhoz, hogy ez sikeresen működjön különböző tanulási technikákat alkalmazhatunk.

2. A KUTATÁS MÓDSZERTANA

A tanítás során az oktató tanárok különböző technikákat alkalmaznak annak érdekében, hogy az anyagot jól átadják, és megfelelő környezetet teremtsenek annak elsajátítására. Ezek a technikák arra szolgálnak, hogy a tanárok motiválják a tanulókat ismereteik elmélyítésére, felhívják a figyelmet a közös munka és az ötletelés fontosságára, valamint a tanárral való kommunikációra.

A tanárokat túl gyakran a kutatások passzív fogyasztóivá degradálják, és arra ösztönzik őket, hogy hajtsák végre az eredményeket anélkül, hogy megértenék, hogyan keletkeznek, vagy, hogy maguk is részt vennének a kutatásban (Wright, 2021). A következőkben bemutatom, miből áll egy ilyen kutatás, valamint összefoglalom az előhívásos tanulási technika fogalmait és jellemzőit. Ez a technika különbözik a tanórák elején gyakran alkalmazott kérdezési és visszakeresési technikáktól, amelyekkel ellenőrizni és mérni lehet a tanultakat. Az ötlet lényege, hogy a rövid távú memóriában tárolt információkat idézzük fel, így aktiválva a megtartást. Kutatások számolnak be arról, milyen mértékkel növekszik a tudásszint azoknál a hallgatóknál, akiknél alkalmazták ezt a technikát. Ennek ellenére én a technika motiváló hatását szeretném kiemelni, kutatni.

2.1.1. Az órák megtervezése

A kutatás folyamatának megtervezése a témavezetőm szakmai segítségével történt. A tervezés és a kivitelezés folyamatos volt, ahogy a felmerülő hibákat kiküszöböltem, vagy adott csoport tematikájára nézve fejlesztettem.

Elsőszámú célom az volt, hogy a felsőoktatásban választott tárgyak esetén ne csak a tervezés és tesztelés, de a tanítás folyamata is felügyeletem alatt legyen. Két irányt határoztam meg, mely a csoportok kis létszámai miatt volt szükséges. Mindkét tárgyra tekintve pedagógiai kutatást terveztem, mely során teszttel támogatott oktatás zajlott az előhívásos tanulás mintájára, óra végi mini e-tesztek segítségével. Továbbá kisebb létszámok esetén interjúval, az utolsó kutatás esetén kérdőív segítségével kerestem a választ a tesztek motiváló hatására.

A választott tárgyak egyike a „Matematikai analízis válogatott fejezetei” volt, tavaszi féléves, 15 kontakt órás tárgyként a teljes tananyagot feldolgoztam az kutatás részeként. A be-és kimeneti e-tesztek között 6 kisebb előhívó e-tesztet alkalmaztam, melyek megírását ütemesen, minden páros óra végére terveztem a hallgatók számára.

Másik matematikai tárgy a „Valószínűségszámítás és matematikai statisztika” volt, melyet a kutatás előrehaladtával többször át kellett formálni. Ebben az esetben nem volt lehetőség hosszabb kutatást tervezni a tárgy teljes lefedésével. Első tervezés folyamán az e-tesztek és a hagyományos tesztek összehasonlítására készítettem a szükséges dokumentumokat. Ezután tértem át ahhoz a kutatáshoz, ahol már csak az e-teszteket alkalmaztam. Az első valószínűségekkel foglalkozó kutatásból arra a következtetésre jutottam, csak az első témakör, a Kombinatorika felel meg a számomra.

A kutatási folyamat általánosítás céljából sor került az alapképzés „Informatika tantárgy-pedagógia” kurzusának záróértékelésére szolgáló mérőeszköz fejlesztésére is. A folyamat során a klasszikus tesztelméleti alapokra támaszkodva hoztunk létre az e-tesztet, kiemelt figyelmet fordítva a kimeneti jószágmutatók objektív elemzésére.

A mérőeszköz megbízhatóságának vizsgálata során – amelyről részletesen Papp (2023a) számol be – a számítások rendhagyó módon negatív reliabilitási értéket mutattak. Noha ez a kísérleti mérés a megbízhatósági együtttható tekintetében nem hozott pozitív statisztikai eredményt, a kutatás szempontjából értékes diagnosztikai tapasztalatokkal szolgált. A kapott adatok rávilágítottak a digitális mérőeszközök korlátaira a gyakorlatorientált tárgyak esetében, egyúttal elősegítették a tesztelméleti ismeretek mélyebb, kritikai integrációját a kutatói gyakorlatba.

2.1.2. A minta

A dolgozatban ismertetett longitudinális vizsgálat a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola (II. RF KMF) különböző csoportjainak bevonásával valósult meg a 2020 és 2024 közötti időszakban. A mintavételezés adaptív módon követte az aktuális tantervi ciklusokat, az adott kutatási fázisnak megfelelő kurzusok indításához igazodva.

A kutatás validitásának és statisztikai erejének növelése érdekében az empirikus szakaszban szükségessé vált a minta extenzív bővítése. Ezt a II. RF KMF további hallgatói csoportjai mellett a komáromi Selye János Egyetem bevonásával sikerült realizálni. Az intézmény kiválasztásakor meghatározó szempont volt a Magyarország határain kívül zajló, azonos szaknyelvi alapokon nyugvó magyar nyelvű felsőoktatási kontextus.

A módszertani kialakítás során – az empirikus fázis kivételével – minden résztvevő csoportot kísérleti csoportként kezeltünk. A kontrollcsoportos elrendezéshez a vizsgált intézményi környezetben lokálisan nem állt rendelkezésre a minta bővítésére és

párhuzamos kontrollálására alkalmas, azonos feltételrendszerű csoport, ami miatt a kontrollcsoportos összehasonlítás nagy valószínűséggel torzított eredményekhez és a statisztikai hibaarány növekedéséhez vezetett volna.

3. táblázat. Kutatásokban résztvevő csoportok

Csoport	Tárgy	Kutatás időpontja	Létszám
Számvitel és adóügy (2. évfolyam) II. RF KMF	Valószínűségyszámítás és matematikai statisztika	2020-2021 ősz félév	19
Matematika (4. évfolyam) II. RF KMF	Matematikai analízis válogatott fejezetei	2020-2021 tavaszi félév	8
Matematika (4. évfolyam) II. RF KMF	Informatika tantárgy-pedagógia	2021-2022 ősz félév	8
Matematika (4. évfolyam) II. RF KMF	Matematikai analízis válogatott fejezetei	2021-2022 tavaszi félév	10
Számvitel és adóügy, Pénzügyek, bank és biztosítás (1. évfolyam) II. RF KMF	Valószínűségyszámítás, matematikai statisztika és ökonometria	2022-2023 tavaszi félév	21
Számvitel és adóügy (1. évfolyam), Informatika (2. évfolyam) II. RF KMF, Alkalmazott informatika (Selye János Egyetem)	Valószínűségyszámítás, matematikai statisztika	2023-2024 tavaszi félév	69
Számvitel és adóügy (1. évfolyam), Informatika (2. évfolyam) II. RF KMF	Valószínűségyszámítás, matematikai statisztika	2024-2025 tavaszi félév	25

A kutatás mérőeszközén a 2024–2025-ös tanév során a további módosításokat hajtottam végre, amelynek keretében szerkezeti és tartalmi optimalizálása valósult meg. Az iteratív fejlesztési folyamatot követően a módosított e-tesztet ismételten alkalmaztam egy kis elemszámú mintán. Az alacsony esetszám ellenére az ismételt mérés célja a pilot-jellegű validálás volt.

3. ELŐZETES KUTATÁSOK

A kutatás elsődleges célja egy olyan módszertani útmutató kidolgozása, amely a tesztelméleti alapkövetelményekre és erős jószágmutatókra építve támogatja a pedagógusok mérési-értékelési gyakorlatát. A munka keretében előzetes kutatásokat végeztünk a munkafelület, valamint az e-tesztek szerkesztésére alkalmas platform kiválasztása érdekében, majd ezek segítségével kipróbáltam az általam létrehozott tesztek megbízhatóságát és alkalmazhatóságát. A cél elérése érdekében a következő feladatokat tűztem ki és végeztem el:

- elektronikus teszteléssel kapcsolatos tapasztalatok feltárása a matematikai diszciplínák oktatásában;
- matematikai tesztelek bázisának létrehozása;
- a létrehozott bázis tesztelése a hallgatók tanulmányi eredményeinek nyomon követésére;
- a hallgatók véleményének vizsgálata az elektronikus értékelési eszközök bevezetésének hatékonyságáról.

3.1. Munkafelület megválasztására vonatkozó felmérések

A kutatási tervezésekor elsődleges szempont volt a tudományos eredmények gyakorlati hasznosíthatósága a szélesebb pedagógiai közösség számára. Ennek érdekében a folyamat egy kiterjedt szükségletfelméréssel és diagnosztikai fázissal vette kezdetét, amely két fő célcsoportra fókuszált: a kárpátaljai magyar tannyelvű intézmények matematikatanáira, valamint a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola (II. RF KMF) Matematika és Informatika Tanszékének hallgatóira. A felmérések központi célkitűzése a kutatásra nézve az volt, hogy megtudjam, milyen online felületet érdemes választani az oktatásszervezésre, pontosabban az e-tesztek kiosztására. A két véleményfelmérést követően megvizsgáltam és összehasonlítottam a népszerűbb platformokat. Megvizsgáltam a platformok előnyeit és hátrányait, szerkesztési és beállítási részletekre hívtam fel az érdeklődők figyelmét.

Az első felmérésről három tanulmányban (Papp, 2020; Papp, 2021b; Papp, 2023b) számoltam be, melynek célcsoportja a Kárpátalján, főként általános vagy középiskolában, de akár magasabb szinten, magyar nyelven oktató matematikatanárok voltak.

A kutatás rávilágított a matematikatanárok eszközhasználati preferenciáira és a digitális mérés-értékelés szubjektív megítélésére. Az oktatásszervezési platformok tekintetében a válaszadók részéről a Facebook Messenger, a GeoGebra csoportok, a Google Tanterem és a Microsoft Teams egyaránt hangsúlyos szerepet kap.

Mindeközben már vizsgáltam a korábban felmerült platformokat, szempontokat kerestem az összehasonlításukra, vizsgáltam a tesztszerkesztő felületeket. Így kerültek erre vonatkozó kérdések is a kérdőívbe. Jellemzően ekkor a Google Űrlap alkalmazás került előtérbe, ezt támasztja alá, hogy 19 válaszadó explicit módon jelölte meg elsődleges eszközként, míg egy további résztvevő a Redmenta mellett kiegészítő opcióként hivatkozott rá. Arra a kérdésre, hogy „Az elektronikus tesztek alkalmazása tudta-e helyettesíteni az iskolai írásbeli teszteket?” a válaszadók pozitívan reagáltak, 3 „teljesen helyettesíti”, 10 „nagyban helyettesíti” és 8 „semleges” jelölés érkezett. Ennek ellenére az utolsó kérdésre – „Ön szerint az e-tesztek motiválják-e a hallgatókat a tanulásban?” – 4 „nem motiválják” és 3 „kicsit motiválják” válasz érkezett.

Mivel a platformok és e-tesztek alkalmazása a jövő oktatói körében pozitívan értékelt a korábban kérdezett tanárok által megnevezett oktatási előnyöktől és hátrányoktól függetlenül, ezért úgy döntöttem, hogy az e-tesztek szerkesztése mellett nem csak tudásintzmérő alkalmazására, de a velük való motiválásra is igyekszem hangsúlyt fektetni munkám során.

3.2.E-tesztek szerkesztése és munkafelületek összehasonlítása

Miközben megismerkedtem a tesztelmélet legfontosabb elemeivel, melyeket figyelembe kell venni az itemek létrehozásakor, elkezdtem vizsgálni elsősorban a kérdőíves felmérésekben említett platformokat. A Facebook Messengert, mint a közösségi média eszközt nem vizsgáltam, de tudomásom szerint hamar ki is került az oktatási környezetből, mivel nem volt kellően követhető és átlátható a különböző feladatok kiküldése és ellenőrzése szempontjából.

Az oktatás részeseként és kutatási tervekkel rendelkezve nem tartottam elegendőnek az egyes felületek tesztszerkesztő lehetőségeinek és beállításainak vizsgálatát. Az oktatás szervezése és lebonyolítása alapján általam fontosnak tartott szempontok összegyűjtésével mutattam be a kérdőíves felmérésben említett három platform összehasonlítását, továbbá előnyeit és hátrányait. A vizsgálat fókuszába a

oktatásban domináns Google Tanterem és Microsoft Teams, valamint a mérés-értékelés alapeszközeként funkcionáló Google Űrlapok került.

A specifikus matematikai igények – különös tekintettel a komplex képletbevitelre és a szakszerű szimbólumkezelésre – kiszolgálása érdekében a kutatásba bevontam az EduBase felületet is. Ezáltal nem csupán elméleti összehasonlítást végeztem, hanem praktikus útmutatást és technológiai alternatívákat kínáltam a digitális átállás során módszertani nehézségekkel küzdő pedagógusok számára. A fejlesztési folyamat negyedik pilléréként az OnlineTestPad rendszert is szisztematikusan elemeztem, amelyet a mérések során magas szintű konfigurálhatósága miatt tudatosan alkalmaztam.

A négy platform (Google Tanterem, Microsoft Teams, EduBase, OnlineTestPad) összehasonlító elemzésének eredményeit, valamint a kidolgozott differenciálási szempontokat két önálló tanulmányban (Papp, 2023; Papp, 2024a) publikáltam. Az eredmények átláthatósága és a gyakorlati alkalmazhatóság elősegítése érdekében ezeket a 4. táblázatban összegeztem, amely referenciaként szolgálhat a pedagógusok számára az optimális digitális környezet kiválasztásához.

4. táblázat. Platformok és beépített teszt szerkesztőik összehasonlítása (Papp, 2023, Papp, 2024a).

Szempont	Google Tanterem	Microsoft Teams	EduBase	Online TestPad
Regisztráció	Google fiókkal	Microsoft 365	Google fiókkal	Google fiókkal
Teszt szerkesztése	Google Űrlapok	Google Űrlapok	EduBase Quiz	OnlineTestPad Tests
Teszt kérdések	Egyéni	Egyéni	Egyéni és publikus kérdésbankból	Egyéni
Matematika képletek	Nem kezeli	Nem kezeli	LaTex alapon	LaTex alapon
Itemek keverése	Nincs	Nincs	Beállításoknál megadható	Beállításoknál megadható
Válasz-lehetőségek keverése	Kérdésenként beállítható	Kérdésenként beállítható	Automatikusan keveri	Beállításoknál megadható
Teszt kiküldése	Beállítástól függően hiperhivatkozás birtokában bárki kitöltheti	Beállítástól függően hiperhivatkozás birtokában bárki kitöltheti	EduBase csoportba vagy importálással külsősöknek	Hiperhivatkozás birtokában bárki csoportba rendezés nélkül kitöltheti

Teszt offline módban	Nyomtatás pdf-be (csak feladat)	Nyomtatás pdf-be (csak feladat)	Nem lehetséges	Letöltés pdf-be (feladat és/vagy megoldókulcs)
Időkorlátozás	igen, kitűzött feladatoknál, de e-teszteknél nincs, csak manuálisan lezárható a válaszfogadás	igen, kitűzött feladatoknál, de e-teszteknél nincs, csak manuálisan lezárható a válaszfogadás	Igen	Igen

A digitális oktatási rendszerek mérés-értékelési és oktatásszervezési funkcióinak elemzése során elengedhetetlen a platformok tudományos igényű, rendszerszintű összehasonlítása. Az elemzés célja volt a matematikát oktatók figyelmét felkelteni egy egyszerűen alkalmazható matematikai alapokkal rendelkező platform alkalmazására. A megvizsgált négy platform közül matematikai tanári szempontból az első kettő inkább az oktatás szervezése miatt fontosabb, míg precíz e-tesztek készítésére az EduBase vagy OnlineTestPad felületek alkalmasabbak. Az alábbiakban a rendelkezésre álló keretrendszereket először funkcionális szempontból vizsgálom, majd részletesen bemutatom az Online Test Pad platformot, mint a kutatásom során alkalmazott elsődleges mérőeszközt.

A digitális pedagógiai folyamatok támogatására szolgáló platformok funkciói három fő dimenzió mentén csoportosíthatók:

- Oktatásszervezési funkciók: Ide tartozik a tanulói csoportok adminisztrációja, a tananyagdisztribúció, a határidők kezelése és a szinkron vagy aszinkron kommunikáció (pl. beépített videóhívások, Drive/OneDrive integráció). Ezen a téren a Google Tanterem és a Microsoft Teams nyújtja a legátláthatóbb struktúrát.
- Mérési-értékelési funkciók: Olyan algoritmusok, amelyek lehetővé teszik különböző itemtípusok (nyitott, zárt, feleletválasztó) létrehozását, automatikus pontozását, manuális felülbírálatát és az eredmények statisztikai elemzését (diagramok, Excel-exportálás).
- Szakspecifikus támogatás (Matematikai tartalomkezelés): A reáltudományok mérése során kritikus szempont a LaTeX-alapú képletszerkesztő megléte. Míg a

Google Űrlapok e tekintetben csak külső médiaelemek (képek) segítségével alkalmazható, addig az EduBase és az Online Test Pad natív megoldást kínál.

E-tesztek szerkesztésére a Google Tanterem és a Microsoft Teams azonos beépített alkalmazásokat használnak, melyek szerkesztési és beállítási lehetőségeinek főbb tudnivalóit egyik megjelent tanulmányban mutattuk be a teljesség igénye nélkül (Papp, 2020).

Az Online Test Pad egy többfunkciós, ingyenes tesztelési és képzési szolgáltatás, amely ötvözi az általános oktatási platformok alapvető tulajdonságait a specifikus méréselméleti igényekkel. Mint minden korszerű mérőeszköz, az Online Test Pad is biztosítja a következő, a Google Űrlapok és az EduBase felületeken is ismert alapfunkciókat:

- A tesztek PC-n, táblagépen és mobilokon is optimálisan jelennek meg;
- Lehetőség van időkorlát beállítására, az itemek keverésére és a kitöltések számának korlátozására;
- Lehetőség van az itemek manuális ellenőrzésére, pontok felülbírálására, valamint a részpontozásra;
- Az adatok diagramok formájában vizualizálhatók, illetve nyers adatsorként (Excel) exportálhatók a mélyebb statisztikai elemzéshez.

Az Online Test Pad kiemelkedik a mezőnyből a matematikai képletek kezelése terén. Míg az EduBase szintén alkalmazza a LaTeX-alapú bevitelt, az Online Test Pad egy vizuális párbeszédablakkal támogatja a szerkesztőt, így a parancskódok ismerete nélkül is precíz képletek illeszthetők be. 17 különböző feladattípust kínál, beleértve a párosítást, a szövegkiegészítést és a fájlcsatolást, ami meghaladja a Google Űrlapok korlátozott itemkészletét. Numerikus válaszok esetén beállítható a megengedett kerekítési hiba, ami drasztikusan csökkenti a manuális javítás szükségességét a matematikai feladatoknál. A platform képes az IP-cím alapú (ami azonos hálózaton történő kitöltéskor hátránya is lehet), valamint a megadott név alapján (Kezdőlapra beállítható a kitöltés kötelezőként) történő korlátozásra a regisztrációt és kitöltések számát illetően. Valamint beállítható a vágólap-másolás tiltása, ami növeli a szummatív mérések validitását. A tesztek PDF-formátumba történő exportálása lehetővé teszi a digitálisan létrehozott feladatsorokkal történő offline mérést is, ez akár javítókulccsal együtt letölthető a szerkesztő számára. További előnye, hogy a generált hiperhivatkozás segítségével csoport létrehozása nélkül bármilyen felületen és csatornánk kiosztható az

elkészített teszt vagy más feladat, és bárki kitöltheti, aki rendelkezik a hiperhivatkozással. A kitöltést és esetleges szükséges felülbírálat követően a kitöltők egyéni feladatlapjai a pdf formában letölthetőek.

Hátrányként említhető, hogy a felület csak néhány nyelven érhető el, és a legtöbb nyelven csak a menüpontokat tudja lefordítani, valamint a beállítások néhány pontját. Ennek ellenére a beviteli nyelv bármilyen nyelv lehet, a magyar ékezetes betűket hiba nélkül megjeleníti. További hátrányát képezi, hogy a diagramok feliratait és magyarázatait nem tudja fordítani a szerkesztés nyelvére.

Bár a kérdőíves felmérések célja volt, hogy megtaláljam a kutatáshoz alkalmas eszközt, a bennük említett platformokon kívül találtam meg azt. A kutatásokat és mérlegelő összehasonlításokat elvégezve az OnlineTestPad bizonyult a legalkalmasabb platformnak kutatásom célkitűzéseinek megvalósításához.

3.3.E-teszt szerkesztése és alkalmazása a felsőfokú matematikaoktatásban (1. és 2. ciklus)

A 2020-2021-es tanév kezdetén nagy lendülettel vágtam bele a tesztelmélet és különböző platformok megismerése mellett az e-tesztek szerkesztésének első ciklusába. Több síkon szerettem volna elindulni annak érdekében, hogy a felsőoktatás matematika tárgyaira nézve vonjak le következtetéseket. Fontosnak tartottam, hogy olyan tárgyat teszteljek, ahol az oktatást is magam végezhetem. A kutatás fókuszában a mérőeszköz-fejlesztési kompetenciák elmélyítése állt, biztosítva azok alkalmasságát a hallgatói tudásszintek objektív meghatározására. A tudományos vizsgálat célkitűzései olyan tárgy vagy annak valamely fejezetének kiválasztását tették szükségessé, amely lehetővé teszi a bemeneti mérést az előzetes tanulmányok során rögzült ismeretanyag felidézésére. Az oktatási intervenciót követően elvégzett kimeneti mérések célja az ismeretek reaktiválásának, valamint az elsajátított új tudásszint mértékének precíz számszerűsítése volt.

A tanév őszi félévében „Valószínűesszámitás és matematikai statisztika” tárgy keretein belül tervezem, majd végeztem el a kutatást. A tárgy gyakorlati foglalkozásaira építettem, melyet a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Számvitel és adóügy szakos, alapképzés 2. évfolyamának 19 hallgatójára terveztem. A tárgy felépítésében az alapfogalmak, és alapproblémák jelennek meg ebben az esetben. 10 gyakorlati óra

anyagára dolgoztam ki a teszteket, melyek segítségével a hallgatók tudását és a tudásszintmérés pontosságát vizsgáltam.

A kutatást úgy építettem fel, hogy két digitális mérőeszköz (e-teszt) és egy hagyományos, papíralapú feladatlap eredményeit vettem össze az órai megfigyelésekkel. Az e-tesztek létrehozásához a kérdőíves felmérések alapján legnépszerűbb Google Űrlapokat választottam. A platform korlátai – különösen a matematikai képletbevitel hiánya és a szöveges válaszok automatikus kiértékelésének reliabilitási problémái – miatt a digitális mérések kizárólag zárt, feleletválasztó itemekre korlátozódtak. A mérések eredményeit két szakaszban publikáltam: a teljes csoportra vonatkozó összefüggéseket Papp és Szegő (2021), míg az egyéni hallgatói teljesítményt három hallgató eredményét kiemelve Papp (2022) összegezte.

A pilot kutatás három szekvenciális mérésből állt, a tanítási módszer pedig az ötletbörze volt. A mérések a tanulási folyamat különböző fázisait monitorozták:

Bemeneti e-teszt (Diagnosztikus mérés): A hallgatók maradéktudásának felmérésére szolgált. A 8 zárt item (4 elméleti és 4 gyakorlati feladat) a klasszikus valószínűségszámítás alapjaira fókuszált. A disztraktorokat (téves válaszlehetőségeket) szisztematikusan a jellemző kognitív hibák alapján hoztam létre, így a teszt alkalmas volt a tanórai differenciálás és a feladatnehezítés szintjének meghatározására.

Köztes e-teszt (Formatív mérés): A negyedik órapár végén végzett 5 ítemes mérés a rövid távú ismeretelsajátítást vizsgálta. A feladatok egy része az órai anyag replikációja volt, így a mérés a hallgatói figyelem és a megoldási algoritmusok rögzülésének indikátoraként is szolgált. A disztraktorokat itt az órákon tapasztalt egyéni tévutak és hibás gondolatmenetek alapján alkottam meg.

Záróteszt (Szummatív mérés): A témakör lezárásaként végzett, nyílt itemekből álló hagyományos, papír alapú feladatsor. Ez a mérési forma lehetővé tette a teljes levezetési folyamat elemzését és a részpontozást (feladatonként 5–7 pont). A követelményrendszer előírta a gondolatmenet explicit dokumentálását, ami betekintést engedett a hallgatók matematikai problémafelismerési és modellezési képességeibe. Ezt a döntést az a módszertani megfontolás indukálta, miszerint a kizárólag zárt itemekből (feleletválasztós kérdésekből) álló e-tesztek nem szolgáltattak kellően árnyalt képet a hallgatók valódi kompetenciaszintjéről és kognitív sémáiról.

A záróteszt alapvető elvárása a hallgatókkal szemben a problémafelismerési készség és a procedurális tudás (részmegoldások generálása) demonstrálása volt a tananyag egésze alapján. A 2. ábra a teszt egyik feladatának az egyik hallgató által teljes

folyamat leírásának vélt megoldását mutatja, ami azt bizonyítja, hogy az adott hallgató nem ismeri fel a feladat típusát, ezért helyes részmegoldásokat sem képes létrehozni.

Feladat: Két egyforma urnában piros és fehér golyók vannak. Az első urnában 4 piros és 5 fehér, a másodikban 5 piros és 7 fehér. Találomra kiveszünk egy golyót valamelyik urnából. (Az urnák közül egyenlő valószínűséggel választunk)

- Mekkora a fehér golyók kihúzásának valószínűsége?
- Mennyi annak a valószínűsége, hogy az első urnából fehér golyót húzunk?

3 1 - Urna - 4 piros és 5 fehér
2 - Urna - 5 piros és 7 fehér
a) $P(A|B_2) = \frac{12}{21} = 0,57 = 57\%$
b) $P(A|B_1) = \frac{5}{9} = 0,56 = 56\%$

2. ábra. Feladatlapos számonkérés egyik feladata és annak hallgatói megoldása (Papp & Szegő, 2021; Papp, 2022).

Szembetűnő, hogy az adott hallgató a fehér golyók húzásának valószínűségét tekinti a feladat teljes megoldásának. Ez arra enged következtetni, hogy nem ismerte fel a feladat szövegéből, hogy a teljes valószínűség, és annak segítségével a Bayes-tétel alkalmazása szükségeltetik a feladat teljesítéséhez.

Nem csak jellegében tért el a 2. teszt, mégis szeretném összehasonlítani őket, amihez a Microsoft Excel által kínált lehetőségeket használtam. A kétmintás t-próba eredményei alapján a digitális előteszt és a papír alapú, ötletbörze alapú oktatást követő záróteszt eredményei között nincs statisztikailag szignifikáns különbség ($t(34) = 0,478, p = 0,636$). A záró feladatlap tartalmazott korábban már megoldott elemeket, a fejlődés elmaradása arra enged következtetni, hogy a tanítási módszer (ötletbörze) szabadsága és a papír alapú számonkérés hagyományos jellege közötti módszertani távolság semlegesítette a várt tudásnövekedést, mivel az alapvető fogalmak nagyobb mértékű elsajátítása volt az elvárásom.

A kutatás eredménye alátámasztja a nyílt végű feladatok jelentőségét a matematikai megértés mérése során. Ennek következtében arra a következtetésre jutottam, hogy kombinálni kell az e-tesztek esetén a nyílt és zárt itemtípusokat a megbízható mérés érdekében.

A valószínűségekkel kapcsolatos 2. ciklus a 2022-2023-as tanévben a „Valószínűségszámítás, matematikai statisztika és ökonometria” kurzus keretében zajlott.

Az 1. ciklusban szerzett tapasztalatok alapján a kombinatorika témakört azonosítottam a kutatások optimális egységeként. Ennek tükrében fejlesztettem ki a speciális e-teszt sorozatot, amely mind oktatási, mind diagnosztikai célokat szolgált. A zárótesztek eredményeinek mélyreható elemzését a Papp (2024) tanulmányban összegeztem.

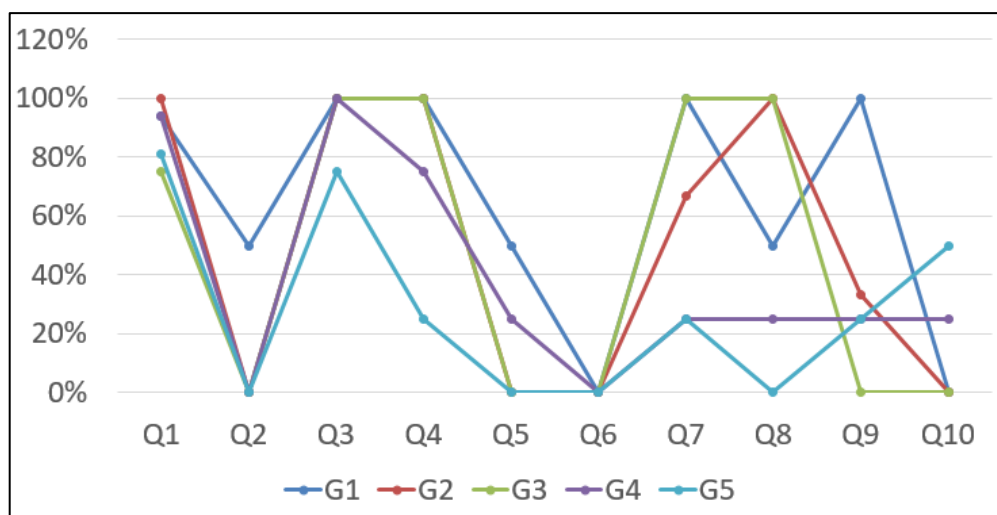
Az említett kurzust a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Számvitel és adóügy, valamint Pénzügyek, bank és biztosítás 1. évfolyamos csoportjaiban, összesen 21 hallgatónak oktattam. A gyakorlati órák keretében végzett mérések az OnlineTestPad platformon keresztül, a Google Tanterembe integrálva zajlottak.

A be- és kimeneti mérések azonosak voltak, 10-10 itemből álló, zárt és nyitott feladatokat egyaránt tartalmazó tesztek alkalmaztam. A kis itemszámot azért véltem indokoltnak, mert az előző kutatás során nehéznek bizonyult a hallgatók számára a kombinatorikához szükséges gondolkodás. A pontosabb mérést az itemek változatosságával szerettem volna elérni, ezért változatosan egy vagy több helyes választ tartalmazó kiválasztó, valamint párosító és rövid választ kifejtő feladatokat tartalmaztak.

Az oktatási folyamatba három köztes mérést iktattam be, amelyek a kombinatorika alappilléreit (permutáció, variáció, kombináció) fedték le. Ezek a 4-4 itemből álló, zárt szerkezetű (feleletválasztó) tesztek voltak. Az alkalmazott módszertan célja az előhívásos tanulás (retrieval practice) szisztematikus integrálása volt, biztosítva a folyamatos diagnosztikai visszacsatolást és a hosszú távú tudásmegőrzés támogatását a matematikai statisztikai alapozás során.

A szummatív záróteszt eredményeinek elemzése során a hallgatói mintát az elért összpontszám alapján öt homogén teljesítménycsoportba (G1–G5) soroltuk. Ez a rétegzett módszertani megközelítés lehetővé teszi az egyes feladatok (Q1–Q10) nehézségi indexének és differenciáló képességének precíz vizsgálatát a különböző képességszintű hallgatók tükrében.

A csoportátlagok eloszlását és az item-szintű sikerességi mutatókat a 3. ábra szemlélteti. A grafikonon a G1 jelöli a legmagasabb teljesítményt nyújtó, míg a G5 a legalacsonyabb pontszámot elért hallgatók csoportját. Az adatok vizualizációja rávilágít az egyes itemek diszkriminációs erejére: a meredekebb különbség a G1 és G5 csoportok között jelzi a feladat hatékony elkülönítő képességét, míg a párhuzamosan alacsony vagy magas értékek az itemek túlzott nehézségére vagy egyszerűségére utalnak. Ez az elemzés elengedhetetlen a mérőeszköz reliabilitásának alátámasztásához, mivel igazolja, hogy a teszt tételei konzisztens módon differenciálnak a hallgatók tényleges tudásszintje szerint.



3. ábra. Hallgatói sikerességi mutatók a záróteszt itemei szerinti bontásban (Papp, 2024).

Az itemek nehézségét vizsgálva a 3. ábrán lévő csoportátlagok segítségével arra a következtetésre jutottam, hogy a differenciáló hatást a Q1, Q3, Q4, Q5 és Q9 itemek érték el, ahol a G1 és G2 csoportok eredményei megelőzik a többiekét. A közepes teszteredménnyel rendelkező G3 csoport szembetűnően ugrál, a csoport tagjai minden itemre vagy egységesen helyesen, vagy helytelenül válaszoltak, így az ő eredményeik alapján nem tudunk következtetéseket levonni az itemek nehézségéről és differenciáló hatásáról.

Szembetűnő, hogy a Q6 kérdés volt a legnehezebb a hallgatóink számára, mivel egyetlen helyes választ sem jelöltek. A Bloom taxonómia eljárási tudás folyamatának és a kognitív folyamatok alkalmazási képességére irányult item nyílt feladat volt, ahol egész megoldás bevitelét vártuk. A feladatot bármely kombinatorikai elemhez csatolás nélkül, logikus gondolkodással ábrázolni lehet, ami segíti a megoldást. A nyílt feladat megoldásához egy egész szám beírására volt szükség, megoldását a 4. ábrán mutatom.

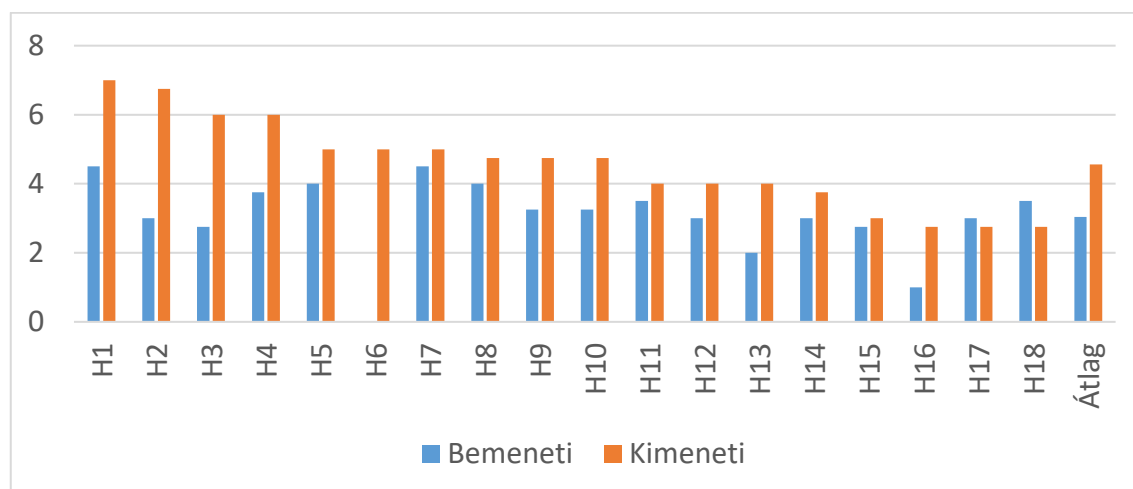
Feladat: Hány ötjegyű páratlan szám képezhető a 0, 1, 2, 3, 4, 5 számjegyekből, ha minden számjegy csak egyszer szerepelhet?

$$\frac{\boxed{}}{4} \cdot \frac{\boxed{}}{4} \cdot \frac{\boxed{}}{3} \cdot \frac{\boxed{}}{2} \cdot \frac{1 \text{ v } 3 \text{ v } 5}{3} = 288$$

4. ábra. Q6 item megoldása.

A platformon a beállításnak köszönhetően minden itemre kötelező a válaszadás ahhoz, hogy a következő feladatot elérjük. Feltételezem, hogy ennek is köszönhető, hogy különböző helytelen válaszok érkeztek. A beérkezett válaszok alapján arra

következtettem, hogy a hallgatók nem fontolták meg kellően a megoldást, inkább az első, de sajnos helytelen művelettel dolgoztak, jellemzően permutációt alkalmaztak, a feladat szempontjából indokolatlanul.



5. ábra. Be- és kimeneti tesztek összehasonlítása.

A kimeneti teszt tudásszintmérő hatását a bemeneti tükrében vizsgáltam. Az 5. ábra mutatja, hogy fejlődés két hallgató kivételével mindenkinél kimutatható. Elvárásom az volt, hogy a bemeneti teszt alacsony pontszámaait több hallgató jóval meghaladja az oktatásnak és az előhívó köztes tesztek rendszerező hatásának következtében.

Az e-teszt differenciáló nehézsége mellett a teszt objektivitására törekedtem azzal, hogy egyértelmű és tesztelméletileg meghatározott helyes válaszok és disztraktorok megadására törekedtem. A pontozás objektivitását elősegítette az e-teszt webes platformon történő beállítása. Ezt úgy végeztem el, hogy minden helyesen megoldott item maximum 1 pontot ért, összesen 10 pontot. A teszt megírására a gyakorlati órán került sor a felügyeletem mellett, hogy a tesztelési folyamat objektivitását biztosítam.

A tesztelmélet alapján tudjuk, hogy a teszt megbízhatóságának értéke több tulajdonságtól is függhet. Bár az itemek szerkesztésekor úgy hoztam létre a tesztet, hogy számuk alapján alacsony megbízhatóságra számíthattam, a kitöltők száma alapján mégis érdemes megvizsgálni azt. A megbízhatóság a helyes megoldások számától is függ, ezért csak a sikeresebb kimeneti tesztet vizsgáltam és elemeztem.

A kimeneti e-teszt megbízhatóságát a Cronbach-alfa (1) segítségével vizsgáltam, amelyet a Microsoft Excel táblázatkezelő program Adatelemzésének Kéttényezős varianciaanalízis (ANOVA) segítségével számoltam ki. A számítást követően a teszt 0,19-es értékkel a 2. táblázat alapján az „Elfogadhatatlan” intervallumába esik.

A be- és kimeneti e-teszteket vizsgáltam továbbá a Microsoft Excel Leíró statisztikájának, valamint a kétmintás t-próba (nem-egyenlő szórásnégyzeteknél) segítségével, mely eredmények alapján a teszteket sikereseknek tekinthetem.

A két mérés adatai között jelentős különbség mutatkozik a fejlődés irányába. A záró e-teszt kitöltése sikeresebb volt, míg a bemeneti teszt terjedelme 4.75 a tesztben elérhető 10 pontból, a kimeneti teszt terjedelme már 7 pont volt. Az átlagérték emelkedése (3,20-ról 4,56-ra) azt jelzi, hogy a csoport teljesítménye javult a két mérés között. A variancia kismértékű növekedése arra utal, hogy a kimeneti tesztnél a hallgatók eredményei valamivel nagyobb szórást mutattak, mint a bemenetinél. A statisztikai elemzés alapján a bemeneti ($M = 3,20$) és a kimeneti ($M = 4,56$) e-tesztek eredményei közötti különbség statisztikailag magasan szignifikáns ($t(34) = -3,41$; $p = 0,0017$). Ez azt jelenti, hogy a hallgatók tudásszintjének növekedése nem a véletlen műve, hanem a vizsgált időszakban történt oktatási folyamat (vagy egyéb intervenció) eredményesnek bizonyult. A 10 tételes mérőeszközök alkalmasak voltak a fejlődés kimutatására, és az elért $p < 0,01$ érték rendkívül erős bizonyítékot szolgáltat a teljesítmény javulására.

A valószínűségi számítás témakörében végzett második kutatási ciklus eredményei megerősítették, hogy a témakör szűkítése a kombinatorikára módszertani szempontból indokolt volt, különösen a bemeneti (diagnosztikus) mérések validitása érdekében. A köztes, előhívó teszteket bár nem elemeztem részletesen, de a kimeneti tesztek eredményeiből a teljesítményjavulás ellenére nem számolhatok be a kombinatorikai elemek sikeres megkülönböztető felismeréséről. Az 5. ábra adatai alapján megállapítható, hogy az órák és a köztes tesztek ellenére sem kimutatható a nagymértékű elsajátítása az adott témának.

Véleményem szerint a kutatás érdekében több változtatást is szükséges eszközölni a teszteken. Elsősorban a be- és kimeneti tesztek jelentős módosítása elengedhetetlen. Bővíteni kell az itemek számát, ki kell egészíteni a meglévő itemeket további differenciáló feladatokkal. Szükség van olyan feladatokra, melyekben egyszerűbb meghatározni, a kombinatorika mely elemét kell alkalmazni, hogy meg tudjuk vizsgálni, melyik hallgató milyen szinten sajátította el a tematikát. Továbbá a megbízhatóság pontos mérésének érdekében szükség van további kitöltőkre, tehát ki kell terjeszteni a kutatás. Amennyiben erre nem találunk megoldást, a felsőoktatás egy másik tárgyára kell koncentrálni.

3.4. E-teszt szerkesztése és alkalmazása a felsőfokú matematikaoktatásban (3. és 4. ciklus)

A kutatási projekt longitudinális szakaszában (3. és 4. ciklus) a vizsgált diszciplínák körét kiterjesztettem a „Matematikai analízis válogatott fejezetei” kurzusra. A tantárgyválasztást a tesztelméleti kutatásaim során azonosított szempontok és a korábbi e-teszt szerkesztési tapasztalatok szintézise alapozta meg. A kurzus a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola matematika alapképzésének záró szemeszterében helyezkedik el, így alkalmas a képzés során elsajátított kognitív sémák stabilitásának mérésére.

A kurzus keretében végzett mérések célja annak igazolása volt, hogy a jól strukturált, tesztelméleti standardokat követő e-tesztek képesek-e érvényesen visszatükrözni a hallgatók legmagasabb szintű szintetizáló tudását a képzési ciklus végén.

A 2020–2021-es és 2021–2022-es tanévek tavaszi féléveiben végzett kutatásaim fókuszában a „Matematikai analízis válogatott fejezetei” kurzus mérési tapasztalatai álltak. A kutatási eredményeket, az itemnehézségi mutatókat és az e-tesztek jóságmutatóit több tanulmányban (Papp, 2021a; Papp, 2023c; Papp, 2024b) publikáltam. Az elemzések során fellépő legfőbb módszertani korlátot a hallgatói csoportok alacsony esetszáma jelentette, ami befolyásolhatja a statisztikai mutatók pontosságát és általánosíthatóságát.

A mérési elrendezés alapját a szigorú tervezési folyamat képezte. A diagnosztikus (bemeneti) és a szummatív (kimeneti) mérésekhez azonos szerkezetű és tartalmú tesztek alkalmaztam a fejlődés pontos nyomon követése érdekében. A mérőeszközök létrehozásához a matematikai képletszerkesztést támogató OnlineTestPad platformot választottam, a tesztek közzététele pedig a Google Tanterem ökoszisztémáján keresztül valósult meg.

Kutatásaimban a köztes beavatkozás további tesztelés volt, melyet a megismert előhívásos tanulásra (lásd. 1.2.1. fejezet) és vele való motiválásra (lásd 1.2.2. fejezet) alapoztam. Hat ilyen mini tesztet készítettem elő, melyeket az adott témák lezárásakor, az óra végén töltöttek ki a hallgatók. Elsősorban motiváló hatására tudtam építeni, mivel késleltetett méréseket nem állt módomban tervezni a tanultak hosszútávú elraktározás vizsgálata érdekében. Ezeket az előhívó e-teszteket Google Űrlapok segítségével, jellemzően feleletválasztó itemekkel hoztam létre, melyek között felmerültek kérdések az előadásból és a gyakorlati órából egyaránt.

A létrehozott be-és kimeneti e-teszt a tesztelmélet által meghatározott minimum 20 itemből állt, melyek zárt és nyitott feladatait a kutatásoknak megfelelő Bloom taxonómiára (1. táblázat) alapoztam. Céлом volt létrehozni egy tudásmérő tesztet, ezért igyekeztem minél pontosabban meghatározni, mely itemek a tudás és kognitív folyamatok dimenzióinak melyik elemeihez tartoznak. Az itemtípusokat próbáltam hozzájuk illeszteni és változatosakká tenni, amennyire azt a téma engedte. A két kutatás között felülvizsgáltam és módosítottam az itemek nyelvezetét és Bloom taxonómia szerinti besorolását (5. táblázat).

5. táblázat. Taxonómiák és itemtípusok összefoglaló táblázata (Papp, 2021a; Papp, 2023c, Papp, 2024b).

Bemeneti teszt		Be- és kimeneti teszt		Kimeneti teszt	
Bloom taxonómia	Darab	Itemtípusok	Darab	Bloom taxonómia	Darab
B1	5	Egyszeres választás	6	B1	5
B2	2	Többszörös választás	1	B2	2
B3	2	Igaz/Hamis	4	B3	2
C1	4	Párosítás	5	C1	4
C2	2	Numerikus	2	C2	2
C3	5	Szövegbevitel	2	C3	4
Összesen	20	Összesen	20	D3	1
				Összesen	20

A két kutatásban a kitöltők száma majdnem megegyezett, a tesztek jóságmutatói viszont eltérőek voltak a megbízhatóság jelentős eltérése miatt. Ez alapján azt a következtetést vontam le, hogy a javítások megfelelőek voltak. Minden item elérhető pontszáma maximálisan 1 volt, részpontokat csak a párosítás itemek esetén állítottam be, valamint a szöveges itemek felülvizsgálatakor adtam.

A kutatás során alkalmazott digitális mérések objektivitását az automatizált OnlineTestPad felület biztosította, amely minimalizálta az értékelői szubjektivitást. A manuális korrekció szükségessége kizárólag a nyitott végű (szövegbeviteli) itemekre korlátozódott, ahol a rendszer bizonyos karakterkészleteket vagy ékezeteket tévesen interpretált. Ezen felülbírálati eljárások során az értékelés objektív fenntartása érdekében a javítási kulcs szigorú betartásával jártam el.

A teszt validitását (érvényességét) a mérési célok és az elért tudásszintek koherenciája igazolta; a visszajelzések és az item-szintű korrekciók (például egy

technikai jellegű jelölési hiba javítása a második ciklusban) tovább fokozták a mérés tartalmi érvényességét.

A mérőeszköz reliabilitásának (megbízhatóságának) vizsgálata során a Cronbach-alfa (1) együttható jelentős ingadozását tapasztaltam a két kutatási ciklus között. Ezeket a mutatókat a Microsoft Excel beépített Adatelemzésének ANOVA varianciaanalízis segítségével újraszámolt eredményekkel mutatom be:

Első szakasz (3. ciklus, $N = 8$): A kimeneti teszt reliabilitása kritikus szintet, $\alpha = 0,52$ -es értéket mutatott, amely a tesztelméleti standardok szerint az „rossz” kategóriába (2. táblázat) sorolható. Ezt az eredményt a minta rendkívül alacsony varianciája és az alacsony esetszám magyarázza.

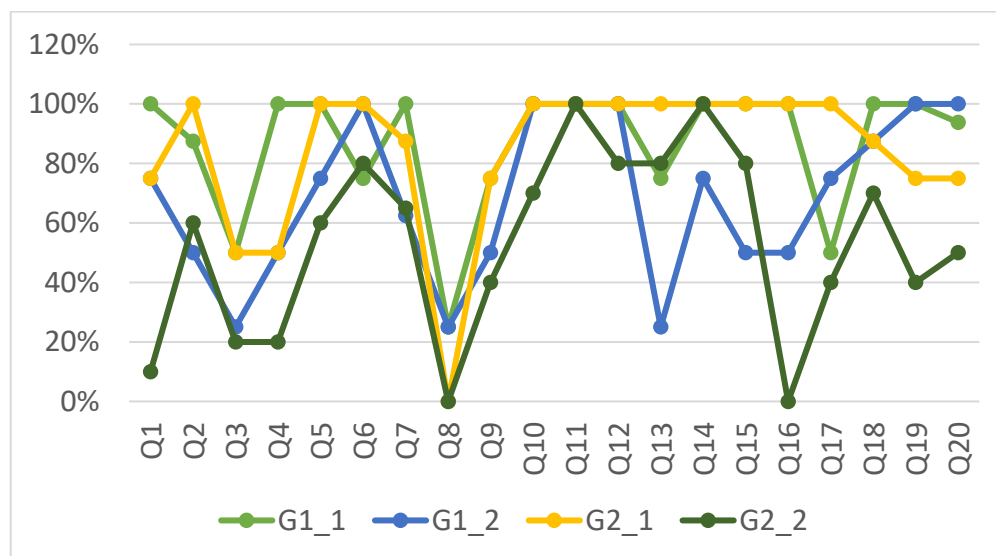
Második szakasz (4. ciklus, $N = 7$): A szisztematikus item-javítások és a tapasztalati úton alapuló módosítások hatására a megbízhatósági mutató $\alpha = 0,78$ -ra emelkedett. Ez az érték az „elfogadható” tartomány felső határán mozog, és szignifikáns módszertani fejlődést jelez az első méréshez képest.

Míg a tanulmány (Papp, 2024b) esetén a be- és kimeneti tesztek közötti fejlődések kimutatása miatt csak 7 hallgató eredményét használtam fel, szerettem volna megvizsgálni csak a kimeneti tesztre nézve a 9 kitöltő esetében milyen megbízhatóságot találok. Ez azért volt fontos számomra, mert az itemek száma mellett a résztvevők számától is függ a megbízhatóság. A két hallgatóval kiegészítve a táblázatot már 0,82 megbízhatóságot állapíthattam meg Cronbach-alfa segítségével. Ez jelentős fejlődés a tesztre nézve, mivel a „jó” kategóriába került.

A Microsoft Excellel végzett leíró statisztika alapján mindkét teszt sikeresnek tekinthető, bár a javított záró teszt a 9 kitöltő eredménye alapján nagyobb eltérést mutat a várható érték és a medián között. A 2020-2021-es tanévben kitöltött teszt terjedelme 7,5, várható értéke 15,5 és mediánja 15,1 értékeket mutat. A felülvizsgálás és javítás ellenére a 2021-2022-es tanévben rosszabb eredményeket kaptam, a kitöltött teszt terjedelme 9,25, ami nagyobb hiányosságot mutat a tesztet legjobban és legrosszabbul kitöltő hallgatók között. Várható értéke is rosszabb eredményt mutat, 12,5-re csökkent, miközben mediánja a csökkenés ellenére is magasabb tőle.

A megjelent tanulmányokban foglalkoztam az itemekkel, főként azt vizsgálva, melyek tűnnek nehezebbeknek a hallgatók számára a kitöltések alapján. Ehhez a hallgatókat 2 csoportba rendeztük, 50 % fölött (G1) és alatt (G2) teljesítőkként. Itemek (Q1, Q2, ..., Q8) kitöltöttségét csoportonként átlagolva ábrázoltuk egy diagram

segítségével. Az összehasonlító diagramot a 6. ábra mutatja, melyen G1_1 és G1_2 az első mérés 2 csoportját, míg G2_1 és G2_2 a másodikat jelölik.



6. ábra. Összehasonlító diagram az itemekre nézve.

Összehasonlításakor szembetűnővé vált, hogy a létrehozott itemeknek differenciáló hatása van. A differenciált tudásról a feladatok vélt vagy valós nehézsége számol be. A nehezebbnek bizonyuló feladatok segítségével tudunk a hallgatók tudásáról pontosabb képet kapni. A 6. ábra – ahol mindkét kutatásban több feladat esetén is magasabb a csoportátlag az 50 % fölött (G1_1 és G2_1) teljesítőknél – mutatja, hogy milyen különbségek születtek az itemek megoldottságában. Az órai aktivitás alapján megfigyelt nagyobb tárgyi tudással rendelkező hallgatók esetén az elvárás az volt, hogy több itemre adnak helyes választ, így nagyobb pontszámot érnek el a tesztek során. Ez mindkét kutatásban több item esetén meg is valósult.

Az egyik nehéznek bizonyult item csak a G2 csoportok számára volt nehezebb mindkét mérésben. Míg az első mérésben a gyengébb tudást mutató csoportban is akadt helyes válasz, addig a későbbi mérésben egy sem.

16
16 from 20

Határozza meg az $f(x) = -x^2 + 6x - 4$ és $g(x) = 1$ függvények által határolt terület nagyságát az $x \in [2, 5]$ intervallumon!

9

7. ábra. A tesztben szereplő Q16 item (Papp, 2024b)

A 7. ábrán a helyes megoldás is szerepel a Q16 itemről. A feladat nehézségét adhatta, hogy számításokat kellett elvégezni, aminek az eredményét meg kellett adni a beviteli mezőben. Itt az üresen hagyott mezőkön kívül születtek helytelen megoldások is, ami arra utal, hogy az eljárás alkalmazása nem alakult ki helyesen a hallgatókban.

A 2021-2022 tanévben lévő teszteléseket követően kvalitatív kutatás keretein belül rövid interjúkat készítettem azokkal a hallgatókkal, akik vállalták azt. Válaszokat kerestem a hallgatói vélemények alapján az előhívásos tanulási technika alkalmazására és annak motivációs hatására. Az interjúkat egyenként, informális, tanórán kívüli beszélgetések során készítettem a teszteket kitöltő hallgatókkal, így a tervezett kérdések mellett a válaszok függvényében változtatásokat is eszközöltem. Az interjút a Papp (2024b) tanulmányban mutattam be.

Nézzünk most néhány fontos, az interjú során felmerült kérdést és az arra adott hallgatói válaszokat, mely a motiválás lehetőségeit vizsgálva a későbbiek során is fontosak lehetnek:

Véleménye szerint a korábban megszokott tanulási technikát jobb-e alkalmazni, vagy azt, amit most tanult?

- „Bár én inkább a régi számonkéréshez vagyok hozzászokva, hiszen korábban a tanárnak feleltünk, vagy tízperces dolgozatokat írtunk az órán, ez nagyon tetszett. Most talán jobban odafigyelek a fontos dolgokra az órán, abban a reményben, hogy majd kérdésként előjönnek (S4).”
- „A tesztek nagyszerűek voltak, legalább nem kellett annyit írnom (S1).”
- „Jobban szeretem a korábbi módszereket, mert jobban fel tudok készülni, most pedig, bár a mini-tesztek jók voltak, úgy éreztem, hogy felkészületlenül kezdtem a tesztet, és könnyen összekeverhettem a dolgokat, és elrontottam a megoldást. De nem így alakult, és ennek mindig örültem, mert végül jól sikerültek (S5).”

Hogyan érzi, a motivációs tesztek sikerei milyen hatással voltak Önre, segítettek-e a záró dolgozatra való felkészülésben?

- „Az első órán nagyon meglepett a teszt eredménye. Azt hittem, hogy jobban emlékszem, de végül nem lett túl jó az eredmény. Aztán a mini-tesztek nem voltak túl nehezek, még ha nem is sikerült mindet teljesen jól megoldanom, de 1-2 óra alatt csak arra kellett emlékezmem, amit tanultam. Természetesen készültem a záróvizsgára, felkészülés közben voltak olyan részletek, melyekre emlékeztem, hogy benne voltak a tesztekben, és megpróbáltam utána nézni (S8).”

- „A mini-tesztek jók voltak, de nem volt teljesen jó eredményük. Talán egy kicsit jobban odafigyeltem rájuk, de nem sikerültek elég jól ahhoz, hogy a teszt jó legyen (S10).”
- „Mindenképpen pozitív hatásként tudnám leírni, hogy a tesztek miatt megpróbáltam jobban koncentrálni az órán. Nem azt mondom, hogy később mindent megjegyeztem, vagy mindent tudtam, ez a teszteken is látszik, de jó volt mindig emlékezni arra, amit nemrég tanultunk. Szerettem volna jól teljesíteni a tantárgyból, és az utolsó teszt nem ezt tükrözi, de legalább jobb volt, mint amit az első órán írtunk (S4).”

Motiválták-e Önt a tesztek és azok eredményei?

- „Nem igazán motiváltak a tesztek. Nem készültem rájuk többet, mint amennyit egyébként tettem volna (S1).”
- „Az első teszt motivált a leginkább, mivel az nagyon rosszul sikerült. A többinél már jobban figyeltem, hogy a saját elvárásaimnak meg tudjak felelni a tanulásban (S8).”

Alkalmazná-e ezt a technikát a jövőbeni munkájában?

- „Szerintem a középiskolában nehezebb, mert kevesebb idő jut mindenre. De jónak tűnik, lehet, hogy kipróbálom (S4).”
- „Sokat kell készülni, ha az e-teszteket akarjuk használni, és lehet, hogy egy iskolában nincs internet, vagy a diákoknak tilos a telefon használata, és nem hozzák magukkal. Ha pedig nem e-tesztről van szó, akkor több időbe telik kijavítani az összes dolgozatot, bár ez a hagyományos tízperceseknél is így van. Nem tudom, hogy belemennék-e, de lehet, hogy igen (S5).”

A Matematikai analízis válogatott fejezetei tárgyra nézve a kutatások alapján azt a következtetést vontuk le, hogy a tesztelmélettel megismerkedve bár szerkesztés szempontjából időigényes, de mindenképp jól alkalmazható az e-teszt mint tudásszintmérő és mint motiváló eszköz. Itemek további módosításával, esetleg cseréjével és a kitöltők számának növelésével a megbízhatóság is tovább növelhető. Mindezek arra engednek következtetni, hogy a felsőoktatás matematika tárgyaira is jól alkalmazhatók az e-tesztek.

4. EMPIRIKUS KUTATÁS

Az előzetes kutatási szakaszok elemzése alapján, bár a „Matematikai analízis válogatott fejezetei” kurzushoz kapcsolódó mérőeszköz mutatott stabilabb reliabilitási mutatókat, a kutatásfejlesztés irányát stratégiai okokból a kombinatorika területére helyeztem át. Ezt a döntést elsősorban a mintabővítés horizontális lehetőségei és az intézményközi együttműködések keretei indokolták.

A mérési megbízhatóság és a validitás növelése érdekében a kutatási design-t az alábbiak szerint módosítottam:

- A matematikai szaknyelv precíz alkalmazása mellett kiemelt figyelmet fordítottam a gyakorlati feladatok szövegezésére. A cél a szemantikai kétértelműségek kiküszöbölése volt, biztosítva, hogy a hallgatói válaszokat ne az értelmezési nehézségek, hanem a tényleges matematikai kompetenciák befolyásolják.

- 2024 tavaszán (5. ciklus) a kutatást kiterjesztettem a komáromi Selye János Egyetem Alkalmazott informatika szakos hallgatóira, valamint a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Informatika szakos csoportjaira. Ezzel a kitöltői létszámot 69-re sikerült emelni. A minta homogenitását az biztosította, hogy minden bevont csoport nem-matematika szakos hallgatókból állt, így a tematika mélysége és az óraszámok összevethetőek maradtak.

- Mivel az 5. ciklus statisztikai mutatói még nem érték el az elvárt reliabilitási szintet, 2025 tavaszán (6. ciklus) további 5 releváns itemmel bővítettem a mérőeszközt. A strukturálisan átdolgozott tesztet a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Számvitel és adóügy, valamint Informatika szakos 25 hallgatója töltötte ki. Ez a mérés a tesztet támogatót oktatást kihagyva a témakör záró akkordja volt, vizsgálva az ismeretek tartósságát.

4.1. E-tesztek létrehozása

A 3.3. alfejezetben kiemeltem, hogy az első ciklus során alkalmazott tematika nehéznek bizonyult a tudásszint vizsgálata szempontjából. Abban a kutatásban nem volt megalapozott az előteszt, mivel korábbi tanulmányok során valószínűségszámítást nem tanultak hallgatóink. Az alfejezetben összehasonlításra nem alkalmas 2. ciklusban alkalmazott, valószínűségszámítás megalapozásához szükséges kutatás tesztjeit továbbfejlesztettem és a fent említett csoportokban az oktatás során alkalmaztam. Ebben

a fejezetben összefoglalom, milyen tervezés és kivitelezés segítségével sikerült létrehozni a kutatási kérdések megválaszolásához szükséges e-teszteket.

A korábbi, 10 íte mből álló e-tesz továbbfejlesztése során a be- és kimeneti e-tesztek nagymértékű átalakításon estek át (lásd 6. táblázat). Alapját a 2. ciklus feladatai alkották, további 10 íte mmel bővítve elértem a tesztelméleti minimum íte mszámot.

6. táblázat. Be- és kimeneti teszt szerkesztési információi.

№	Téma			Bloom-taxonómia	Íte mtípusok						Nehézség		
	Perm., ism. perm.	Var., ism. var.	Komb., ism. komb.		Feleletválasztó (1 helyes)	Feleletválasztó (több helyes)	Igaz-hamis	Párosítás	Rövid válasz	Bepótlás	Könnyű	Közepes	Nehéz
1	+			B1						+	+		
2			+	C4			+				+		
3	+	+	+	B1				+			+		
4	+			B3					+		+		
5	+			C3	+							+	
6	+			B2					+			+	
7	+			C3	+								+
8		+		C3					+			+	
9		+		C3	+							+	
10		+		C2		+						+	
11		+		C4					+		+		
12			+	C3	+							+	
13			+	B3		+					+		
14			+	C3	+								+
15			+	C3	+								+
16			+	C4			+				+		
17			+	C3					+			+	
18			+	B1						+	+		
19			+	C4	+							+	
20			+	D3					+				+
Ö	6	5	11	20	7	2	2	1	6	2	8	8	4

A kutatás mérési elrendezése a bemeneti (diagnosztikus, Függelék A) és kimeneti (szummatív, Függelék E) e-tesztek elemzésére. A mérőeszközök fejlesztése során kettős célkitűzést követtem: egyrészt a hallgatók kombinatorikai maradéktudásának vizsgálatát az oktatási szakasz kezdetén, másrészt a tudásszint-változást a kurzus lezárásakor. A validitás és a reliabilitás biztosítása érdekében a két mérőeszközt szigorú módszertani ekvivalencia mentén strukturáltam, így a be- és kimeneti tesztek itemtípus-eloszlása, becsült nehézségi indexe, valamint a Bloom-féle kognitív taxonómia szerinti besorolása teljesen megegyezik (6. táblázat). Míg a bemeneti mérésnél az eltérő oktatási háttér miatt az alacsonyabb pontszámok is elfogadhatók a kiindulási szint meghatározásához, a kimeneti teszt már az oktatási folyamat és az e-tesztekkel támogatott tanulás hatékonyságának szummatív indikátoraként szolgál.

A rövid idejű kutatás miatt feltételezhető, hogy az információk nagy része még a rövidtávú memóriában is megtalálható. Ennek kiküszöbölése érdekében az egyes itemek szövegkörnyezetén kisebb változtatásokat eszközöltem. Figyelmet fordítottam arra, hogy a változások a feladat újbóli megoldására készítsék a hallgatókat, nem pedig emlékezetből történő válaszok megadására. A változások például a Q5 feladatban megnevezett gyerekek a bemeneti tesztben egymás mellé szeretnének ülni, míg a kimeneti tesztben nem, Q8 itemben kezdeti felmérő teszt a páratlan számokat, míg a záró teszt a páros számokat keresi, de vannak olyan jellegű szövegkörnyezeti változások, melyek a helyes választ nem érintik.

Az be- és kimeneti e-tesztek teljesítésekor maximálisan 20 pontot szerezhettek a hallgatók. Az itemek helyes megoldás esetén 1, helytelen megoldás esetén 0 pontot értek. A büntető pontok, pontlevonások beállítását nem tartottam indokoltnak, valamint a feladatok nehézségi szintjei nem differenciáltak a pontozás szempontjából. Részpontok a több helyes válasz megadása és a párosítás típusoknál voltak szereshetőek.

A tanulási folyamatban a hallgatók a valószínűségszámítás alapjaként sajátítják el a kombinatorika elemeit, a gyakorlati feladatok során felelevenítve korábbi ismereteiket. Az teszttel támogatott tanulás órai környezetét az előhívásos tanulás (retrieval practice) modelljére alapoztam: a tanórák végén alkalmazott formatív mérések célja a hallgatói motiváció fenntartása és az aktív kognitív figyelem ösztönzése volt. Ezek a tesztek az információk rövid távú memóriából történő előhívásával segítik az ismeretanyag tartós rögzülését.

Az előhívást támogató minitesztek – melyek szerkesztési adatait a 7. táblázat tartalmazza – egységesen 5 itemből állnak (Függelék B, C, D), lefedve a kombinatorika alaptípusait és azok ismétléses változatait. A tervezés során az itemeket a Bloom-féle taxonómia szerint klasszifikáltam, ügyelve a becsült nehézségi szintek szisztematikus elosztására. A vizsgálat rávilágított, hogy a hallgatók számára a tesztek záró, nyílt végű feladatai jelentették a legnagyobb kognitív kihívást.

7. táblázat. Előhívó mini tesztek szerkesztési információi.

№	Téma						Bloom-taxonómia	Itemtípusok			Nehézség		
	Permutáció	Ism. permutáció	Variáció	Ism. variáció	Kombináció	ism. kombináció		Feleletválasztó (1 helyes)	Feleletválasztó (több helyes)	Rövid válasz	Könnyű	Közepes	Nehéz
1	+						B1	+			+		
2	+						B1	+			+		
3	+						B2	+				+	
4		+					B2	+				+	
5	+						C3			+			+
Ö	4	1					5	4		1	2	2	1
1			+				B1	+			+		
2			+	+			B1		+		+		
3			+				B2	+				+	
4				+			B2	+				+	
5			+				C3			+			+
Ö			4	2			5	3	1	1	2	2	1
1					+		B1	+			+		
2					+		B1	+			+		
3						+	B2	+				+	
4					+		B2		+			+	
5					+		C3			+			+
Ö					4	1	5	3	1	1	2	2	1

4.2. E-teszt, mint mérőeszköz

Az e-tesztek kitöltésének 5. ciklusára a 2023-2024-es tanév tavaszi félévében került sor az említett csoportokban. A be- és kimeneti teszt elvégzésére 45 perc állt a hallgatók számára, melyet az OnlineTestPad felületen időzítettem, míg az előhívó mini teszteknel eltekintettem az időzítéstől. Minden tesztet az óra keretén belül kaptak meg és töltöttek ki a hallgatók, így felügyelni tudtam (ahogy a Selye János Egyetem oktatója is), hogy sem egymástól, sem jegyzetből vagy internetről nem használnak segítséget, önálló gondolatok alapján számolnak és válaszolják meg az előkészített itemeket.

A 69 résztvevőből eltérő számban töltötték ki a különböző teszteket. Ez részben szándékos volt, mivel a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola informatika szakos hallgatók esetén kontroll csoportként terveztem meg a kitöltést, így ők az előhívó mini teszteket nem kapták meg. Bár kis létszámú résztvevőkről van szó, ezzel szerettem volna vizsgálni, milyen hatása van az előhívó teszteknek a hallgatók teljesítményére. A kontrollcsoport tervezésekor az volt a hipotézisem, hogy azoknak a hallgatóknak, akik az előhívó tesztekkel is rendszeresen találkoztak nagyobb lesz a motiváltságuk az aktív figyelem fenntartására a gyakorlati órák folyamán, aminek hatására sikeresebb lesz a kimeneti teszt eredménye. A kitöltések további eltéréseit az órai hiányzások okozták, mivel a teszteket nem tartottam célszerűnek utólagosan pótolgatni a kutatás szempontjából. Emiatt az elemzés egyes részeinél különböző mintáról beszélek, miközben a következő kérdésekre keresem a választ:

H1. Az e-tesztekben a hallgatók teljesítménye szignifikánsan romlik a Bloom-taxonómia magasabb kognitív szintjeit (pl. alkalmazás, elemzés) igénylő feladatoknál, függetlenül az előzetesen meghatározott nehézségi szinttől, valamint a mini tesztek kitöltésétől.

K1. Mely kognitív szinthez tartozó feladatok bizonyultak a legnehezebbnek a kombinatorika tesztben?

K2. Összhangban van-e a tanári előzetes nehézségi becslés a hallgatói eredményekkel (nehézségi index)?

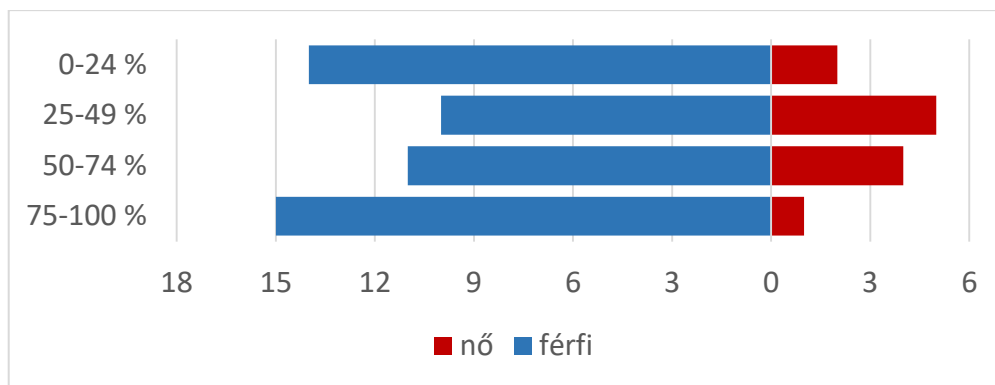
K3. Milyen hatással van az előhívó e-teszt egy tudásmérő tesztre?

A kutatást úgy terveztem meg, hogy az e-teszt objektivitásáért az alkalmazott OnlineTestPad felel, mely a szerkesztőtől és oktató tanártól függetlenül, az előre beállított helyes válaszok vizsgálatával javítja a hallgatók eredményeit. Ezt erősíti meg az a tudatos

szerkesztés, mely során a különböző itemeket a Bloom taxonómia dimenzióiba soroltam, nyílt és zárt típusaikat szükség szerint és gondos odafigyeléssel megválasztottam, így hozva létre olyan teszteket, melyek objektíven mérik az előrehaladást. A tesztelési folyamat objektivitását a segédeszközök alkalmazásának kiküszöbölésével értem el, ennek érdekében tantermi körülmények között, tanári felügyelet mellett töltötték azt ki a hallgatók.

A pedagógiai mérés eredményeinek validálása során párosított kétmintás t-próbát végeztem, amellyel a tanulók kombinatorikai kompetenciájának változását vizsgáltam a kezdeti és a záró teszt eredményei alapján. A vizsgálatban részt vevő $N = 37$ fős minta (akik minden tesztet kitöltöttek) eredményei egyértelmű és statisztikailag alátámasztható fejlődést mutatnak. A leíró statisztikai adatokból kiolvasható, hogy míg a kezdeti teszt során az átlagpontszám 16,15 pont volt, addig a záró tesztre ez az érték jelentősen, 27,85 pontra emelkedett. Ez az elmozdulás a hallgatók tudásszintjének markáns javulását jelzi. Érdeemes megfigyelni a variancia alakulását is: a kezdeti mérésnél tapasztalt magasabb szórásnégyzet (71,15) a záró mérésre lecsökkent (50,33). Ez azt sugallja, hogy az oktatási folyamat végére a tanulók tudása nemcsak magasabb szintűvé, hanem homogénebbé is vált, tehát a csoporton belüli különbségek mérséklődtek, a gyengébb teljesítményt nyújtók felzárkóztak. A két mérés közötti kapcsolatot a Pearson-féle korrelációs együttható ($r = 0,53$) közepesen erős pozitív értéke igazolja, ami a mérőeszköz megbízhatóságát és a tanulási folyamat folytonosságát támasztja alá. A statisztikai próba legfontosabb mutatója, a t-érték (-15,39) és a hozzá tartozó p-érték ($p < 0,001$) messze elmarad a kritikus 0,05-ös szignifikanciaszinttől. Ez tudományos szempontból azt jelenti, hogy a tapasztalt fejlődés rendkívül nagy valószínűséggel nem a véletlennek, hanem a célzott oktatási módszertannak és a tananyag sikeres elsajátításának köszönhető. Az eredmények tükrében kijelenthető, hogy a vizsgált minta esetében a kombinatorikai ismeretek átadása hatékonyan ment végbe, és a diákok szignifikánsan jobb eredményt értek el a kognitív szinteket érintő záró mérés során.

A továbbiakban a szummatív e-tesztet az adott teszt összes kitöltőjének eredménye által ($N=62$) vizsgálom, melynek pedagógiai célja volt a helyzetfeltáró minősítés, ahol differenciáltan tudom értékelni az egyes hallgatók tudásszintjét a témakörben. Ezt a célt a teszt validitása (érvényesség) vizsgálata szempontjából tűztük ki. Ennek sikerét mutatja be a 8. ábra, melyen a hallgatók teljesítményük alapján a négy csoportba sorolva közel azonos létszámban (15-16 fő) szerepelnek.



8. ábra. Hallgatók eloszlása különböző teljesítési intervallumban.

A 8. ábra a záró teszt differenciáló erejét mutathatja be. A tudásszint mérőeszközeként kijelenthetem, a kitűzött célt elértem a teszt segítségével. Az e-teszt egyidejűleg gyors visszajelzést kínál mind az oktató, mind a hallgató számára az eredmény megjelenítésével a kitöltést követően. Ez a funkció nagyobb létszám esetén válik fontossá, mikor a nagyszámú, nyílt itemekből álló felmérő dolgozat kijavításának ideje nem akadályozza a visszajelzést, melyre a hallgatóknak is szükségük van a tanulás folyamán.

Fontos megjegyezni, hogy az e-teszt érvényességét az is biztosítja, hogy a különböző oktatóktól, gyakorlati foglalkozások alatt elhangzott feladatoktól és magyarázatoktól függetlenül minden csoportnál kimutatható a tájékoztató jellegű eredmény a hallgatók különböző tudásszintjéről.

A kombinatorika záró teszt eredményeinek feladatszintű vizsgálata során a Bloom-féle kognitív taxonómia mentén (1. táblázat) világos mintázat rajzolódik ki. Az N=62 fős mintán végzett elemzés adatai alapján megállapítható, hogy a teszt belső szerkezete szoros korrelációt mutat a kognitív igénybevétel és a feladatok tényleges nehézsége között. A 8. táblázat adatai szerint a legalacsonyabb kognitív szintet képviselő emlékezés kategóriába tartozó feladatoknál a hallgatók kimagasló hatékonysággal teljesítettek, ami jelzi, hogy a definíciók és alapképletek felismerése nem okozott nehézséget. A tanulók magabiztosan azonosították a permutáció és a faktoriális fogalmát, ami stabil elméleti alapokra utal.

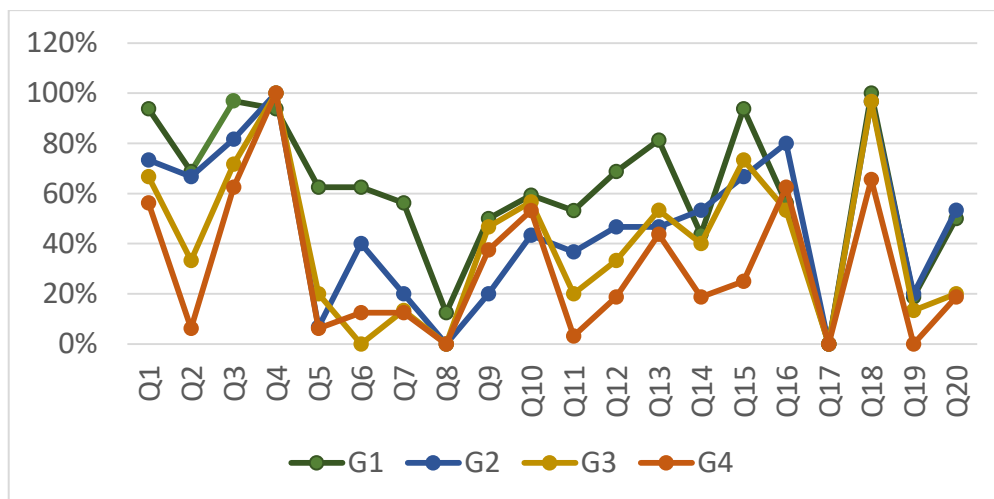
8. táblázat. A feladatok nehézségi mutatóinak és a tanári előrejelzések összehasonlítása

Taxonómia	Item	Tanári nehézségi szint	Hallgatói nehézségi szint	Feladat leírása
B1	1	könnyű	0,73 (könnyű)	Ismeret felidézése
	3	könnyű	0,78 (könnyű)	Ismeret felidézése

	18	könnyű	0,89 (könnyű)	Ismeret felidézése
B2	6	közepes	0,29 (nehéz)	Képlet használata
B3	4	könnyű	0,98 (könnyű)	Képlet használata
	13	könnyű	0,56 (közepes)	Modellválasztás
C2	10	közepes	0,53 (közepes)	Modellválasztás
C3	5	közepes	0,24 (nehéz)	Modellválasztás
	7	nehéz	0,25 (nehéz)	Képlet használata
	8	közepes	0,03 (nehéz)	Modellválasztás
	9	közepes	0,39 (közepes)	Modellválasztás
	12	közepes	0,42 (közepes)	Képlet használata
	14	nehéz	0,39 (közepes)	Összetett probléma
	15	nehéz	0,65 (közepes)	Összetett probléma
	17	közepes	0 (nehéz)	Képlet használata
C4	2	könnyű	0,44 (közepes)	Értelmezés, felismerés
	11	könnyű	0,28 (nehéz)	Képlet használata
	16	könnyű	0,63 (közepes)	Értelmezés, felismerés
	19	közepes	0,12 (nehéz)	Értelmezés, felismerés
D3	20	nehéz	0,35 (közepes)	Értelmezés, felismerés

A **K1** kutatási kérdésre válaszolva, a legnehezebbnek bizonyult kognitív szintek egyértelműen az elemzés és az értékelés/alkotás kategóriái voltak. Különösen az 5., 7., 8., 11., 17. és a 15. feladatok mutatják a legalacsonyabb megoldottságot ($p < 0,35$). Ezek a tételek nemcsak a képletek ismeretét, hanem speciális logikai kényszerfeltételek kezelését is megkövetelték, ami a hallgatók többségénél modellválasztási nehézségekhez vezetett.

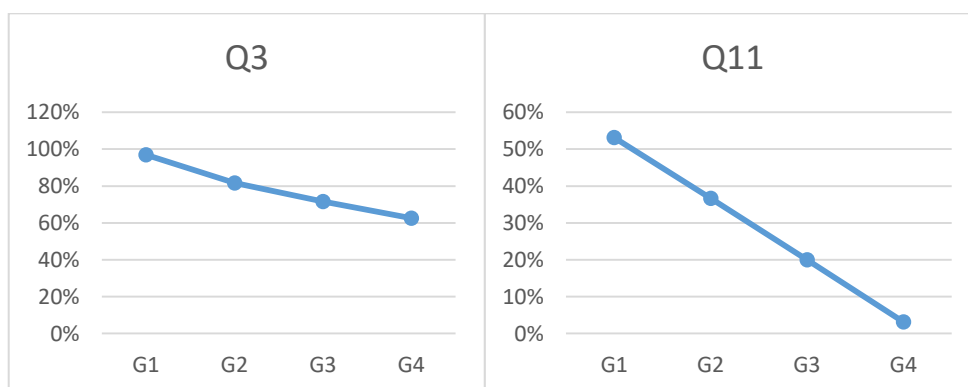
A csoportátlagok eloszlását vizsgáló 9. ábra megerősíti, hogy a hallgatók teljesítménye a kognitív hierarchiában felfelé haladva folyamatosan csökken. A 8. ábrán lévő intervallumokban Microsoft Excel segítségével meghatározott csoportátlag alapján vizsgáltam, majd a 9. ábra segítségével szemléltetem, mely feladatok hozták létre a differenciáló hatást. A G1, G2, G3 és G4 a különböző csoportokat, Q1, Q2, ... Q20 a tesztben szereplő itemeket, míg a hozzájuk tartozó csoportátlagot százalékban megadva az egyes csomópontok jelzik. Az ábra alapján arra a következtetésre jutottam, hogy a módosítások következtében az itemek megfelelő differenciáló hatással bírnak a tudásszint mérése érdekében.



9. ábra. Az itemek vizsgálata csoportátlag alapján.

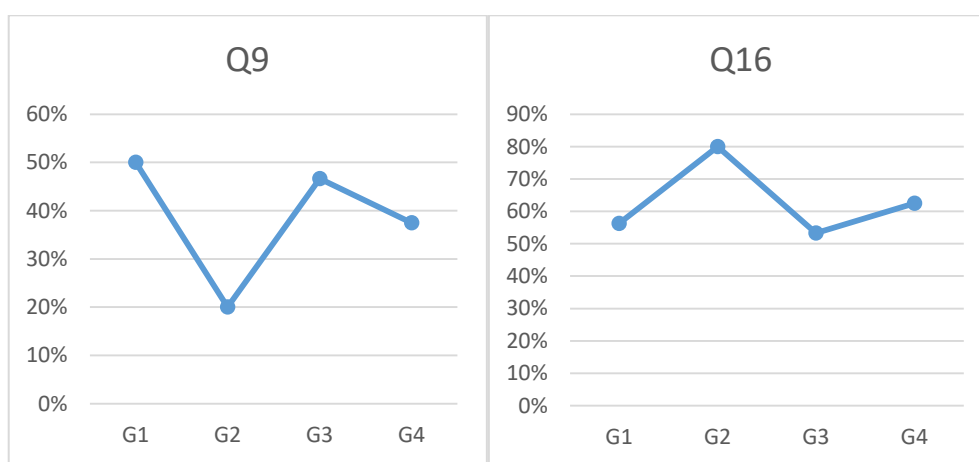
Sikeresen differenciáló hatásúnak tekintem azokat az itemeket, melyek megoldottsága a csoportban elvárt, vagy közel megfelelő rangsorban helyezi el a hallgatókat. Eszerint, aki a teszteredmény alapján a G1 csoportba került, az a hallgató tudja legjobban megoldani a feladatokat, míg a G4 csoport tagjai a legkevésbé. A csoportok átlaga alapján ez a meghatározás a Q1, Q2, Q3, Q11, Q12, de véleményem szerint még a Q19 és Q20 item esetén is igaz, melyek különböző kognitív szinteken helyezkednek el. Az itemek nehézségi indexe eltérő, a kitöltést követően ezen itemek közül legnehezebbnek a Q19 bizonyul 0,12 értékkel, míg legkönnyebb 0,78-cal a Q3.

A 10. ábrán ebből a csoportból szemléltetem a Q3 és Q11 itemeket, melyeket sikeresen alkotott itemnek tekintünk, mivel differenciáló hatása kimutatható. A Q3 esetén mind a négy csoport a hipotézisnek megfelelő rendben helyezkedik el 60% fölött teljesítve, a becslült és kitöltés utáni nehézségi indexe pedig megegyező. Míg a Q11 esetén nehézségi index becslése szintén a könnyű besorolást kapta, ennek ellenére a kitöltést követően 0,28 értékű, amit a 10. ábrán úgy láthatunk, hogy 4-ből 3 csoport 50% alatt teljesített, differenciáló hatása erős.



10. ábra. Q3 és Q11 – megfelelően differenciáló itemek

Néhány feladat esetén kisebb vagy nagyobb csoportugrásokkal keverednek a különböző csoportok teljesítményei az elvárthoz képest. Ezek az itemek a 9. ábra alapján Q5, Q6, Q7, Q9, Q10, Q13, Q14, Q15, Q16 és Q18, melyeket kevésbé sikeresnek tekintek. A nehézségi index alapján az ebbe a csoportba került legkönnyebb a Q18 item 0,89 értékkel, míg a legnehezebb a Q5 item 0,24-gyel, melyek különböző kognitív szintekhez tartoznak. Szemléltetés céljából a Q9 és Q16 itemeket ábrázolom (11. ábra) ebből a csoportosításból. A Q9 esetén 50% körüli teljesítést mutat a G1, és hasonló a 2 gyengébb csoport, míg 20%-os átlagot ér el a G2. A másik kevésbé sikeres itemben pedig épp fordítva, minden csoport 50% fölött teljesít, de a G1 és G3 csoportok kisebb eredményről számolnak be, mint azok a párjaik, akiktől ez elfogadottabb lett volna.



11. ábra. Q9 és Q16 – kevésbé sikeres itemek

A Q9 item esetén a rossz értelmezés lehet a rossz válaszok oka, amelyért feladatírókként felelős vagyok, a szöveg ugyanis nem utal egyértelműen arra, hogy számít a sorrend. A zárt, egy helyes válasz megjelölését engedélyező item Bloom taxonómiája szerint az eljárási tudás és a megérteni folyamat vizsgálatára vonatkozott, melyet a előzetesen a könnyű feladatok közé soroltam. A bemeneti teszt esetén a szövegben még fagylaltokról van szó, mely esetén rendszerint a legsavanyúbbtól a legédesebbig rendezik tölcsérbe a cukrászdákban. Szerkesztéskor a gondolatmenet ugyanez lett volna, mikor a fagylaltot joghurtra cseréltem, a helyes megoldásokhoz viszont szükséges lett volna az ízesítés sorrendjének meghatározása a biztos megoldáshoz. A hibám ellenére a hallgatók mindegyike csak az egyszerű kiválasztás és az általam is elvárt meghatározott sorrend szerinti kiválasztás lehetőségét tekintették helyesnek. Pontosítva 36 hallgató a kombinációval számolható, és 24 az általam elvárt meghatározott sorrendű kiválasztást jelölte variációval számolva. Mivel csak a 2 kiválasztási mód helyes megoldásai jelentek meg a válaszokban, nem tudtam pontos következtetést levonni az item differenciáló

hatásáról. Hibásnak fogadom el a feladatot, az elemzés szempontjából pedig csak a beérkezett, általam helyesként beállított válasszal számoltam.

A harmadik csoportot alkotja az a három item, melyek esetén közel azonosan teljesített minden csoport. Ezek közül Q4-et túl könnyűnek bizonyult, nehézségi index alapján is, mely 0,98 értékkel az e-teszt legkönnyebb itemje. Továbbá Q8-as és Q17-es itemeket nehezek a beérkezett válaszok alapján. A Q8 item nehézségi indexe 0,03 értéket mutat, mely a 3.3 alfejezetben is a nehéz feladat jelzőt kapta. Fentebb a Q6 itemként a 4. ábra mutatta be, valamint ugyanígy szerepelt a bemeneti tesztben. Továbbra is nehéznek bizonyult, de már a legerősebb csoport részéről érkezett helyes megoldás az itemre, mely a bemeneti tesztben szereplő módosított párja. A nyílt item, melyet Bloom taxonómiájának eljárás tudás és alkalmazási folyamatok dimenziójának vizsgálatára szolgált, általam a közepes nehézségi szintű besorolást kapta.

Hasonlóan az előzőhöz, a másik legnehezebbnek tűnő Q17 item, melyre egy helyes válasz sem érkezett, így nehézségi indexe 0. Az item szintén nyílt, melyet Bloom taxonómiájának eljárás tudás és alkalmazási folyamatok dimenziójának vizsgálatára hoztam létre. Arra támaszkodva, hogy ebben ismétléses kombinációt kell alkalmazni, amihez a képlet ismerete is szükséges úgy becsültem, ez is a közepes nehézségi szintet erősíti a hallgatók szemszögéből, bár a képlet ismeretében inkább a könnyű kategóriába sorolnánk. A feladatot, melyet *Denkinger Géza: Valószínűségi számítási gyakorlatok* c. könyvéből választottam 12. ábra szemlélteti megoldásával. A megoldásból kiderül, hogy az ismétlést mindenképp felismerték azok a hallgatók, akik válaszolni próbáltak, és nem szókozt vagy egyjegyű számot írtak annak érdekében, hogy tovább engedje őket a rendszer a következő feladat felé. A 62 kitöltő közül 42 választotta az ismétléses variációval történő megoldást a dobáshármasok megtalálására, ellenben itt nem számít a sorrend, ezért ismétléses kombinációval lehet eljutni a helyes válaszhoz (12. ábra).

Feladat: Három dobókockával dobva 3 számjegyből álló dobáshármast kapunk. Hányféle eredmény adódhat?

$$C_6^{3,i} = \binom{6+3-1}{3} = 56$$

12. ábra. Q17 item és megoldása.

A harmadik jóságmutató, mely szoros kapcsolatban áll az e-teszt érvényességével a reliabilitás, más néven megbízhatóság. Ez az a tulajdonsága, mely számszerűsíthető, így

a legpontosabb képet mutatja a tesztről. Értéke függ az itemek és kitöltők számától, valamint a beérkezett helyes megoldások alapján vizsgálható a Cronbach-alfa kiszámításával és besorolásával.

A **K2** kutatási kérdés kapcsán kijelenthető, hogy a tanári előzetes nehézségi becslés és a hallgatói eredmények között eltérő az összhang. A általam elméletileg nehéznek minősített, magasabb taxonómiai szintű feladatok a gyakorlatban az alacsonyabb nehézségi mutatókat produkálták és esetenként fodítva. Ugyanakkor megfigyelhető egyfajta "kognitív szakadék": a hallgatók számára a váltás a rutinszerű behelyettesítésről a probléma mélyebb strukturális elemzésére sokkal nagyobb teljesítményromlással jár, mint amit a feladatok matematikai komplexitása önmagában indokolna. Ez igazolja, hogy az e-teszt környezetben a technikai segítség nem tudja áthidalni az absztrakt műveletek során fellépő nehézségeket.

A kutatás során alkalmazott diagnosztikus (bemeneti) és szummatív (kimeneti) e-tesztek összehasonlító elemzése rávilágít a hallgatói teljesítmény változására, valamint a mérőeszközök fejlesztésének irányaira.

A bemeneti mérés során ($N = 57$) a hallgatók átlagteljesítménye alacsony szinten mozgott, amit az ANOVA módszerrel számított 0,34-es Cronbach-alfa érték kísért. Ez a mutató statisztikai szempontból az „elfogadhatatlan” kategóriába esik. A mérés eredménye arra utal, hogy a bemeneti szakaszban a hallgatók előzetes tudása még rendkívül heterogén és töredezett volt, vagy a teszt tételei nem alkottak koherens mérési egységet a vizsgált populáció aktuális ismeretszintjéhez mérten. A Microsoft Excel Leíró statisztikája alapján a tesztet alkalmazhatónak tekintjük a tudás mérésére, mivel várható értéke 5,55 és mediánja 6 értékkel elég közel állnak egymáshoz, a standard hiba 0,29, miközben a terjedelme 10,25.

A záró mérés adatai ($N = 62$) ezzel szemben jelentős javulást mutatnak mind a teljesítmény, mind a megbízhatóság terén. A kimeneti e-teszt esetében a Cronbach-alfa értéke 0,54-re emelkedett. Bár a szakirodalmi besorolás szerint ez az érték még a „gyenge/rossz” megbízhatósági tartományba tartozik, a bemeneti teszthez (0,34) és a korábbi tesztverziókhoz képest tapasztalt növekedés figyelemre méltó módszertani fejlődést jelez. A Microsoft Excel Leíró statisztikája alapján a záró tesztünket alkalmazhatónak tekintjük a tudás mérésére, mivel várható értéke 8,97 és mediánja 8,87 értékkel elég közel állnak egymáshoz, a standard hiba 0,32, miközben a terjedelme 12,75.

A két mérési pont közötti megbízhatósági növekedés, valamint az itemek differenciáló erejének vizsgálata igazolja, hogy a teszt tételeinek finomítása és az oktatási

folyamatba való illesztése jó irányba halad. Az elemzés rávilágít, hogy a teszt megbízhatóságát jelenleg elsősorban a rendkívül nehéznek bizonyult itemek (alacsony megoldottsági mutatók) korlátozzák.

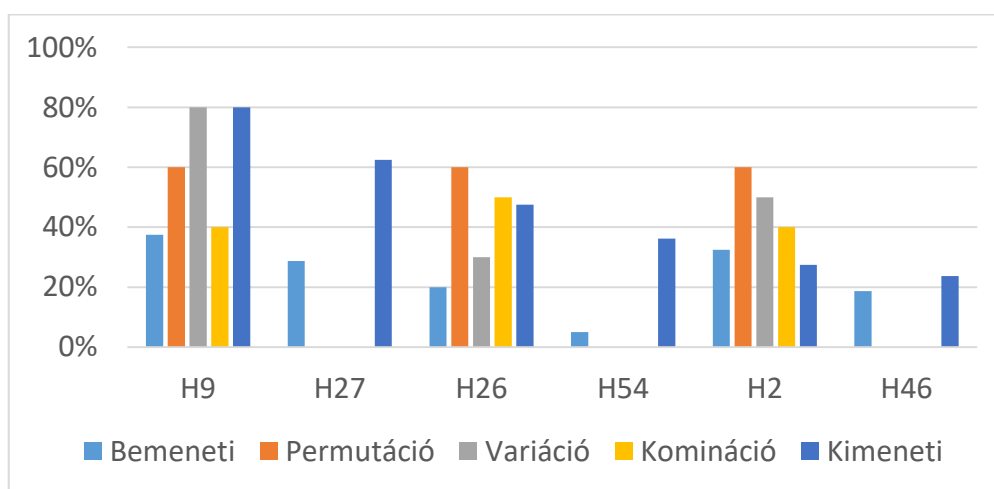
A kutatás 6. ciklusában a teszttel támogatott oktatási intervenció elhagyásával vizsgáltam a módosított és 5 további releváns itemmel bővített záróteszt reliabilitását. A mérést N=25 informatika, valamint számvitel és adóügy szakos hallgató körében, szummatív záródolgozat keretében valósítottam meg. A varianciaanalízis (ANOVA) alapján számított Cronbach-alfa mutató 0,73-as értéket eredményezett, amely a tesztelméleti standardok szerint a mérőeszköz belső konzisztenciájának „elfogadható” kategóriájába sorolható.

Az. 5. ciklus be- és kimeneti tesztje mellett az előhívó teszteket is vizsgáltam, mint mérőeszközöket. Kis itemszámuk miatt reliabilitásukat nem vizsgáltam, de ugyanolyan figyelemmel állítottam össze őket, ahogy a be- és kimeneti teszteket. Az objektivitásuk véleményem szerint megfelelő, mivel csak olyan itemeket tartalmaz, melyeket a platform egyértelműen javít. Validnak tekintem, mivel előhívó jelleggel tudást mér, annak ellenére, hogy longitudinális hatásának vizsgálatát az előhívósos tanulási technikának nem tudtam tervbe venni.

A három mini tesztből a permutáció előhívó tesztje volt a legsikeresebb, ahol az 51 kitöltőből bár 1 hallgatótól érkezett be maximális megoldottság (5 pont), további 42-en érték el 3 vagy 4 pontot. A kombinatorika másik két eleme már nagyobb fejtörést okozott a hallgatók számára. A Variáció mini tesztjében 52 kitöltőből 29 ért el 3, vagy annál nagyobb pontszámot, míg a Kombináció előhívó tesztje esetén a 45 kitöltőből mindössze 18 érte el azt.

Az előhívó tesztek longitudinális hatásának vizsgálata nélkül kerestem a választ a **K3** kérdésre. Az eredmény pontosabb vizsgálataként a letisztított mintát alkalmaztam, és külön táblázatba rendeztem a kontroll csoportot alkotókat. Így kaptam 37 hallgatót, aki mind az 5 tesztet teljesítette, és további 7 hallgatót a kontroll csoportban, ahol a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola informatikus hallgatóit (eredeti kontroll csoportunk) 1 további hallgatóval egészítettem ki a komáromi Selye János Egyetemről, akik csak a be- és kimeneti tesztet töltötték ki. A nagy létszám miatt 3 hallgató eredményét mutatom be, akik mindhárom előhívó tesztet kitöltötték, és 3 hallgatót a kontrollcsoportból. A kiválasztás alapját a kimeneti teszt eredménye adja, mindkét csoportból a legtöbb, közepes és legkevesebb pontot szerzett hallgatót választva. A 13. ábrán egymás mellé állítottam a csoportokból kiválasztott hallgatók eredményeit.

Az ábrán a H9 és H27 hallgatók érték el a csoportokban a legnagyobb pontszámot a kimeneti tesztben. Mindkét esetben nagymértékű fejlődés mutatható ki a be- és kimeneti e-tesztek között. A H9 hallgató esetében ez több mint 40%-os annak ellenére, hogy a kombináció előhívó tesztje nem lett túl sikeres. A kontroll csoportban lévő H27 hallgató eredménye csak kicsivel több, mint 30%-os előrehaladást mutat, miközben azon 37 kitöltő közül, akik minden tesztet hiány nélkül beküldtek 7 hallgató haladja meg ezt.



13. ábra. Hallgatói eredmények.

Közepes teszteredményekkel rendelkezve további 8 hallgató ért el 20-30% közötti fejlődést a két teszt között, amit nem tudok határozott kijelentéssel az előhívó tesztek hatásának megállapítani. Ide tartozik a kimeneti teszt szerint közepes sikerességgel teljesítő H26 hallgató, aki a 28%-os különbséggel a kontroll csoport hallgatójával szinte megegyező fejlődést mutat. A 37 kitöltő további több, mint fele 20% alatti, vagy kis százalékban, de negatív fejlődést mutat. Ez utóbbival rendelkezik a H2 hallgató, aki az egyre csökkenő mini teszteredmények után a bemeneti teszthez képest kevesebb helyes megoldást ad ellenben a kontroll csoportból a H6-os hallgatóval, aki hasonlóan kismértékű, de fejlődést mutat (13. ábra).

Összességében az előhívó e-tesztek pozitív hatását a tudásmérő kimeneti tesztre nem lehet biztosan kimutatni. Arra a következtetésre jutottam, hogy ennek oka az lehet, hogy a felsőfokú oktatás tematikája és óraszama miatt túl gyorsan zajlott a felmérés. A gyors lefolyás miatt számítani lehetett erre az eredményre, ezért az előhívásos tanulással, csak mint tesztrel támogatott oktatás eszközével foglalkoztam, annak motiváló hatásával és hosszabb távú figyelem fenntartásával.

Nem kutatási céllal, de mindenképp fontosnak tartjom megemlíteni a kutatás során észlelt előnyeit és hátrányait. Bár a tudatos és a tesztelmélet szerint is szabályszerű

tesztek létrehozása időigényes, amit a hátrányukként említhetünk, legalább ugyanannyi időt spórolunk meg azzal, hogy a nagy létszámú csoportokban nem kell mindezt saját magunknak javítani, ami határozottan pozitív előnye.

Az e-tesztek felhasználásának pozitív hatását látom többek között abban, hogy az eszközhasználat egy új, oktatási színteret kap azáltal, hogy nemcsak engedéllyel, de szükség szerint alkalmazják a tanórák bizonyos szakaszán. Emellett beállítástól függően parametrikus megadással azonos típusból különböző feladatok kiosztásával lehet egyedivé tenni a teszteket, ezzel kiküszöbölve, hogy az eredményeket közöljék egymással a hallgatók, valamint ugyanezt nehezíti meg, ha a feladatok és válaszlehetőségek véletlenszerű keverésével tesszük egyedivé a tesztet. További pozitív hatása, hogy azonnali visszajelzést ad mind az oktató, mind a hallgató számára az elért eredményről, amit ha egyes itemek esetén felül is kell vizsgálni, mindenképp gyorsabb visszajelzés adható.

Az e-tesztek felhasználásának egyik negatív hatása véleményem szerint, hogy kitöltéséhez internet elérhetőségre van szükség. Ez abban az esetben is problémát jelenthet, ha az intézményben van és elég erős az elérhetőség, de az óra olyan teremben zajlik, mely részben vagy teljesen holtterbe kerül. További negatív hatásaként említhetem meg, hogy a feleletválasztó itemek teret engednek a megfontolás nélküli jelölésnek, aminek következtében nem a hallgató tudásszintjét mérjük.

A **H1** hipotézis a mérési adatok alapján beigazolódott. Amint a feladatok az alkalmazás szintjére léptek, a nehézségi indexek csökkenni kezdtek. Míg az egyszerűbb alkalmazások még a 0,70 feletti tartományban maradtak, addig az összetettebb, több lépésből álló elemzést igénylő feladatoknál drasztikus törés tapasztalható. Ez alátámasztja, hogy a hallgatók teljesítménye szignifikánsan romlik a magasabb szintű kognitív műveleteknél, függetlenül az előhívó mini tesztek kitöltésétől, valamint attól, hogy a tanári becslés azt mennyire tartotta kivitelezhetőnek.

4.3. A pontszerzés mint külső motivációs dominancia

A kutatás folyamán a tudás mérése mellett azt is vizsgáltam, a teszttel támogatott oktatás mennyire motiváló a hallgatók számára. A Google-alkalmazások segítségével Űrlapokban (Függelék F) készítettem elő a felmérés kérdőívét, mely kitöltése önkéntes alapú (N=32) volt. A beérkezett válaszok feldolgozása segítségével keresem a választ a következő kérdésekre.

H2. A hallgatók számára az e-tesztek motivációs erejét elsősorban nem az azonnali visszacsatolás (önértékelés), hanem a végső értékelésbe való beszámíthatóság (pontgyűjtés) adja.

K1. Mennyire tudatos a válaszok megjelölése a tesztek kitöltése során?

K2. Mely funkciót tartják a hallgatók az e-tesztek legfőbb előnyének: a tanulást segítő visszajelzést vagy a jegyszerzés megkönnyítését?

K3. Van-e összefüggés a hallgatók motivációs szintje és a pontgyűjtési lehetőség között?

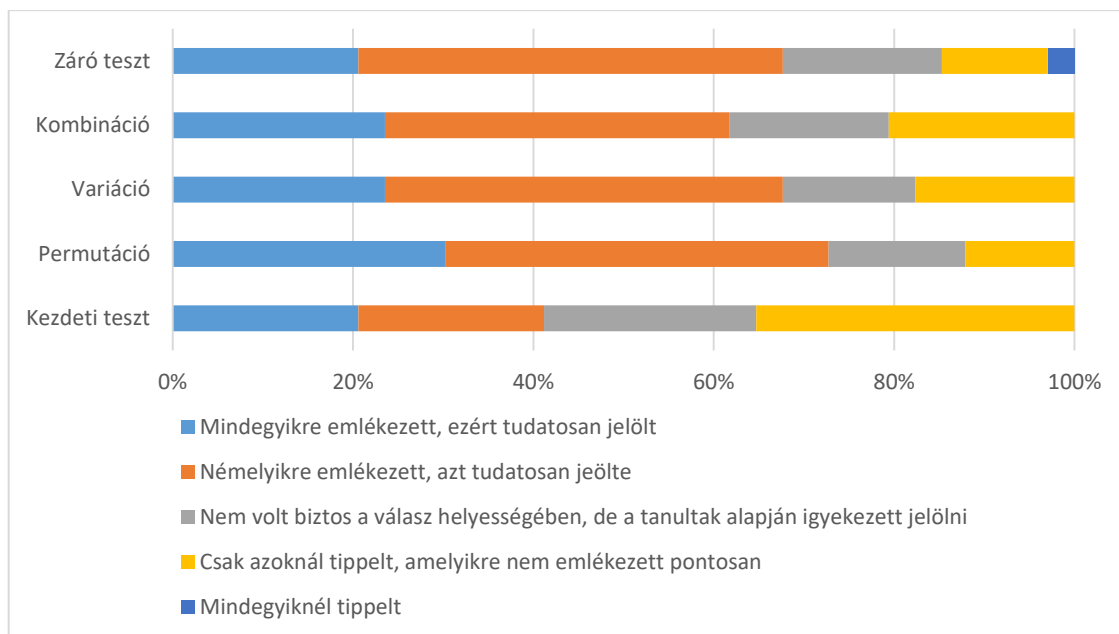
A hallgatók véleménykifejtő kérdőíve az alapvető regisztrációs adatokon kívül 2 fő szakaszból áll: az e-tesztek kitöltése és a motiváció vizsgálata. A hallgatók zárt kérdések segítségével fejthették ki véleményüket motiváltságukról, az óra végén használt mini tesztek hatásáról a motiváltságukra, valamint a tudásszintmérő tesztre, eredményeikre.

A kérdőív készítésekor a hallgatók véleményét szerettem volna megtudni, okozott-e nehézséget a tesztek kitöltése, amire elég nagy számban pozitívan válaszoltak. Mindössze 7 hallgató adta azt a választ, hogy nehézséget okozott számára a tesztek kitöltése. Közülük 1 személy szerint nehezek voltak a kérdések, 3 a szöveges válaszok helytelen javítását nevezte meg, 2 a kérdésben felhozott példából nevezett meg választ, míg 1 fő nem válaszolt.

A **K1** kérdésben felvetett válaszadási tudatosságot illetően az elemzés ennek ellenére pozitív képet mutat. Ezt a szakaszt azért tartottam fontosnak, hogy választ kapjak a hallgatóktól, milyen módszerrel jelölték a különböző feladatok megoldásait. A kérdések esetén Bloom-taxonómia fogalmi tudást vizsgáló itemeire úgy hivatkoztam, mint elméleti kérdések, míg az eljárási tudás esetén a gyakorlati feladatok szókapcsolatot használtam.

A gyakorlati feladatoknál még erősebb tudatosság figyelhető meg: 2 hallgató a „Mindegyiket kiszámolta”, míg a kitöltők jelentős része minden feladatot megpróbált kiszámolni, ha nem is mindegyiket sikerült. Ez a válaszlehetőség engedélyezi, hogy némely esetben előfordulhatott tippelés, ennek ellenére pozitívan értékeltem, hogy az önbevallás során a tippelésre utaló opciók egyáltalán nem merültek fel.

Az elméleti kérdéseknél a hallgatók többsége önbevallásuk alapján tudatosan jelölt, a tippelés csak elvétve fordult elő; mindössze egyetlen hallgató jelölte, hogy a záró tesztnél minden esetben tippelt (14. ábra).



14. ábra. Hallgatók válaszadása a tesztek elméleti kérdéseinél

A 14. ábra szembeűnően mutatja, hogy a kezdeti teszt kivételével szinte mindegyik esetben a legtöbb jelölést kapta a „Némelyikre emlékezett, azt tudatosan jelölte”, miközben alacsony mennyiségben van a „Nem volt biztos a válasz helyességében, de a tanultak alapján igyekezett jelölni”, valamint a „Csak azoknál tippelt, amelyekre nem emlékezett pontosan” válasz, ami mindenképp erős tudatosságra enged következtetni. A jelölésekre utaló válaszok alapján arra következtettem, hogy a tesztek eredménye a valós tudást mutatják, így alkalmasak a tudásszint mérésére.

A hallgatók motiválását kutatva a 9. táblázatban összefoglalt kérdéseket is tartalmazta a kérdőív. Ez alapján segítette őket az előhívó tesztek kitöltése, ennek ellenére néhány esetben ellentétbe ütközünk a környező kérdések kapcsán.

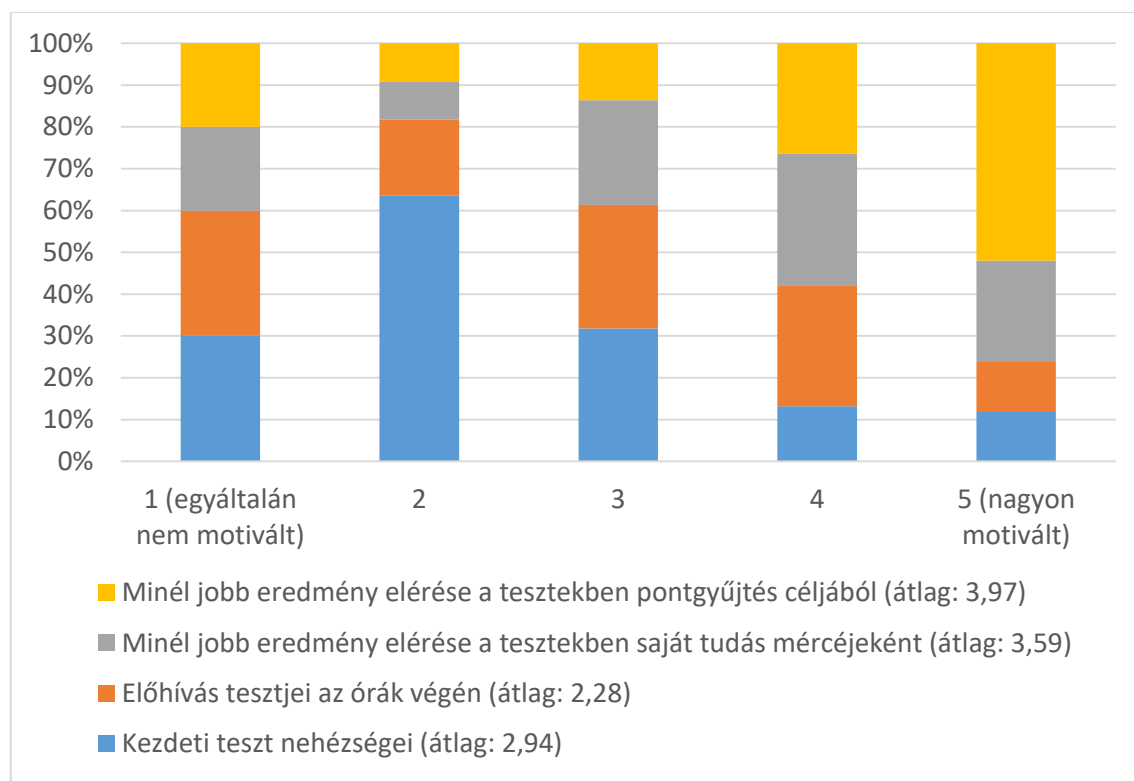
9. táblázat. Véleménykifejtő információk összegyűjtő táblázata.

Kérdés	Igen	Nem	Talán
Ön szerint megbízhatóan alkalmazható-e az e-teszt, mint a tudásszint mérőeszköze?	16	9	7
Hasznosnak érezte-e a témák végén írt előhívó teszteket?	21	3	8
Ön szerint segíti-e az előhívó tesztek kitöltése az adott témakörre vonatkozó önértékelést?	22	4	6
Ön szerint az előhívó tesztek sikerei segítettek-e a modulzáró dolgozatra való felkészülésben?	19	7	6

Miközben 22 hallgató szerint egyértelműen segítette őt az előhívó teszt abban, hogy meg tudja vizsgálni saját tudását a témában a záró teszt előtt, érdekes képet mutat

számomra, hogy csak 16 hallgató véli úgy, megbízhatóak a tesztek, mint mérőeszközök. Kicsi ellentmondást látunk itt, mivel a logika alapján, ha a hallgató szerint megbízhatóan mér a teszt, akkor az önértékelésében is megbízhat, mivel folyamatos képet lát tudásáról. 4 válasz esetén nem megbízhatóak a tesztek, mint mérőeszközök, az előhívók mégis segítettek az önértékelésben, de nem segítettek a záró tesztre való felkészülésben. Felmerül a kérdés, jól értelmezték-e a kérdőív adott kérdéseit és megfontoltan válaszoltak-e? Ha a hallgató úgy véli, nagyobb (vagy kisebb) a tudása, mint azt a tesztek mutatják, akkor milyen képet mutatnak számára az önértékelésben? Ahogy minden ember más, minden hatás is eltérő rá nézve, ezért bár sokak véleménye szerint önértékelésében, témakörre vonatkozó tudása felmérésében segített az előhívó e-teszt, elfogadom, hogy ezt nem vonatkoztatták a tudásszintmérő tesztekre.

A kérdőívben a pontosabb következtetések érdekében kicsit eltértünk a fő kérdésünktől, így nemcsak az előhívó teszt motiváló hatásáról kérdeztem a hallgatókat. 5 fokú Likert-skálás (15. ábra) értékeléssel több esemény motiváló hatását kellett pontozniuk 1 (egyáltalán nem motivált) és 5 (nagyon motivált) pont között az aktív figyelem fenntartására nézve. Ennek segítségével kaptam választ K2 és K3 kérdésekre, valamint további órai eseményekre, melyek véleményem szerint motiválóak lehetnek a hallgatók számára az aktív figyelem fenntartása érdekében.



15. ábra. Motiváció vizsgálata 5 fokú Likert-skálával

A hallgatók e-tesztekkel kapcsolatos attitűdjét vizsgálva, a **K2**-es kérdésre az elemzés azt a választ adja, hogy bár mindegyik esemény erősen motiváló hatással bírt számukra a legfőbb előnynek a jegyszerzés megkönnyítését tekintik a tanulást segítő funkcióval szemben. Ezt igazolja, hogy a legnagyobb átlaggal (3,97) az a válaszlehetőség rendelkezik, amely a tesztek sikerességét a megszerzett pontok beszámíthatóságához köti. Ezzel szemben a kizárólag önértékelésre szolgáló funkciót – kezdeti teszt nehézségei és az előhívó mini tesztek – jóval kisebb átlaggal jelölték.

A **K3** kérdés az összefüggést vizsgálja a motivációs szint és a pontgyűjtés között, melyet stíltén a 15. ábra Likert-skálás eredményei támasztanak alá. A hallgatók számára az aktív figyelem fenntartása érdekében a pontok gyűjtése bizonyult a legerősebb motiváló tényezőnek. Bár az előhívó tesztek önmagukban is sikeresen fenntartották az érdeklődést – a hallgatók többsége 4-es vagy 5-ös szintű motiváltságról számolt be –, ez az elkötelezettség akkor vált igazán aktívvá, ha a megszerzett pontok a tantárgy teljesítéséhez is hozzájárultak. Összességében a pontszerzési lehetőség nemcsak a részvételi kedvet, hanem a feladatmegoldás alaposságát is közvetlenül fokozta.

Ezt a kérdőív kitöltése során maguk is megerősítették a *Mitől lesz motiváló egy e-teszt?* kérdésre válaszolva. Nem szerettem volna kifejtő itemként szerkeszteni, mivel tapasztalatok alapján ezeket inkább kihagyják a válaszadásból, így több válasz kiválasztását engedő zárt kérdésként fogalmaztam meg. A hallgatóknál itt is nagy szerepet kapott a pontgyűjtés. A legtöbb, 23 jelölést az a válasz kapott „Ha azok gyakorló tesztek, melyeket helytől és időtől függetlenül lehet kitölteni, de sikere esetén a megszerzett pontok beszámíthatóak”. A csupán önértékelésre szolgáló előhívó tesztek mindössze 2 hallgató jelölte. A 15. ábra és az adott kérdés alapján arra a következtetésre jutottam, hogy a hallgatók számára a tantárgyakban szerzett pontszámok a leginkább motiválóak.

A hallgatók motivációját vizsgálva megállapítható, hogy a **H2** hipotézis beigazolódott: az e-tesztek ösztönző erejét elsősorban nem a belső igényű önértékelés, hanem a végső értékelésbe való beszámíthatóság (pontgyűjtés) adja. Ezt támasztja alá a **K2** és **K3** kérdésre adott válaszok megoszlása is, mely szerint a hallgatók az e-tesztek legfőbb előnyének a jegyszerzés megkönnyítését tartják a tanulást segítő visszajelzéssel szemben. A kérdőív adatai alapján a legtöbb, összesen 23 jelölés arra a válasza érkezett, amely a pontok beszámíthatóságát nevezte meg fő motivációként, miközben a kizárólag önértékelésre szolgáló funkciót mindössze két hallgató választotta.

KÖVETKEZTETÉSEK, TOVÁBBI KUTATÁSI LEHETŐSÉGEK

A dolgozatot az oktatóként is túl kevés háttérinformációval rendelkező e-tesztek szerkesztése ihlette, melyekkel a diákok és hallgatók tudásszintjét kívánjuk mérni. Ezért arra törekedtem, hogy felkutassam a teszthez szükséges szakirodalmi alapot, és megvizsgáljam a létrehozott tesztet a kitöltés eredményére támaszkodva. A kutatást a felsőoktatásra irányítottam, amit mind az előzetes mérések, mind az empirikus kutatás során sikerült kivitelezni más-más minta segítségével.

Az empirikus kutatást a komáromi Selye János Egyetem, valamint a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola további csoportjának bevonásával 69 fős mintával sikerült elvégezni, alapját pedig a 3.3. alfejezetben bemutatott 2. ciklus képezi, melyet a tapasztalatok alapján tovább módosítottam, fejlesztettem. Két hipotézist, és hozzájuk tartozó kutatási kérdéseket fogalmaztam meg, melyekre a kutatás eredményeinek vizsgálatával válaszolok meg.

Ezek alapján a jól alkalmazható e-teszt létrehozásához figyelembe kell venni a Bloom-taxonómia dimenzióit, és a lehető legpontosabban megválasztani, melyiknek megfelelő az adott item. Fontos előre meghatározni a teszt pedagógiai céljait (például differenciálás), betartani a feladatírás 1.1.3. alfejezetben említett szabályait, mely segítségével gondosan odafigyelünk az itemtípusok megválasztására, a nyelvezetre, disztraktorok megfogalmazására és a feladat nehézségére. Jóságmutatók meglétét igyekeztem biztosítani a szerkesztés folyamán, megbízhatóság szempontjából figyeltem az itemek számára, de csak a kitöltést követően vonhattam le a következtetés, mennyire sikerült ezen tulajdonságokkal rendelkező teszt létrehozása.

Objektivitása a kitöltést követően sem sérült, mivel a gondosan megtervezett és létrehozott itemeket a platform javította, így független a szerkesztőtől. Érvényessége bár függ a megbízhatóságától, ennek ellenére azt a következtetést vontam le, hogy a pedagógiai célt – mint differenciáló tudásszintmérés az adott témában – elértem. Fő elemzésben vizsgált kimeneti e-teszt megbízhatósága Cronbach-alfa kiszámítását követően még alacsony, mindössze 0,54 értékkel „gyenge/rossz” megbízhatósági tartományba tartozik, ami a 6. ciklusban elvégzett módosításokat követően a kisebb mintán 0,73-as értékkel az „elfogadható” tartományba kerül. Differenciáló hatását az itemek nehézségi indexe és a teljesítés alapján létrejött egyenletes csoportok támasztják alá. Ezen eredményekre támaszkodva válaszoltam meg a következő kutatási kérdést:

A kutatás eredményei alapján a **K1** kérdésre válaszolva megállapítható, hogy a

hallgatók számára az elemzés, valamint az értékelés és alkotás kognitív szintjei bizonyultak a legnehezebbnek. A leggyengébb megoldottságot mutató feladatok (5., 7., 8., 11., 15., 17.) esetében a kudarcot nem a képletek ismeretének hiánya, hanem a speciális logikai kényszerfeltételek kezelése és az ebből adódó modellválasztási nehézségek okozták. Bár a kognitív hierarchiában felfelé haladva a teljesítmény folyamatosan csökkent, a módosított tesztitemek a csoportátlagok eloszlása alapján megfelelő differenciáló hatással bírnak a tudásszint mérésére.

A **K2** kérdés kapcsán kiderült, hogy a tanári előzetes nehézségi becslés és a tényleges hallgatói eredmények nem vágtak egybe teljes mértékben. Fény derült egy jelentős „kognitív szakadékra”: a rutinszerű behelyettesítésről a mélyebb strukturális elemzésre való áttérés jóval nagyobb teljesítményromlást idézett elő, mint amit a feladatok matematikai összetettsége indokolt volna.

A **K3** kérdés vizsgálatához egy letisztított mintát használtam (37 rendszeresen tesztelő és 7 kontrollcsoportos hallgató), amelyből a kimeneti teszt alapján a legjobb, közepes és leggyengébb eredményt elérő hallgatók teljesítményét emeltem ki. Összességében az előhívó e-tesztek kimeneti tesztre gyakorolt közvetlen pozitív hatását nem sikerült egyértelműen kimutatni. Ennek hátterében a felsőoktatási tematika és a feszes óraszám miatti gyors lefolyású felmérés állhat.

Végül a **H1** hipotézis maradéktalanul beigazolódott: az alkalmazás szintjére lépve a feladatok nehézségi indexe romlani kezdett, az összetettebb, több lépésből álló elemzést igénylő feladatoknál pedig drasztikus törés tapasztalható. Ezt bizonyítja, hogy a magasabb szintű kognitív műveleteknél a hallgatói teljesítmény szignifikánsan visszaesik, függetlenül a tanári becsléstől vagy a mini tesztek kitöltésétől.

A **motiváló hatás** kutatásának eredményei alapján a **K1** kérdésre válaszolva megállapítható, hogy a hallgatók tudatosan és felelősségteljesen álltak a feladatmegoldáshoz. Az önbevallások alapján a tippelés lehetősége szinte egyáltalán nem merült fel a válaszokban; a diákok többsége mind a Bloom-féle taxonómia szerinti fogalmi tudást vizsgáló elméleti kérdéseknél, mind az eljárási tudást mérő gyakorlati feladatoknál törekedett a valós megoldásra és a számítások elvégzésére. Csupán egyetlen hallgató jelezte, hogy a záró teszt elméleti kérdéseinél szisztematikusan tippelt.

A **K2** és **K3** kutatási kérdések, valamint az ezekhez kapcsolódó Likert-skálás elemzések a hallgatók e-tesztekkel kapcsolatos attitűdjét és motivációját vizsgálták. Az eredmények alapján a tesztek alapvetően érdeklődést generálóak és fenntartják az aktív figyelmet, a legfőbb ösztönző erő nem a tanulást segítő visszajelzés vagy a belső igényű

önértékelés jelenti, hanem a pontgyűjtés és a jegyszerzés megkönnyítése. A hallgatók elkötelezettsége és a feladatmegoldás alapossága akkor vált igazán intenzívvé, ha a megszerzett pontok beszámíthatóak voltak a tantárgy végső teljesítésébe (3,97-es átlag). Ezt igazolja a zárt kérdések megoszlása is: a helytől és időtől független, de pontszerzésre alkalmas tesztek lehetősége kapta a legtöbb (23) jelölést, miközben a kizárólag önértékelésre szolgáló funkciókat mindössze két hallgató tartotta motiválónak.

Mindezek alapján a **H2** hipotézis maradéktalanul beigazolódott: az e-tesztek valódi ösztönző erejét a külső elismerés – azaz a végső értékelésbe való beszámíthatóság – adja, háttérbe szorítva a tisztán önfejlesztő és tanulást támogató funkciókat.

Mindezek ismeretében meg tudom válaszolni a fő kérdét:

Mennyire alkalmas a Bloom-taxonómia kognitív szintjeire épülő e-teszt a hallgatók differenciált tudásszintjének objektív mérésére a felsőfokú matematikaoktatásban — a jelenlegi reliabilitási mutatók tükrében —, és e mérési folyamatban miként befolyásolja a pontszerzési lehetőség az e-tesztek motivációs hatékonyságát?

A Bloom-taxonómia kognitív szintjeire épülő e-teszt kifejezetten alkalmasnak bizonyult a hallgatók differenciált tudásszintjének objektív mérésére a felsőfokú matematikaoktatásban, adott kutatási csoportban. A jelenlegi reliabilitási és statisztikai mutatók (a csoportátlagok eloszlása és az itemek nehézségi indexei) igazolják, hogy a tesztitemek megfelelő differenciáló hatással bírnak.

Ugyanakkor a mérési folyamatban a pontszerzési lehetőség alapvetően meghatározza az e-tesztek motivációs hatékonyságát. Bár a tesztek önmagukban is képesek fenntartani a hallgatói érdeklődést, az elkötelezettség és a feladatmegoldás alapossága akkor válik igazán aktívvá, ha a megszerzett pontok beszámítanak a tantárgy végső teljesítésébe (jegyszerzésbe). A kutatás adatai szerint az e-tesztek ösztönző erejét elsősorban nem a belső igényű önértékelés, hanem ez a külső, pontgyűjtő funkció adja, amely így közvetlenül és pozitívan fokozza a hallgatók részvételi kedvét és munkájuk alaposságát.

További kutatási lehetőségeket képez az alkalmazott tesztek javítása és továbbfejlesztése annak érdekében, hogy megbízhatósága magasabb értéket mutasson. Emellett az előhívó tesztek vizsgálata során nemcsak teszttel támogatott oktatás kapcsán, hanem ismeretet elmélyítő hatásának vizsgálata érdekében érdemes késleltetett méréseket tervezni és kivitelezni. Ennek következtében a tudásmérő teszt is nagyobb szerepet kaphat.

FELHASZNÁLT IRODALOM

- Anderson, L. W., & Krathwohl, D. R. et al. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
https://www.researchgate.net/publication/235465787_A_Taxonomy_for_Learning_Teaching_and_Assessing_A_Revision_of_Bloom's_Taxonomy_of_Educational_Objectives
- Armstrong, P. (2010). Bloom's Taxonomy. *Vanderbilt University Center for Teaching*.
<https://health.ucdavis.edu/mdprogram/curriculum/pdfs/blooms-taxonomy-vanderbilt.pdf> (online proxy segítségével nyitható)
- Awopeju, O. A., & Afolabi, E. R. I. (2016). Comparative Analysis of Classical Test Theory and Item Response Theory Based Item Parameter Estimates of Senior School Certificate Mathematics Examination. *European Scientific Journal, ESJ*, 12(28), 263. <https://doi.org/10.19044/esj.2016.v12n28p263>
- Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Szerk.), *Practical aspects of memory II* (o. 396–401). London: Wiley. <https://bjorklab.psych.ucla.edu/wp-content/uploads/sites/13/2016/07/Bjork1988ReRetrieval.pdf>
- Brame, C. J., & Biel, R. (2015). Test-Enhanced Learning: The Potential for Testing to Promote Greater Learning in Undergraduate Science Courses. *CBE—Life Sciences Education*, 14(2), 1–12. <https://doi.org/10.1187/cbe.14-11-0208>
- Ceglédy I. (2011). *A matematika tanításának pedagógiai – pszichológiai vonatkozásai*. EKF.
https://dtk.tankonyvtar.hu/bitstream/handle/123456789/8131/0038_matematika_Cegledi1.pdf?sequence=2&isAllowed=y
- Cho, E., & Kim, S. (2015). Cronbach's Coefficient Alpha: Well Known but Poorly Understood. *Organizational Research Methods*, 18(2), 207–230.
<https://doi.org/10.1177/1094428114555994>
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98–104.
<https://doi.org/10.1037/0021-9010.78.1.98>

- Crehan, K., & Haladyna, T. M. (1991). The Validity of Two Item-Writing Rules. *The Journal of Experimental Education*, 59(2), 183–192.
<https://doi.org/10.1080/00220973.1991.10806560>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334. <https://doi.org/10.1007/BF02310555>
- Cronbach, L. J., & Hartmann, W. (1954). A Note On Negative Reliabilities. *Educational and Psychological Measurement*, 14(2), 342–346.
<https://doi.org/10.1177/001316445401400213>
- Csapó B. (1988). *A tanulói teljesítmények értékelésének mérési módszerei*. Művelődési Minisztérium Vezetőképző és Továbbképző Intézete; Veszprém Megyei Pedagógiai Intézet. <https://publicatio.bibl.u-szeged.hu/11835/>
- Csapó B. (2004). Tudásszintmérő tesztek. In Falus I. (Szerk.), *Bevezetés a pedagógiai kutatás módszereibe* (4. kiad, o. 277–316). Műszaki K.
<https://core.ac.uk/download/pdf/84775002.pdf>
- Csapó, B., Molnár, G., & R. Tóth, K. (2008). A papíralapú tesztektől a számítógépes adaptív tesztelésig. *Iskolakultúra*, 3–4, 3–16. <http://real.mtak.hu/id/eprint/5635>
- Csíkos, C., & B. Német, M. (2002). A tesztekkel mérhető tudás. In *Csapó Benő (szerk.) Az iskolai tudás* (o. 91–123). Osiris Kiadó. https://publicatio.bibl.u-szeged.hu/11931/1/CsBenoi_Iskolai_tudas_2002.pdf
- Demkanin, P., Hajdúk, M., Hanuljakova, H., Kubiš, T., Lapitka, M., & Malčík, M. (2015). *Metodika tvorby testových úloh a testov* (Mgr. Timotej Kubiš).
https://www.researchgate.net/publication/349279914_Metodika_tvorby_testovych_uloh_a_testov
- Ebel, R., & Frisbie, D. (1991). *Essentials educational measurement*. Prentice Hall.
https://ebookppsunp.wordpress.com/wp-content/uploads/2016/06/robert_l-ebel_david_a-frisbie_essentials_of_edbookfi-org.pdf
- Einhorn Á. (2012). *Feladatkönyv*. Nemzeti Tankönyvkiadó.
<https://mek.oszk.hu/17800/17864/17864.pdf>
- Eluwa, O. I., Eluwa, A. N., & Abang, B. K. (2011). Evaluation of Mathematics Achievement Test: A Comparison Between Classical Test Theory (CTT) and Item Response Theory (IRT). *Journal of Educational and Social Research*, 1(4).
<https://www.richtmann.org/journal/index.php/jesr/article/view/11750>

- Filgona, J., Sakiyo, J., Gwany, D. M., & Okoronka, A. U. (2020). Motivation in Learning. *Asian Journal of Education and Social Studies*, 16–37. <https://doi.org/10.9734/ajess/2020/v10i430273>
- Fuqoha, A. A. N., Budiyono, B., & Indriati, D. (2018). Motivation in Mathematics Learning. *Pancaran Pendidikan*, 7(1). <https://doi.org/10.25037/pancaran.v7i1.151>
- Fűzné Dr. Kószó, M. (2016). *A természetismeret tanításának módszertana*. Szegedi Tudományegyetem; Juhász Gyula Pedagógusképző Kar; Alkalmazott Pedagógiai Intézet; Tanítóképző Tanszék. <https://eta.bibl.u-szeged.hu/2204/2/index.html>
- Goforth, C. (2015). Using and Interpreting Cronbach's Alpha. *University of Virginia Library*. <https://library.virginia.edu/data/articles/using-and-interpreting-cronbachs-alpha>
- Gonzalez, J. (2017). *Retrieval Practice: The Most Powerful Learning Strategy You're Not Using*. Cult of Pedagogy. <https://www.cultofpedagogy.com/retrieval-practice/>
- Guncaga, J., Korenova, L., & Hvorecky, J. (2019). Calculators as Facilitators of Understanding Computational and Mathematical Contexts. *Open Education Studies*, 1(1), 177–183. <https://doi.org/10.1515/edu-2019-0012>
- Hajdu, G. (2018). A kvantitatív és a kvalitatív társadalomtudományi kutatás módszerei – dióhéjban. *Forum Sententiarum Curiae*, 2. https://kuria-birosag.hu/sites/default/files/fsc_tanulmany/2018_2_szam_hajdu.pdf
- Haladyna, T. M., & Downing, S. M. (1989). A Taxonomy of Multiple-Choice Item-Writing Rules. *Applied Measurement in Education*, 2(1), 37–50. https://doi.org/10.1207/s15324818ame0201_3
- Haladyna, T. M., & Downing, S. M. (1993). How Many Options is Enough for a Multiple-Choice Test Item? *Educational and Psychological Measurement*, 53(4), 999–1010. <https://doi.org/10.1177/0013164493053004013>
- Hercz, M. (2007). A pedagógiai értékelés gyakorlata. In *Bábosik István és Torgyik Judit (szerk.): Pedagógusmesterség az Európai Unióban* (o. 191–215). Eötvös Kiadó. https://www.researchgate.net/publication/314225676_A_pedagogiai_ertekeles_gyakorlata
- Hidegkuti, I., & Balázs, K. (2015). *Tesztelmélet*. https://www.researchgate.net/publication/305678594_Tesztelmelet

- Horovčák, P., & Stehlíková, B. (2007). Electronic testing of knowledge and factors influencing its results. *Acta Montanistica Slovac*, 12(2), 102–113.
https://www.researchgate.net/publication/26462944_Electronic_testing_of_knowledge_and_factors_influencing_its_results
- Huang, R., Spector, J. M., & Yang, J. (2019). *Educational Technology: A Primer for the 21st Century*. Springer Singapore. <https://doi.org/10.1007/978-981-13-6643-7>
- Konnova, L. P., Lipagina, L. V., Olekhova, E. F., Rylov, A. A., & Stepanyan, I. K. (2021). Problems of assessing students' knowledge in mathematics using LMS Moodle. *SHS Web of Conferences*, 125, 05011.
<https://doi.org/10.1051/shsconf/202112505011>
- Korenova, L. (2013). Usage possibilities of e-tests in a digital mathematical environment. In *Usta ad Albim BOHEMICA* (Köt. 13, o. 77–83).
https://www.researchgate.net/publication/317866496_Usage_possibilities_of_e-tests_in_a_digital_mathematical_environment
- Korenova, L. (2015). What to use for mathematics in high school: PC, tablet or graphing calculator? *International Journal for Technology in Mathematics Education*, 16(2), 59–64. https://doi.org/10.1564/tme_v22.2.03
- Magyar A., & Molnár G. (2013). Számítógép alapú adaptív és rögzített formátumú tesztelés összehasonlító hatékonyságvizsgálata. *Magyar Pedagógia*, 113(3), 181–193.
https://pedagogus.edia.hu/sites/default/files/public/3/Magyar_Molnar_2013.pdf
- Molnár, E. K., & Vigh, T. (2013). *A tantervelmélet és a pedagógiai értékelés alapjai*. „Mentor(h)áló 2.0 Program” TÁMOP-4.1.2.B.2-13/1-2013-0008 projekt.
http://www.jgypk.hu/mentor_halo/tananyag/Tantervelmlet_s_a_pedagogiai_rtke_ls_alapjai/
- Molnár G. (2003). Az ismeretek alkalmazásának vizsgálata modern tesztelméleti (IRT) eszközökkel. *Magyar Pedagógia*, 103(4), 423–446.
<https://www.magyarpedagogia.hu/index.php/magyarpedagogia/article/view/294/293>
- Moreno, R., Martínez, R. J., & Muñoz, J. (2015). Guidelines based on validity criteria for the development of multiple choice items. *Psicothema*, 27(4), 388–394.
<https://doi.org/10.7334/psicothema2015.110>
- Moreno-Guerrero, A.-J., Aznar-Díaz, I., Cáceres-Reche, P., & Alonso-García, S. (2020). E-Learning in the Teaching of Mathematics: An Educational Experience in

- Adult High School. *Mathematics*, 8(5), 840.
<https://doi.org/10.3390/math8050840>
- Muzsnay, A., Szabó, C., & Szeibert, J. (2024). Retrieval practice – a tool to be able to retain higher mathematics even 3 months after the exam. *Annales Mathematicae et Informaticae*, Accepted manuscript, 218–227.
<https://doi.org/10.33039/ami.2024.05.002>
- N. Tóth, Á. (2015). A tudásmérés. *A pedagógia adósságai. Savaria University Press*, 135–140. https://www.researchgate.net/publication/317551826_A_tudasmeres
- Négyesi, P. (2021). Adaptivitást támogató elektronikus tanulási környezet fogalmi keretei a számelmélet speciális témaköreinek tanításában. In E. K. Nagy & I. Zagyváné Szűcs (Szerk.), *Kihívások és megoldások a XXI. század pedagógiájában: Válogatás a Pedagógiai Szakbizottság tagjainak a munkáiból* (o. 89–101). Eszterházy Károly Katolikus Egyetem Líceum Kiadó.
<https://doi.org/10.46403/Kihivasokesmegoldasok.2021.89>
- Niemierko, B. (2012). Educational diagnostics for contemporary school systems : measuring and assessing growth of student human capital. Part I: main concepts and the scope. *COLLOQUIUM WNHIS*, 2(6), 123–144.
<https://www.semanticscholar.org/paper/Boles%C5%82aw-Niemierko-Educational-diagnostics-for-%3A-of-Niemierko/a4077f5327958f68857c1970d6dffde0f0eb951a>
- Niermeyer, R., & Seyffert, M. (2011). *Motivation*. Haufe-Lexware.
- Orosz G. (1994). Motiváció a matematikaórákon. In *Az Eszterházy Károly Tanárképző Főiskola tudományos közleményei (Új sorozat 22. köt.) Tanulmányok a matematikai tudományok köréből = Acta Academiae Paedagogicae Agriensis. Sectio Mathematicae* (o. 155–162). <http://publikacio.uni-eszterhazy.hu/id/eprint/4200>
- Panayides, P. (2013). Coefficient Alpha: Interpret With Caution. *Europe's Journal of Psychology*, 9(4), 687–696. <https://doi.org/10.5964/ejop.v9i4.653>
- Papp G. (2020). Az e-tesztek szerkesztése és alkalmazása tudásszintmérés céljából a karanténi távoktatás ideje alatt. In Abonyi-Tóth A., Zsákó L., & Stoffová V. (Szerk.), *Proceedings of XXXIII. DidMatTech 2020 Conference, New Methods and Technologies in Education, Research and Practice* (o. 203–212). ELTE Informatikai Kar. <http://didmattech.inf.elte.hu/wp->

content/uploads/2020/09/Didmattech2020_Proceedings_XXXIII_v20200921.pdf

- Papp, G. (2021a). A pandémia ideje alatt alkalmazott e-teszt megbízhatóságának vizsgálata. In A. Abonyi-Tóth, V. Stoffová, & L. Zsakó (Szerk.), *Proceedings of XXXIV. DidMatTech 2021 Conference. New Methods and Technologies in Education, Research and Practice* (o. 282–293). ELTE Informatikai Kar. <http://didmattech.inf.elte.hu/proceedings-2021/>
- Papp G. (2021b). Online felületek alkalmazása a matematika tanításában Kárpátalján. *Limes: a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola tudományos évkönyve*, VIII, 219–223. https://kmf.uz.ua/wp-content/uploads/2022/04/limes-2021-issue-viii_2021-12-30.pdf
- Papp G. (2022). Tudásszintmérés e-tesztek segítségével. In Marosi I. (Szerk.), *Hitkeresés. Az Ortutay Elemér Görögkatolikus Szakkollégium szervezésében 2021. március 18-án megrendezett VII. Keresztény Tudományos Diákköri Konferencia tanulmányainak gyűjteménye* (o. 39–46). Kálvin Nyomda. https://karpataljaiadatbank.com/wp-content/uploads/2022/08/IV.-kotet_Hitkereses.pdf
- Papp, G. (2023). Comparison of platforms used in online education. *Актуальні питання у сучасній науці*, (11(17)). [https://doi.org/10.52058/2786-6300-2023-11\(17\)-712-721](https://doi.org/10.52058/2786-6300-2023-11(17)-712-721)
- Papp G. (2023a). Informatika tárgyi e-teszt megbízhatóságának vizsgálata. In Tóth A. & Marosi I. (Szerk.), *Hittel a jövőbe. Az Ortutay Elemér Görögkatolikus Szakkollégium szervezésében 2022. május 5-én rendezett VIII. Keresztény Tudományos Diákköri Konferencia tanulmányainak gyűjteménye* (o. 49–63). RIK-U. <https://dspace.kmf.uz.ua/jspui/handle/123456789/2505>
- Papp, G. (2023b). Usage of Online Platforms in Education of Mathematics in Transcarpathia at the Beginning of Quarantine. In U. Kähler, M. Reissig, I. Sabadini, & J. Vindas (Szerk.), *Analysis, Applications, and Computations* (o. 155–162). Springer International Publishing. https://doi.org/10.1007/978-3-031-36375-7_9; <https://m2.mtmt.hu/gui2/?mode=browse¶ms=publication;34328493>
- Papp, G. (2023c). Using e-test in online mathematics education during a pandemic for Ferenc Rakoczi II Transcarpathian Hungarian College of Higher Education students. In *DisCo 2021: Active Learning in Digital Era: How Digital Tools*

- promote a Conscious, Open-minded, Creative and Social-Oriented Thinking* (o. 275–283). Centre for Higher Education Studies. <https://www.disconference.eu/wp-content/uploads/2023/08/16thconference-Reader-DisCo2021.pdf>; <https://m2.mtmt.hu/gui2/?mode=browse¶ms=publication;34325692>
- Papp, G. (2024). Examining e-test and its goodness indicators. *Academis Notes. Series: Pedagogical Sciences, Pedagogical Sciences*, 10, 136–139. https://doi.org/10.59694/ped_sciences.2024.10.136
- Papp G. (2024a). Online felületek összehasonlítása oktatásszervezés és e-teszt szerkesztése céljából. In Berghauer-Olasz E., Csopák É., Greba I., & Lizák K. (Szerk.), *Krízishelyzetek hatása és kihívásai az oktatásban. Nemzetközi tudományos konferencia Beregszász, 2023. március 30-31. Tanulmánykötet* (o. 351–358). II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola. <https://dspace.kme.org.ua/items/d4e632b1-b1da-4dc6-a99c-8c821daa81ee>
- Papp, G. (2024b). The use of e-tests in education as a tool for retrieval practice and motivation. *Teaching Mathematics and Computer Science*, 22(1), 59–76. <https://doi.org/10.5485/TMCS.2024.13495>
- Papp, G., & Szegő, D. (2021). Traditional test versus e-test. *EDULEARN21. 13th International Conference on Education and New Learning Technologies July 5th-6th, 2021*, 5029–5033. <https://doi.org/10.21125/edulearn.2021.1040>
- Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, 17(4), 382–395. <https://doi.org/10.1037/a0026252>
- Roediger, H. L., & Karpicke, J. D. (2006). Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM – Mathematics Education*, 55(2), 249–267. <https://doi.org/10.1007/s11858-022-01463-2>
- Schunk, D. H., & Usher, E. L. (2012). Social Cognitive Theory and Motivation. In R. M. Ryan (Szerk.), *The Oxford Handbook of Human Motivation* (1. kiad., o. 13–27).

<https://doi.org/10.1093/oxfordhb/9780195399820.013.0002>

- Shabatura, J. (2022). *Using Bloom's Taxonomy to Write Effective Learning Objectives*. <https://tips.uark.edu/using-blooms-taxonomy/>
- Siswaningsih, W., Firman, H., Zackiyah, & Khoirunnisa, A. (2017). Development of Two-Tier Diagnostic Test Pictorial-Based for Identifying High School Students Misconceptions on the Mole Concept. *Journal of Physics: Conference Series*, 812, 012117. <https://doi.org/10.1088/1742-6596/812/1/012117>
- Sokolova, M., & Totkov, Assoc. Prof. Dr. G. (2005). About Test Classification in E-Learning Environment. *International Conference on Computer Systems and Technologies*. https://www.researchgate.net/publication/251757911_About_Test_Classification_in_E-Learning_Environment
- Székely, J. (2014). *Mérés-értékelés a pedagógiában oktatási segédanyag integrációs szakmai feladatokra és pedagógus-szakvizsgára felkészítő szakirányú továbbképzési szak programhoz*. TÁMOP-jegyzet <https://docplayer.hu/29678336-Meres-ertekeles-a-pedagogiaban.html>; <https://aries.ektf.hu/~hz/pdf-tamop/index.html>
- Tóthné Parázsó, L. (2012). Online teszt és tudásszintmérés. In *Agria Media 2011: „Az információs társadalom az alkotó tudás társadalmának digitális előtere”*. (o. 117–125). Eszterházy Károly Főiskola Líceum Kiadó. <http://publikacio.uni-eszterhazy.hu/id/eprint/5952>
- Tóthné Parázsó, L. (2020). Teljesítményértékelés elektronikus tanulási környezetben. In T. Lengyelne Molnár (Szerk.), *A kultúraváltás hatása az egyéni kompetenciákra: A digitális kompetencia modelljei* (o. 119–142). Eszterházy Károly Egyetem Líceum Kiadó. <https://doi.org/10.46403/Akulturavaltashatasaegyeni.2020.119>
- Usiskin, Z. (2019). Beauty and Serendipity in Teaching Mathematics. *Educational Designer*, 3(12). <https://www.educationaldesigner.org/ed/volume3/issue12/article47/>
- Varga V. (2011). *Feleletválasztós tesztek feltöltésére, tárolására alkalmas online keretrendszer fejlesztése az oktatás elősegítése érdekében*. Szent István Egyetem Gazdaság- és Társadalomtudományi Kar Tudományos Diákköri Konferencia.
- Wigfield, A., Cambria, J., & Eccles, J. S. (2012). Motivation in Education. In R. M. Ryan (Szerk.), *The Oxford Handbook of Human Motivation* (1. kiad., o. 463–478).

Oxford

University

Press.

<https://doi.org/10.1093/oxfordhb/9780195399820.013.0026>

Wright, P. (2021). Transforming mathematics classroom practice through participatory action research. *Journal of Mathematics Teacher Education*, 24(2), 155–177.
<https://doi.org/10.1007/s10857-019-09452-1>

ÁBRÁK ÉS TÁBLÁZATOK JEGYZÉKE

1. ÁBRA. A FELADATOK RENDSZEREZÉSE A TANULÓI TEVÉKENYSÉG ALAPJÁN (EINHORN, 2012).....	26
2. ÁBRA. FELADATLAPOS SZÁMONKÉRÉS EGYIK FELADATA ÉS ANNAK HALLGATÓI MEGOLDÁSA (PAPP & SZEGŐ, 2021; PAPP, 2022).....	49
3. ÁBRA. HALLGATÓI SIKERESSÉGI MUTATÓK A ZÁRÓTESZT ITEMEI SZERINTI BONTÁSBAN (PAPP, 2024).	51
4. ÁBRA. Q6 ITEM MEGOLDÁSA.	51
5. ÁBRA. BE- ÉS KIMENETI TESZTEK ÖSSZEHASONLÍTÁSA.....	52
6. ÁBRA. ÖSSZEHASONLÍTÓ DIAGRAM AZ ITEMEKRE NÉZVE.	57
7. ÁBRA. A TESZTBEN SZEREPLŐ Q16 ITEM (PAPP, 2024B).....	57
8. ÁBRA. HALLGATÓK ELOSZLÁSA KÜLÖNBÖZŐ TELJESÍTÉSI INTERVALLUMBAN.	66
9. ÁBRA. AZ ITEMEK VIZSGÁLATA CSOPORTÁTLAG ALAPJÁN.	68
10. ÁBRA. Q3 ÉS Q11 – MEGFELELŐEN DIFFERENCIÁLÓ ITEMEK	68
11. ÁBRA. Q9 ÉS Q16 – KEVÉSBÉ SIKERES ITEMEK	69
12. ÁBRA. Q17 ITEM ÉS MEGOLDÁSA.....	70
13. ÁBRA. HALLGATÓI EREDMÉNYEK.	73
14. ÁBRA. HALLGATÓK VÁLASZADÁSA A TESZTEK ELMÉLETI KÉRDÉSEINÉL	76
15. ÁBRA. MOTIVÁCIÓ VIZSGÁLATA 5 FOKÚ LIKERT-SKÁLÁVAL.....	77
1. táblázat. A felülvizsgált Bloom-taxonómia összefoglaló táblázata (Anderson & Krathwohl, 2001; Demkanin et al., 2015).	16
2. TÁBLÁZAT. CRONBACH-ALFA OSZTÁLYOZÁSA (SISWANINGSIH ET AL., 2017).....	22
3. TÁBLÁZAT. KUTATÁSOKBAN RÉSZTVEVŐ CSOPORTOK	41
4. TÁBLÁZAT. PLATFORMOK ÉS BEÉPÍTETT TESZTSZERKESZTŐIK ÖSSZEHASONLÍTÁSA (PAPP, 2023, PAPP, 2024A).....	44
5. TÁBLÁZAT. TAXONÓMIÁK ÉS ITEMTÍPUSOK ÖSSZEFoglaló TÁBLázATA (PAPP, 2021A; PAPP, 2023C, PAPP, 2024B).....	55
6. TÁBLÁZAT. BE- ÉS KIMENETI TESZT SZERKESZTÉSI INFORMÁCIÓI.....	61
7. TÁBLÁZAT. ELŐHÍVÓ MINI TESZTEK SZERKESZTÉSI INFORMÁCIÓI.....	63
8. TÁBLÁZAT. A FELADATOK NEHÉZSÉGI MUTATÓINAK ÉS A TANÁRI ELŐREJELZÉSEK ÖSSZEHASONLÍTÁSA	66
9. TÁBLÁZAT. VÉLEMÉNYKIFEJTŐ INFORMÁCIÓK ÖSSZEGYŰJTŐ TÁBLázATA.	76

ÖSSZEFOGLALÁS

A tudásszintmérés és motiválás mindennapos tevékenység az oktatási folyamat során. Az oktatók órai megfigyelés, szóbeli feleltetés és különböző tesztek segítségével mérik és vizsgálják diákjaik tudását, miközben az óra folyamán különböző tanítási technikák és praktikák segítségével igyekeznek fenntartani a tanulók motiváltságát az anyag megértésére. Ebben a felgyorsult világban, ahol az információk tömkelegét tudjuk megtalálni a világhálón, nemcsak az általános és középiskolai oktatásban nehéz feladat ez, de a felsőoktatásban is. Főképp, ha ez nem a hallgatók szakjának megfelelő főtantárgy, csak egy szükséges és általánosan nehéznek elfogadott tárgy teljesítésére irányul.

A munkám fő célja volt egy olyan módszertani útmutató összeállítása, amely gyakorlati segítséget nyújt a pedagógusoknak tesztkészítési és értékelési gyakorlatuk fejlesztéséhez. E cél elérése érdekében, alapos háttérkutatásra és a tesztelmélet alapvető követelményeire támaszkodva, egy erős jószágmutatókkal rendelkező mérőeszközt hoztam létre a célcsoport tudásszintjének objektív mérésére, magában az útmutatóban pedig ennek a fejlesztési folyamatnak a lépéseit és az e-teszt sikerességének kiértékelését mutatom be. További kiemelt cél volt kérdőíves vizsgálattal feltérképezni a hallgatók véleményét a tesztelés oktatási-motivációs szerepéről, alátámasztva azt a feltevést, hogy a megfelelően kidolgozott tesztek a pusztá számonkérésen túlmutatva a tanulási folyamat szerves részévé válhatnak, ezáltal támogatva a diákok önértékelését, fejlődését.

Az 1. fejezetben a témában meghatározó szakirodalom segítségével végeztem háttérkutatást, és foglaltam azt össze a teljesség igénye nélkül. Első részében meghatároztam a teszt fogalmát (Csapó, 1988; Csapó et al, 2008; Demkanin et al., 2015; Hidegkuti & Balázs, 2015; Székely, 2014), Demkanin és mtársai (2015) nyomán megvizsgáltam és összehasonlítottam Bloom- és Niemierko-taxonómiáját (Anderson & Krathwohl, 2001; Demkanin et al., 2015). Ezután összefoglaltam a tesztek alapvető és egyéb jellemzőit (Demkanin et al., 2015), melyeket jószágmutatóknak (Cronbach, 1951; Csapó, 2004; Ebel & Frisbie, 1991; Molnár & Vigh, 2013; Siswaningsih et al., 2017) is nevezünk. Továbbiakban összefoglaltam az itemtípusokat (Csapó, 2004; Einhorn, 2012; Haladyna & Downing, 1993; Varga, 2011) és a feladatírás fontosabb szabályait (Haladyna & Downing, 1989; Shabatura, 2022; Varga, 2011). Végül az e-tesztekkel támogatott oktatást foglaltam össze (Csapó, 1988; Csapó et al, 2008; Demkanin et al., 2015; Huang et al., 2019; Korenova, 2015; Ususkin, 2019; Négyesi, 2021), mely az információs és kommunikációs technológiákon alapszik. Figyelmet fordítottam az

eszközként alkalmazott előhívásos tanulásra (Gonzalez, 2017; Horovčák & Stehlíková, 2007; Roediger & Karpicke, 2006), valamint a motivációra (Filgona et al., 2020; Fuqoha et al., 2018; Schukajlow et.al, 2023), amit megpróbáltam megvizsgálni a matematikaoktatás irányából (Cegléd, 2011; Orosz, 1994; Wigfield et al., 2012).

A 2. fejezet a kutatás módszertanát írja le, alfejezetei az órák megtervezését, valamint a különböző ciklusokban résztvevő mintákat foglalják össze.

A 3. fejezetben az előzetes kutatásokat foglaltam össze négy alfejezetre bontva. Az első két alfejezet a munkafelület megválasztását (Papp, 2020; Papp, 2021b, Papp, 2023b), összehasonlítását (Papp, 2023; Papp, 2024a) és az e-tesztek szerkesztését (Papp, 2020) mutatja be, melyek alapja oktatók és hallgatók kérdőíves felmérése, valamint a platformok önálló vizsgálata volt. A további alfejezetek a gyakorlatba beültetett e-tesztek szerkesztésének és alkalmazásának sikerességét vizsgálják két matematika tárgy esetén négy ciklusban.

A „Valószínűségszámítás és matematikai statisztika” tárgy keretében zajló előzetes (1. és 2. ciklus) kutatások (Papp & Szegő, 2021; Papp, 2022; Papp, 2024) összehasonlítási alapot nem képeztek, de annál többet tudtam tanulni, fejlődni, segítségével pontosabban tervezni és a hibák kiküszöbölésével fejleszteni az e-tesztet. A folyamatos fejlődés okán a „Matematikai analízis válogatott fejezetei” tárgy esetén az előzetes (3. és 4. ciklus) kutatások (Papp, 2021a; Papp, 2023c; Papp, 2024b) segítségével már kimutatható előrehaladásról számolhattam be az összehasonlítás során. A kutatásokról megjelent tanulmányokban a kezdeti és záró teszt megegyezett, valamint előhívó e-teszteket alkalmaztam. Az elemzés során vizsgáltam az záró tesztben lévő itemek nehézségét a kitöltések helyes válaszai alapján valamint az e-teszt jószágmutatóit. Egyik kutatás során interjú segítségével vizsgáltam a hallgatók véleményét a tesztekéről, motivációs hatásukról.

A 4. fejezetben az empirikus kutatást (5. és 6. ciklus) és elemzését mutatom be, melyet a komáromi Selye János Egyetem, valamint a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola további csoportjának bevonásával sikerült kivitelezni, alapját pedig a 3.3. alfejezetben bemutatott kutatás képezi, melyet a tapasztalatok alapján tovább módosítottam és fejlesztettem. A fejezetet a kutatási kérdések kategorizálását hangsúlyozva bontottam 3 alfejezetre.

A 4.1. alfejezet az e-tesztek létrehozását hivatott bemutatni. Meghatározom a különböző tesztek céljait, táblázatba foglalva bemutatom felépítését itemtípusok és

nehézsége alapján, valamint kiemelem azok Bloom-taxonómia dimenzióiba való besorolását.

Az alfejezetből kiderül, hogy az e-teszt létrehozásához figyelembe vettem a Bloom-taxonómia dimenzióit, és a lehető legpontosabban igyekeztem megválasztani, melyiknek megfelelő az adott item. Fontosnak tartom betartani a feladatírás említett szabályait, mely segítségével gondosan odafigyelhetünk az itemtípusok megválasztására, a nyelvezetre, disztraktorok megválasztására és megfogalmazására.

A 4.2. alfejezetben térek ki a jóságmutatókra, ahol részletesen elemeztem azok meglétét, miközben a vizsgált e-tesztet, mint mérőeszközt vizsgálom. Objektivitást, érvényességet és megbízhatóságot kutatom a nehézségi index, differenciálás és Cronbach-alfa segítségével. Az eredményekre támaszkodva válaszolhatom meg a **H1** hipotézist és hozzá tartozó kutatási kérdéseket:

K1. *Mely kognitív szinthez tartozó feladatok bizonyultak a legnehezebbnek a kombinatorika tesztben?*

A kérdés kapcsán az elemzés, az értékelés és az alkotás kognitív szintjei bizonyultak a legnehezebbnek. A leggyengébb eredményű feladatoknál (5., 7., 8., 11., 15., 17.) a kudarcot nem a képlettudás hiánya, hanem a speciális logikai kényszerfeltételek kezelése és a modellválasztási nehézségek okozták. Bár a kognitív hierarchiában felfelé haladva a teljesítmény folyamatosan csökkent, a módosított itemek differenciáló hatása megfelelőnek bizonyult.

K2. *Összhangban van-e a tanári előzetes nehézségi becslés a hallgatói eredményekkel (nehézségi index)?*

A kérdés rámutatott, hogy a tanári nehézségi becslés és a hallgatói eredmények nem vágtak egybe teljes mértékben. Fény derült egy jelentős „kognitív szakadékra”: a rutinszerű behelyettesítésről a mélyebb strukturális elemzésre való áttérés jóval nagyobb teljesítményromlást idézett elő, mint amit a feladatok matematikai összetettsége önmagában indokolt volna.

K3. *Milyen hatással van az előhívó e-teszt egy tudásmérő tesztre?*

A kérdés vizsgálatához használt letisztított mintából (37 tesztelő és 7 kontrollcsoportos hallgató) a szélsőértékeket és a középmezőnyt emeltem ki. Az előhívó e-tesztek kimeneti eredményekre gyakorolt közvetlen pozitív hatása nem volt egyértelműen kimutatható, aminek hátterében a felsőoktatási tematika és a feszes óraszám miatti gyors lefolyású felmérés állhat.

H1. *Az e-tesztekben a hallgatók teljesítménye szignifikánsan romlik a Bloom-taxonómia magasabb kognitív szintjeit (pl. alkalmazás, elemzés) igénylő feladatoknál, függetlenül az előzetesen meghatározott nehézségi szinttől, valamint a mini tesztek kitöltésétől.*

A **H1** hipotézis maradéktalanul beigazolódott: az alkalmazás szintjétől a feladatok nehézségi indexe romlani kezdett, az összetettebb elemzéseknél pedig drasztikus törés tapasztalható. Ez bizonyítja, hogy a magasabb szintű kognitív műveleteknél a hallgatói teljesítmény szignifikánsan visszaesik, függetlenül a tanári előrejelzéstől vagy a mini tesztek kitöltésétől.

A 4.3. alfejezetben a hallgatók által kitöltött kérdőív elemzését mutatom be, mely segítségével a hallgatók motiválásával kapcsolatos **H2** hipotézisre, valamint a hozzá tartozó kutatási kérdésekre kaptam választ.

K1. *Mennyire tudatos a válaszok megjelölése a tesztek kitöltése során?*

A kérdés kapcsán az elemzés pozitív képet mutat: a hallgatók véleménynyilvánítása alapján tudatosan és felelősségteljesen oldották meg a feladatokat. Az önbevallások szerint a tippelés szinte egyáltalán nem fordult elő, mindössze egyetlen diák jelölte, hogy a záró teszt elméleti részénél szisztematikusan tippelt.

K2. *Mely funkciót tartják a hallgatók az e-tesztek legfőbb előnyének: a tanulást segítő visszajelzést vagy a jegyszerzés megkönnyítését?*

A kérdés a hallgatók e-tesztekkel kapcsolatos attitűdjét vizsgálta. Az eredmények alapján a diákok az e-tesztek legfőbb előnyének a jegyszerzés megkönnyítését tekintik a tanulást támogató funkcióval szemben. Ezt igazolja, hogy a legmagasabb átlagot (3,97) az a válasz kapta, amely a sikerességet a megszerzett pontok beszámíthatóságához köti, miközben a kizárólag önértékelésre szolgáló lehetőségeket jóval kisebb átlaggal értékelték.

K3. *Van-e összefüggés a hallgatók motivációs szintje és a pontgyűjtési lehetőség között?*

A kérdés a motivációs szint és a pontgyűjtés közötti összefüggésre mutatott rá. A pontszerzés a részvételi kedv mellett a feladatmegoldás alaposságát is fokozta. A zárt kérdések adatai ezt megerősítik: a helytől és időtől független, de beszámítható pontokat adó gyakorló tesztek opciója kapta a legtöbb (23) jelölést, míg a tisztán önértékelő funkciót mindössze két hallgató választotta.

H2. *A hallgatók számára az e-tesztek motivációs erejét elsősorban nem az azonnali visszacsatolás (önértékelés), hanem a végső értékelésbe való beszámíthatóság (pontgyűjtés) adja.*

Mindezek alapján a **H2** hipotézis beigazolódt: az e-tesztek ösztönző erejét elsősorban nem a belső igényű önértékelés, hanem a végső értékelésbe való beszámíthatóság (a pontgyűjtés) adja.

Mindezek ismeretében meg tudom válaszolni a fő kérdést:

Mennyire alkalmas a Bloom-taxonómia kognitív szintjeire épülő e-teszt a hallgatók differenciált tudásszintjének objektív mérésére a felsőfokú matematikaoktatásban — a jelenlegi reliabilitási mutatók tükrében —, és e mérési folyamatban miként befolyásolja a pontszerzési lehetőség az e-tesztek motivációs hatékonyságát?

A Bloom-taxonómia kognitív szintjeire épülő e-teszt a csoportátlagok eloszlása és az itemek nehézségi indexei alapján kifejezetten alkalmasnak bizonyult a hallgatók differenciált tudásszintjének objektív mérésére a felsőfokú matematikaoktatásban, miközben a mérési folyamat során a pontszerzési lehetőség alapvetően meghatározta a rendszer motivációs hatékonyságát. Bár a tesztek önmagukban is képesek voltak fenntartani a hallgatói érdeklődést, a feladatmegoldás alapossága és az elkötelezettség akkor vált igazán intenzívvé, ha a megszerzett pontok beleszámítottak a jegyszerzésbe. A kutatás statisztikai így egyértelműen igazolják, hogy az e-tesztek valódi ösztönző erejét nem a belső igényű önértékelés, hanem ez a külső, pontgyűjtő funkció adja, amely közvetlenül és pozitívan fokozza a hallgatók részvételi kedvét, valamint a feladatokhoz való hozzáállását.

További kutatási lehetőségeket képez az alkalmazott tesztek javítása és továbbfejlesztése annak érdekében, hogy megbízhatósága magasabb értéket mutasson. Emellett az előhívó tesztek vizsgálata során nemcsak teszttel támogatott oktatás kapcsán, hanem ismeretet elmélyítő hatásának vizsgálata érdekében érdemes késleltetett méréseket tervezni és kivitelezni. Ennek következtében a tudásmérő teszt is nagyobb szerepet kaphat.

SUMMARY

Measuring knowledge levels and motivating students are daily activities in the educational process. Instructors measure and examine their students' knowledge through classroom observation, oral examinations, and various tests, while employing different teaching techniques and practices during class to maintain students' motivation to understand the material. In this fast-paced world, where a wealth of information can be found on the internet, this is a difficult task not only in primary and secondary education but also in higher education. This is especially true when the subject is not a core course related to the students' major, but is merely aimed at fulfilling a requirement for a subject that is universally accepted as difficult.

The main objective of my work was to compile a methodological guide that provides practical assistance to educators in developing their test creation and evaluation practices. To achieve this goal, relying on thorough background research and the fundamental requirements of test theory, I created a measurement tool with strong quality indicators to objectively measure the knowledge level of the target group. In the guide itself, I present the steps of this development process and the evaluation of the e-test's success. Another key objective was to map students' opinions on the educational-motivational role of testing through a questionnaire survey, supporting the assumption that properly developed tests can go beyond mere assessment and become an integral part of the learning process, thereby supporting students' self-assessment and development.

In Chapter 1, I conducted a background research using the definitive literature on the subject and summarized it without claiming to be exhaustive. In the first part, I defined the concept of the test (Csapó, 1988; Csapó et al., 2008; Demkanin et al., 2015; Hidegkuti & Balázs, 2015; Székely, 2014) and, following Demkanin et al. (2015), examined and compared Bloom's and Niemierko's taxonomies (Anderson & Krathwohl, 2001; Demkanin et al., 2015). Following this, I summarized the fundamental and other characteristics of tests (Demkanin et al., 2015), which are also referred to as psychometric or quality indicators (Cronbach, 1951; Csapó, 2004; Ebel & Frisbie, 1991; Molnár & Vigh, 2013; Siswaningsih et al., 2017). Furthermore, I summarized the item types (Csapó, 2004; Einhorn, 2012; Haladyna & Downing, 1993; Varga, 2011) and the key rules of item writing (Haladyna & Downing, 1989; Shabatura, 2022; Varga, 2011). Finally, I summarized e-test-supported education (Csapó, 1988; Csapó et al., 2008; Demkanin et al., 2015; Huang et al., 2019; Korenova, 2015; Ususkin, 2019; Négyesi, 2021), which is

based on information and communication technologies. Particular attention was paid to retrieval practice applied as a tool (Gonzalez, 2017; Horovčák & Stehlíková, 2007; Roediger & Karpicke, 2006), as well as to motivation (Filgona et al., 2020; Fuqoha et al., 2018; Schukajlow et al., 2023), which I attempted to examine from the perspective of mathematics education (Ceglédý, 2011; Orosz, 1994; Wigfield et al., 2012).

Chapter 2 describes the methodology of the research; its subchapters summarize the planning of the lessons, as well as the samples participating in the different cycles.

In Chapter 3, I summarized the preliminary research, divided into four subchapters. The first two subchapters present the selection (Papp, 2020; Papp, 2021b; Papp, 2023b) and comparison (Papp, 2023; Papp, 2024a) of the measurement platforms, as well as the editing of e-tests (Papp, 2020), which were based on questionnaire surveys of instructors and students, alongside an independent evaluation of the platforms. The subsequent subchapters investigate the success of editing and implementing e-tests in practice across four cycles for two mathematics subjects.

The preliminary research (Cycles 1 and 2) conducted within the framework of the "Probability calculation and mathematical statistics" course (Papp & Szegő, 2021; Papp, 2022; Papp, 2024) did not serve as a basis for comparison; however, it allowed me to learn, improve, plan more precisely, and develop the e-test by eliminating errors. Due to this continuous improvement, in the case of the "Selected chapters of mathematical analysis" course, I was already able to report demonstrable progress during the comparison of the preliminary research (Cycles 3 and 4) (Papp, 2021a; Papp, 2023c; Papp, 2024b). In the studies published on these researches, the pre- and post-tests were identical, and I utilized retrieval e-tests. During the analysis, I examined the item difficulty in the post-test based on the correct answers of the submissions, as well as the quality indicators of the e-test. In one of the studies, I used interviews to investigate students' opinions on the tests and their motivational effects.

In Chapter 4, I present the empirical research (Cycles 5 and 6) and its analysis, which was carried out with the involvement of J. Selye University in Komárno and an additional group from the Ferenc Rákóczi II Transcarpathian Hungarian College of Higher Education. This research is based on the study presented in Subchapter 3.3, which I further modified and developed based on the experiences gained. Emphasizing the categorization of the research questions, I divided this chapter into three subchapters.

Subchapter 4.1 is dedicated to presenting the creation of the e-tests. I define the objectives of the various tests, present their structure in a tabular format based on item

types and difficulty, and highlight their classification into the dimensions of Bloom's taxonomy.

This subchapter reveals that for the creation of the e-test, I took into account the dimensions of Bloom's taxonomy and endeavored to select the most appropriate dimension for each given item as precisely as possible. I consider it essential to adhere to the aforementioned rules of item writing, which allow us to pay careful attention to the choice of item types, language, and the selection and formulation of distractors.

In Subchapter 4.2, I address the goodness indicators, where I analyzed their presence in detail while examining the investigated e-test as a measurement tool. I investigate objectivity, validity, and reliability using the difficulty index, discrimination power, and Cronbach's alpha. Based on the results, I was able to address hypothesis H1 and its related research questions:

Q1. *Which tasks, belonging to which cognitive levels, proved to be the most difficult in the combinatorics test?*

In relation to this question, the cognitive levels of analyzing, evaluating, and creating proved to be the most difficult. For the tasks with the weakest results (items 5, 7, 8, 11, 15, 17), failure was caused not by a lack of formula knowledge, but by the handling of specific logical constraints and difficulties in model selection. Although performance continuously declined moving upward in the cognitive hierarchy, the discrimination power of the modified items proved to be appropriate.

Q2. *Is the teacher's preliminary difficulty estimation consistent with the students' results (difficulty index)?*

This question pointed out that the teacher's difficulty estimation and the students' results did not fully coincide. A significant "cognitive gap" was revealed: the transition from routine substitution to deeper structural analysis caused a much greater decline in performance than would have been justified by the mathematical complexity of the tasks alone.

Q3. *What effect does a retrieval-based e-test have on a summative knowledge test?*

From the cleaned sample used to investigate this question (37 students in the experimental group and 7 in the control group), I highlighted the extreme values and the middle range. A direct positive effect of retrieval e-tests on the final outcomes could not be clearly demonstrated, which may be due to the fast-paced nature of the assessment dictated by the higher education curriculum and the tight number of class hours.

H1. *In e-tests, students' performance declines significantly in tasks requiring higher cognitive levels of Bloom's taxonomy (e.g., applying, analyzing), regardless of the predetermined difficulty level or the completion of mini-tests.*

Hypothesis **H1** was fully confirmed: starting from the application level, the difficulty index of the tasks began to deteriorate, and a drastic drop was observed in the case of more complex analyses. This proves that student performance declines significantly at higher-level cognitive operations, regardless of the teacher's prediction or the completion of the mini-tests.

In Subchapter 4.3, I present the analysis of the questionnaire completed by the students, through which I obtained answers to hypothesis **H2** regarding student motivation, as well as its related research questions.

Q1. *How conscious is the selection of answers during the completion of the tests?*

In relation to this question, the analysis shows a positive picture: based on the students' feedback, they solved the tasks consciously and responsibly. According to self-reports, guessing barely occurred at all; only a single student indicated that they guessed systematically in the theoretical part of the post-test.

Q2. *Which function do students consider the main advantage of e-tests: feedback that aids learning, or making it easier to obtain a grade?*

This question investigated students' attitudes toward e-tests. Based on the results, students view making grade acquisition easier as the primary advantage of e-tests, as opposed to their learning-support function. This is confirmed by the fact that the highest mean (3.97) was given to the response that links success to the accountability of the points earned, while options serving solely for self-assessment were rated with a much lower mean.

Q3. *Is there a correlation between students' motivation levels and the opportunity to collect points?*

This question highlighted the correlation between motivation levels and point accumulation. Earning points enhanced not only the willingness to participate but also the thoroughness of task-solving. Data from the closed-ended questions confirm this: the option of practice tests providing points that count toward the grade, independent of place and time, received the most selections (23), while the purely self-assessment function was chosen by only two students.

H2. *For students, the motivational power of e-tests is primarily derived not from immediate feedback (self-assessment), but from their countability toward the final evaluation (point accumulation).*

Based on all the above, Hypothesis **H2** was confirmed: the incentive power of e-tests is driven primarily not by an intrinsic need for self-assessment, but by their accountability in the final evaluation (point accumulation).

In light of all this, I can answer the primary research question:

To what extent is an e-test based on the cognitive levels of Bloom's taxonomy suitable for the objective measurement of students' differentiated knowledge levels in higher education mathematics—in view of current reliability indicators—and how does the opportunity to earn points influence the motivational effectiveness of e-tests within this measurement process?

Based on the distribution of group means and the difficulty indices of the items, the e-test built upon the cognitive levels of Bloom's taxonomy proved to be highly suitable for the objective measurement of students' differentiated knowledge levels in higher mathematics education, while the opportunity to earn points fundamentally determined the motivational efficiency of the system throughout the measurement process. Although the tests were capable of maintaining student interest on their own, the thoroughness of task-solving and commitment became truly intensive when the points earned counted toward their final grade. The statistics of the research thus clearly demonstrate that the true incentive power of e-tests is driven not by an intrinsic need for self-assessment, but by this extrinsic, point-accumulating function, which directly and positively enhances students' willingness to participate as well as their attitude toward the tasks.

Further research opportunities include the refinement and development of the applied tests to ensure their reliability demonstrates a higher value. Furthermore, during the investigation of retrieval tests, it is worthwhile to design and implement delayed post-tests—not only in relation to test-supported education but also to examine their impact on knowledge retention. Consequently, the knowledge assessment test itself could play a more prominent role.

FÜGGELÉK

A. Bemeneti e-teszt

Elérhető: <https://onlinetestpad.com/o3klqpuvrwbue>

Kombinatorika - Kezdeti teszt

Leírás

Kedves Hallgatók!

A teszt célja az előzetes tudás mérése a Kombinatorika témakörben. Célunk a hosszú távú memóriából előhívni és feleleveníteni azt a tanult ismeretet, mellyel az oktatás kezdeti szintjén rendelkeznek a témában.

A teszt kitöltésére 45 perc áll rendelkezésükre. A feladatoknál visszalépésre nincs lehetőség és a felület csak akkor enged tovább a következő kérdésre, ha válaszadás történt. Kérjük figyeljenek az időbeosztásra, a számláló jelzi majd a fennmaradt időt.

Kérjük mindenki a saját tudása szerint töltse ki a tesztet!

Kérem töltse ki a mezőket!

Szak	
Név	
Intézmény	
Hozzájárul anoním adatainak felhasználásához?	

Author: Papp Gabriella

Kombinatorika - Kezdeti teszt

1

1 from 20

Pótolja a hiányzó szót!

Adott n különböző elem ismétlés nélküli _____ az elemek egy meghatározott sorrendjét értjük.

2

2 from 20

Határozza meg az állítás igazságtartalmát!

" n különböző elemből tetszőlegesen választott k különböző elem nem meghatározott sorrendjét az n elem k -ad osztályú ismétlés nélküli variációjának nevezzük."

- ☐ Igaz
☐ Hamis

3

3 from 20

Párosítsa a fogalmakat!

Permutáció	
Variáció	
Kombináció	
Ismétléses kombináció	

- 1 $n!$
- 2 $\frac{n!}{k_1!k_2!\dots k_r!}$
- 3 $\frac{n!}{(n-k)!}$
- 4 n^k
- 5 $\frac{n!}{k!(n-k)!}$
- 6 $\frac{(n+k-1)!}{k!(n-1)!}$

4

4 from 20

Egy 8 fős társaságot hányféleképpen tudunk leültetni egy padra? (válaszként egy természetes számot írjon)

5

5 from 20

Egy 10 fős csoportban Janika és Sanyika NEM ülhet egymás mellé a mozi székeire egy sorba. Hányféleképpen ültetheti le a csoportot a tanítónéni?

- ☐ $10! - 2 \cdot 9!$
☐ $2 \cdot 9!$
☐ $2 \cdot 8!$
☐ $10! - 2 \cdot 8!$

6

6 from 20

Hány 7 jegyű telefonszámot képezhetünk a 2, 2, 2, 3, 3, 5, 5 számjegyek felhasználásával?

7

7 from 20

F	Ö	I	S
Ö	I	S	K
I	S	K	O
S	K	O	L
K	O	L	A

Hányféleképpen olvasható ki a következő ábrából a FŐISKOLA szó, ha a táblázat bal felső sarkából indulunk és minden lépésben csak jobbra vagy lefelé haladhatunk?

- ☐ 35
- ☐ 128
- ☐ 256
- ☐ 56

8

8 from 20

Hány ötjegyű páratlan szám képezhető a 0, 1, 2, 3, 4, 5 számjegyekből, ha minden számjegy csak egyszer szerepelhet?

9

9 from 20

A pultban 8 különböző ízű fagyalt van. Hányféleképpen lehet kiválasztani háromgombócos fagyaltot, ha mindhárom gombóc különböző ízű?

- ☐ 336
- ☐ 24
- ☐ 56
- ☐ 512

10

10 from 20

Egy dobókockával hétszer dobunk egymás után. Hányféle dobássorozat fordulhat elő, amelyben pontosan egyszer szerepel 6-os? (Válaszát kétféleképpen adja meg)

- ☐ 326592
- ☐ 109375
- ☐ $7 \cdot 6^6$
- ☐ $7 \cdot 5^6$

11

11 from 20

Melyik kombinatorikai elem segítségével oldható meg a feladat?

A matek vizsgán 10 tesztkérdést tesznek fel, mindegyikre négy válasz közül kell megjelölni egyet. Hányféle válasz születhet?

12

12 from 20

Hányféleképpen rakhatunk be 6 levelet 13 rekeszbe, ha a levelek között nem teszünk különbséget és egyrekeszbe maximum egy levelet teszünk?

- ☐ 78
- ☐ $13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$
- ☐ 1716
- ☐ 18564

13

13 from 20

100 csavar közül, amelyek között 10 darab selejtes, kiválasztunk 5 csavart. Hány olyan eset van, amelyben a kiválasztott mind hibátlan?

- ☐ C_{90}^5
- ☐ V_{90}^5
- ☐ 43949268
- ☐ 5273912160

14

14 from 20

Mityu bácsi a hétszemélyes tütáján 18 embert át szeretne szállítani a folyó másik oldalára, de a szabályok betartásával ezt 3 fordulóval tudja csak teljesíteni. Hányféleképpen teheti?

- ☐ $C_{18}^6 \cdot C_{12}^6 \cdot C_6^6$
- ☐ $C_{18}^7 \cdot C_{11}^7 \cdot C_4^4$
- ☐ $V_{18}^6 \cdot V_{12}^6 \cdot V_6^6$
- ☐ $V_{18}^7 \cdot V_{11}^7 \cdot V_4^4$

15

15 from 20

Egy cégnél három osztályvezető, hat csoportvezető és harminc beosztott dolgozik. Hányféleképpen választhatunk ki közülük egy küldöttséget, melyben egy osztályvezető, két csoportvezető és tíz beosztott szerepel, ha a kereskedelmi osztály vezetője és az áruforgalmi csoport vezetője mindenképpen a küldöttek között kell, hogy legyenek?

- ☐ $1 + C_5^1 + C_{30}^{10}$
- ☐ $1 \cdot C_5^1 \cdot C_{30}^{10}$
- ☐ $1 + V_5^1 + V_{30}^{10}$
- ☐ $1 \cdot V_5^1 \cdot V_{30}^{10}$

16

16 from 20

Egy tánciskolában 7 fiú és 5 lány közül hányféleképpen alkotnak 5 egyszerre táncoló párt?

Határozza meg az állítás igazságtartalmát!

A feladat megoldásában a lányok sorbarendezésének számossága is lényeges elem.

- ☐ Igaz
- ☐ Hamis

17

17 from 20

Három dobókockával dobva 3 számjegyből álló "dobáshármast" kapunk. Hányféle eredmény adódhat?

18

18 from 20

Pótolja a hiányzó szavakat!

n különböző elemekből k elemet kiválasztva úgy, hogy egy elem többször is szerepelhet, és nem vagyunk tekintettel a kiválasztás sorrendjére, az n elem k-ad osztályú _____ nevezzük.

19

19 from 20

Egy 32 fős osztályban egy verseny folyamán négy ugyanolyan tárgyat sorsolnak ki. Hányféleképpen történhet ez, ha egy tanuló több tárgyat is elnyerhet?

- ☐ $\frac{35 \cdot 34 \cdot 33 \cdot 32}{4!}$
- ☐ 4^{32}
- ☐ $\frac{32 \cdot 31 \cdot 30 \cdot 29}{4!}$
- ☐ 32^4

20

20 from 20

Adott számú könyvből négyet 210-féleképpen lehet kiválasztani. Mennyi a könyvek száma?

B. Permutáció

Elérhető: <https://onlinetestpad.com/hyqovxz3ohq44>

Permutáció

Leírás

Kedves Hallgatók!

A teszt célja a felelevenített ismeretet előhívni a rövidtávú memóriából. Célunk, hogy motiváljuk Önöket az órai aktivitásra, az aktív figyelem fenntartására, valamint az előhívással segítsük az információ hosszútávú memóriában való elhelyezését.

Kérem mindenki a saját tudása szerint töltse ki!

Kérem töltse ki a mezőket!

Intézmény

Szak

Hozzájárul anoním adatainak felhasználásához?

Név

Permutáció

1

1 from 5

Mi a permutáció?

☐ n különböző elem egy meghatározott sorrendje

☐ n különböző elemből tetszőlegesen kiválasztott k különböző elem egy meghatározott sorrendje

☐ n különböző elemből tetszőlegesen kiválasztott k különböző elem, ahol a kiválasztott elemek sorrendje nem számít

2

2 from 5

Válassza ki a permutáció képletét!

☐ $n!$

☐ $(n - k)!$

☐ $\frac{n!}{k_1!k_2!\dots k_r!}$

☐ $\frac{n!}{(n - k)!}$

3

3 from 5

Hányféleképpen készíthetsz egy listát 5 tanulóból?

☐ 45

☐ 120

☐ 240

☐ 25

4

4 from 5

Hány hétjegyű telefonszám képezhető a 2, 2, 2, 3, 3, 5, 5 számjegyekből?

☐ 120

☐ 210

☐ 5040

☐ 6

5

5 from 5

Egy 8 főből álló társaság egy kör alakú asztalnál akar helyet foglalni. Hányféleképpen történhet ez meg?

C. Variáció

Elérhető: <https://onlinetestpad.com/3unqo6e6qnnqu>

Variáció

Leírás

Kedves Hallgatók!

A teszt célja a felelevenített ismeretet előhívni a rövidtávú memóriából. Célunk, hogy motiváljuk Önöket az órai aktivitásra, az aktív figyelem fenntartására, valamint az előhívással segítsük az információ hosszútávú memóriában való elhelyezését.

Kérem mindenki a saját tudása szerint töltse ki!

Kérem töltse ki a mezőket!

Név

Intézmény

Szak

Hozzájárul anoním adatainak felhasználásához?

Variáció

1

1 from 5

Mi a variáció?

☐ n különböző elem egy meghatározott sorrendje

☐ n különböző elemből tetszőlegesen kiválasztott k különböző elem egy meghatározott sorrendje

☐ n különböző elemből tetszőlegesen kiválasztott k különböző elem, ahol a kiválasztott elemek sorrendje nem számít

2

2 from 5

Válassza ki a variáció és ismétléses variáció képletét!

☐ $n!$

☐ $\frac{n!}{(n-k)!}$

☐ $\frac{n!}{k!(n-k)!}$

☐ n^k

3

3 from 5

6 szabad hely maradt az asztalnál. Hányféleképpen tud 4 ember helyet foglalni?

☐ 4^6

☐ 696

☐ 360

☐ $4!$

4

4 from 5

A Morze ABC-ben kétféle jel van: pont és vonás. Hányféle legfeljebb 5 elemű jelsorozat képezhető?

☐ 25

☐ 62

☐ 2^5

☐ 20

5

5 from 5

Egy focibajnokságon 8 csapat vett részt. Hányféleképpen osztható ki az arany, ezüst és bronz érem a csapatok között?

D. Kombináció

Elérhető: <https://onlinetestpad.com/gprizjwigaynk>

Kombináció

Leírás

Kedves Hallgatók!

A teszt célja a felelevenített ismeret előhívni a rövidtávú memóriából. Célunk, hogy motiváljuk Önöket az órai aktivitásra, az aktív figyelem fenntartására, valamint az előhívással segítsük az információ hosszútávú memóriában való elhelyezését.

Kérem mindenki a saját tudása szerint töltse ki!

Kérem töltse ki a mezőket!

Név

Intézmény

Szak

Hozzájárul anoním adatainak felhasználásához?

Kombináció

1

1 from 5

Mi a kombináció?

- ☐ n különböző elem egy meghatározott sorrendje
- ☐ n különböző elemből tetszőlegesen kiválasztott k különböző elem egy meghatározott sorrendje
- ☐ n különböző elemből tetszőlegesen kiválasztott k különböző elem, ahol a kiválasztott elemek sorrendje nem számít

2

2 from 5

Válassza ki a kombináció képletét! / Виберіть формулу комбінації!

- ☐ $n!$
- ☐ $\frac{n!}{(n-k)!}$
- ☐ $\frac{n!}{k!(n-k)!}$
- ☐ $\frac{(n+k-1)!}{k!(n-1)!}$

3

3 from 5

Hányféleképpen készíthető 8 süteményből álló csomag, ha 4 különböző fajta sütemény van?

- ☐ 70
- ☐ 165
- ☐ 1680
- ☐ 4096

4

4 from 5

Három autóba 9 ember száll be úgy, hogy minden autóba 3 ember jut. Hányféleképpen történhet ez?

- ☐ $C_9^3 \cdot C_6^3 \cdot C_3^3$
- ☐ $C_9^3 + C_6^3 + C_3^3$
- ☐ 1680
- ☐ 105

5

5 from 5

A sakkörbe 5 fiú és 3 lány jár. Három meghívót kaptunk a nagymesterrel való találkozáshoz. Hányféleképpen oszthatod szét a meghívókat, ha legalább 1 lány kell eljusszon a találkozóra?

E. Kimeneti e-teszt

Elérhető: <https://onlinetestpad.com/ftdpzvj7ul54y>

Kombinatorika - Záró teszt	45:00
Leírás	
Kedves Hallgatók!	
A teszt célja az megszerzett tudás mérése a Kombinatorika témakörben. Célunk az oktatás végén megvizsgálni a tudás elmélyítésének mértékét, valamint megvizsgálni az előhíásos tesztelés hatását az eredményre.	
A teszt kitöltésére 45 perc áll rendelkezésükre. Kérem mindenki a saját tudása szerint töltse ki!	
Kérem töltse ki a mezőket!	
Név	
Intézmény	
Szak	
Hozzájárul anoním adatainak felhasználásához?	
Author: Papp Gabriella	

Kombinatorika - Záró teszt

1

1 from 20

Pótolja a hiányzó szót!

Adott n különböző elem ismétlés nélküli _____ az elemek egy meghatározott sorrendjét értjük.

2

2 from 20

Határozza meg az állítás igazságtartalmát!

" n különböző elemből tetszőlegesen választott k különböző elem nem meghatározott sorrendjét az n elem k -ad osztályú ismétlés nélküli variációjának nevezzük."



Igaz



Hamis

3

3 from 20

Párosítsa a fogalmakat!

Permutáció

Variáció

Kombináció

Ismétléses variáció

1 $n!$

2 $\frac{n!}{k_1!k_2!\dots k_r!}$

3 $\frac{n!}{(n-k)!}$

4 n^k

5 $\frac{n!}{k!(n-k)!}$

6 $\frac{(n+k-1)!}{k!(n-1)!}$

4

4 from 20

Egy 8 fős társaságot hányféleképpen tudunk leültetni egy padra? (válaszként egy természetes számot írjon)

5

5 from 20

Egy 10 fős csoport egy sorban foglal helyet a moziban. Hányféleképpen ülhetnek le, ha Peti és Rozi egymás mellett akarnak ülni?



$10! - 2 \cdot 9!$



$2 \cdot 9!$



$2 \cdot 8!$



$10! - 2 \cdot 8!$

6

6 from 20

Hány 7 jegyű jelszót képezhetünk a 3, 3, 3, 5, 5, 7, 7 számjegyek felhasználásával?

7

7 from 20

F	Ö	I	S
Ö	I	S	K
I	S	K	O
S	K	O	L
K	O	L	A

Hányféleképpen olvasható ki a következő ábrából a FŐISKOLA szó, ha a táblázat bal felső sarkából indulunk és minden lépésben csak jobbra vagy lefelé haladhatunk?

- ☐ 35
- ☐ 128
- ☐ 256
- ☐ 56

8

8 from 20

Hány ötjegyű páros szám képezhető a 0, 1, 2, 3, 4, 5 számjegyekből, ha minden számjegy csak egyszer szerepelhet?

9

9 from 20

A hűtőben 8 különböző ízű joghurt van. Hányféleképpen lehet kiválasztani három joghurtot, ha mindhárom különböző ízű?

- ☐ 336
- ☐ 24
- ☐ 56
- ☐ 512

10

10 from 20

Egy dobókockával hétszer dobunk egymás után. Hányféle dobássorozat fordulhat elő, amelyben pontosan egyszer szerepel 6-os? (Válaszát kétféleképpen adja meg)

- ☐ 326592
- ☐ 109375
- ☐ $7 \cdot 6^6$
- ☐ $7 \cdot 5^6$

11

11 from 20

Melyik kombinatorikai elem segítségével oldható meg a feladat?

A matek vizsgán 10 tesztkérdést tesznek fel, mindegyikre négy válasz közül kell megjelölni egyet. Hányféle válasz születhet?

12

12 from 20

Hányféleképpen rakhatunk be 6 golyót 13 dobozba, ha a golyók között nem teszünk különbséget és dobozba maximum egy golyót teszünk?

- ☐ 78
- ☐ $13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$
- ☐ 1716
- ☐ 18564

13

13 from 20

100 fülhallgató közül, amelyek között 10 darab selejtes, kiválasztunk 5 fülhallgatót. Hány olyan eset van, amelyben a kiválasztott mind hibátlan?

- ☐ C_{90}^5
- ☐ V_{90}^5
- ☐ 43949268
- ☐ 5273912160

14

14 from 20

Imre bácsi egy hétszemélyes autóval 18 embert szeretne szállítani a szomszéd faluba, de a szabályok betartásával ezt 3 fordulóra tudja csak teljesíteni. Hányféleképpen teheti?

- ☐ $C_{18}^6 \cdot C_{12}^6 \cdot C_6^6$
- ☐ $C_{18}^7 \cdot C_{11}^7 \cdot C_4^4$
- ☐ $V_{18}^6 \cdot V_{12}^6 \cdot V_6^6$
- ☐ $V_{18}^7 \cdot V_{11}^7 \cdot V_4^4$

15

15 from 20

Egy cégnél három osztályvezető, hat csoportvezető és harminc beosztott dolgozik. Hányféleképpen választhatunk ki közülük egy küldöttséget, melyben egy osztályvezető, két csoportvezető és tíz beosztott szerepel, ha a kereskedelmi osztály vezetője és az áruforgalmi csoport vezetője mindenképpen a küldöttek között kell, hogy legyenek?

- ☐ $1 + C_5^1 + C_{30}^{10}$
- ☐ $1 \cdot C_5^1 \cdot C_{30}^{10}$
- ☐ $1 + V_5^1 + V_{30}^{10}$
- ☐ $1 \cdot V_5^1 \cdot V_{30}^{10}$

16

16 from 20

Egy sakkversenyen 7 fiú és 5 lány közül hányféleképpen alkotnak 5 egyszerre sakkozó párt, ha azonos neműek nem sakkozhatnak egymással?

Határozza meg az állítás igazságtartalmát!

A feladat megoldásában a lányok sorbarendezésének számossága is lényeges elem.

- ☐ Igaz
- ☐ Hamis

17

17 from 20

Három dobókockával dobva 3 számjegyből álló "dobáshármast" kapunk. Hányféle eredmény adódhat?

18

18 from 20

Pótolja a hiányzó szavakat!

n különböző elemekből k elemet kiválasztva úgy, hogy egy elem többször is szerepelhet, és nem vagyunk tekintettel a kiválasztás sorrendjére, az n elem k-ad osztályú _____ nevezzük.

19

19 from 20

Egy iskolai szüreti bálon 32 tombolajegy kelt el. A sorsolás folyamán négy ugyanolyan tárgyat sorsolnak ki. Hányféleképpen történhet ez, ha egy tanuló több tárgyat is elnyerhet?

- ☐ $\frac{35 \cdot 34 \cdot 33 \cdot 32}{4!}$
- ☐ 4^{32}
- ☐ $\frac{32 \cdot 31 \cdot 30 \cdot 29}{4!}$
- ☐ 32^4

20

20 from 20

Adott számú könyvből négyet 210-féleképpen lehet kiválasztani. Mennyi a könyvek száma?

F. Kérdőív

Kérdőív

Kedves Hallgatók!

2 hosszabb és 3 rövidebb teszt kitöltése után szeretnénk, ha válaszolnának a kérdőívben szereplő kérdésekre. A célunk információt szerezni arról, mi a véleményük az előhívó tesztekéről, azok hatásáról a motivációjukra a tudás elmélyítésében.

A kérdőív kitöltőinek adatait anoním formában szeretnénk felhasználni a kutatás értékelése során. A név regisztrálására csak addig van szükségünk, míg az előhívó és tudásmérő tesztek eredményeivel kötjük össze a kitöltőt.

** Kötelező kérdés*

1. Név *

2. Nem *

Soronként csak egy oválist jelöljön be.

☐ Nő

☐ Férfi

3. Intézmény *

Soronként csak egy oválist jelöljön be.

☐ Selye János Egyetem, Komárom

☐ II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola, Ukrajna

4. Szak *

5. Hozzájárul anoním adatainak felhasználásához? *

Soronként csak egy oválist jelöljön be.

☐ Igen

☐ Nem

E-tesztek

A következő kérdések az e-tesztek kitöltésével kapcsolatosak

6. Hogyan jelölte az e-tesztek elméleti kérdéseit? *

Válassza ki az összeset, amely érvényes.

	Mindegyikre emlékezett, ezért tudatosan jelölt	Némelyikre emlékezett, azt tudatosan jeölte	Nem volt biztos a válasz helyességében, de a tanultak alapján igyekezett jelölni	Csak azoknál tippelt, amelyekre nem emlékezett pontosan	Mindegyiknél tippelt
Kezdeti teszt	☐	☐	☐	☐	☐
Permutáció	☐	☐	☐	☐	☐
Variáció	☐	☐	☐	☐	☐
Kombináció	☐	☐	☐	☐	☐
Záró teszt	☐	☐	☐	☐	☐

7. Hgyan jelölte az e-tesztek gyakorlati feladatait? *

Soronként csak egy oválist jelöljön be.

- ☐ Mindegyiket kiszámolta
☐ Mindegyiket megpróbálta kiszámolni, de nem mindegyiket sikerült
☐ Ahol lehetett, tippelt, csak ott számolt, ahol saját eredményt kért a feladat
☐ Mindegyiket tippelte

8. Ön szerint megbízhatóan alkalmazható-e az e-teszt, mint a tudásszint mérőeszköze?

Soronként csak egy oválist jelöljön be.

- ☐ Igen
☐ Nem
☐ Talán

9. Okozott-e önnek bármilyen nehézséget a tesztek kitöltése? *

Soronként csak egy oválist jelöljön be.

- ☐ Igen
☐ Nem [Ugrás a\(z\) 11. kérdésre](#)

Okozott-e önnek bármilyen nehézséget a tesztek kitöltése?

Ez a kérdés csak akkor jelenik meg, ha az előző kérdésre IGENnel válaszolt

10. Amennyiben nehézségekbe ütközött a tesztek folyamán, kérem jelezze, mik voltak azok! (pl. időhiány, szövegezés, számolási hibák)

Motiváció

A következő kérdések az e-tesztek motivációs hatásával kapcsolatosak

11. Kérjük pontozza 1 és 5 között, mennyire motiválták Önt a következő események az aktív figyelem fenntartására! *

Soronként csak egy oválist jelöljön be.

	1 (egyáltalán nem motivált)	2	3	4	5 (nagyon motivált)
Kezdeti teszt nehézségei	☐	☐	☐	☐	☐
Előhívás tesztjei az órák végén	☐	☐	☐	☐	☐
Minél jobb eredmény elérése a tesztekben saját tudás mértékjeként	☐	☐	☐	☐	☐
Minél jobb eredmény elérése a tesztekben pontgyűjtés céljából	☐	☐	☐	☐	☐

12. Véleménye szerint Ön mitől lesz motiváltabb az e-tesztek elvégzésére? *

Válassza ki az összeset, amely érvényes.

- ☐ Ha azok gyakorló tesztek, melyeket helytől és időtől függetlenül lehet kitölteni és sikere csak a tudás elmélyítésére és önértékelésre szolgál
- ☐ Ha azok gyakorló tesztek, melyeket helytől és időtől függetlenül lehet kitölteni, de sikere esetén a megszerzett pontok beszámoltathatóak
- ☐ Ha azok előhívó tesztek, de a megszerzett pontszám csak önértékelésre szolgál
- ☐ Ha azok előhívó tesztek, de a megszerzett pontszám a tárgy teljesítésébe beszámolható
- ☐ Ha azok tudásmérő, modulzáró tesztek a tárgy teljesítése céljából

13. Hasznosnak érezte-e a témák végén írt előhívó teszteket? *

Soronként csak egy oválist jelöljön be.

- ☐ Igen
- ☐ Nem
- ☐ Talán

14. Ön szerint segíti-e az előhívó tesztek kitöltése az adott témakörre vonatkozó önértékelést? *

Soronként csak egy oválist jelöljön be.

- ☐ Igen
- ☐ Nem
- ☐ Talán

15. Ön szerint az előhívó tesztek sikerei segítettek-e a modulzáró dolgozatra való felkészülésben? *

Soronként csak egy oválist jelöljön be.

- ☐ Igen
- ☐ Nem
- ☐ Talán

16. Hogyan segítették Önt az előhívó tesztek a modulzáróra való felkészülésben?

Soronként csak egy oválist jelöljön be.

- ☐ Megvizsgálta az elkövetett hibákat és a felkészülés során megpróbálta kiküszöbölni azokat
- ☐ Nem sikerültek a saját értékrendjének megfelelően az előhívó tesztek, ezért még többet készült a modulzáróra
- ☐ Könnyebb volt szétválasztani az egyes elemeket, mivel az előhívó tesztekben külön voltak
- ☐ Mivel minden órán tesztelvel lett az önálló feladatmegoldás adott témában, könnyebbnek érezte ugyanezt egy összefoglaló modulzáró esetén

17. Amennyiben az előző válaszlehetőségek egyike sem igaz Önre, kérem fogalmazza meg, hogyan segítették Önt az előhívó tesztek a modulzáróra való felkészülésben.

Ezt a tartalmat nem a Google hozta létre, és nem is hagyta azt jóvá.

Google Űrlapok