

SZABÓ GÁBOR – KISZELY ZOLTÁN

Államilag elismert nyelvvizsgarendszerek, illetve az emelt szintű érettségi összevetése próbavizsgázói teljesítmények tükrében német és angol nyelvből

Comparing nationally accredited language examinations and the advanced level school-leaving exam in the light of exam results in English and German

Nationally accredited language examinations in Hungary have been required to align their tests' levels to those of the Common European Framework of Reference (CEFR) and to provide evidence of having done so. Little empirical research, however, has been conducted so far on comparing aligned exams' results as well as on the alignment of exam levels to the CEFR concerning the advanced level school-leaving exam, which is aimed at CEFR level B2.

This paper presents a project by the Hungarian Accreditation Board for Foreign Language Examinations and the Accreditation Centre for Foreign Language Examinations aiming at collecting empirical data in the course of a comparative analysis of exam results of five accredited language exams and the school-leaving exam in English and German, focusing on the Reading, Writing and Grammar and Usage components of the exams. The analyses are based on candidates' results who took the live exams, except for the school-leaving exam, which was taken under simulated conditions.

The results of the analyses indicate that occasionally there appear to be considerable differences between certain components of some of the exams. Yet, owing to the limitations of the study no ultimate conclusions can be drawn. Instead, further research is encouraged in order to gather more data.

Bevezetés

A *Közös Európai Referenciakeret (KER)* (2002) nemzetközi térhódítása természetesen vezetett el a különféle vizsgarendszerek szintjeinek a KER-hez történő illesztéséhez. Ennek precíz végrehajtása érdekében az Európa Tanács egy (első megjelenése óta jelentősen átdolgozott) kézikönyvet is megjelentetett, amely pontosan meghatározza az illesztéshez szükséges lépéseket (North és mtsai 2006; 2009).

A szintillesztés Magyarországon az államilag elismert vizsgarendszerek esetében jogszabályi kötelezettségként is megjelent. A kétszintű érettségi kidolgozása során világossá vált, hogy az új vizsga, különösen annak emelt szintje a KER releváns szintjét lesz hivatott megjeleníteni. Mindebből az is következik, hogy a KER-hez kapcsolt különféle vizsgák immáron egy minden eddiginél pontosabban definiált szintleírás alapján a korábbi állapothoz képest objektívebb módon egymáshoz képest is összevethetővé váltak.

Mind az államilag elismert vizsgák, mind az érettségi esetében sor került különféle belső és külső minőség-ellenőrzési lépések végrehajtására; az akkreditált vizsgák ese-

tében az Oktatási Hivatal Nyelvvizsgáztatási Akkreditációs Központja (OH-NYAK) a Nyelvvizsgát Akkreditáló Testülettel (NYAT) együtt szintillesztési felülvizsgálatokat is lefolytatott. Mindeztidáig azonban a különféle vizsgák összehasonlítása kapcsán csak kvalitatív adatok álltak rendelkezésre.

Ezért indította útjára az OH-NYAK és a NYAT azt a próbavizsga-projektet, amelynek egyik fő célja empirikus adatok gyűjtése volt. Az adatok alapján objektívebben, tapasztalati tények tükrében lehetett összevetni a különféle vizsgarendszereket. A törekvés úttörő jellegűnek mondható, hiszen a külföldi és a magyarországi szakirodalom jónéhány szintillesztési projektről ugyan beszámol (pl. Tannenbaum és Wylie 2008; Dávid 2009; Dunlea és Matsudaira 2009; Taylor 2009), azonban kutatás még nem vállalkozott arra, hogy szintillesztettnek minősített nyelvvizsgák nyelvi szintjeit vesse össze egymással.

Jelen tanulmány az OH-NYAK és a NYAT projektjének empirikus adatokon alapuló eredményeit hivatott bemutatni. Először a kutatás módszereit tekinti át, majd ismerteti az adatgyűjtés menetét. Ezt az adatok több szempontú elemzése követi, s végül a következtetések levonására kerül sor.

Adatvédelmi okok miatt a tanulmány nem közli a részt vevő akkreditált vizsgaközpontok nevét.

1. A kutatás módszerei

1.1. A kutatási terv

A kutatás abból a feltételezésből indult ki, hogy a leghitelesebb adatok egy vizsga működéséről akkor gyűjthetők, ha azok a vizsgafeladatok éles felhasználása kapcsán keletkeznek. Természetesen az államilag elismert vizsgarendszerek esetében az OH-NYAK és a NYAT jogosult az éles vizsgák kapcsán keletkezett statisztikák vizsgálatára, ám ezek a vizsgálatok a vizsgarendszerek önmagukhoz képesti működését képesek csak feltárni, azaz különféle vizsgák között ilyen módon kapcsolat nem teremthető.

A próbavizsga-projekt ezért olyan kutatási tervre épült, amelyben egy csoport próbavizsgázó viszonylag rövid idő alatt több vizsgát is tényleges vizsgakörülmények között tesz le. Így az érintett vizsgázók teljesítménye a különböző vizsgarendszerekben elért eredményeik alapján összehasonlíthatóvá válik, s ezáltal maguk a vizsgák is empirikus adatok alapján válnak összevethetővé.

Három tényezőt azonban ennek a kutatási tervnek a gyengéiként már ezen a ponton azonosítanunk kell.

Először is, bár a vizsgák letételére előirányzott idő viszonylag rövid volt (2007. november – 2008. május), nem zárható ki teljes mértékben, hogy a vizsgázók tudásszintjében ez alatt az idő alatt változás állt be. Ha ez valóban megtörtént volna, az eredmények értelemszerűen megkérdőjelezhetőek lennének. A próbavizsgázók azonban nem aktív nyelvtanulók voltak, így az idő rövidege miatt tudásszintjük szignifikáns emelkedése és lényegi csökkenése egyaránt valószínűtlen. Vélhetően tehát az eredmények a fenti okból nem torzultak. Az is lényeges körülmény, hogy ennél rövidebb idő alatt a vizsgákat tényleges vizsgakörülmények között nem is lehetett volna lebonyolítani.

A második tényező az, hogy a projekt (praktikus okokból) nem törekedhetett teljeségre. Ez négy dolgot jelentett.

- a) Nem kerülhetett sor adatgyűjtésre valamennyi lehetséges vizsganyelven. A projekt a hazai nyelvoktatási kontextus számára két leglényegesebb nyelv, az angol és a német körében vizsgálódott.
- b) Ugyancsak gyakorlati korlátok miatt a vizsgálatnak nem lehetett része az összes vizsgarendszer, amelyben a két vizsgált nyelven vizsga tehető. Ehelyett az angol nyelvű vizsgák közül öt (ANGOL1, ANGOL2, ANGOL3, ANGOL4, ANGOL5), míg a német nyelvű vizsgák közül négy (NÉMET1, NÉMET2, NÉMET3, NÉMET4) vizsgarendszer került kiválasztásra. A válogatás figyelembe vette azt a két szempontot, hogy egynyelvű és kétnyelvű, illetve magyarországi és külföldi központú vizsgarendszerek is szerepeljenek a vizsgálatban. A teljesebb kép érdekében került be még a vizsgálati körbe mindkét nyelvből az emelt szintű érettségi vizsga is (a 2007. májusi feladatok), amely ugyan nem államilag elismert nyelvvizsga, ám számos vizsgázó számára a nyelvvizsgával egyenértékű dokumentum megszerzésének államilag finanszírozott lehetőségét képviseli. Ily módon tehát, bár a kép kétségtelenül nem teljes, a minta a lehetőségekhez mérten elfogadhatónak mondható.
- c) A vizsgálat az érintett vizsgáknak csupán egy szintjét (B2) vette górcső alá. Ennek természetesen az az oka, hogy több szint vizsgálatához több próbavizsgázói csoportra lett volna szükség. A nyelvvizsga-kontextust figyelembe véve ez három szinten két nyelvből összesen hat csoportot jelentett volna. A projekt volumene ekkora vizsgálatot nem tett lehetővé. Megjegyzendő azonban, hogy a statisztikák tanúsága szerint a két érintett nyelvből a B2-es vizsgákra jelentkeznek a legtöbben, és az emelt szintű érettségi is ezt a szintet veszi célba. Azaz a vizsgálat szempontjából a B2-es szint kiválasztása tűnt a leglogikusabbnak.
- d) A teljesség hiánya a próbavizsgázók számában is megnyilvánult. Angol nyelvből 20, német nyelvből 16 fő vett részt a projektben. Kétségtelen, hogy ez a két viszonylag kis vizsgázói csoport nem ad alapot messzemenő következtetések levonására. Ez a jelen projektnek nem is lehetett célja. Ahhoz azonban, mint látni fogjuk, ez a vizsgázói létszám is elegendőnek bizonyult, hogy bizonyos lényegesnek tűnő tendenciákra ráirányítsa a figyelmet.

A harmadik kérdőjeles pont az érettségi vizsgák beemelése a projektbe. Mint fentebb hangsúlyoztuk, a cél az volt, hogy a vizsgaeredményeket „éles” helyzetben, tényleges vizsgakörülmények között gyűjtsük be. Ez azonban az érettségi vizsgák esetében praktikus okokból nem volt lehetséges. Ezeket a vizsgafeladatokat a próbavizsgázók külön időpontban, a vizsga eredeti céljának ismerete nélkül oldották meg. Az eljárás természetesen bizonyos aggályokat kelthet, ám a jelen körülmények között ez volt az egyetlen módja annak, hogy a vizsgálatba az emelt szintű érettségi vizsga is integrálható legyen. Arra azonban természetesen fel kell hívnunk a figyelmet, hogy a vizsga eltérő körülményei miatt az érettségit illető következtetések a többi vizsgához képest is csak igen óvatosak lehetnek.

Végezetül megemlítendő, hogy a vizsgált vizsgarendszerek abban is eltértek, hogy némelyek csak komplex formában ajánlanak vizsgát. Ilyen esetben a próbavizsgázók hallásértési és szóbeli feladatokat is megoldottak, míg más vizsgák esetében csak az írásbeli vizsgarészen vettek részt. A kutatás keretei nem tették lehetővé a szóbeli teljesítmények összevetését, így az elemzések csak az írásbeli vizsgarészen elért eredményekre terjedtek ki.

1.2. A kutatás résztvevői

A próbavizsgálók, mint fentebb jeleztük, nyelvenként alkottak egy-egy csoportot. Az angolos csoport 20, a németes 16 főből állt, valamennyien huszoneves fiatalok. Kiválasztásukkor az volt a fő szempont, hogy „B2-es szinthez közeli” vizsgázókból álljanak a csoportok: természetesen azért, hogy az eredményekből tényleges következtetéseket lehessen levonni. Ha például egy B1-es szinthez közeli vizsgázó valamennyi vizsgán gyenge eredményt ért volna el, abból igen kevés érdemi információ származott volna (kivéve, hogy a vizsgák nem B1-es szintűek). Mint látni fogjuk, a tényleges vizsgateljesítményekben így is lényeges eltéréseket mutatott a próbavizsgálói csoport, s ezt vélhetően a résztvevők önértékelésének pontatlansága, illetve a vizsgák közötti esetleges különbségek indokolhatják. Fontos megemlíteni, hogy a próbavizsgálók a projekt megelőzően nem tettek vizsgát az érintett nyelvvizsga-rendszerekben, így teljesítményüket a rendszer, illetve a feladattípusok előzetes ismerete nem befolyásolta.

1.3. Kutatási kérdések

A projekt általános célkitűzései a következő három konkrét kutatási kérdésben öltöttek testet:

1. Milyen egymáshoz viszonyított alapvető leíró statisztikai mutatói vannak az egyes vizsgáknak az összesített írásbelit, illetve a részteszteket (olvasáskészség, íráskészség, nyelvismeret) illetően?
2. Milyen erősségű korreláció mutatható ki az egyes vizsgarendszerekben elért eredmények között az olvasáskészség, az íráskészség, a nyelvismeret (nyelvtan, szókincs), illetve az összesített írásbeli eredmények tekintetében?
3. Kimutatható-e szignifikáns különbség a vizsgázók tényleges teljesítményében a különféle vizsgák különféle készségterületein?

A három kérdés közül külön figyelmet érdemel a harmadik. Ennek az az oka, hogy bár a projekt alapvetően a próbavizsgálók által elért eredmények alapján kívánta az összehasonlítást elvégezni, a harmadik kérdés lehetőséget adott egy elvileg ennél sokkal szélesebb körű s így sokkal megbízhatóbb vizsgálatra is. Az előzetes tervek szerint a próbavizsgákat követően az OH-NYAK bekérte az érintett vizsgaközpontoktól a releváns éles vizsgaadatokat, különös tekintettel az objektíven és szemi-objektíven értékelt feladatok nyers adatfájljaira. Ennek célja az volt, hogy a próbavizsgálói teljesítményeket horgonyteljesítményként használva a teljes vizsgázói populációk eredményeit összekapcsolva IRT¹ elemzést lehessen végrehajtani. Ennek során lehetőség nyílt volna az egyes vizsgázók készségterületenkénti képességindexének megállapítására, illetve a különböző vizsgák azonos készségterületet mérni hivatott feladatainak nehézségiindex-vizsgálatára is. Ezek az adatok meglehetősen nagy biztonsággal jelezték volna a vizsgák közötti esetleges nehézségi különbségeket. Minderre azonban, mint látni fogjuk, csak részben kerülhetett sor.

2. Az adatgyűjtés menete

A projekt előzetes tervének megfelelően a próbavizsgálók 2007 novembere és 2008 májusa között tettek le a nyelvtől függően öt, illetve négy nyelvvizsgát. Ezt követte egy külön alkalommal az érettségi feladatok megoldása.

¹ IRT = Item Response Theory.

Fontos hangsúlyozni, hogy az adatok használhatósága érdekében a próbavizsgálóknak az azonos vizsgákat azonos vizsgaidőszakban kellett letenniük annak érdekében, hogy eredményeik garantáltan azonos feladatokra vonatkozzanak. Az esetek nagy többségében ez nem is okozott problémát, ám helyenként előfordult, hogy előre nem látható okokból egy-egy próbavizsgáló nem tudott megjelenni az adott vizsgaalikalmón, s az illető vizsgát egy másik alkalommal tette le. Bár elméletileg a vizsgarendszerek a különböző vizsgaidőszakokban nem alkalmaznak lényegesen eltérő nehézségű feladatsorokat, a kutatás szerkezetének védelmében ezeket az utólagosan letett vizsgákat az elemzések során nem vettük figyelembe. Ebből adódóan az eredetileg meghatározott próbavizsgálói létszám esetenként alacsonyabb lett. Ezek a változások azonban, tekintettel a próbavizsgálói csoportok eredetileg is viszonylag alacsony létszámára, nem ásták alá a kutatás egészének szerkezetét.

3. Elemzések

A vizsgázók a különböző vizsgák kapcsán eltérő itemszámú feladatokat oldottak meg. Ebből következően teljesítményük közvetlen összevetéséhez a teljesítményeket közös skálára kellett konvertálni. Az egyértelműség és könnyebb átláthatóság kedvéért ez a teljesítmények százalékos eredménnyé történő alakítását jelentette. A továbbiakban tehát a próbavizsgálók által elért valamennyi eredményt százalékos teljesítmény formájában jelenítjük meg.

Az alábbiakban az adatok elemzését a projekt kutatási kérdéseinek struktúráját követve közöljük.

3.1. Leíró statisztikai mutatók

Az elemzésnek ebben a szakaszában a teljes írásbeli vizsgák, illetve a készségterületenként bontott eredmények kapcsán vizsgáljuk meg a különböző vizsgák alapvető leíró statisztikai mutatóit. A nagyobb próbavizsgálói létszám, illetve a több elemzett vizsga okán először az angol nyelvű vizsgákat vesszük közelebről szemügyre.

Az 1. ábra az angol nyelvű vizsgák teljes írásbeli eredményeinek leíró statisztikai mutatóit tartalmazza. Az érintett vizsgák a leíró statisztikák kapcsán nem mutatnak markáns eltéréseket. Kivételt képez azonban az ANGOL1 markánsan alacsonyabb átlageredménye. Ez az érték azt sugallja, hogy ez a vizsga nehezebb volt a többinél, ám ennek kimondásához szükség lesz a további elemzési eredményekre is. Ugyancsak figyelmet érdemel az ANGOL2 esetében megjelenő alacsonyabb szórás érték, amely azt sejteti, hogy a teszt, összevetve a többi vizsgával, kevésbé tett különbséget a vizsgázók között. Érdekes az ANGOL1 és az érettségi kapcsán megjelenő alacsony minimum-teljesítmény is, ám ezek egyedi eseteket is jelezhetnek, így messzemenő következtetéseknek aligha adnak alapot.

	N	Átlag	Szórás	Minimum	Maximum
ANGOL1	20	53,9000	19,89684	6,00	83,00
ANGOL2	20	69,1500	11,86448	41,00	89,00
ANGOL3	14	63,7857	16,19201	15,00	80,00
ANGOL4	19	63,8947	20,85637	12,00	91,00
ANGOL5	19	66,2105	18,78409	24,00	95,00
Érettségi	14	66,5714	21,70684	9,00	93,00

1. ábra: Leíró statisztikák (angol, teljes írásbeli)

A 2. ábrán ugyanezen vizsgák olvasási komponensének leíró statisztikáit láthatjuk. Érdekes módon az átlagok látszólag három csoportra bontják a vizsgákat. Az elsőben a magas átlageredményű vizsgák szerepelnek (ANGOL2, ANGOL4, érettségi), a másodikban a közepes eredményűek (ANGOL3, ANGOL5), míg az ANGOL1 itt is látványosan alacsonyabb eredményt mutat. Bár kínálkozik a következtetés, ezen adatok alapján korai volna még a vizsgák tényleges nehézségére következtetnünk. A szórás értékek két csoportot látszanak megjeleníteni, de itt is korai volna még ennek jelentőségét hangsúlyozni.

	N	Átlag	Szórás	Minimum	Maximum
ANGOL1	20	55,6000	24,79049	,00	92,00
ANGOL2	20	77,2000	15,37804	46,00	100,00
ANGOL3	14	66,7857	15,26776	25,00	85,00
ANGOL4	19	71,4474	24,64069	15,00	100,00
ANGOL5	19	65,9474	18,42545	33,00	97,00
Érettségi	14	76,0000	23,15167	17,00	100,00

2. ábra: Leíró statisztikák (angol, olvasás)

A 3. ábrán látható az írásfeladatok eredményeinek leíró statisztikája. Ismét megfigyelhető az ANGOL1 jelentősen alacsonyabb átlaga, bár itt az ANGOL4 és az érettségi is alacsonyabb értékeket mutat a többi vizsgához képest. A szórások tekintetében itt alacsony értéket mutat az ANGOL2, s a többihez képest magasat az ANGOL5.

	N	Átlag	Szórás	Minimum	Maximum
ANGOL1	20	49,5000	18,48897	,00	70,00
ANGOL2	20	61,5500	12,29238	37,00	81,00
ANGOL3	14	65,0000	20,56883	,00	85,00
ANGOL4	19	56,8947	20,52341	,00	82,00
ANGOL5	19	68,0000	26,45961	,00	86,00
Érettségi	14	59,0714	23,38462	,00	93,00

3. ábra: Leíró statisztikák (angol, írás)

A 4. ábrán szereplő nyelvismereti komponensek eredményei az eddigieknél is markánsabb különbséget mutatnak az ANGOL1 kapcsán. Különösen figyelemreméltó az eltérés az ANGOL1 és az ANGOL2 között, s a különbség a szórásadatokban is megmutatkozik. Figyelembe kell azonban venni a viszonylag kis minta hatásait is.

	N	Átlag	Szórás	Minimum	Maximum
ANGOL1	20	38,3500	21,10693	13,00	80,00
ANGOL2	20	68,8500	10,81556	40,00	89,00
ANGOL3	14	59,2857	16,39150	20,00	80,00
ANGOL4	19	66,1316	17,94114	33,00	93,00
ANGOL5	19	62,1053	24,53545	20,00	100,00
Érettségi	14	65,0000	22,26069	10,00	97,00

4. ábra: Leíró statisztikák (angol, nyelvismeret)

Úgy tűnik tehát, hogy az angol nyelvű vizsgák esetében a különféle mérések nagyjából hasonló eredményekre vezettek. Bár számos különbség figyelhető meg, ezek jelentősége pusztán a leíró statisztikai adatok alapján nem állapítható meg. Feltűnő ugyanakkor, hogy az ANGOL1 tesztjei konzekvensen nehezebbnek tűnnek a többi vizsgáénál, ám még ezek a látványos különbségek is további elemzést igényelnek.

A német nyelvű vizsgák esetében is először a teljes írásbeli eredményeket vehetjük szemügyre az 5. ábrán. Nincs olyan komolyabbnak látszó eltérés, mint az angol nyelvű vizsgák esetében, bár az érettségi szórása valamelyest nagyobbak tűnik.

	N	Átlag	Szórás	Minimum	Maximum
NÉMET1	16	67,0625	15,84285	32,00	96,00
NÉMET2	7	72,5714	14,18752	53,00	91,00
NÉMET3	15	69,1333	12,87227	49,00	90,00
NÉMET4	15	66,3333	15,39790	40,00	87,00
Érettségi	11	73,7273	21,29362	32,00	98,00

5. ábra: Leíró statisztikák (német, teljes írásbeli)

Ehhez képest a 6. ábrán látható olvasásteszt-adatok már kissé eltérő képet nyújtanak. Magas átlagértéket mutat az érettségi és különösen a NÉMET2. Utóbbi esetében azonban fel kell figyelnünk az igen alacsony számú mintára, amely könnyen okozhat ilyesfajta torzulást, s ez potenciálisan igaz lehet az érettségre is.

	N	Átlag	Szórás	Minimum	Maximum
NÉMET1	16	68,5000	15,44884	36,00	88,00
NÉMET2	7	84,2857	14,55695	60,00	100,00
NÉMET3	15	70,2667	15,91705	33,00	90,00
NÉMET4	15	73,0667	20,02665	28,00	100,00
Érettségi	11	79,4545	23,85525	20,00	100,00

6. ábra: Leíró statisztikák (német, olvasás)

A 7. ábrán látható írásteszt-eredmények alapján ismét az érettségi tűnik a legmagasabb átlagpontszámot generáló vizsgának, ám érdekes ellenpontja ennek a NÉMET2, amely ezt a résztesztet tekintve a legnehezebbnek tűnik. Az olvasás kapcsán legmagasabb átlagérték (NÉMET2) az írás esetében szembeszökően a legalacsonyabb, ami (bár a minta mérete itt is lényeges lehet) azt sejteti, hogy a vizsgán belül különös viszony áll fenn a két vizsgarész között. A német vizsgák esetében, mivel nem minden vizsgában szerepel különálló nyelvismereti komponens, ezen terület vizsgálatától eltekintünk.

	N	Átlag	Szórás	Minimum	Maximum
NÉMET1	16	60,5000	20,64946	24,00	100,00
NÉMET2	7	55,8571	19,65052	33,00	90,00
NÉMET3	15	67,6000	14,04991	46,00	86,00
NÉMET4	15	61,6667	20,41241	20,00	100,00
Érettségi	11	71,7273	29,64824	23,00	100,00

7. ábra: Leíró statisztikák (német, írás)

Összességében a német nyelvű vizsgák kapcsán a leíró statisztikák kevésbé jeleznek érdemi eltéréseket a vizsgák között. Jóllehet látványos eltérések léteznek, a minta kisebb mérete miatt egyelőre ezzel kapcsolatban csak óvatosan fogalmazhatunk.

A fenti adatok mindkét nyelv esetében figyelemreméltó jelenségekre irányították figyelmünket, ám a leíró statisztikai adatok nem elégségesek komoly következtetések levonására. Lépünk tehát tovább a 2. kutatási kérdéshez, amely újabb részletekkel fogja gazdagítani a képet.

3.2. Korreláció-elemzések

Az elemzések ezen szakaszában azt vizsgáljuk meg, milyen erős korrelációkat találunk a vizsgák, illetve a vizsgakomponensek között. Előjáróban azonban két fontos megállapítást kell tennünk.

Egyrészt az esetleges magas korrelációs együttható önmagában nem bizonyítja, hogy a két vizsgált teljesítmény egyforma vagy közel egyforma értékelést kapott. Előfordulhat az is, hogy a vizsgázók az egyik teszten magas pontszámokat érnek el, a másikon alacsonyakat, a korrelációs együttható mégis magas lesz, mivel az adatSORokban megjelenő trendek hasonlóak. Ebből adódóan önmagában a korrelációs együttható sem ad elegendő információt a teljes körű összehasonlításhoz.

Másrészt lényeges megállapítanunk, hogy melyik típusú korrelációs együtthatók alkalmazhatóak a jelen kontextusban. Fontos leszögeznünk, hogy a látszat ellenére a tesztpontszámok nem intervallum-skálán helyezkednek el, hanem rangsor-skálán. Ráadásul a viszonylag alacsony létszámok miatt az adatok normál eloszlása is kétséges. Mindezek alapján a következőkben a Spearman-féle korrelációs együtthatók kerülnek bemutatásra.

A 3.1. szakasz logikáját követve először az angol nyelvű vizsgák eredményeit vesszük sorra. Elsőként a teljes írásbeli vizsgák összehasonlítása látható a 8. ábrán. Az esetek többségében a vizsgaeredmények viszonylag erős összefüggést mutatnak, azaz a különböző vizsgaeredmények hasonló trendeket tartalmaznak. Lényeges körülmény az is, hogy a korrelációs együtthatók mindegyike statisztikailag szignifikáns összefüggést mutat, azaz az eredmények nem tekinthetők véletlenszerűnek. Ez alól igazából csak az ANGOL1 és az érettségi viszonya kivétel, ahol is valamivel gyengébb, bár szignifikáns kapcsolat látható.

Spearman's rho		ANGOL1	ANGOL2	ANGOL3	ANGOL4	ANGOL5	Érettségi
ANGOL1	Korrelációs együttható	1,000	,824(**)	,727(**)	,827(**)	,721(**)	,691(**)
	Szignifikanciaszint	.	,000	,003	,000	,000	,006
	N	20	20	14	19	19	14
ANGOL2	Korrelációs együttható	,824(**)	1,000	,862(**)	,842(**)	,778(**)	,776(**)
	Szignifikanciaszint	,000	.	,000	,000	,000	,001
	N	20	20	14	19	19	14
ANGOL3	Korrelációs együttható	,727(**)	,862(**)	1,000	,862(**)	,789(**)	,868(**)
	Szignifikanciaszint	,003	,000	.	,000	,001	,001
	N	14	14	14	14	13	11

ANGOL4	Korrelációs együttható	,827(**)	,842(**)	,862(**)	1,000	,860(**)	,861(**)
	Szignifikanciaszint	,000	,000	,000	.	,000	,000
	N	19	19	14	19	18	13
ANGOL5	Korrelációs együttható	,721(**)	,778(**)	,789(**)	,860(**)	1,000	,824(**)
	Szignifikanciaszint	,000	,000	,001	,000	.	,001
	N	19	19	13	18	19	13
Érettségi	Korrelációs együttható	,691(**)	,776(**)	,868(**)	,861(**)	,824(**)	1,000
	Szignifikanciaszint	,006	,001	,001	,000	,001	.
	N	14	14	11	13	13	14

** Szignifikáns korreláció ($p < 0,01$).

8. ábra: Korrelációs együtthatók (angol, teljes írásbeli)

A 9. ábrán az olvasásteszték eredményei közötti összefüggést vehetjük szemügyre. Az olvasásteszték kapcsán már korántsem olyan kedvező a kép, mint az összesített eredmények esetében. Az olyan változók esetében, ahol erőteljes összefüggést várunk el, a korrelációs együtthatók legalacsonyabb elfogadható értéke 0,7–0,8 körül van. Jelen esetben egyértelműen magas korrelációt várnánk, hiszen ha ugyanazon vizsgázók ugyanazon képességét (olvasás) ugyanazon a szinten mérik a különböző vizsgák, az eredményeknek meglehetősen hasonlóságot kellene mutatniuk. Ehhez képest 0,7 feletti értéket mindössze egy vizsga-pár mutat (ANGOL2 és ANGOL3), az összes többi esetben meglehetősen gyenge, 0,5–0,6 körüli, esetleg a 0,7-es értéket közelítő adatokkal találkozunk. Mindez arra utal, hogy a teljes írásbeli vizsgák relatív hasonlóságának ellenére az olvasástesztékben igen komoly különbségek rajzolódnak ki a vizsgák között. Fontos azonban arra is kitérnünk, hogy az ábrán szereplő alacsony értékek egy része nem szignifikáns, azaz a véletlen műve is lehet. A különös eredmények háttérében ráadásul újra feltűnik az alacsony elemszám mint potenciális magyarázat. Mindennek ellenére a kutatás fókuszát tekintve az olvasásteszték eredményei igen figyelemreméltóak.

Spearman's rho		ANGOL1	ANGOL2	ANGOL3	ANGOL4	ANGOL5	Érettségi
ANGOL1	Korrelációs együttható	1,000	,464(*)	,531	,497(*)	,446	,683(**)
	Szignifikanciaszint	.	,039	,051	,030	,055	,007
	N	20	20	14	19	19	14
ANGOL2	Korrelációs együttható	,464(*)	1,000	,796(**)	,677(**)	,627(**)	,608(*)
	Szignifikanciaszint	,039	.	,001	,001	,004	,021
	N	20	20	14	19	19	14
ANGOL3	Korrelációs együttható	,531	,796(**)	1,000	,611(*)	,549	,554
	Szignifikanciaszint	,051	,001	.	,020	,052	,077
	N	14	14	14	14	13	11

ANGOL4	Korrelációs együttható	,497(*)	,677(**)	,611(*)	1,000	,585(*)	,511
	Szignifikanciaszint	,030	,001	,020	.	,011	,074
	N	19	19	14	19	18	13
ANGOL5	Korrelációs együttható	,446	,627(**)	,549	,585(*)	1,000	,639(*)
	Szignifikanciaszint	,055	,004	,052	,011	.	,019
	N	19	19	13	18	19	13
Érettségi	Korrelációs együttható	,683(**)	,608(*)	,554	,511	,639(*)	1,000
	Szignifikanciaszint	,007	,021	,077	,074	,019	.
	N	14	14	11	13	13	14

* Szignifikáns korreláció ($p < 0,05$).

** Szignifikáns korreláció ($p < 0,01$).

9. ábra: Korrelációs együtthatók (angol, olvasás)

A 10. ábrán az írástesztek korrelációs együtthatóit vizsgálhatjuk meg. Jól kirajzolódik, hogy számos vizsga igen gyenge kapcsolatot mutat a többi vizsgával. Leggyengébbnek az ANGOL4 és az ANGOL5 látszik, amelyek semelyik más vizsgával nem mutatnak elfogadható mértékű összefüggést, még egymással sem. Ehhez képest az ANGOL1 vagy az érettségi a többi vizsgával elfogadható szintű összefüggést mutat, az ANGOL4 és az ANGOL5 kivételével. Itt azonban ismét meg kell állnunk egy pillanatra. Bár az adatok alapján az ANGOL4 és az ANGOL5 nincs összhangban a többi vizsgával, ez nem jelenti feltétlenül, hogy nem jól mérnek. Ennek eldöntéséhez tudnunk kellene, hogy a vizsgált tesztek közül melyik az etalon, a célkésztséget feltétlenül érvényes módon mérő vizsga. Ezt azonban nem tudjuk. Az adatok mindössze annyit sejtetnek, hogy a vizsgák között némi diszharmonia áll fenn; bár itt is igaz, hogy számos index nem statisztikailag szignifikáns, s a látszólag problematikus értékek elméletileg betudhatóak az alacsony elemszámnak is.

Spearman's rho		ANGOL1	ANGOL2	ANGOL3	ANGOL4	ANGOL5	Érettségi
ANGOL1	Korrelációs együttható	1,000	,732(**)	,749(**)	,535(*)	,590(**)	,723(**)
	Szignifikanciaszint	.	,000	,002	,018	,008	,003
	N	20	20	14	19	19	14
ANGOL2	Korrelációs együttható	,732(**)	1,000	,503	,612(**)	,664(**)	,720(**)
	Szignifikanciaszint	,000	.	,067	,005	,002	,004
	N	20	20	14	19	19	14
ANGOL3	Korrelációs együttható	,749(**)	,503	1,000	,305	,412	,809(**)
	Szignifikanciaszint	,002	,067	.	,290	,161	,003
	N	14	14	14	14	13	11

ANGOL4	Korrelációs együttható	,535(*)	,612(**)	,305	1,000	,322	,609(*)
	Szignifikanciaszint	,018	,005	,290	.	,192	,027
	N	19	19	14	19	18	13
ANGOL5	Korrelációs együttható	,590(**)	,664(**)	,412	,322	1,000	,585(*)
	Szignifikanciaszint	,008	,002	,161	,192	.	,036
	N	19	19	13	18	19	13
Érettségi	Korrelációs együttható	,723(**)	,720(**)	,809(**)	,609(*)	,585(*)	1,000
	Szignifikanciaszint	,003	,004	,003	,027	,036	.
	N	14	14	11	13	13	14

* Szignifikáns korreláció ($p < 0,05$).

** Szignifikáns korreláció ($p < 0,01$).

10. ábra: Korrelációs együtthatók (angol, írás)

A 11. ábrán az angol nyelvű vizsgák utolsó elemeként a nyelvismereti tesztek eredményeinek összefüggését tekinthetjük át. Először is azt érdemes megjegyeznünk, hogy valamennyi korrelációs együttható statisztikailag szignifikáns, azaz az összefüggések valószínűleg nem a véletlen művei. A többi részvizsgálóhoz képest nagyobb egyezés látszik a vizsgák között; az érettségi például egy kivétellel (ANGOL3) valamennyi más vizsgával megfelelő erősségű kapcsolatot mutat. Tény ugyanakkor, hogy továbbra is akadnak 0,7 alatti értékek, amelyek a kívánatosnál gyengébb összefüggésre utalnak.

Spearman's rho		ANGOL1	ANGOL2	ANGOL3	ANGOL4	ANGOL5	Érettségi
ANGOL1	Korrelációs együttható	1,000	,654(**)	,708(**)	,823(**)	,645(**)	,846(**)
	Szignifikanciaszint	.	,002	,005	,000	,003	,000
	N	20	20	14	19	19	14
ANGOL2	Korrelációs együttható	,654(**)	1,000	,668(**)	,623(**)	,725(**)	,845(**)
	Szignifikanciaszint	,002	.	,009	,004	,000	,000
	N	20	20	14	19	19	14
ANGOL3	Korrelációs együttható	,708(**)	,668(**)	1,000	,877(**)	,717(**)	,667(*)
	Szignifikanciaszint	,005	,009	.	,000	,006	,025
	N	14	14	14	14	13	11
ANGOL4	Korrelációs együttható	,823(**)	,623(**)	,877(**)	1,000	,759(**)	,845(**)
	Szignifikanciaszint	,000	,004	,000	.	,000	,000
	N	19	19	14	19	18	13
ANGOL5	Korrelációs együttható	,645(**)	,725(**)	,717(**)	,759(**)	1,000	,820(**)
	Szignifikanciaszint	,003	,000	,006	,000	.	,001
	N	19	19	13	18	19	13

Érettségi	Korrelációs együttható	,846(**)	,845(**)	,667(*)	,845(**)	,820(**)	1,000
	Szignifikanciaszint	,000	,000	,025	,000	,001	.
	N	14	14	11	13	13	14

* Szignifikáns korreláció ($p < 0,05$).

** Szignifikáns korreláció ($p < 0,01$).

11. ábra: Korrelációs együtthatók (angol, nyelvismeret)

A német nyelvű vizsgák esetében is célszerű először a teljes írásbeli eredményeket megvizsgálni. A 12. ábrán láthatók ezek az adatok. A vizsgák többsége erős, helyenként igen erős összefüggést mutat egymással. Ez alól egyedül az érettségi kivétel, amely ugyan tökéletesen korrelál a NÉMET2-vel, ám itt igen kicsi a közös vizsgázók száma, ellenben a többi vizsgával gyenge, illetve helyenként gyenge és nem szignifikáns az összefüggés. Ismét fontos azonban felhívni a figyelmet arra, hogy az eredmények háttérében a kutatás ismert korlátai is állhatnak.

Spearman's rho		NÉMET1	NÉMET2	NÉMET3	NÉMET4	Érettségi
NÉMET1	Korrelációs együttható	1,000	,857(*)	,909(**)	,813(**)	,419
	Szignifikanciaszint	.	,014	,000	,000	,200
	N	16	7	15	15	11
NÉMET2	Korrelációs együttható	,857(*)	1,000	,964(**)	,929(**)	1,000(**)
	Szignifikanciaszint	,014	.	,000	,003	.
	N	7	7	7	7	4
NÉMET3	Korrelációs együttható	,909(**)	,964(**)	1,000	,783(**)	,636(*)
	Szignifikanciaszint	,000	,000	.	,001	,035
	N	15	7	15	15	11
NÉMET4	Korrelációs együttható	,813(**)	,929(**)	,783(**)	1,000	,565
	Szignifikanciaszint	,000	,003	,001	.	,070
	N	15	7	15	15	11
Érettségi	Korrelációs együttható	,419	1,000(**)	,636(*)	,565	1,000
	Szignifikanciaszint	,200	.	,035	,070	.
	N	11	4	11	11	11

* Szignifikáns korreláció ($p < 0,05$).

** Szignifikáns korreláció ($p < 0,01$).

12. ábra: Korrelációs együtthatók (német, teljes írásbeli)

Az olvasástereszték kapcsán már itt is árnyaltabb a kép, amint az a 13. ábrán látható. A vizsgák egy része erős összefüggést mutat, bár ez a NÉMET2 esetében az alacsony elemszámnak is betudható. Számos esetben azonban az összefüggés gyenge és/vagy nem szignifikáns, azaz ebben a vizsgarészben nagyobbak tűnik a bizonytalanság.

Spearman's rho		NÉMET1	NÉMET2	NÉMET3	NÉMET4	Érettségi
NÉMET1	Korrelációs együttható	1,000	,953(**)	,716(**)	,457	,642(*)
	Szignifikanciaszint	.	,001	,003	,086	,033
	N	16	7	15	15	11
NÉMET2	Korrelációs együttható	,953(**)	1,000	,743	,833(*)	,943
	Szignifikanciaszint	,001	.	,056	,020	,057
	N	7	7	7	7	4
NÉMET3	Korrelációs együttható	,716(**)	,743	1,000	,434	,673(*)
	Szignifikanciaszint	,003	,056	.	,106	,023
	N	15	7	15	15	11
NÉMET4	Korrelációs együttható	,457	,833(*)	,434	1,000	,256
	Szignifikanciaszint	,086	,020	,106	.	,447
	N	15	7	15	15	11
Érettségi	Korrelációs együttható	,642(*)	,943	,673(*)	,256	1,000
	Szignifikanciaszint	,033	,057	,023	,447	.
	N	11	4	11	11	11

* Szignifikáns korreláció (p<0,05).

** Szignifikáns korreláció (p<0,01).

13. ábra: Korrelációs együtthatók (német, olvasás)

Némiképp hasonló a helyzet az írástesztek esetében. A 14. ábra adatai tanúsítják, hogy a vizsgák többnyire megfelelő vagy ahhoz közeli összefüggést mutatnak egymással, kivéve az érettségit, amely kizárólag alacsony és nem szignifikáns értékeket mutat.

Spearman's rho		NÉMET1	NÉMET2	NÉMET3	NÉMET4	Érettségi
NÉMET1	Korrelációs együttható	1,000	,844(*)	,698(**)	,730(**)	,548
	Szignifikanciaszint	.	,017	,004	,002	,081
	N	16	7	15	15	11
NÉMET2	Korrelációs együttható	,844(*)	1,000	,774(*)	,618	,316
	Szignifikanciaszint	,017	.	,041	,139	,684
	N	7	7	7	7	4
NÉMET3	Korrelációs együttható	,698(**)	,774(*)	1,000	,736(**)	,399
	Szignifikanciaszint	,004	,041	.	,002	,225
	N	15	7	15	15	11

NÉMET4	Korrelációs együttható	,730(**)	,618	,736(**)	1,000	,392
	Szignifikanciaszint	,002	,139	,002	.	,233
	N	15	7	15	15	11
Érettségi	Korrelációs együttható	,548	,316	,399	,392	1,000
	Szignifikanciaszint	,081	,684	,225	,233	.
	N	11	4	11	11	11

* Szignifikáns korreláció ($p < 0,05$).

** Szignifikáns korreláció ($p < 0,01$).

14. ábra: Korrelációs együtthatók (német, írás)

Összességében megállapítható, hogy a vizsgák, illetve a vizsgakomponensek korrelációs együtthatóinak vizsgálata mindkét nyelv esetében csak részleges eredményekkel szolgál. Bár számos esetben erős szignifikáns korreláció mutatható ki, ez a korreláció természetéből adódóan még nem bizonyítja a különböző tesztek hasonló működését. Másrészt ugyancsak számos esetben a kívánatosnál gyengébb, bár szignifikáns, illetve esetenként nem szignifikáns összefüggések is kimutathatók voltak. Ezek az esetek többféle módon is magyarázhatóak, így szükséges az elemzés további fázisainak áttekintése.

3.3. A tényleges vizsgázói teljesítménykülönbségek elemzése

Mint fentebb már jeleztük, a leíró statisztikák, illetve a korrelációanalízisek nem nyújtanak elegendő információt annak eldöntésére, vajon a vizsgázók tényleges teljesítménye milyen mértékben különbözik a különböző vizsgák és vizsgarészek esetében. A megoldást ilyen esetekben az jelenti, ha a tényleges vizsgázói teljesítmények közötti eltéréseket statisztikai eszközökkel vizsgáljuk meg. A jelen kontextusban azt kellett figyelembe venni, hogy (amint azt a korrelációtípusok kapcsán már jeleztük) a teszt-pontszámok nem intervallum-, hanem rangsor-skálán helyezkednek el, a próbavizsgázói csoport mérete pedig kétséges teszi a normáeloszlás meglétét. Ennek a két feltételnek a megléte volna szükséges a parametrikus statisztikák alkalmazásához. Mivel a jelen esetben erre nem volt lehetőség, nemparametrikus statisztikai eljárást kellett alkalmazni, konkrétan a Wilcoxon-próbát (*Wilcoxon Signed Ranks Test*).

A következőkben a Wilcoxon-próbák eredményeit tekintjük át a vizsgák és vizsgarészek tekintetében. Az eddigi rendszert követve először az angol tesztekkel vesszük sorra. A teljes tesztek összevetését a 15. ábra mutatja. Látható, hogy a vizsgák nagy részének esetében a vizsgázói teljesítmények között nincs szignifikáns különbség a különböző vizsgákon. Kivétel azonban az ANGOL1, amely minden egyéb vizsgától szignifikánsan eltérő eredményeket produkált. Ez az eredmény egybecsengeni látszik a leíró statisztikákban tapasztalt eltéréssel, amely szerint az átlagpontszám az ANGOL1 esetében lényegesen alacsonyabb volt. Úgy tűnik tehát, az eltérés szignifikáns, s így tényleges eltérést jelez. Mivel a többi vizsga egymáshoz képest semmilyen kombinációban nem mutat szignifikáns eltérést, az ANGOL1 viszont minden kombinációban eltérő eredményt mutat, ez azt sejteti, hogy (legalábbis a teljes írásbeli eredm-

nyek szintjén) az ANGOL1 lényegesen különbözik a többi vizsgától. A végleges következtetések levonása előtt azonban itt is ellenőrizni kell a vizsgakomponensekre vonatkozó adatokat is.

	Z	Szignifikanciaszint
ANGOL2 – ANGOL1	-3,744	0,000181086
ANGOL3 – ANGOL1	-2,828	0,004687947
ANGOL4 – ANGOL1	-3,443	0,000575729
ANGOL5 – ANGOL1	-3,280	0,001036593
Érettségi – ANGOL1	-2,622	0,008754057
ANGOL3 – ANGOL2	-1,887	0,059150276
ANGOL4 – ANGOL2	-,785	0,432603646
ANGOL5 – ANGOL2	-,981	0,326627227
Érettségi – ANGOL2	-1,163	0,245028035
ANGOL4 – ANGOL3	-1,479	0,139070673
ANGOL5 – ANGOL3	-1,333	0,182624303
Érettségi – ANGOL3	-,581	0,561572387
ANGOL5 – ANGOL4	-,305	0,760036899
Érettségi – ANGOL4	-,946	0,344266548
Érettségi – ANGOL5	-1,571	0,116101744

15. ábra: Wilcoxon-próba (angol teljes írásbeli)

A 16. ábrán az angol olvasásteszteken elért teljesítményekre vonatkozó Wilcoxon-próba eredményei láthatóak. Itt már nem annyira egyértelmű a helyzet. Az ANGOL1 az olvasásteszt esetében is szignifikáns különbséget mutat három vizsgához képest (ANGOL2, ANGOL4, érettségi), ám kettővel nem (ANGOL3, ANGOL5). Igaz, ez utóbbi két eset is súrolja a szignifikáns különbség határát. Még érdekesebb azonban, hogy itt már az ANGOL2 is szignifikáns különbséget mutat az ANGOL1-n kívül két másik vizsgához képest is (ANGOL3, ANGOL5).

	Z	Szignifikanciaszint
ANGOL2 - ANGOL1	-3,323	0,000889
ANGOL3 - ANGOL1	-1,923	0,054428
ANGOL4 - ANGOL1	-2,859	0,004255
ANGOL5 - ANGOL1	-1,771	0,076588
Érettségi - ANGOL1	-3,045	0,002329
ANGOL3 - ANGOL2	-2,669	0,007601
ANGOL4 - ANGOL2	-,893	0,371747
ANGOL5 - ANGOL2	-2,810	0,004954
Érettségi - ANGOL2	-1,119	0,263057
ANGOL4 - ANGOL3	-1,957	0,050382
ANGOL5 - ANGOL3	-,078	0,937473
Érettségi - ANGOL3	-1,378	0,168253

ANGOL5 - ANGOL4	-1,329	0,183772
Érettségi - ANGOL4	-,489	0,624489
Érettségi - ANGOL5	-,801	0,423367

16. ábra: Wilcoxon-próba (angol, olvasás)

A kép további árnyalása érdekében lépünk tovább az írástesztekhez.

A 17. ábra az erre vonatkozó Wilcoxon-próba adatait mutatja be. Az ANGOL1 ismét három vizsgálóhoz képest mutat szignifikáns eltérést (ANGOL2, ANGOL3, ANGOL5), bár ezek (egy kivétellel) nem azonosak az olvasásteszt kapcsán hasonló tendenciát mutató vizsgákkal. Érdekes, hogy az érettségi a ANGOL3-hez, a ANGOL5 pedig a ANGOL4-hez képest mutat szignifikáns különbséget.

	Z	Szignifikanciaszint
ANGOL2 - ANGOL1	-3,501	0,000463
ANGOL3 - ANGOL1	-2,598	0,009363
ANGOL4 - ANGOL1	-1,873	0,061023
ANGOL5 - ANGOL1	-3,269	0,001078
Érettségi - ANGOL1	-1,015	0,310009
ANGOL3 - ANGOL2	-1,601	0,109421
ANGOL4 - ANGOL2	-,915	0,359968
ANGOL5 - ANGOL2	-1,711	0,087084
Érettségi - ANGOL2	-,565	0,571797
ANGOL4 - ANGOL3	-1,295	0,195367
ANGOL5 - ANGOL3	-,785	0,432413
Érettségi - ANGOL3	-2,456	0,014053
ANGOL5 - ANGOL4	-2,227	0,025942
Érettségi - ANGOL4	-,044	0,964524
Érettségi - ANGOL5	-1,728	0,084021

17. ábra: Wilcoxon-próba (angol, írás)

Az angol nyelvű vizsgák utolsó elemzésre váró komponense a nyelvismereti teszt. Az ehhez kapcsolódó Wilcoxon-próba eredményeit a 18. ábra tartalmazza. Az ANGOL1 elkülönülése a többi vizsgától ismét egyértelmű: az eredmények valamennyi vizsgálóhoz viszonyítva szignifikánsan különböznek. Hasonlóan szignifikáns eltérést mutat azonban az ANGOL3 is az ANGOL2-hoz és az ANGOL4-hez képest.

	Z	Szignifikanciaszint
ANGOL2 - ANGOL1	-3,920	0,000088
ANGOL3 - ANGOL1	-2,972	0,002959
ANGOL4 - ANGOL1	-3,829	0,000129
ANGOL5 - ANGOL1	-3,399	0,000677
Érettségi - ANGOL1	-3,051	0,002278
ANGOL3 - ANGOL2	-2,702	0,006891

ANGOL4 - ANGOL2	-,479	0,631658
ANGOL5 - ANGOL2	-1,449	0,147252
Érettségi - ANGOL2	-1,375	0,168989
ANGOL4 - ANGOL3	-3,188	0,001435
ANGOL5 - ANGOL3	-1,256	0,209076
Érettségi - ANGOL3	-,223	0,823669
ANGOL5 - ANGOL4	-,501	0,616408
Érettségi - ANGOL4	-1,766	0,077328
Érettségi - ANGOL5	-1,260	0,207846

18. ábra: Wilcoxon-próba (angol, nyelvismeret)

Az angol nyelvű vizsgák áttekintése után most áttérünk a német nyelvű vizsgákhoz kapcsolódó Wilcoxon-próbák eredményeinek elemzésére. A 19. ábra a teljes írásbeli vizsgákhoz kapcsolódó eredményeket mutatja be. Megállapítható, hogy a vizsgák sehol nem mutatnak szignifikáns különbséget egymáshoz képest. A vizsgálat szempontjából ezt mindenképpen pozitívként könyvelhetjük el, hiszen a különféle vizsgák hasonló működésére utal. A résztesztek elemzése azonban itt is igen lényeges lehet.

	Z	Szignifikanciaszint
NÉMET2 - NÉMET1	-,848	0,396333
NÉMET3 - NÉMET1	-1,015	0,310009
NÉMET4 - NÉMET1	-,057	0,954644
Érettségi - NÉMET1	-,711	0,476907
NÉMET3 - NÉMET2	-,138	0,89023
NÉMET4 - NÉMET2	-,315	0,752494
Érettségi - NÉMET2	-1,841	0,0656
NÉMET4 - NÉMET3	-1,414	0,157504
Érettségi - NÉMET3	-,713	0,476033
Érettségi - NÉMET4	-1,112	0,266166

19. ábra: Wilcoxon-próba (német, teljes írásbeli)

Következő lépésként tehát az olvasástesztek elemzését végezzük el a 20. ábra segítségével. A vizsgák nagy része nem mutat egymáshoz képest szignifikáns eltérést, kivéve a NÉMET2-t, amely a NÉMET1 és a NÉMET3 viszonylatában szignifikánsan különbözik. Ez újfent a vizsgák nagy részének hasonló működését sejteti, a NÉMET2 kapsán pedig ismét felvetődik az alacsony elemszám mint magyarázat.

	Z	Szignifikanciaszint
NÉMET2 - NÉMET1	-2,384	0,017148
NÉMET3 - NÉMET1	-,739	0,459849
NÉMET4 - NÉMET1	-1,721	0,085319
Érettségi - NÉMET1	-1,778	0,075368
NÉMET3 - NÉMET2	-2,197	0,027992

NÉMET4 - NÉMET2	-1,016	0,309629
Érettségi - NÉMET2	-1,134	0,256839
NÉMET4 - NÉMET3	-,880	0,378884
Érettségi - NÉMET3	-1,468	0,141974
Érettségi - NÉMET4	-,979	0,327586

20. ábra: Wilcoxon-próba (német, olvasás)

A 21. ábra az írásteszték eredményeit mutatja be. Ismét úgy tűnik, hogy a vizsgák nagy része egymáshoz képest hasonló módon működik, leszámítva a NÉMET3 és a NÉMET2 között fennálló szignifikáns különbséget. Mint korábban többször is utaltunk rá, a NÉMET2 az alacsony elemszáma miatt nem feltétlenül mérvadó eredményt mutat.

	Z	Szignifikanciaszint
NÉMET2 - NÉMET1	-1,521	0,12819
NÉMET3 - NÉMET1	-1,648	0,099267
NÉMET4 - NÉMET1	-,157	0,875261
Érettségi - NÉMET1	-1,290	0,197105
NÉMET3 - NÉMET2	-2,205	0,027434
NÉMET4 - NÉMET2	-1,947	0,051495
Érettségi - NÉMET2	-1,826	0,067889
NÉMET4 - NÉMET3	-1,363	0,172762
Érettségi - NÉMET3	-,979	0,327586
Érettségi - NÉMET4	-,889	0,373945

21. ábra: Wilcoxon-próba (német, írás)

Az elemzések közül már csupán egy van hátra, a közös vizsgázókon keresztül összekapcsolt különböző vizsgák eredményeinek IRT elemzése. Mint fentebb jeleztük, ehhez az OH-NYAK bekérte az érintett vizsgaközpontoktól a releváns nyers adatfájlokat. Sajnálatos módon erre azonban többszöri ismételt felszólítás ellenére is csak három vizsgaközpont reagált (ANGOL1, ANGOL4, ANGOL2), s még ezek az adatok is csak egy nyelv (angol) és egy készségterület, az olvasásértés kapcsán voltak értékelhetőek.

Természetesen ez a körülmény igen hátrányosan érintette az elemzést, hiszen így csupán egyetlen szempont tényleges elemzésére volt mód: vajon a három vizsga olvasástesztjeiben szereplő itemek nehézségi (logit) indexei mutatnak-e szignifikáns különbséget. Bár az eredeti elképzelésekhez képest ez igen csekély mozgásteret jelent, így is érdemes megvizsgálnunk, mit mutatnak az adatok. Az elemzés ezen szakaszában a különböző feladatokban szereplő itemek nehézségi indexeit varianciaanalízis segítségével hasonlítottuk össze. Ennek eredményei láthatóak a 22. ábrán. Az ANGOL1 olvasástesztjében szereplő itemek szignifikánsan nehezebbnek mutatkoztak a másik két vizsgával (ANGOL4, ANGOL2) összehasonlítva, amelyek egymáshoz viszonyítva ugyanakkor nem mutatnak szignifikáns különbséget.

	(A)	(B)	Átlaglogit különbség (A-B)	Standard hiba	Sznif.	95% konfidencia intervallum	
						Alsó határ	Felső határ
Scheffe	ANGOL1	ANGOL2	2,060	0,521	0,001	0,737	3,383
		ANGOL4	1,598	0,550	0,021	0,203	2,993
	ANGOL2	ANGOL1	-2,060	0,521	0,001	-3,383	-0,737
		ANGOL4	-0,462	0,348	0,420	-1,344	0,420
	ANGOL4	ANGOL1	-1,598	0,550	0,021	-2,993	-0,203
		ANGOL2	0,462	0,348	0,420	-0,420	1,344

22. ábra: Varianciaanalízis post-hoc tesztek eredményei

Összegzés

A próbavizsga-projekt empirikus adatai alapján lehetőségünk volt a korábban megfogalmazott kutatási kérdésekre választ adni. Áttekintettük a különféle vizsgák leíró statisztikai adatait, s megállapítható volt, hogy azok helyenként igen eltérő képet mutatnak, ami potenciálisan a vizsgák eltérő működésére utalhat. Ezt a feltételezést több ponton megerősítette a vizsgák, illetve vizsgarészek közötti korrelációs mutatók vizsgálata, amely ugyancsak arra utalt, hogy helyenként lényegi különbség lehet az egyes vizsgák, illetve vizsgafeladatok között. A képet tovább árnyalta a tényleges vizsgázói teljesítmények összevetése, amely ugyancsak számos különbséget mutatott ki. Az empirikus vizsgálatok legmegbízhatóbb elemeként elvégzett IRT elemzés, illetve az itemnehézségi indexek összevetése pedig újfent megmutatta, hogy lényegi különbségek állnak fenn egyes esetekben.

Összességében megállapítható, hogy az írásbeli teszt egészét, illetve a részteszteket figyelembe véve szisztematikusan leginkább az ANGOL1 vizsga látszik különbözni a többi vizsgától. Úgy tűnik, az ANGOL1 vizsga szignifikánsan nehezebb a többinél. Természetesen ez nem jelenti azt, hogy emiatt más KER-szintet mérne, mint a többi vizsga, hiszen ezen adatok alapján ez a kérdés nem dönthető el. Arra pedig végképp nem kapunk választ ezekből az adatokból, hogy vajon az ANGOL1 távolabb vagy éppen közelebb áll-e a szignifikáns különbségek okán a tényleges KER-szinthez.

Ugyancsak figyelemre méltó az angol nyelvű olvasásteztek kapcsán tapasztalt látványos különbségek sora. A korrelációs együtthatók alacsony volta mindenképpen figyelmeztető jelzés, még ha a vizsgázói teljesítmények közvetlen összevetése nem is erősíti meg mindenben a különbségek létének feltételezését.

Hasonló a helyzet az angol nyelvű írásteztekkel is. Az alacsony korrelációk mellett több esetben szignifikáns különbségek látszanak a tényleges vizsgázói teljesítményekben.

Különös képet mutat a német nyelvű olvasásteztek elemzése, ahol a korrelációk számos esetben problematikus viszonyt jeleztek, a tényleges vizsgázói teljesítmények ugyanakkor kevésbé tűnnek riasztónak. A különbség magyarázata vélhetően a kutatás általános korlátaiban kereshető, amennyiben a próbavizsgázók létszáma – különösen a NÉMET2 esetében – viszonylag alacsony volt. Hasonló jelenség figyelhető meg a német nyelvű írásteztek kapcsán is.

Mindent egybevetve megállapítható, hogy a projekt eredményei számos igen lényeges összefüggésre, illetve potenciális problémára derítettek fényt. Az adatok alapján feltételezhető, hogy a vizsgák között helyenként érdemi tartalmi, illetve nehézségi szintet érintő különbségek állnak fenn.

Mindemellett a végkövetkeztetéseket körültekintően kell kezelnünk. Mint arra többször utaltunk, a próbavizsgázók viszonylag alacsony létszáma miatt nem állíthatjuk teljes bizonyossággal, hogy a feltárt anomáliák ténylegesen jelen vannak. Éppen ezért szükségesnek látszik a kutatás kiterjesztése és nagyobb mintán történő reprodukálása.

IRODALOM

- Dávid Gergely (2009): A KER az idegennyelv-tudás mérésében: Lehetőségek a nyelvvizsgák szintillesztésében. In: Frank Tibor – Károly Krisztina (szerk.): *Anglisztika és amerikanisztika. Magyar kutatók az ezredfordulón*. Budapest: Tinta Könyvkiadó, pp. 383–394.
- Dunlea, J. – T. Matsudaira (2009): Investigating the relationship between the EIKEN tests and the CEFR. In: Figueras, N. – J. Noijons (szerk.): *Linking to the CEFR levels: Research perspectives*. Arnhem: Cito, EALTA, pp. 103–109.
- Európa Tanács (2002): *Közös Európai Referenciakeret: Nyelvtanulás, nyelvtanítás, értékelés*. Pedagógus-továbbképzési Módszertani és Információs Központ.
- North, B. – P. Avermaet – N. Figueras – S. Takala – N. Verhelst (2006): *Nyelvvizsgák szintillesztése a Közös Európai Referenciakerethez. Kézikönyv*. Előkészítő kísérleti változat. Budapest: NYAT-NYAK.
- (2009): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual*. Letöltve a következő helyről: <http://www.coe.int/t/dg4/Linguistic/Source/Manual%20Revision%20-%20proofread%20-%20FINAL.pdf>
- Tannenbaum, R.J. – E.C. Wylie (2008): *Linking English-language test scores onto the Common European Framework of Reference: An application of standard-setting methodology*. Princeton, NJ: Educational Testing Service. Letöltve a következő helyről: http://www.ea.etsglobal.org/fileadmin/free_resources/ETS_Global_master/CEFR_complete_study.pdf
- Taylor, C. (2009): Relating the Trinity College London International ESOL examinations to the CEFR. In: Figueras, N. – J. Noijons (szerk.): *Linking to the CEFR levels: Research perspectives*. Arnhem: Cito, EALTA, pp. 69–74.