



# Értelmezések, viták

**Klein Balázs**

## Pontosítás

A 2012-es magyar irodalom érettségi szövegértési feladatsorának adatelemzése<sup>1</sup>

*A 2012-es középszintű magyar nyelv és irodalom érettségi feladatok javítókulcsa miatt polémia alakult ki a sajtóban a Magyar tanárok Egyesülete és a tételkészítő bizottság között.<sup>2</sup> Vizsgálatunkban mintegy 250 kitöltött érettségi feladatsor statisztikai feldolgozása alapján azt találtuk, hogy a Magyar tanárok Egyesületének észrevételei statisztikailag megalapozottak voltak, és azok figyelembevételével (az egyébként alapvetően jó javítókulcs) még jobb lehetett volna. A magyar tanárok észrevételeit azonban a központi javítókulcs végül nem vette figyelembe, és ezért az intézmények egy része saját – nem hivatalos – kulcs alapján végezte a feladatsorok javítását. Ez a javítási technika az egyes intézményeken belül igazságosabbá tette az eredményeket, ugyanakkor az így megnövekedett pontszámok jelentősen előnyösebb helyzetbe hozták a „kreatív” pontozást alkalmazó intézmények diákjait, a központi kulcsot szigorúan követő intézmények diákjaival szemben. Mivel az érettségi eredmények tételes statisztikai feldolgozása tudomásunk szerint korábban nem történt meg, ezért ez a vizsgálat elsőként szembesíti az érintetteket az adatokkal.*

<sup>1</sup> Ez a cikk mindenféle intézményi támogatás nélkül, kizárólag önkéntes munkából jött létre. Köszönöm elsősorban Horváth Andrásnak, hogy biztatott a cikk megírására, kigyűjtötte az elektronikus sajtó hivatkozásait, oroszlánrészt vállalt az adatbevitelből, és folyamatosan segített a tanácsaival. Nélküle nem jött volna létre ez a cikk.

Köszönöm továbbá azoknak a tanároknak a segítségét, akik lehetővé tették, hogy hozzáférjünk az adatokhoz. Ez bátor, felelősségteljes döntés volt a részükről, amivel személy szerint sem nyervek, ám (az adatok ismeretében) sokat kockáztattak. A vizsgálatban részt vevő intézmények nevének elhallgatásával elsősorban őket szeretnénk védeni.

## Előjáróban

A Magyar tanárok Egyesülete és a tételkészítő bizottság közötti vita során a Magyar tanárok Egyesülete konkrétan azt javasolta, hogy

- az 1-es feladatcsoportban ne csak tökéletes megoldásért érdemljenek maximális pontot, hanem az egyes időbeli viszonyok felismeréséért is járhasson részpontszám,
- az 5-ös feladatcsoportban legyen több jó megoldás, és
- a 10-es feladatcsoportban legyen alternatív megoldás is.

A kifogások hatására a központi javítókulcs nem változott.

Az adatbevitel ellenőrzése során – meglepetésünkre – azt találtuk, hogy az általunk vizsgált három intézmény közül kétfőben a tanárok a központilag megadott javítókulcs-tól szisztematikusan eltérő módon javították a feladatsorokat. Erre a továbbiakban még többször visszatérünk.

## Munkahipotéziseink

### 1. A tanárok jól azonosították a problematikus feladatokat

- a. A tanároknak a javítókulccsal kapcsolatos, a médiában hangoztatott kifogásaik statisztikailag is azonosítható problémákra mutatnak rá.
- b. A megoldókulcstól való eltéréssel a tanárok a megoldókulccsal való egyet nem értésüket fejezték ki. Összefüggés van az egyes feladatok pszichometriai minősége és között, hogy a tanárok mennyiben használtak a központitól eltérő kulcsot.

### 2. A tanárok által helyesnek tartott javítási módszer javított volna a feladatsor minőségén

- a. A tanárok által javasolt javítási mód segítségével az érettségi feladatsor statisztikai mutatói javultak volna.
- b. A tanárok a központitól eltérő „kurucos” javításukkal jobb pszichometriai minőségű itemeket állítottak elő.

- 
- 2 [http://koronascimer.blog.hu/2012/05/08/hibas\\_a\\_magyar\\_erttsegi\\_es\\_a\\_javitokulcsa](http://koronascimer.blog.hu/2012/05/08/hibas_a_magyar_erttsegi_es_a_javitokulcsa)  
[http://eduline.hu/erttsegi\\_felveteli/2012/5/7/Magyar\\_nyelv\\_es\\_irodalom\\_erttsegi\\_2012\\_fel\\_SG5EKM](http://eduline.hu/erttsegi_felveteli/2012/5/7/Magyar_nyelv_es_irodalom_erttsegi_2012_fel_SG5EKM)  
<http://www.hir24.hu/belfold/2012/05/07/a-magyarerttsegi-megoldasai/>  
[http://index.hu/belfold/2012/05/09/a\\_hires\\_irok\\_nem\\_mertek\\_igy\\_mi\\_erttsegiztunk\\_le\\_magyarbol/](http://index.hu/belfold/2012/05/09/a_hires_irok_nem_mertek_igy_mi_erttsegiztunk_le_magyarbol/)  
<http://www.fovarosi-hirhatar.hu/hir/nem-modosit-az-oktatasi-hivatal-2012-magyar-erttsegi>  
[http://index.hu/belfold/2012/05/11/hibas\\_a\\_magyarerttsegi\\_javitokulcsa/](http://index.hu/belfold/2012/05/11/hibas_a_magyarerttsegi_javitokulcsa/)  
<http://www.fovarosi-hirhatar.hu/hir/a-magyartanarok-egyesulete-felszolitotta-az-oktatasi-hivatal-a-valtoztatásra-magyar-erttsegi-2012>  
<http://www.oh.gov.hu/magyar-nyelv-irodalom-120516>

## Az adatok

2012 októberében és novemberében 3 középiskolában összesen **246 szövegértési feladatsor** adatait vittük számítógépre statisztikai feldolgozás céljából, név, vagy bármilyen más személyes azonosító adat nélkül.

### A feladatsor adatai

#### *Bevitt adatok*

Az adatbevitel során a szövegértési feladatsorok minden feladatát feldolgoztuk. Az egyes feladatokat a feladatcsoport számával és a feladatnak azon belül elfoglalt helyével jelöltük (pl. **i\_01\_2** – az első feladatcsoport második feladata).

Az adatbevitel során számítógépre vittük:

- a feleletválasztásos feladatokban a ténylegesen kapott válaszokat (ilyen feladatcsoportok voltak a következők: **i\_01, i\_05, i\_07, i\_10**);
- a szabad szöveges feladatokban a pontozó tanár által adott helyes/helytelen értékelést (ilyen feladatcsoportok voltak a következők: **i\_02, i\_03, i\_04, i\_06, i\_08, i\_09**);
- a pontozó tanárok által a feladatcsoportokra adott összpontszámokat. Ezeket a következőképpen jelöltük: pl. **i\_01\_GM**. Ebben a jelölésben a változónév második tagja a feladatcsoportot, harmadik tagja pedig azt mutatja, hogy a pontszám kiszámolása **manuálisan** történt.

#### *Számolt adatok*

A bevitt adatokból kiszámoltuk az egyes feladatcsoportokra a kulcs mechanikus alkalmazásával kapott pontszámokat is. Ezeket a következőképpen jelöltük: pl. **i\_01\_GA**. Ebben a jelölésben a változónév második tagja a feladatcsoportot, harmadik tagja pedig azt mutatja, hogy a pontszám kiszámolása **automatikusan** történt.

#### *Háttér adatok*

##### *Az iskolák*

Az adatokat három budapesti iskolából gyűjtöttük. Az iskolák kiválasztásának szempontja mindössze az volt, hogy mely intézményben volt olyan ismeretségünk, ami miatt az intézmény hozzáférést engedett nekünk az adatokhoz. Az intézmények számára anonimitást ígértünk, ezért a tanulmányban csak a, b, c betűvel jelöljük őket, és a nehezebb azonosíthatóság végett a minta nagyságát is csak enyhén torzítva adjuk meg (ezért az összegük is enyhén eltér a fent említett 246-os elemszámtól).

1. Iskola „a”. Típusa: szakközépiskola. Minta nagysága: **51**. Feldolgozottság: minden 2012-es feladatlap.

2. Iskola „b”. Típusa: gimnázium. Minta nagysága: **116**. Feldolgozottság: minden 2012-es feladatlap.
3. Iskola „c”. Típusa: gimnázium. Minta nagysága: **82**. Feldolgozottság: 3 osztály feladatlapjai.

## Az elemzés során fel nem használt háttéradatok

Az adatbevitel során jelöltük az egyes osztályokat, illetve az adatbevivő személyét, de ezeket az adatokat az adatfeldolgozás során nem használtuk.

### *Az eredmények*

Az eredmények feldolgozásakor elsősorban a feladatsornak mint mérőeszköznek a **statisztikai** minőségére voltunk kíváncsiak, és számos más – érdekes, de a kérdéskörhöz szorosan nem kapcsolódó – kérdést nem vizsgáltunk.

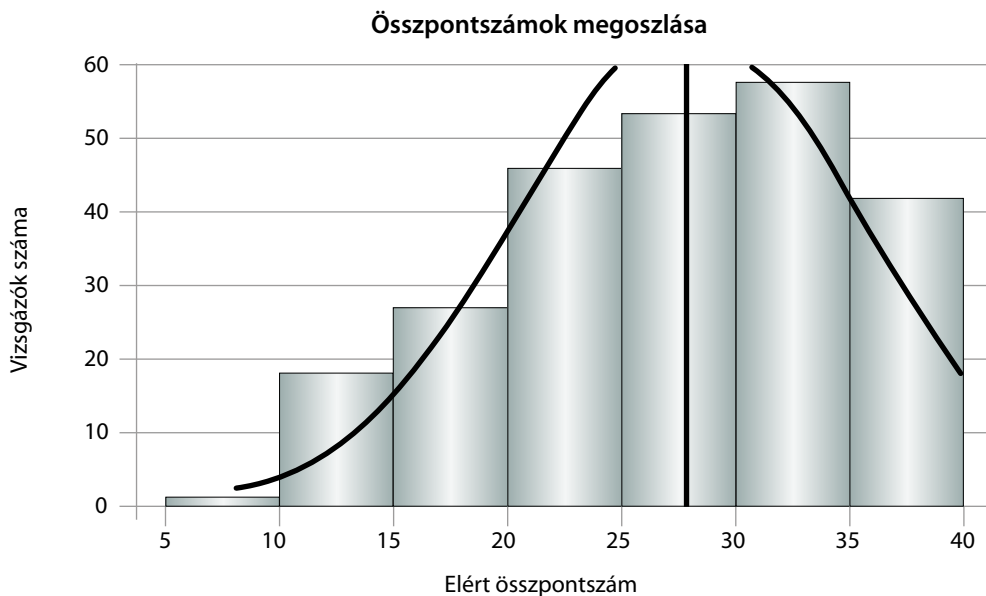
E munkában nem végeztük el a statisztikai mutatóknak a feladatok tartalmával és logikájával történő egybevetését, ami kimutathatná, hogy milyen **tartalmi és logikai okai** lehetnek az adott statisztikai mutatók kialakulásának, és hozzásegíthetne minket ahhoz, hogy egy következő alkalommal jobb eszközt készítsünk. Ez a lépés egy pszichometriai mérőeszköz kipróbálásának, bemérésének és fejlesztésének elengedhetetlen része lenne – jelen munkánk kereteit azonban meghaladja. Ebben a munkánkban igyekszünk a statisztika adta keretek között maradni és ezért következtetéseink is inkább leíró jellegűek.

A vizsgálat során a manuálisan és az automatikusan számított pontszámok között jelentős (erősen szignifikáns) eltérést tapasztaltunk, amiről a továbbiakban még többször esik majd szó. Nyilván felmerül a kérdés, hogy ez hogyan fordulhatott elő, és általában milyen visszajelzések érkeztek a feladatsor javításával kapcsolatban a **vizsgaelnököktől**. Jelen munkánkban ezzel a kérdéskörrel nem foglalkozunk. Ahol külön nem említjük, e tanulmányban az automatikusan számolt pontszámokat adjuk meg.

## Az összpontszám

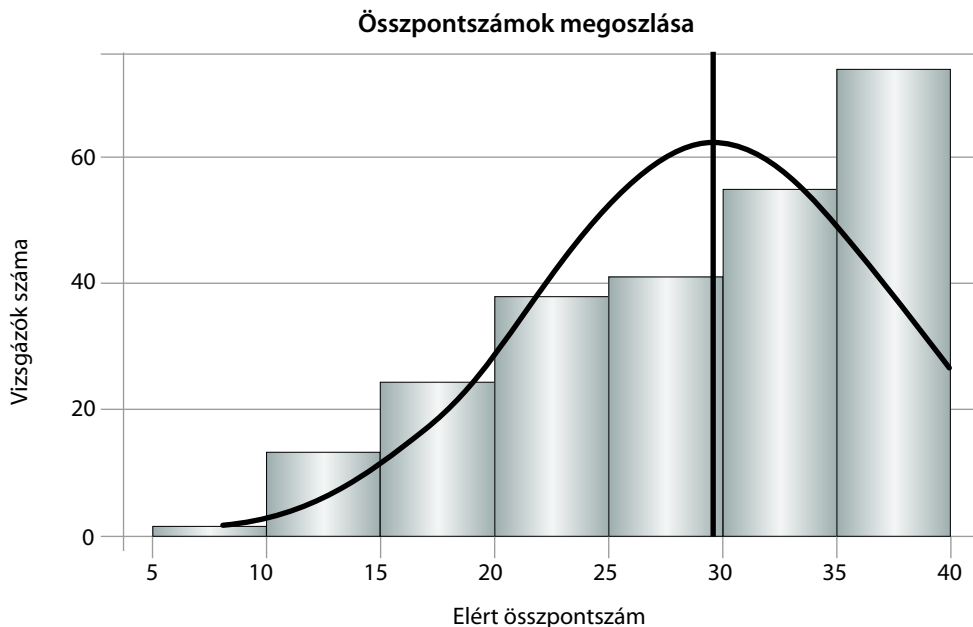
### *Az automatikusan számolt összpontszámok megoszlása*

Az átlag **27,68**. Az eloszlás a normál eloszláshoz képest (szignifikánsan) jobbra ferde, ami arra utal, hogy az általunk vizsgált tanulók számára a feladatsor túl könnyű volt. Az elérhető maximális pontszám 40 volt, így a tanulók átlagosan az elérhető pontszám kb. 70%-át szerezték meg. Ebből következik, hogy a feladatsor kevésbé volt képes különbséget tenni az általunk vizsgált, jól teljesítő vizsgázók között.



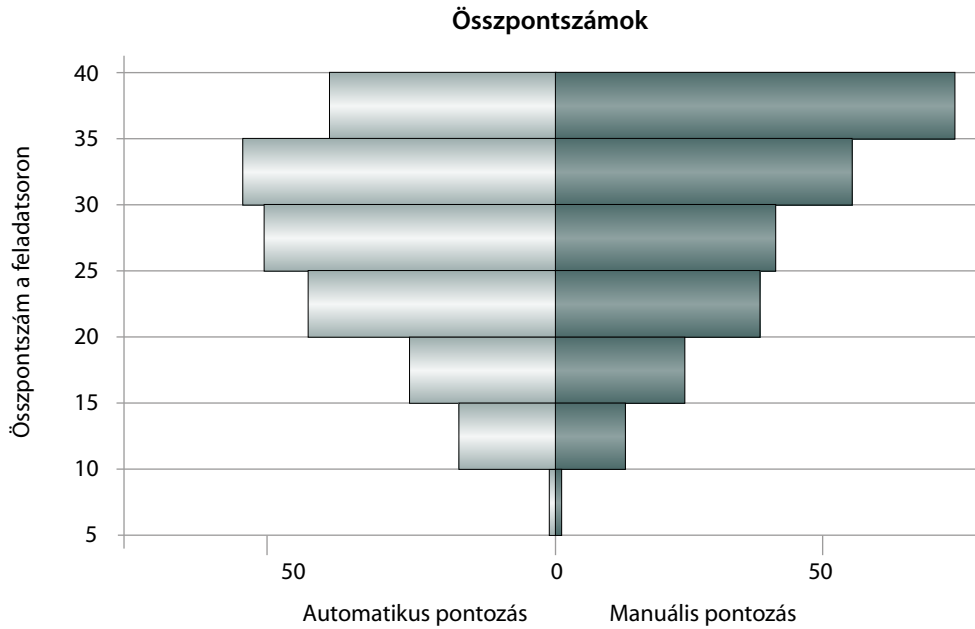
### *A tanárok által adott összpontszámok megoszlása*

Az átlag **29,62**. Az eloszlás még az automatikusan számolt pontszámoknál is erősebben jobbra ferde.



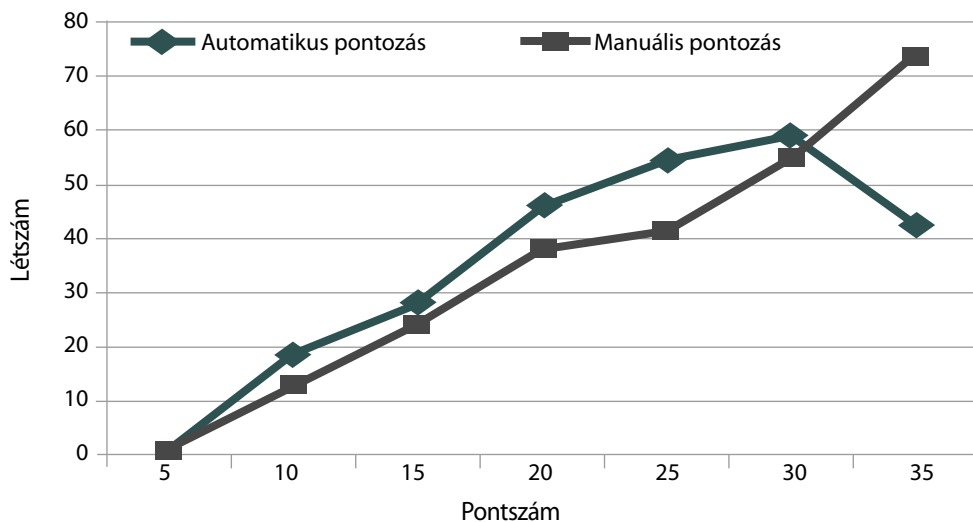
### Az automatikus és a manuális pontozás összehasonlítása

A következő ábra mutatja a kétfajta pontozás összehasonlítását. Az ábrán jól látható, hogy a manuális pontozás (szignifikánsan) magasabb összpontszámokat eredményezett.



Az egyes kategóriák létszámát megfigyelve jól látszik az is, hogy manuális pontozás esetén leginkább a nagyon magas (35-40 pont) pontszámot elérők száma növekedett.

Pontszám	Automatikus pontozás	Manuális pontozás	Különbség
35	42	74	32
30	58	55	-3
25	54	41	-13
20	46	38	-8
15	27	24	-3
10	18	13	-5
5	1	1	0



Összességében megállapíthatjuk tehát, hogy a tanárok számos esetben **nem a hivatalos pontozókulcs szerint pontoztak**. Az eltérés nem véletlenszerű (nem hiba történt), hanem látható, hogy a tanárok által követett pontozási módszer következtében a nagyon magas pontszámot elérők létszáma jelentősen megnőtt, és ez szignifikáns pontszám-növekedést eredményezett.

## A tanári pontozás következményei

Ahogy azt korábban megállapítottuk, a tanári pontozás és az automatikus pontozás közötti eltérés nem véletlen hiba eredménye, ezért különösen érdekes megvizsgálni, hogy ez az eltérés hogyan jelentkezik az egyes intézményekben és az egyes feladatokban.

## Intézmények szerinti eltérések a tanári pontozásban

Míg az „a” középiskolában nem találtunk eltérést a mechanikusan számolt kulcstól, addig a „b” és „c” középiskolákban a tanári pontozás átlagosan **2-3 ponttal magasabb** összpontszámot eredményezett. (Ez azt jelenti, hogy a tanári pontozás miatt mind a „b”, mind a „c” iskolában szignifikánsan több pontot kaptak az érettségizők, mint az „a” középiskolában; a „b” és „c” iskolák közötti ilyen típusú eltérés nem szignifikáns.)

Tudjuk, hogy az „a” iskola típusa szakközépiskola, míg a „b” és „c” iskolák gimnáziumok. Látható, hogy a „b” és „c” iskolákban érettségizők jobban jártak a tanári pontozással, mint az „a” iskolába járók (akiknek ez nem jelentett plusz pontokat). Érdekes

lenne megvizsgálni, hogy az iskola típusa és a tanári pontozásból adódó eltérés általában is összefügg-e, de erre az általunk gyűjtött minta mérete nem alkalmas.

Leszögezhetjük tehát, hogy az egyes középiskolák pontozó tanárai különböző mértékben függetlenítették magukat a központilag megadott értékelő kulcstól.

## Feladatonkénti eltérések a tanári pontozásban

Mivel az előzőekben megállapítottuk, hogy az „a” iskolában a tanárok szigorúan követték a központilag megadott megoldókulcsot, és így nincs eltérés a tanári pontozás következményeként, ezért e pont elemzésében csak a „b” és „c” iskolákból származó eredményeket használjuk fel.

A vizsgált feladatcsoportok közül a tanári pontozás a következő feladatokban okozott szignifikáns változást:

Feladatcsoport	Össz változás	„b” iskola	„c” iskola	Iskolák közötti különbség szignifikáns?	Feladatcsoport szövege
1	1.3	1.20	1.41	Nem	Állítsa időrendi sorrendbe a Nemzeti olvasókönyv létrejöttének menetét!
5	0.4	.77	-.06	Igen	Rendelje a Nemzeti olvasókönyv alábbi fejezetcímeihez a következő, a szövegben szereplő tárgykörök sorszámát!
10	0.4	0.00	.80	Igen	Állapítsa meg, hogy az alábbi idézetek a Nemzeti olvasókönyv melyik tárgyköréből származnak!
9	0.3	.63	0.00	Igen	A szövegből vett kifejezésekkel fejezze be az alábbi mondatokat!

A táblázatból jól látszik, hogy az 1-es feladatcsoport javításánál a két iskola tanárai körülbelül ugyanannyira tértek el a kulcstól, míg a többi feladat esetében a két iskolában eltérő mértékű különbséget okozott a tanári javítás. A tanári javítás az 1-es feladat esetében okozta magasan a legnagyobb pontszámbeli különbséget.

## Az érettségi feladatsor – mint pszichometriai mérőeszköz – minősége

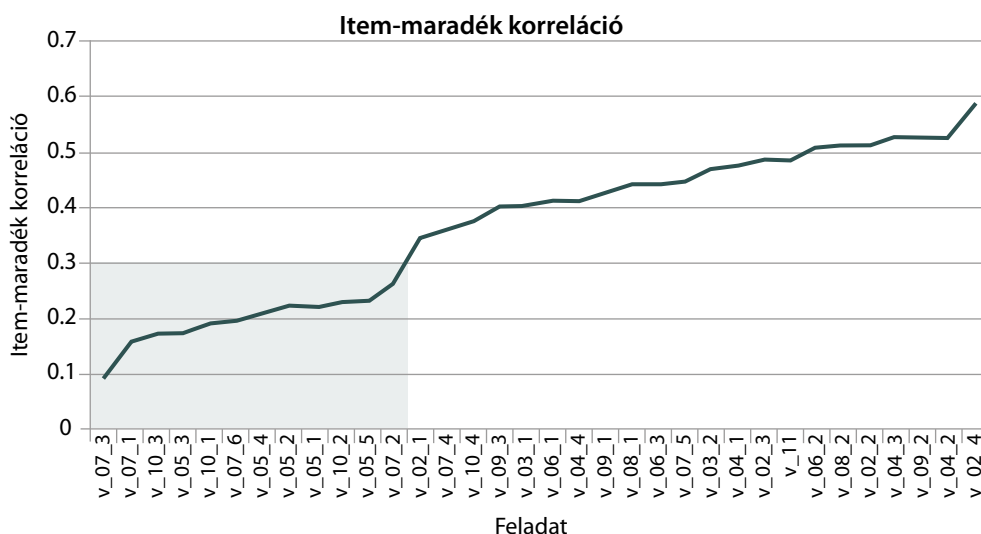
A maximális teljesítmények mérésére szolgáló eszközök (tudás-, képességtesztek) egyik leggyakrabban használt megbízhatósági mutatója a Cronbach alfa.

Az általunk vizsgált feladatsor Cronbach alfája **0,8**. Ez azt jelzi, hogy az eszköz összességében megbízhatóan méri az adott tulajdonságot.

(Mivel az 1-es feladatcsoport feladatainak értékelése nem független egymástól, ezért a további elemzésekben az 1-es feladat feladatait összevonva, egy feladatként értelmezzük, és 11-esnek jelöljük.)

### Az egyes itemek minősége

Az egyes itemek minőségének vizsgálatához kiszámoltuk és sorba rendeztük az értékelő kulcsnak megfelelően, automatikusan számolt item-maradék korrelációkat.



Általában nem szoktuk elfogadhatónak tekinteni a 0.1 item-maradék korrelációval rendelkező itemeket, rendkívül gyengének tekintjük a 0.2 alatti itemeket és gyengének tekintjük a 0.3 alatti itemeket. E vizsgálatban 12 olyan itemet találtunk, melynek item-maradék korrelációja nem érte el a 0,3-at.

A 12 feladat 3 feladatcsoportból származik. A gyenge feladatok majdnem fele (5 feladat) az 5-ös feladatcsoportból származik, 4 feladat származik a 7-es, 3 feladat pedig a 10-es csoportból.

Ha megfigyeljük ezeknek a feladatoknak a nehézségi szintjét, akkor azt találjuk, hogy a 12 gyengén „teljesítő” feladat közül 7-ben az átlageredmény 0,8 fölött van, tehát rendkívül könnyű feladat.

Mivel a jelen elemzést elsősorban a megoldókulccsal kapcsolatos polémiák statisztikai vizsgálatára készítettük, ezért a továbbiakban nem vizsgáljuk a túl könnyű feladatokat – a vita nem az ilyen feladatok kapcsán bontakozott ki.

A fennmaradó, problémás feladatok:

- az **5-ös feladatcsoport 1., 3., 4. és 5.** feladata;
- a **10-es feladatcsoport 3.** feladata.

Az 5-ös feladatcsoport (egy kivételével) összes feladata gyengén teljesített (és nem azért, mert a feladatok túl könnyűek voltak). Ez egybecseng azzal, hogy a tanári kifogások jórésze ezzel a feladatcsoporttal volt kapcsolatos.

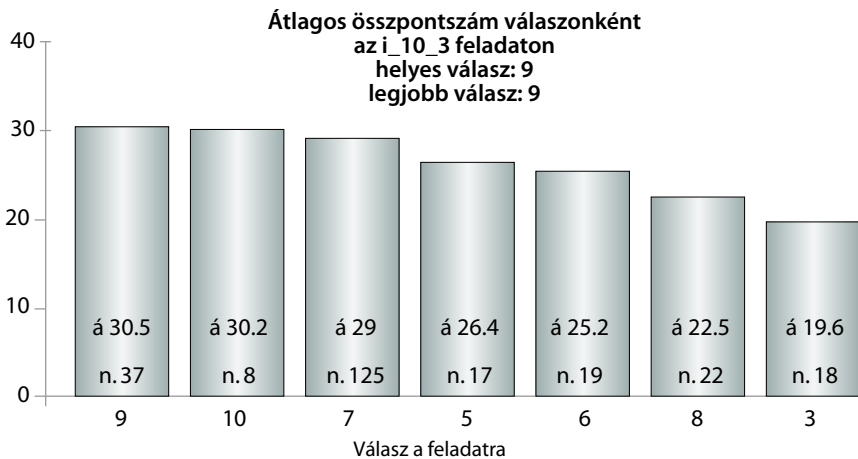
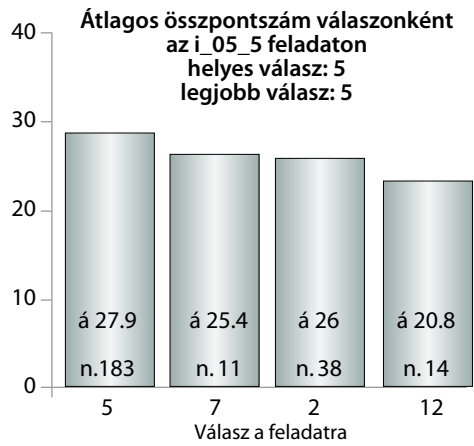
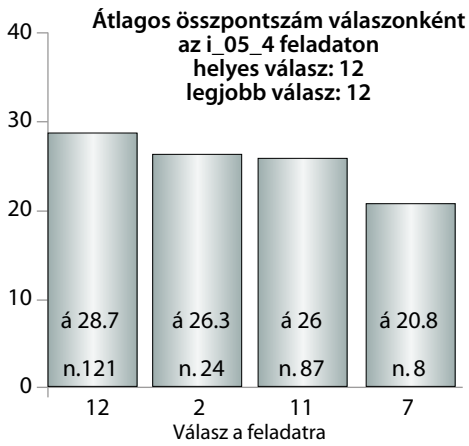
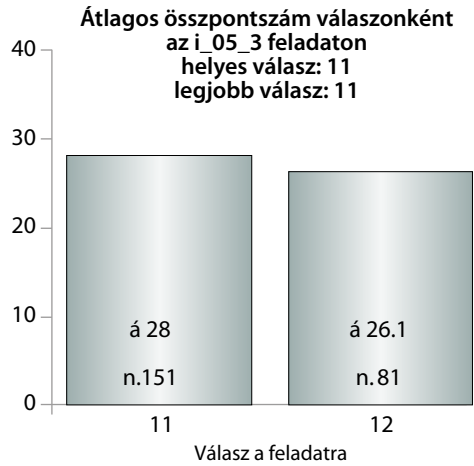
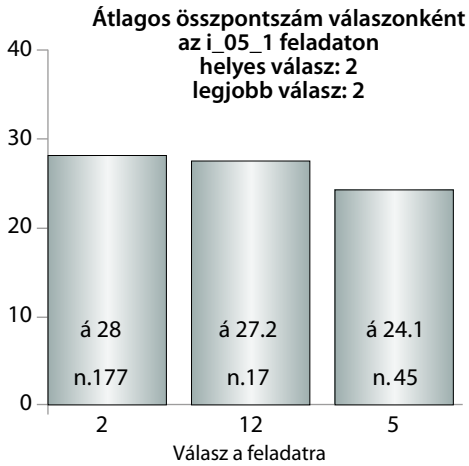
Ugyanakkor nem találjuk itt az 1-es (v\_11) feladatot. Bár a tanárok e feladat javítását is kifogásolták, statisztikailag ez a feladat nem teljesített rosszul. Az 1-es feladattal kapcsolatos problémákat a későbbiekben tekintjük át – alapvetően arra koncentrálva, hogy a tanárok által felvetett javaslatok javítottak volna-e e feladat minőségén. Előljáróban annyit, hogy a feladat javítása nem dichotóm (helyes/helytelen), így ezzel az alapvetően statisztikailag jó feladattal kapcsolatban is érdemes megvizsgálni, hogy a feladaton elért **részpontoszámok** jól tükrözik-e a vizsgált tulajdonságot.

## Alternatív megoldások

Miután azonosítottuk a gyengén teljesítő itemeket, érdemes megnéznünk, hogy mi lehetett a statisztikailag gyenge teljesítmény oka.

Az alábbi ábrák azt mutatják, hogy az egyes feladatra adott választ adók hányan voltak és átlagosan hány pontot értek el. (Az ábrákon csak azok a válaszopciók látszanak, amelyeket legalább öt vizsgázó választott, az átlagpontoszámot minden esetben a maradék itemekből számoltuk.) Egy jól szerkesztett feladaton a helyes választ adók átlagpontoszáma kiugróan magasabb, mint a helytelen választ adók átlagpontoszáma. Más szavakkal: a feladatsoron jól teljesítők által helyesnek tartott válasz megegyezik a kulcs szerint helyes válasszal.

Az ábrákon jól látszik, hogy az adott kérdések azért teljesítettek statisztikailag gyengén, mert bár minden esetben a kulcs szerint helyes választ adók összpontoszáma bizonyult a legmagasabbnak, a helytelenként pontozott válaszok között e válasznak jó alternatívái voltak. Ezeknél a feladatoknál érdemes megvizsgálni, hogy a statisztikailag jó alternatívaként megjelenő, ám helytelenként pontozott válaszok tartalmilag valóban helytelenek-e.



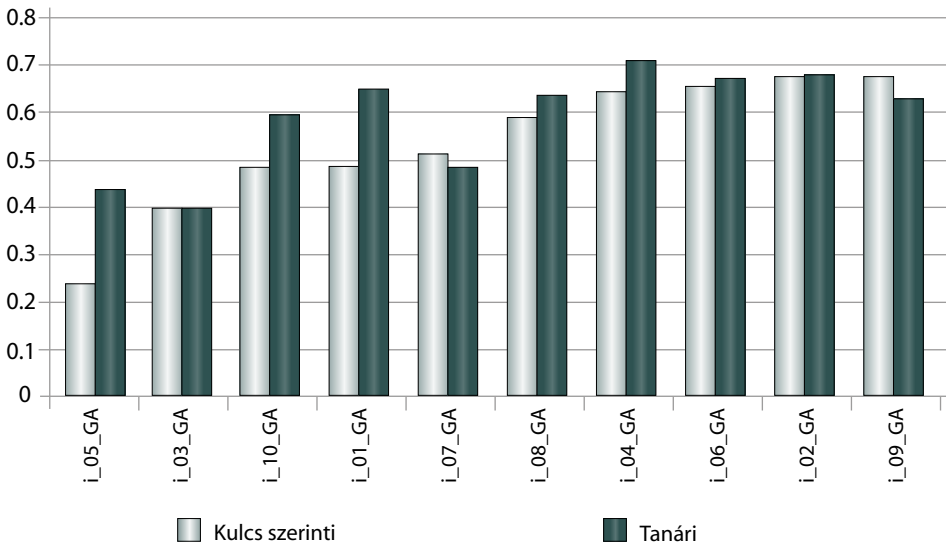
## A feladatcsoportok pszichometriai minősége

Mivel a tanári javítások pontszámai csak feladatcsoport-szinten állnak rendelkezésünkre, ezért a továbbiakban elsősorban a feladatcsoportokra – és nem az egyes feladatokra koncentrálnak az elemzésben. Úgy véljük, hogy ez a következtetések levonásánál nem okoz majd különösebb problémát, mivel az alapul szolgáló polémia is elsősorban az egyes feladatcsoportokhoz kapcsolódott.

Az alábbiakban feladatcsoportonként tekintjük át az érettségi feladatsor pszichometriai minőségét és azt, hogy azokban milyen változást hozott a tanári javítás.

Korábban már megállapítottuk, hogy a tanári pontozás az 1, 5, 9 és 10-es feladatcsoportokban okozott szignifikáns pontszámváltozást. Az alábbi ábra bemutatja, hogy az egyes feladatcsoportok pszichometriai minőségében milyen változást hozott a tanári javítás.

**Az egyes feladatcsoportok item-maradék korrelációja kulcs szerinti és tanári javítás esetén**



Az ábráról jól látszik, hogy a 10 feladatcsoport közül 3 pontozásában okozott a tanári javítás szignifikáns minőségi változást, és ezek mindegyikében a változás következtében a feladatcsoportok pontszám-eredményei jelentősen javultak.

E javításoknak köszönhetően a feladatcsoportok alapján számolt Cronbach alpha 0.79-ről 0.87-re változott, azaz a feladatsor belső megbízhatósága a tanári javítások következtében megnőtt.

**A tanári változtatás eredményeképpen jelentősen javuló 3 feladatcsoport az 1-es, az 5-ös és a 10-es.** Az alábbiakban ezeket tekintjük át.

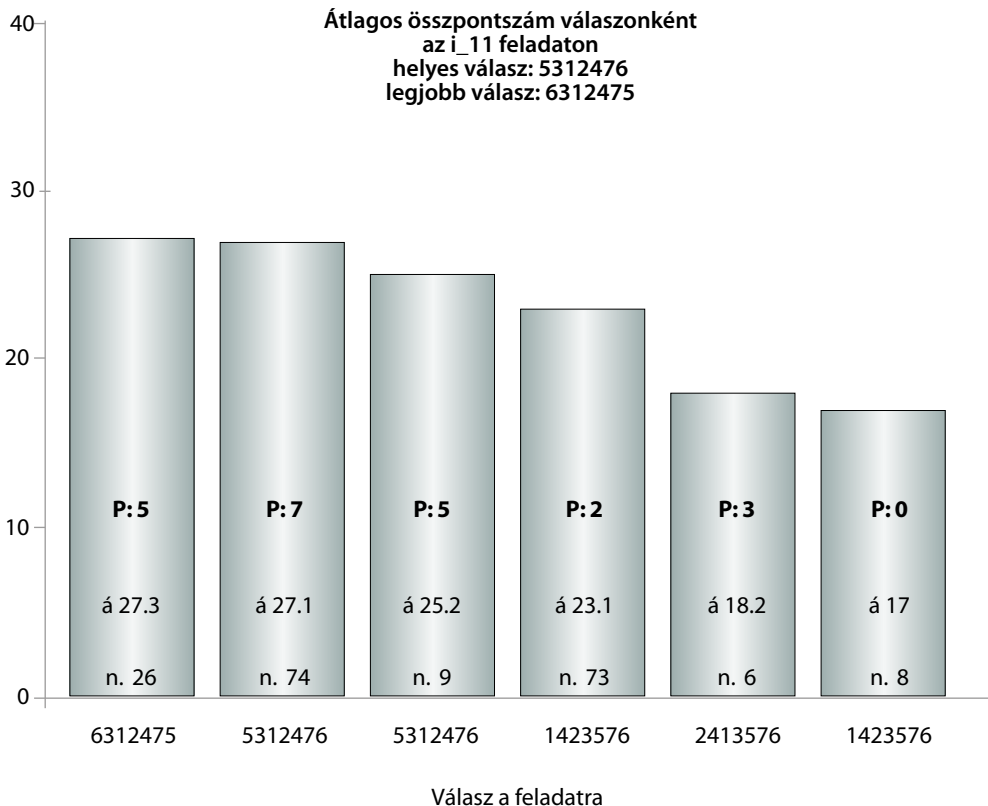
## Az 1-es feladatcsoport

Ahogy azt korábban említettük, az 1-es feladatcsoport feladatainak értékelése nem volt független egymástól, ezért ezt a feladatcsoportot összevonva, egy feladatként értelmeztük és helyenként 11-es feladatként jelöljük – a 11-es feladat és az 1-es feladatcsoport ugyanazt jelöli.

E feladat item-maradék korrelációja (0.48) jónak tekinthető, mégis, jellegénél fogva (rendezze sorba az eseményeket) ez a feladat – és ennek pontozási módja – igen sok vitát váltott ki.

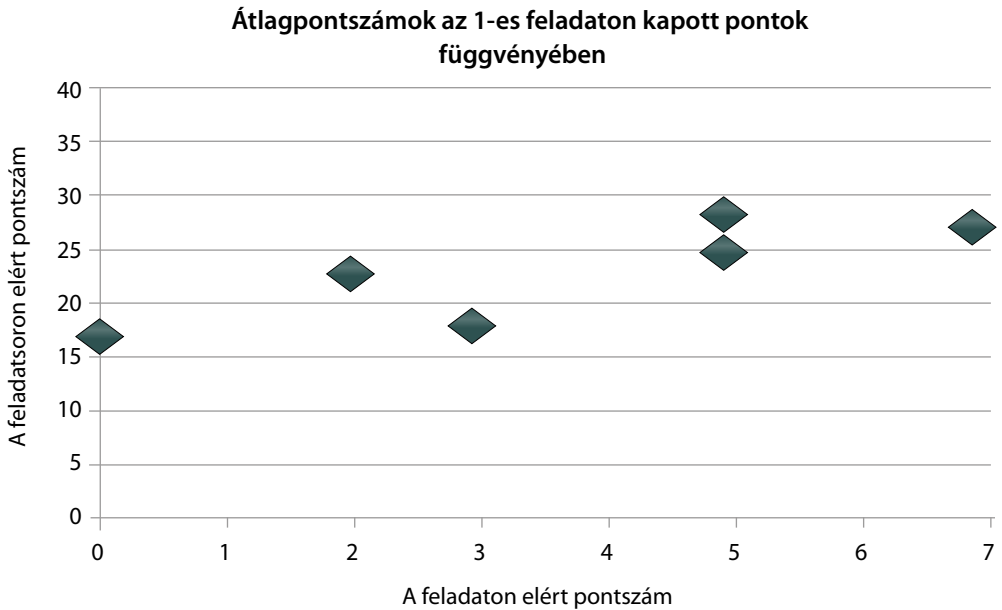
Az alábbi ábrán az egyes megoldási sorrendek és az adott sorrendet választók átlagpontszáma látható (P – a feladaton kapott pontszám, á – a feladatsoron kapott átlagos pontszám, n – az adott választ választók száma).

Az ábráról leolvasható, hogy e feladat esetében **nem** a kulcs szerinti legjobb választ adók átlagpontszáma a legmagasabb. E feladat esetében mindenképpen célszerű lenne megvizsgálni, hogy a statisztikailag jó alternatívaként megjelenő, ám helytelennek pontozott sorrendek tartalmilag valóban helytelenek-e.



Ahogy azt már említettük, ezen a feladaton – a többivel ellentétben – részpontszámokat is lehetett kapni.

A fenti adatokat másként elrendezve megfigyelhetjük, hogy a feladaton kapott pontszámok mennyiben követik az összpontszámok megoszlását.

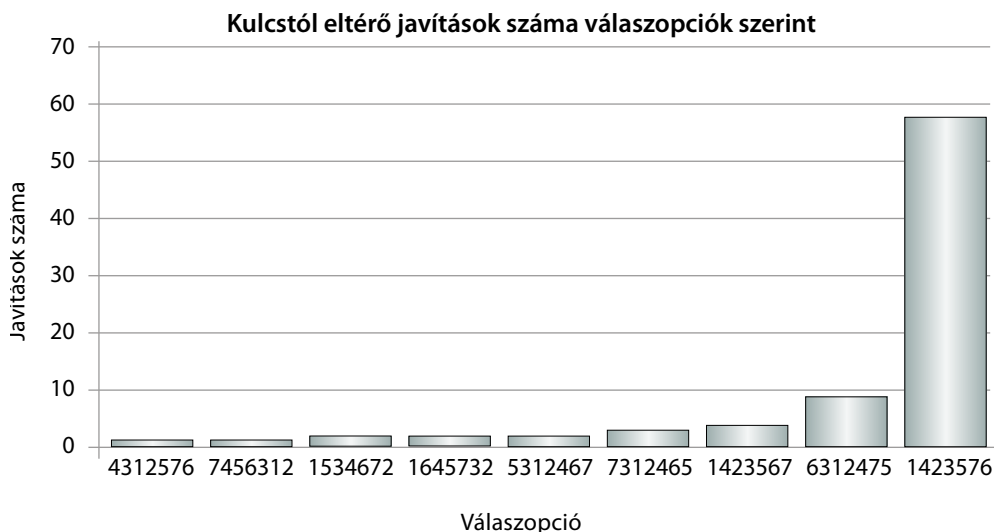


A fenti két ábra összevetéséből látható, hogy a 6. 3. 1. 2. 4. 7. 5, illetve az 5. 3. 1. 2. 4. 6. 7. sorrendek (5 pont) statisztikailag körülbelül ugyanolyan jók, mint a kulcs szerinti válasz (7 pont), és hogy az 1. 4. 2. 3. 5. 7. 6. válaszra (2 pont) aránytalanul kevés pontszámot lehetett kapni (az azt választók teljesítményéhez mérten).

## Az 1-es feladat opcióinak tanári javítása

Az „a” és „b” iskolában a 196 esetből 93-ban – tehát közel az esetek felében – kapott az érettségiző a kulcstól eltérő pontszámot az első feladatban (ebből 5 alkalommal a kulcs szerint járó pontszámnál kevesebbet – ezeket az eseteket elírásnak tekintjük és a továbbiakban nem vizsgáljuk).

A maradék 88 esetben a kulcs szerint járó pontszámnál 1-4 ponttal nagyobb pontszámot kapott a vizsgázó.

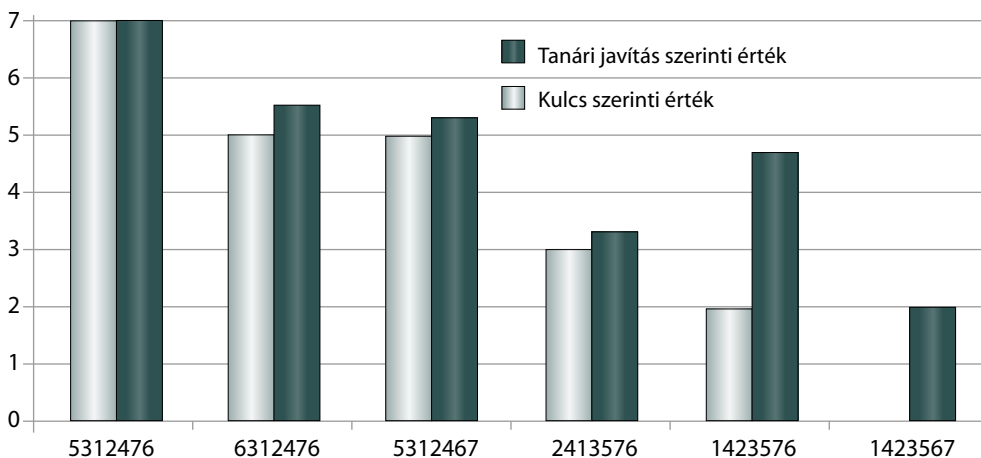


A fenti ábrából jól látszik, hogy a kulcstól eltérő javítások (az 1-es feladaton) döntő többségében az 1423576 válaszopciónál fordultak elő. A második leggyakrabban felülbírált opció a 6312475, és az 5312467 opciót is két esetben plusz ponttal jutalmazták a tanárok. Összességében az egyes feladat javításának tanári felülbírálása a vizsgált esetek 80%-ában az általunk feltételezett 3 válaszopció egyikének esetében történt.

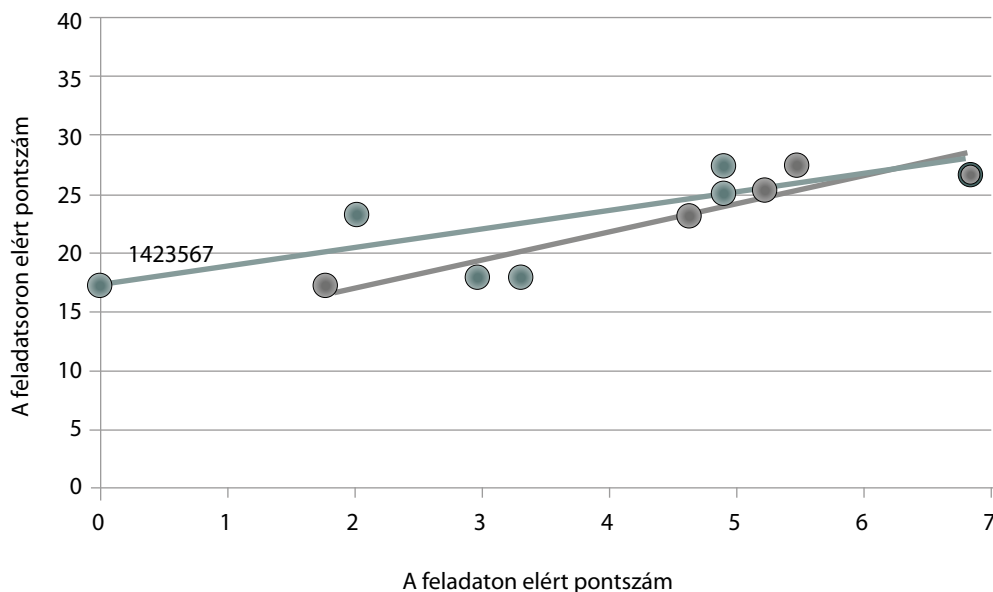
Az 1-es feladatban tehát a kulcstól eltérő tanári javítás az alulértékelt opciókat (6312475, 5312467, 1423576) jobban érintette, mint az egyéb opciókat.

## Az 1-es feladat tanári javításának hatása a feladat pszichometriai minőségére

### Az egyes feladat leggyakoribb opcióinak kulcs szerinti és tanárok szerinti értéke



Látjuk tehát, hogy a tanári javítások azokon a helyeken fordultak elő leggyakrabban, ahol az pszichometriailag is indokolt lett volna. Az alábbiakban megvizsgáljuk, hogy a kulcstól eltérő értékelések javították-e a feladat minőségét.



A fenti ábrán jól látszik, hogy a feladat tanári javítása során

- minden gyakran előforduló opció magasabb értéket kapott, mint a kulcs szerint meghatározott (a maximális pontszámot érő opció kivételével, ami a kulcs szerinti maximális pontszámot kapta);
- a tanárok a kulcshoz képest legtöbb plusz pontot az 1423567 opcióért adták;
- a tanárok a kulcshoz képest jelentősen több pontot adtak az 1423567 opcióért is.

A fenti ábrán látszik, hogy a tanári javításnak köszönhetően a gyakran választott válaszopciókért kapható pontszámok és a feladatsoron elért összpontszám közötti összefüggés szorosabbá vált ( $r=0.85$ -ről  $r=0.93$ -ra), és ezzel a feladatcsoport **belső** pszichometriai minősége javult.

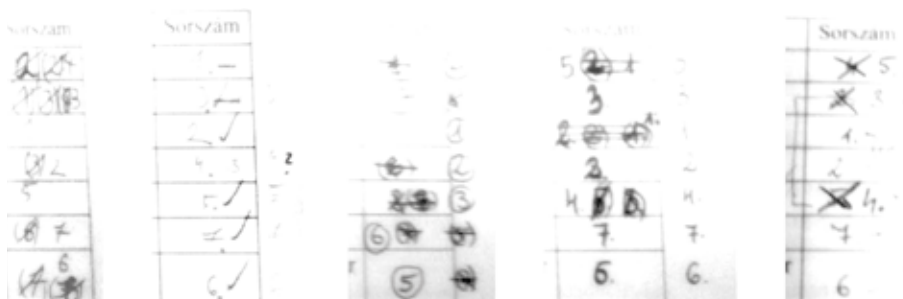
A feladatcsoportokból számolt item-maradék korrelációk vizsgálatából az is látszik, hogy **a tanári javítás jelentősen növelte az 1-es feladatcsoport pszichometriai minőségét a feladatsoron belül** ( $r=0.49$ -ről  $r=0.65$ -re).

Ha most külön megvizsgáljuk a leggyakrabban előforduló, 1423567 opció tanári javítását, akkor azt is láthatjuk, hogy az egyes **intézmények következetesen, de egymástól eltérő módon javították** ezt az opciót: ez a válasz az „a” iskolában mindig a kulcs szerinti 2 pontot kapta, a „b” iskolában mindig 6 pontra értékelték, míg a „c” iskolában mindig 5 pontot ért.

## Az 1-es feladatcsoporttal kapcsolatos egyéb tapasztalatok

A fenti – számszerű – mutatókon túl az adatbevitel során azt tapasztaltuk, hogy a vizsgázók sokkal nagyobb számban javították ezt a feladatot, mint bármely másikat. Véleményünk szerint ez azt mutatja, hogy a sorbarendezen alapuló feladatok utólagos megváltoztatása a hagyományos papír-ceruza módszerrel túlságosan nehézkes. Véleményünk szerint a sorbarendezenes feladatok mind a kitöltés, mind a javítás és pontozás tekintetében aggályosak papír-ceruza kitöltés mellett, ám valószínűleg sokkal jobban működhetnek egy interaktív, számítógépes környezetben.

Bár az utólagos javítások számát nem számszerűsítettük, álláspontunk illusztrálására bemutatunk néhány példát.



*A sorbarendezen utólagos módosításának nehézségeit illusztráló fényképek az 1-es feladatcsoportról*

## Az 5-ös feladatcsoport

Megvizsgálva, hogy az 5-ös feladatcsoport egyes feladataiban mely opciók esetén tértek el a tanárok szignifikánsan nagyobb számban a javítókulcstól, a következőket kapjuk:

A tanárok az 5-ös feladatcsoport

- 3. feladatában elfogadták a 12-es opciót is,
- és 4. feladatában elfogadták a 11-es opciót is.

Ez azt jelenti, hogy a megoldókulcs szerinti 2., 7., 12., 11., 5. válaszain túl a tanárok elfogadták a 12-11 opciók megcserélését is. Ez egybeesik a Magyar tanárok Egyesülete által az 5. feladattal kapcsolatban megfogalmazott kérdésnek, melyben a következőket javasolják: „... javasoljuk az 5. feladatnál a 2., 7., 12., 11., 5. sorrend helyes megoldásként való elfogadását is”.

A tanári javításban helyesként elfogadott alternatívák mindkét feladat esetében magas átlagpontoszámokkal jártak (jó alternatívák voltak). **A tanári javítás összességében az 5-ös feladatcsoport igen gyenge item-maradék korrelációját ( $r=0.24$ ) nagyban megemelve, azt elfogadható szintre (0.44) hozta.**

## A 10-es feladatcsoport

A 10-es feladatcsoport tanári javítása gyakorlatilag annyiban tér el a kulcs szerinti javítástól, hogy a tanári javítás elfogadta a 10/3-as feladatban a kulcs szerinti 9-es opció mellett a válaszadók messze legnagyobb része által választott 7-es opciót is.

Ez szintén megegyezik a Magyar tanárok Egyesülete által a 10. feladattal kapcsolatban megfogalmazott észrevételnek, mely szerint: „szükséges lenne elfogadni [...] a 7. »szabadság és demokrácia« tárgykörhöz való hozzárendelését is”.

**A tanári javítás a 10-es feladatcsoport pszichometriai minőségét (item-maradék korrelációját  $r=0.40$ -ról  $r=0.59$ -re) jelentősen megemelte.**

## Befejezés

### A munkahipotézisek ellenőrzése

#### Igaz-e, hogy a tanárok jól azonosították a problematikus feladatokat?

- a. **Igaz-e**, hogy a tanároknak a javítókulccsal kapcsolatos, a médiában hangoztatott kifogásaik statisztikailag is azonosítható problémákra mutatnak rá?

**Igen**, a Magyar tanárok Egyesülete által kifogásolt mindhárom feladatcsoport (1, 5, 10) benne van a statisztikailag leggyengébbnek bizonyult 4 feladatcsoport között. A maradék egy feladatcsoport (3) gyenge statisztikai teljesítménye annak tudható be, hogy a feladatcsoport túlságosan könnyű volt.

- b. **Igaz-e**, hogy összefüggés van az egyes feladatok pszichometriai minősége és aközött, hogy a tanárok mennyiben használtak a központitól eltérő kulcsot?

**Igen**, a feladatsort javító tanárok azokat a feladatokat javították a központitól eltérő módszerrel, amelyek pszichometriailag problémásnak bizonyultak.

#### Igaz-e, hogy a tanárok által helyesnek tartott javítás jobb feladatsort eredményez?

- a. **Igaz-e**, hogy a tanárok által a médiában javasolt javítási mód segítségével az érettségi feladatsor statisztikai mutatói javultak?
- b. **Igaz-e**, hogy a tanárok a központitól eltérő, „kurucos” javításukkal jobb pszichometriai minőségű itemeket állítottak elő?

**Igen**, a tanárok azt a javítási módszert alkalmazták „nem hivatalosan”, amit hivatalosan is javasoltak, és ez mindig (amikor jelentősebb változást okozott) javított a feladatsor minőségén.

## Az eredmények összefoglalása

Vizsgálati adathalmazunk természetesen csak nagyon kis merítés a teljes állományból. Ez nyilván korlátozza a levonható tanulságok érvényességének. Másrészt, ezt a kevés adatmennyiséget is fáradságos munkával, kézi adatbevitellel, szívevények útján gyűjtöttük össze, ami önmagában is figyelmeztet arra, hogy milyen hiányos az érettségi vizsgák jelenlegi minőségellenőrzése.

Megállapításunk szerint a 2012-es magyar irodalom érettségi szövegértési feladatso-ra alapvetően jól mérte az érettségizők szövegértési képességét. Ugyanakkor a Magyar-tanárok Egyesülete által azonosított módszertani problémák valószínűleg bizonyultak, és a javasolt megoldások általában valóban tovább javították volna az eszköz statisztikai hatékonyságát.

Mivel a tételkészítő bizottság nem vette figyelembe a tanárok javaslatát, a tanárok egy része (mintánkban az intézmények kétharmada – ám erősen kérdéses, hogy ez, a pusztán három intézmény vizsgálatából kapott arány mennyire általánosítható) maga vette kezébe az igazságos javítókulcs kidolgozásának feladatát. Az így kidolgozott javítókulcsok az adott intézményen belül elég állandónak bizonyultak, a különböző intézmények között azonban eltértek egymástól. Az alternatív javítókulcsok használata tipikusan javította ugyan a mérőeszköz minőségét, de igazságtalanná tette az eredményeket azokkal a diákokkal szemben, akiknek tanárai ragaszkodtak a hivatalos megoldókulcshoz.

## Javaslatunk

A jövőre nézve javasoljuk

- az egyszerű, feleletválasztásos, dichotóm értékelésű (helyes/helytelen) kérdések használatát;
- az eredmények központi feldolgozását;
- a javítókulcsnak az adatok és vélemények ismeretében történő, utólagos felülvizsgálatát;
- a végső pontozásnak a bevizsgált javítókulccsal történő elvégzését
- és az eredmények kutathatóvá tételét.