

A 2012. MÁJUS-JÚNIUSI ÉRETTSÉGI FELADATSOROK
MEGOLDÁSAINAK ITEMSZINTŰ RÖGZÍTÉSE,
A FELADATSOROK ÉS AZ EGYES FELADATOK
MÉRÉSMETODIKAI VIZSGÁLATA
A TÁMOP-3.1.8-09/1-2010-0004
AZONOSÍTÓ SZÁMÚ PROJEKT KERETÉBEN

NÉMET NYELV

A produktum elkészítésében közreműködött:

Petneki Katalin
Vígh Tibor

Készítette:
Expanzió Humán Tanácsadó Kft.

Budapest, 2013. szeptember

Tartalom

A 2012. MÁJUS-JÚNIUSI ÉRETTSÉGI FELADATSOROK MEGOLDÁSAINAK ITEMSZINTŰ RÖGZÍTÉSE, A FELADATSOROK ÉS AZ EGYES FELADATOK MÉRÉSMETODIKAI VIZSGÁLATA	1
Előszó.....	5
I. BEVEZETŐ	7
1. A projekt célja, lebonyolítása	8
2. A német nyelvi érettségi általános jellemzése	9
3. A 2012. középszintű német nyelvi érettségi vizsga a korábbi eredmények tükrében	14
4. Az emelt szintű német nyelvi érettségi.....	18
4.1. Az emelt szintű vizsga általános jellemzői	18
4.2. A 2012. évi emelt szintű német nyelvi érettségi vizsga a korábbi eredmények tükrében.....	20
5. A kutatás módszerei.....	24
5.1. A kutatás céljai, kérdései és hipotézisei.....	24
5.2. Az adatbázisok jellemzői.....	26
5.3. Az adatelemzés módszerei	28
Irodalom	31
II. A 2012. MÁJUSI NÉMET NYELV KÖZÉPSZINTŰ VIZSGA	
ELSŐ HÁROM FELADATSORÁNAK ELEMZÉSE	36
1. A 2012. májusi vizsga általános jellemzői.....	38
2. Készségenkénti (vizsgarész-összetevőnkénti) elemzés.....	40
2.1. Olvasott szöveg értéke (I. vizsgarész-összetevő).....	40
2.2. Nyelvhelyesség (II. vizsgarész-összetevő)	52
2.3. Hallott szöveg értéke (III. vizsgarész-összetevő).....	60
3. A három, itemalapú vizsgarész-összetevő elemzésének összefoglalása	69
Irodalom	72
III. A 2012. MÁJUSI NÉMET NYELV KÖZÉPSZINTŰ	
ÍRÁSKÉSZSÉG FÜZET/VIZSGARÉSZ (ÖSSZETEVŐ) ELEMZÉSE.....	73
1. Az íráskészség vizsgarész-összetevő felépítése.....	75
2. Az íráskészség vizsgarész-összetevő alapstatisztikai jellemzői.....	78
3. Feladatszintű elemzés.....	80
4. Az értékelési rendszer elemzése	85
4.1. Az értékelési szempontok jellemzői és teljesítésének nehézsége	85
4.2. A vizsgázók íráskészség-szintje és az értékelési szempontok nehézsége	86
4.3. Az értékelési skálák működése.....	89
5. Az íráskészség vizsgarész-összetevő elemzésének összefoglalása	92
Irodalom	95
IV. A 2012. MÁJUSI NÉMET NYELV EMELT SZINTŰ VIZSGA	
ELSŐ HÁROM FELADATSORÁNAK ELEMZÉSE	96
1. A 2012. májusi vizsga általános jellemzői.....	98
2. Készségenkénti (vizsgarészenkénti) elemzés	101
2.1. Olvasott szöveg értéke (I. vizsgarész)	101
2.2. Nyelvhelyesség (II. vizsgarész).....	109
2.3. Hallott szöveg értéke (III. vizsgarész)	123
3. A három, itembázisú vizsgarész összefoglalása	132

Irodalom	135
V. A 2012. MÁJUSI NÉMET NYELV EMELT SZINTŰ ÍRÁSKÉSZSÉG FÜZET/VIZSGARÉSZ (ÖSSZETEVŐ) ELEMZÉSE.....	137
1. Az íráskészség vizsgarész felépítése	139
2. Az íráskészség vizsgarész alapstatisztikai jellemzői	141
3. Feladatszintű elemzés.....	143
4. Az értékelési rendszer elemzése	148
4.1. Az értékelési szempontok jellemzői és teljesítésének nehézsége	148
4.2. A vizsgázók íráskészség-szintje és az értékelési szempontok nehézsége	150
4.3. Az értékelési skálák működése.....	152
5. Az íráskészség vizsgarész elemzésének összefoglalása.....	159
Irodalom	162
VI. A KÖZÉP- ÉS EMELT SZINTŰ NÉMET NYELV ÉRETTSÉGI MÉRÉSI-ÉRTÉKELÉSI RENDSZERÉNEK MÓDOSÍTÁSÁRA VONATKOZÓ JAVASLATOK	163
1. Mi az, ami bevált, ami a vizsga továbbfejlesztésekor ne változzon	164
2. Mi az, amit a feladatokkal és feladattípusokkal kapcsolatban érdemes újragondolni	166
3. Mi az, amit az értékelési rendszerben újra kellene gondolni	174
4. Mi az, amit az egész koncepción érdemes továbbgondolni és változtatni	177
Irodalom	179

A munka a címben megnevezett projekt keretében az Expanzió Kft megbízásából készült. A szerződés 1. és 10. pontja arra kötelezte a szerzőket, hogy a német érettségi továbbfejlesztésére vonatkozó javaslataikat tartalmazó tanulmányt három-három dokumentumban készítsék el. Az Oktatási Hivatal honlapján a projekt leírása az alábbi feladatokat fogalmazta meg: „A kutatás során vizsgatárgyanként és szintenként legalább 3-3 elemzésre kerül sor, az eredeti és a javaslatot követően módosított értékelési rendszerrel. (...) A kutatás eredményeit egy összegző tanulmány foglalja össze.”¹

Ennek megfelelően a két szerző közösen egy összefüggő, egységes tanulmányt hozott létre, összesen hat dokumentumba rendezve. Ezért felhívjuk a Felhasználó figyelmét arra, hogy a dokumentumok önállóan nehezen értelmezhetőek, mert összefüggő egységet alkotnak. A dokumentumok megegyeznek a tartalomjegyzék római számmal jelölt fő fejezeteivel. Tájékoztatásul és segítségül az alábbi összefoglalót nyújtjuk:

A dokumentum neve	A dokumentum fő témája, tartalma	Terjedelem
nemet_dok_I	Bevezető (a projektről, a vizsgáról, kutatási módszerekről)	29 oldal
nemet_dok_II	A 2012. májusi német középszintű vizsga első három feladatsorának elemzése	38 oldal
nemet_dok_III	A 2012. májusi német nyelv középszintű íráskészség elemzése	23 oldal
nemet_dok_IV	A 2012. májusi német nyelv emelt szintű vizsga első három feladatsorának elemzése	41 oldal
nemet_dok_V	A 2012. májusi német nyelv emelt szintű íráskészség elemzése	27 oldal
nemet_dok_VI	A közép- és emelt szintű német nyelv érettségi mérési-értékelési rendszerének módosítására vonatkozó javaslatok	21 oldal

¹ http://www.oktatas.hu/koznevelés/projektek/tamop318_minosegfejl/projekthirek/tamop318_tanuloi_teljesitmenymeres?itemNo=2
(letöltve: 2013. 06. 25.)

PETNEKI KATALIN, VÍGH TIBOR

BEVEZETŐ

1. A projekt célja, lebonyolítása

Az Oktatási Hivatal által gondozott „Átfogó fejlesztés a közoktatásban című” TÁMOP-3.1.8. kiemelt projekt keretében az Expanzió Kft. az Oktatási Hivatal megbízásából „A 2012. május-júniusi érettségi feladatsorok megoldásainak itemszintű rögzítése, a feladatsorok és az egyes feladatok mérésmethodikai vizsgálata a TÁMOP-3.1.8-09/1-2010-0004 azonosító számú projektben” kért fel szakértőként a német nyelv vizsgatárgy elemzésére. A projekt célkitűzései: (1) a közép- és emelt szintű írásbeli vizsgarészek feladatainak mérésmethodikai vizsgálata és értékelése, (2) továbbfejlesztési javaslatok megfogalmazása az értékelési rendszer módosítására vonatkozóan, valamint (3) ezeknek az egyes vizsgázók teljesítményére és a teljes vizsga teljesítménystruktúrájára gyakorolt hatásának elemzése.

A bevezető további részében először bemutatjuk a német nyelvi érettségi általános jellemzőit, majd a közép- és az emelt szintű német nyelvi érettségien elért eredményeket. Végül a kutatás módszereit részletezzük.

2. A német nyelvi érettségi általános jellemzése

A 2005-ben bevezetett érettségi vizsga egységes, azaz a korábbiakkal ellentétben minden iskolatípusra egyformán érvényes, kétféle mérési céllal rendelkezik: középszinten azt méri, hogy a vizsgázó a középiskolai tanulmányait eredményesen lezárta-e, az emelt szint pedig arra fókuszál, hogy a vizsgázó alkalmas-e a tárgy továbbtanulására felsőfokú szinten. Harmadik fontos és általános jellemzője a vizsgának, hogy standardizált, azaz egységes útmutató alapján történik a teljesítmények értékelése (Horváth és Lukács, 2006: 20-23).

„Az idegen nyelvi érettségi vizsgájában [a korábbihoz képest] új elem, hogy a vizsga mind a négy alapvető kommunikatív készséget külön vizsgarésszel méri, a feladatok autentikus szövegekre épülnek, a feladatsorok teljesen egynyelvűek, és számtalan új feladattípust tartalmaznak. A vizsga értékelési rendszere is megváltozott. Az eddigi gyakorlattal ellentétben a feladatokhoz megoldási kulcs készül, és ezt a javító tanárnak követnie kell. A produktív készségeket mérő vizsgarészeknél (írás, beszéd) értékelési skálákat kell alkalmazni, azaz az értékelőnek különböző szempontok alapján kell különböző minőségi szintekbe besorolnia a létrejött szöveget” (Horváth és Lukács, 2006: 24). A továbbiakban ezeket a jellemzőket részletezzük.

Az idegen nyelvi érettségi célja a kommunikatív kompetencia mérése. Ezeket a kompetenciákat a Közös Európai Referenciakeret² (2002) az alábbiak szerint határozza meg:

0.1. táblázat
Kommunikatív nyelvi tevékenységek a Közös Európai Referenciakeretben
(KER, 2002: 275; Einhorn, 2006)

Receptív tevékenységek	Interaktív tevékenységek	Produktív tevékenységek	Közvetítés a nyelvek között
Hallott szöveg értése	Interakció beszédben	szóban	szóban
Olvasott szöveg értése	Interakció írásban	írásban	írásban

Az érettségi vizsga a fenti készségek közül nem méri a nyelvek közötti közvetítést, mert ehhez már nem elegendő az elementáris nyelvtudás. Ezért elsősorban a négy alapkészségre irányul a vizsga: a beszédértésre, olvasott szövegértésre, íráskészségre és beszédkészségre. A vizsgafejlesztés során – főleg a tanárok elvárásai miatt – bekerült egy olyan vizsgarész/vizsgarész-összetevő, ami a nyelvhelyesség tudatos használatát elsősorban a tanárok elvárásaihoz igazodva célzottan méri.

A német nyelvi érettségi vizsga követelményei megegyeznek az Idegen nyelvi érettségi vizsga általános követelményrendszerével, amely az Európa Tanács idegennyelv-

² A későbbiekben használni fogjuk a Közös európai referenciakeret megnevezésére a KER rövidítést is.

oktatással kapcsolatos ajánlásai alapján készült. Az idegen nyelvi érettségi vizsga szintmeghatározásai igazodnak az Európa Tanács skálájához. A vizsga középszintje átfogja az A2–B1 szinteket, az emelt szint pedig a B2 szintnek felel meg (100/1997 korm. rend, 40/2002 OM rendelet).

0.2. táblázat:

Az érettségi vizsgaszintek viszonya a Közös Európai Referenciakeret szintjeihez

Európa Tanács	Érettségi vizsga
Mesterszint	
Haladó szint	
Középszint	<i>Emelt szint</i>
Küszöbszint	<i>Középszint</i>
Alapszint	
Minimumszint	

0.3. táblázat:

Az egyes szintek globális leírása a vizsgakövetelményekben (KER; 2002: 33)

Szint	Leírás
B2	Megérti a változatos, konkrét vagy elvont témájú szövegek fő gondolatmenetét, követni tudja a hosszabb, összetettebb érveléseket is. Folyamatos és természetes módon tud a célnyelven interakciót folytatni. Világos és részletes szöveget tud létrehozni különböző témákról. Véleményét indokolni tudja, részletezni tudja a különböző lehetőségekből adódó előnyöket és hátrányokat.
B1	Megérti a fontosabb információkat olyan egyszerű, hétköznapi szövegekben, amelyek gyakori élethelyzetekhez kapcsolódnak (pl. iskola, szabadidő, munka). Képes külföldiekkel kommunikálni mindennapi helyzetekben. Egyszerű, összefüggő szöveget tud alkotni olyan témákban, amelyeket ismer, vagy amelyek az érdeklődési körébe tartoznak. Be tud számolni eseményekről, élményeiről, érzelmeiről és törekvéseiről. Rövid magyarázatot tud fűzni eseményekhez, jelenségekhez, indokolni tud különböző álláspontokat és terveket.
A2	Megért olyan mondatokat és rövid szövegeket, amelyek az őt közvetlenül érintő területekhez kapcsolódnak. Tud kommunikálni olyan mindennapi, begyakorolt helyzetekben, amelyekben egyszerű és közvetlen információcserére van szükség. Tud egyszerű nyelvi eszközöket használva beszélni önmagáról, családjáról és szűkebb környezetéről.

A szintekhez tartozó követelményeket mind a Közös Európai Referenciakeret, mind a vizsgaszabályzat tartalmazza. A 0.3. táblázat az Európa Tanács A2, B1, B2 szintjeinek általános leírását tartalmazza:

A vizsgaleírás részletesen is kifejti, hogy mit jelentenek a szintleírások közép- és emelt szinten az egyes készségekre lebontva, mire legyen képes a vizsgázó. A leírt szintektől függ továbbá, hogy a vizsgán milyen típusú és terjedelmű szövegek szerepelhetnek, milyen feladattípusok használhatóak. A szintek az értékelés módját is meghatározzák, de

nem befolyásolják az értékelési alapelveket. Ezek az alábbiakban foglalhatók össze (Einhorn, 2007b):

- Az egyes vizsgarészek/vizsgarész-összetevők értékelése egymástól független.
- Minden vizsgarész/vizsgarész-összetevőben csak a célzottan mért készséget értékeljük.
- Az értékelési útmutató tartalmazza a lehetséges válaszokat.
- Az íráskészség és a beszéd-készség értékeléséhez analitikus skálákat alkalmazunk.

Ma már magától értetődőnek tartjuk a készségmérő nyelvvizsgát. Mi volt ebben a vizsgában akkoriban újszerű? A korábbi érettségi vizsgafeladatokról nem lehet elmondani, hogy a célnyelvi készségeket mérték volna, nem volt hozzá sem egységes vizsgaleírás, sem egységes értékelési útmutató, a vizsgáztatók leginkább saját tapasztalataikra hagyatkoztak. Az egyik legfőbb újdonság az alaposan kidolgozott vizsgadokumentáció volt. A részletes vizsgaleírást hat élő nyelvre dolgozták ki a vizsgafejlesztők (Einhorn, 2007b: 74). Míg a NAT már első változatában a célnyelvi kommunikatív kompetencia elérését tűzte ki célul (Petneki, 2007: 56), az érettségi vizsgán az új vizsga bevezetéséig ez a cél háttérbe szorult, hisz az eredményes vizsgához inkább grammatikai és lexikai tudásra volt szükség. Ezért volt nagy jelentősége annak, „hogy az idegen nyelvekből készségmérő típusú vizsga lépett a korábbi, tradicionálisabb helyébe, és mivel az új rendszerben mindegyik kommunikatív készség mérésére sor kerül (olvasás, hallás, írás, beszéd), a középiskolai idegennyelv-tanításban megkerülhetetlenül kötelezővé vált a készségfejlesztésen alapuló nyelvtanítás” (Einhorn, 2007b: 74).

A 2005-ben bevezetett német nyelv érettségi alapvetően új szemléleten alapul, amit így lehet összefoglalni (Einhorn, 2007b: 74–75; Petneki, 2007: 119–120):

- a) **Autentikus szövegek, ill. helyzetek:** A vizsgán autentikus szövegek alapján készülnek a feladatok, a produktív feladatoknál pedig inputként szerepelnek autentikus szövegek (pl. hirdetés), és a feladatmeghatározás igyekszik egy lehetséges autentikus helyzetet szimulálni. Ez azt jelenti, hogy a feladat meghatározza, hogy a vizsgázó milyen szerepben, milyen helyzetben kivel beszél (ill. kinek ír) és mit kell ennek során elérnie.
- b) **Releváns tartalmak:** Mind a receptív, mind a produktív feladatok csak olyan témákat érintenek, amiket egy érettségiző fiatalnak tudnia ill. ismernie kell. Nem tartalmaz semmi olyat, ami a korosztálynak nem megfelelő. Nem kérdez rá tárgyi tudásra, csak a nyelvi teljesítményt méri.
- c) **Feladatok:** Az egyes vizsgarészek/vizsgarész-összetevők több (legalább két) feladtból állnak. A vizsgaleírásban található bármelyik feladattípus előfordulhat. Ez azt jelenti, hogy a vizsgára való felkészítés nem működhet úgy, mint eddig, azaz nem

lehet egy kaptafára kialakított tesztfüzetekkel készülni³. Ha a tanuló több éven keresztül kommunikatív szemléletű nyelvoktatásban részesül, azaz a tanulás során változatos feladatokkal találkozik, akkor az érettségi idejére már elsajátította azokat a technikákat, amikkel ilyen feladatokat meg lehet oldani.

- d) **Értékelés:** A vizsgán elért eredményeket egységes javítási és értékelési útmutató alapján mérik. A produktív feladatoknál részletesen kidolgozott analitikus értékelési skála áll a tanárok rendelkezésére. Ezzel biztosítja a vizsga a teljesítmények összehasonlíthatóságát.

A továbbiakban a vizsga mérési-értékelési szemléletéből adódóan összegezzük a jelenlegi érettségi alapvető jellemzőit (Petneki, 2007: 121–122):

- a) **Kétszintű:** a szintek a Közös Európai Referenciakerethez igazodnak
- b) **Egynyelvű:** mind a feladat-meghatározásokat, mind a feladatokat a célnyelven fogalmazták meg.
- c) **Kommunikatív:** valós nyelvhasználatot mér, a feladatok autentikus forrásokra épülnek:
- kommunikatív szándékok és helyzetek,
 - „új” készség a hallott szöveg értése,
 - autentikus szövegek,
 - változatos, szövegközpontú feladattípusok,
 - életkornak megfelelő témák.
- d) **A nyelvhelyesség szerepe:**
- a vizsga nem használ nyelvtani terminológiát,
 - a nyelvtan középszinten nem játszik döntő szerepet,
 - emelt szinten a nyelvhelyesség vizsgarész egyenrangú a többi vizsgarésszel.
- e) **Standardizált:**
- A vizsgakövetelmények egységesen érvényesek, leírásuk kellően részletes.
 - A javítási és az értékelési útmutató kellően részletes, támpontot ad az objektív és árnyalt értékeléshez, valamint lehetővé teszi a teljesítmények jobb összehasonlítását.
- f) Középszinten a szinten belül is differenciál: az írásbeli vizsgarész-összetevőben a **lépcsőzetesség elve** érvényesül azaz az itemek és a feladatok nehezednek (40/2002. OM rendelet).

³ Sajnos a gyakorlat erre rácsófol. Azóta számos olyan „gyakorló” tesztkönyv készült, amivel továbbra is feladatközpontúan lehet a diákokat a vizsgára felkészíteni – a kommunikatív kompetencia fejlesztése helyett.

A vizsga mind a négy kommunikatív készséget célzottan méri. Emellett középszinten – kisebb súllyal – méri a grammatikai és lexikai kompetenciát is egy külön vizsgarész-összetevőben. Mindegyik készséget két szinten méri: az A2 szint az elementáris nyelvtudást mutatja, míg a B1 szint már az önálló nyelvhasználó szintje – jóllehet, ez is még elég alacsony nyelvi szintet jelent. Emelt szinten minden vizsgarész azonos a súllyal szerepel.

0.4. táblázat:

**A 2005-ben bevezetett német nyelvi érettségi vizsga szerkezete
(40/2002 OM rendelet, Vizsgaleírás: 1. és 8)**

Vizsgarész/ vizsgarész-összetevő	Készség	Középszint		Emelt szint	
		Arány (%)	Vizsgapont	Arány (%)	Vizsgapont
Írásbeli	Olvasott szövegértés	22	33	20	30
	Nyelvhelyesség	12	18	20	30
	Hallott szövegértés	22	33	20	30
	Íráskészség	22	33	20	30
Szóbeli	Beszédkésztség	22	33	20	30
Összesen:		100	150	100	150

A vizsga bevezetését követően még néhány évig volt mód arra, hogy a követelmények megvalósulását elemzésekkel kövessük (lásd. pl. Einhorn 2006, 2007a, 2007b, 2008a, Vígh 2010, 2012), de azóta nem volt mód a vizsga bevalásának behatóbb elemzésére. Jelen munka ezt kívánja pótolni.

3. A 2012. középszintű német nyelvi érettségi vizsga a korábbi eredmények tükrében

A 2012. májusi írásbeli feladatsor – a vizsgaleírásnak megfelelően – négy vizsgarész-összetevőből áll: olvasott szövegértés, nyelvhelyesség, hallott szövegértés és íráskészség. Az érettségi ötödik része, a beszéd-készség mérése szóbeli vizsgán történik.

Az OH honlapján elérhető összesített adatok (Érettségi 2012 május-június, 2012) szerint német nyelvből középszinten 22280-an vizsgáztak a 2012. évi május-júniusi érettségi időszakban, amely az angolhoz képest (N=54945) jóval kevesebb. A társadalom számára fontos visszajelzés a vizsgákról, hogy mennyien tettek panaszt. Az 0.5. táblázatból az derül ki, hogy a legtöbb beadványt 2009-ben adták be, a legkevesebbet pedig 2008-ban. A többi tárggyal összevetve ez a legjobb százalékos arány: az összes többi tárgyból ennél több panasz, ill. beadvány érkezett a vizsgált időszakban.

0.5. táblázat:

Tanulói észrevételek alakulása német nyelvből

Beadványok	Vizsgaév				
	2008	2009	2010	2011	2012
száma (db)	53	104	71	101	93
aránya (%)	3,00	5,11	3,20	4,13	3,57

A következőkben a vizsgán elért eredményekkel foglalkozunk az összes vizsgázóra vonatkoztatva. A 0.6. táblázat azt mutatja, hogy az új érettségi bevezetése előtti időszakhoz képest ugyan némileg csökkent az átlagosztályzat, de megállapítható, hogy aránylag egyenletes szinten, a 3,5 átlag körül mozognak az eredmények. Kissé kiugró azonban a 2011. év 3,23 átlaga – nyilvánvalóan itt meg kell vizsgálni ennek okát.

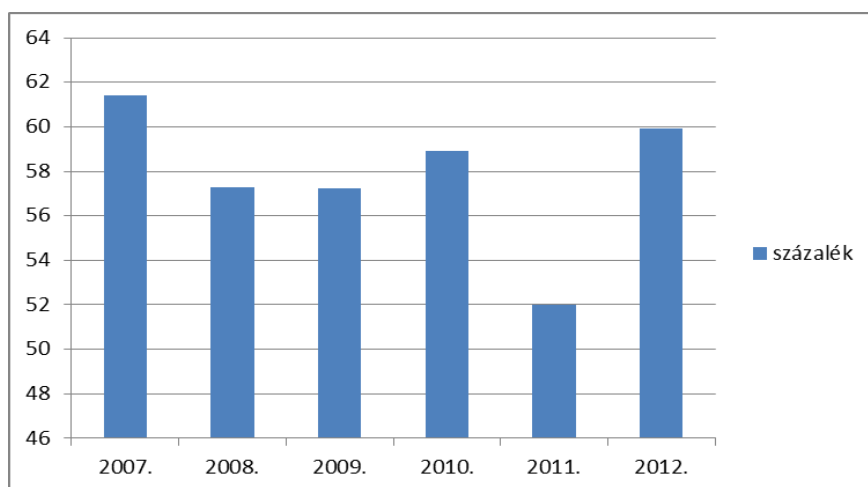
0.6. táblázat:

Tanulói átlagosztályzatok alakulása középszinten

2001-03 átlag	2008	2009	2010	2011	2012
3,74	3,46	3,46	3,53	3,23	3,57

A 0.1. ábra hasonló képet nyújt, mint a 0.6. táblázat. Itt még jobban látszik a 2011. évi érettségi vizsgán elért gyengébb átlagteljesítmény, majd a 2012-es év viszonylag magasabb átlageredménye. A 2007-2012 közötti időszakban a teljes vizsga megoldottságában megjelenő különbségek az azonos készséget mérő feladatsorok feltehetően eltérő nehézségéből, különböző belső szerkezetéből adódik, valamint abból, hogy a vizsgára való felkészülés során a diákok megismerik a vizsgán várható feladattípusokat, ami befolyásolhatja teljesítményüket. Az Oktatási Hivatal honlapján található összesített adatok alapján tehát meg lehet állapítani, hogy a 2012. május-júniusi vizsga eredményei nem térnek el

jelentős mértékben a korábbi átlagoktól, az ezt megelőző évhez viszonyítva pedig jobb volt.



0.1. ábra:

A középszintű német nyelvi érettségi átlagának alakulása százalékban

A 0.7. táblázat azt mutatja, hogy miképp oszlik meg a százalékos teljesítmény a képzési forma függvényében. Ez megerősíti azt az általános vélekedést, hogy a nappali tagozatos képzésben a legjobb teljesítményt a gimnáziumi érettségizők hozzák. A felnőttoktatásban nagyobb a különbség: a középszintű érettségit gyenge eredménnyel teszik le.

0.7. táblázat:

Német nyelvi érettségi átlagok eloszlása középszinten %-ban vizsgázói rétegek szerint

Év	Gimnázium	Szakközépiskola	Felnőttoktatás	Tanulói jv. ⁴ kívül
2010	73,10	50,90	42,21	58,51
2011	66,83	43,93	36,58	51,79
2012	73,63	52,85	43,72	54,00

A továbbiakban a nappali tagozatosok körében az egyes vizsgarész-összetevők megoldottságát mutatjuk be a korábbi vizsgaeredmények tükrében a május-júniusi vizsgaidőszakok alapján⁵. Azért ezt a tanulói csoportot választottuk ki, mert jelen projektben célunk, hogy erre a vizsgázói csoportra vonatkoztatva legyen reprezentatív a minta. Mindezek mellett pedig a vizsgázók teljes összetétele 2007 és 2012 között jelentősen megváltozott, amelynek oka, hogy a felnőttoktatásban magasabb arányban vesznek részt a vizsgázók. Ugyanakkor a nappali képzésben végzetek száma és aránya nem változott jelentősen, ezért az egyes feladatsorok megoldottsága alapján a tendenciák jobban összehasonlíthatók. Az egyes vizsgarész-összetevők teljesíthetőségét a 0.8. táblázat tartalmazza, melyek-

⁴ jv.=jogviszony

⁵ A fejezet további része Vígh (2013: 22–25. o.) cikke alapján készült.

ben a vizsgapontszámot százalékos skálára transzformáltuk, megadtuk a teljesítményát-
lagot és a szórást. Közvetlenül összehasonlíthatjuk az egyes vizsgaéveken belül az ered-
ményeket, mivel a tesztek ugyanazon vizsgázók oldották meg, így képet kaphatunk a
vizsgarész-összetevőkben mért készségek fejlettségéről. Közvetetten vizsgálhatók az
egyed-évek közötti teljesítmények, mivel az adatbázisokat nem köti össze közös személy
vagy item, de minden feladatsor a vizsgakövetelményekre épül, a feladatírás alapelve
azonos, a nappali képzésben végzett vizsgázói populáció nagy, és összetétele nem válto-
zott jelentősen.

A vizsgázók átlagteljesítményét két, egymással összefüggő tényező határozza meg:
a diákok nyelvtudásszintje és a feladatok nehézsége. Ez utóbbi főleg az átlagos képes-
ségi szintű tanulók teljesítményét befolyásolja. Az egyes vizsgaévekben az átlagokban
megjelenő ingadozások eltérő nehézségű feladatsorokra utalnak, a teljesítménynöveke-
désben a felkészítés hatása jelenhet meg, a csökkenő teljesítményátlagok azt jelezhetik,
hogy a vizsgafeladatokat összeállító tantárgyi bizottság nehezíti a feladatsorokat. Ugya-
nakkor az okok pontos megállapítására feladat- és itemszintű elemzésekre van szükség,
ami a jelenlegi projekt egyik célja a 2012. évi érettségire vonatkozóan. Először vizsgaévek-
ken belül elemezzük az eredményeket, majd az átlagok változásai alapján vonunk le kö-
vetkeztetéseket.

0.8. táblázat:

A német nyelv középszintű vizsgák eredményei a nappali tagozaton végzett diákok körében
(százalékpontban megadott átlagok, zárójelben a szórások) (Vígh, 2013: 24)

Vizsgarész-összetevő	Vizsgaév					
	2007	2008	2009	2010	2011	2012
Olvasott szöveg értéke	74 (20)	63 (21)	61 (23)	58 (24)	69 (19)	62 (19)
Nyelvhelyesség	48 (18)	53 (20)	57 (21)	52 (24)	58 (20)	61 (19)
Hallott szöveg értéke	71 (20)	55 (23)	51 (18)	66 (16)	26 (21)	67 (17)
Íráskészség	50 (33)	54 (33)	56 (33)	60 (32)	56 (34)	57 (33)
Beszéd-készség	64 (27)	66 (27)	67 (27)	67 (28)	68 (27)	66 (28)
Teljes vizsga	62 (21)	59 (22)	58 (22)	61 (21)	51 (22)	62 (21)

Az olvasott szöveg értésénél megfigyelhető ingadozás: 2007 és 2010 között e feladat-
sor teljesítettsége csökkent, míg 2011-ben a legkönnyebb vizsgarész-összetevő volt. Mé-
résmetodikai szempontból a legproblémásabb vizsgarész-összetevő a hallott szöveg érté-
se. E feladatsor a többihez képest 2012-ben volt a legkönnyebb, a teljesítettsége 2007-
ben és 2010-ben a második legmagasabb, 2008-ban a harmadik, míg 2009-ben és 2011-
ben a többi vizsgarész-összetevőhöz képest ezt teljesítették a legkevésbé az érettségizők.
Különösen a 2011-es feladatsor volt nehéz: kiugróan alacsony volt az átlag. E két készség
esetében tehát a feladatsorok átlagos megoldottsága ingadozik. A középszintű érettségi
célja a referenciakeret A2–B1 szintjének mérése, amit úgy valósítanak meg, hogy az írás-
beli feladatsorokon belül a lépcsőzetesség elve érvényesül. Az eredmények viszont arra

utalnak, hogy a feladatsorok nehézsége nem állandó. Ennek oka, hogy ezek a tesztek kevés, általában 18–25 itemet tartalmaznak, ezért a mérni kívánt nyelvi szintben minimális arányeltolódás eltérő nehézségű feladatsort eredményez (Einhorn 2007b: 83).

A nyelvhelyesség vizsgarész-összetevőnél a 2010-es vizsgaévet kivéve folyamatosan javultak az eredmények, vagyis itt a fejlesztés hatása jelenhet meg. A szórás alapján e vizsgarész-összetevőnél kisebbek a tanulók teljesítményében megjelenő különbségek.

A szórás alapján minden évben heterogén a populáció a produktív készségekhez. Ez különösen az íráskészségre jellemző. 2005 és 2010 között folyamatosan javultak az átlagteljesítmények (a korábbi vizsgán elért eredményekről lásd Einhorn 2008a). Ebből arra következtetünk, hogy az idegen nyelvi érettségi bevezetését követően volt jellemző a vizsga közvetlenebb hatása e készség esetében a felkészülési folyamatra, melyben a diákok megismerik az írásművek megírásának folyamatát, az értékelési rendszert. Hazai vizsgálatok (vö. Nikolov, Pércsich és Szabó 2000; Józsa és Nikolov 2005) azt mutatták, hogy a feladatok előzetes ismertsége, a nyelvórákon való gyakorlásuk mértéke és nehézségének megítélése erőteljesebben befolyásolja a teljesítményeket, mint az olvasásnál.

A legkönnyebben teljesíthető vizsgarész a beszéd-készség, amelynél az elért átlag minden évben szignifikánsan ($p < 0,001$) magasabb az írásbeli vizsgarésznél és a teljes vizsgánál. E vizsgarész mérési-értékelési eljárásai jelentős mértékben különböznek a többitől: a feladatsorokat a tanárok helyi szinten állítják össze a vizsgakövetelmények alapján, a szóbeli produktumokat a tanárok értékelik központi analitikus értékelési skálákkal. A vizsgán elért eredményt számos tényező befolyásolhatja: a vizsgázók beszéd-készség-szintje, az értékelési szempontok és skálák jellemzői, az értékelők szigorúsága, a feladatok minősége. E tényezők közül az OH adatbázisaiban az értékelés eredményéről szerezhetünk információkat. A 2007–2009. évi érettségien az adatbázisokban egyértelműen azonosítható hibát találunk: ennél a vizsgarésznél több esetben fél pont szerepel, amely sem az értékelési skálák alkalmazásakor, sem a vizsgapontok átszámításakor nem keletkezhetne. Ez alapján feltételezzük a korábbi vizsgára jellemző értékelési gyakorlatot, melyben a vizsgázók írásbeli érettségien elért eredményét korrigálták a szóbeli pontszámok kialakításakor, amely így nem hozott összemérhető eredményeket (Einhorn 2008b: 211).

4. Az emelt szintű német nyelvi érettségi

4.1. Az emelt szintű vizsga általános jellemzői

A vizsgafejlesztés alapvető és általános kiindulási pontja az volt, hogy legyen a vizsga egynyelvű, standardizált és kétszintű, ezért azt kellett megvizsgálni, hogy a két különböző szint milyen tartalmi különbségeket jelent. A vizsgafejlesztés kezdetekor (1996) ugyan létezett az akkori Európai Unió alapvető szintleírása (Ek és Trim, 1991), de a Közös Európai Referenciakeret angol nyelven csak 2001-ben jelent meg. Tehát a vizsgafejlesztés első fázisában csak hipotézisekre lehetett támaszkodni a szintet illetően. A legfontosabb kérdés az volt, hogy kinek szól az emelt szintű vizsga? A célcsoportnak milyen szintű nyelvtudásra van szüksége, csak ez után kezdődhetett meg a vizsgafejlesztés. A hipotézis a következő volt:

Emelt szintű vizsgát azok tesznek (Einhorn, 2007b: 81), akik

- az adott nyelvet a felsőoktatásban szeretnék továbbtanulni (német nyelv esetében elsősorban itt a germanisztika szak volt az irányadó)
- egyéb (nem nyelvszakos) felsőoktatásba kívánnak jelentkezni.

Ez a várakozás a vizsga bevezetésekor nem valósult meg. Kizárólag a nyelvszakra jelentkezőknél várta el a felsőoktatás az emelt szintű érettségi vizsgát. Ez a vizsga ugyanis átvette a felvételi helyét, így ennek egyértelműen tudásszintmérő vizsgának kell lennie.

0.9. táblázat:

A lezáró típusú és a tudásszintmérő vizsga különbségei (Einhorn, 2007b: 89)

Szempon	Középiskolát lezáró vizsga	Tudásszintmérő vizsga
Tartalmi, szintbeli igazodás	A középiskola lehetőségeihez igazodik	Standardokhoz vagy a felhasználók igényeihez
Vizsgafeladatok készítése	Belső (iskolai szint)	Külső (központi)
Értékelés	Belső (saját tanár)	Külső (képzett értékelő)
Vizsgaszintek	A lehető legkevesebb szint	A lehető legtöbb szint
Vizsgalehetőség	A tanulmányok lezárásakor	Bármikor (korábban is)

Az emelt szintű vizsga magában foglalja a lezáró és a tudásszintmérő vizsgát is, jóllehet, a kettő egymásnak sok tekintetben ellentmond (Einhorn, 2007b: 89). Korábban is volt ilyen: 2005 előtt azok a diákok, akik felsőfokú oktatásba jelentkeztek, közös írásbeli érettségi-felvételi vizsgát tettek a választott szaktárgyból – így német nyelvből is. Ez azt feltételezi, hogy mindegyik felsőoktatási intézménynek azonosak voltak az elvárásaik. De 1997-ben például a német nyelv közös érettségi-felvételi feladatsort 23 intézmény használta: nemcsak bölcsész- és tanárszakok, hanem kereskedelmi, pénzügyi, számviteli, idegenforgalmi és mérnöki szakok is (Einhorn, 1998: 128). Ez azt jelenti, hogy a vizsga ilyen formá-

ja nem vehette figyelembe az adott szaktárgy nyelvi igényeit. Az új vizsga fejlesztése során is elsősorban az általános, a Közös Európai Referenciakeretben leírt B2 szintű nyelvtudás mérése volt a cél. De hogy igazán fel lehessen mérni a különbséget a korábbi közös érettségi-felvételi feladatsorral, érdemes a kettőt röviden összehasonlítani.

0.10. táblázat:

Az 1995. évi közös érettségi-felvételi feladatsor szerkezete (Einhorn, 1998: 132)

Feladat	Feladattípus	Mérési cél	Pont / %
A) Szövegértést tesztelő feladatok			
A/1	Hozzárendelés	Globális szövegértés	5
A/2	Igaz-hamis	Részletes szövegértés	20
A/3	Adott szöveghely értelmezése	Részletes szövegértés	6
A/4	Feleletválasztó	Lexikai kompetencia, részletes szövegértés	9
Összesen			40
B) Nyelvtani, lexikai ismereteket és koherens szövegalkotási kompetenciát tesztelő feladatok			
B/1	Kiegészítés	Nyelvtani kompetencia	5
B/2	Átalakítás	Nyelvtani és lexikai kompetencia	5
B/3	C-teszt	koherens szövegalkotási kompetencia	10
Összesen			20
C) Fordítási feladatok			
C/1	Fordítás német nyelvre	Közvetítés	20
C/2	Fordítás magyarra	Közvetítés	20
Összesen			40
Mindösszesen:			100

A feladatsorról megállapítható (0.10. táblázat), hogy a kommunikatív készségek közül csak az olvasott szövegértést és a közvetítést mérte, a többi kommunikatív készséget nem. Erre a két készségre egyenként 40-40 % jutott. A maradék 20% nyelvtani és lexikai ismeretekre kérdezett. A vizsgán az A és B részben döntően zárt feladattípusok fordultak elő. Ez a feladatsor már egy újítás volt a korábbihoz képest, amely egy feladatlapot döntően nyelvtanra és lexikára vonatkozó feladatokkal (40%) és két fordítást tartalmazott. Itt a (mindkét irányban egy-egy), amiből a magyarra fordítás kevesebbet ért (20%), mint a német nyelvre (40%) (lásd. Einhorn, 1998: 129–131). 1995 előtt tehát kizárólag a fordítás szerepelt, mint készség, de a kommunikatív kompetencia semmilyen szerepet nem játszott. Ezért jelentett rendkívüli nagy változást a 2005-ben bevezetett emelt szintű érettségi. Ebben ugyanis mind a négy kommunikatív készség egyenlő súllyal szerepel (lásd. 0.4. táblázat). Emellett – főleg a tradicionális oktatási módszerek máig ható működése miatt – van egy, a nyelvtani és lexikai ismereteket is mérő rész. A vizsga során mindegyik vizsgarész egyforma súllyal szerepel. Az eddigi tapasztalatok azt mutatták, hogy az emelt szintű vizsgák általában jó eredménnyel zárulnak. Nincs azonban visszajelzés a felsőoktatásból, hogy számukra mennyire vált be a felvételit helyettesítő funkció. Márpedig ha nincs vissza-

jelzés, akkor a változtatáskor is újra csak tapogatózás lesz. Fontos lenne, hogy a felsőoktatás is adjon értékelhető visszajelzést a vizsga működéséről.

4.2. A 2012. évi emelt szintű német nyelvi érettségi vizsga a korábbi eredmények tükrében

A 2012. májusi emelt szintű írásbeli feladatsor – a vizsgaleírásnak megfelelően – négy vizsgarészből áll (lásd. 0.4. táblázat). A beszédképesség mérése szóbeli vizsgán történik.

Német nyelvből az elemzett vizsgaidőszakban 2291-en tettek emelt szintű vizsgát, míg angolból nemcsak középszinten, hanem emelt szinten is jóval nagyobb számban vizsgáztak (N=7117) (Érettségi 2012 május-június, 2012). Ugyanakkor minden évben kb. 100 fővel nő a vizsgázók száma (0.11. táblázat), amit vélhetően a szintemelő és a javítóvizsgára jelentkezők számának emelkedése okoz. A középszinthez képest, emelt szinten a jelentkezők jóval alacsonyabb száma azzal magyarázható, hogy az emelt szintű érettségit a felsőoktatásban leginkább a nyelvszakok kérik, más intézmények az államilag elismert nyelvvizsgákat preferálják. Pedig az emelt szintű vizsga bevezetésének egyik fő célja az volt, hogy kiváltsa a közoktatáson kívüli nyelvvizsgázás lehetőségét.

Az emelt szintű vizsga megoldottságáról korábbi adatokat a 2005-2007. évi vizsgák elemzését vizsgáló tanulmányból (Einhorn, 2008b), a többről pedig az Oktatási Hivatal oldalán közzétett adatokból⁶ nyerhetünk.

0.11. táblázat:

A vizsgázók száma és eredménye 2005–2012. évi május-júniusi emelt szintű német nyelv érettségien (Einhorn, 2008b: 208; Vígh, 2010: 62; Érettségi 2012 május-június, 2012)

Jellemzők	Vizsgaév							
	2005	2006	2007	2008	2009	2010	2011	2012
Vizsgázók száma	1470 ⁷	1645	1812	1679	1942	2031	2135	2219
Átlag (%)	70,52	69,39	69,45	73,44	74,96	71,66	66,79	67,63

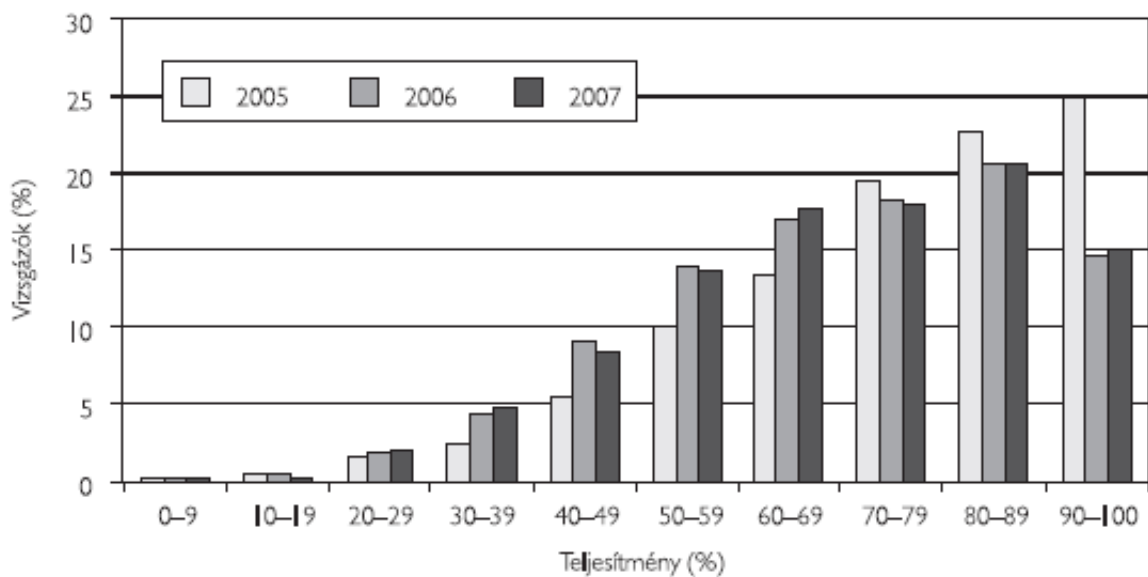
A fenti adatok szerint az emelt szintű érettségi vizsgát 2005–2007 között 70% körüli átlagon teljesítették a vizsgázók, 2008 és 2009-ben némiképp e fölött, míg 2011 és 2012-ben alacsonyabb átlageredményt értek el a vizsgázók. A 2012. május-júniusi vizsga eredményei tehát átlagosan nem térnek el jelentős mértékben a korábbi átlagoktól. Ahhoz, hogy ezt a vizsgát államilag elismert középfokú nyelvvizsga-eredményként lehessen elfo-

⁶ http://www.oktatas.hu/koznevelas/erettsegi/prezentaciok_tanulmanyok

⁷ Az összes emelt szinten érettségiző száma (2005): 15.601, de ebből 14.131 vizsgázó nyelvvizsga-bizonyítvánnyal váltotta ki az érettségi vizsgát. Az átlageredmény csak a valóban vizsgázókra vonatkozik.

gadtatni, a vizsgázónak 60% fölött kell teljesíteni. Ezt minden vizsgált vizsgaidőszakban teljesítette a vizsgázók többsége. Az emelt szintű vizsgát 2005-ben még ki lehetett váltani középfokú nyelvvizsgálóval, de 2006-tól már nem.

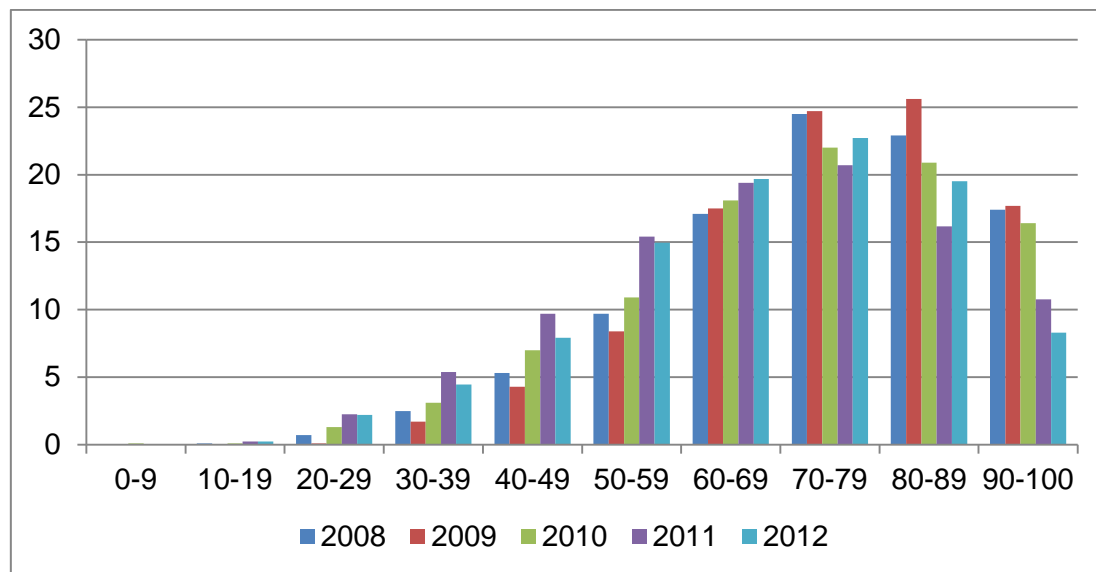
A teljesítmény-eloszlások alapján (0.2. ábra) a legtöbb 90-100% közötti teljesítmény az első évben, 2005-ben volt emelt szinten. A 2006. és 2007. évi eredmények közelebb állnak egymáshoz. A 2007. év adatai azért fontosak különösen, mert az volt a felsőoktatásban a bolognai rendszer bevezetésének az éve. Ettől kezdve a nyelvszakokra kizárólag emelt szintű érettségivel lehetett jelentkezni, s ennek eredménye döntött a felvételtől.



Forrás: OH prezentáció (https://www.ketszintu.hu/publicstat.php?stat=_2006_1)

0.2. ábra:
A teljesítmények eloszlása az emelt szintű vizsgán 2005-ben,
2006-ban és 2007-ben (Ein-horn, 2008b: 214)

Az Oktatási Hivatal 2012-es érettségi vizsgáról közzétett összegzése (Érettségi 2012 május-június, 2012) alapján, a teljesítmények jobbra tolódnak (0.3. ábra): 2008-ban a vizsgázók 82%-a, 2009-ben 87%-a, 2010-ben 77%-a teljesített 60% fölött és jeles érdemjegyet szerzett. Ez az arány az utolsó két évben csökkent: 2011-ben az emelt szinten érettségizők 67%-a, 2012-ben 70%-a teljesített 60% felett. Ez vélhetően a vizsgára jelentkezők számának emelkedésével (0.11. táblázat), illetve az egyes vizsgarészek eltérő nehézségével magyarázható. Ez utóbbit a továbbiakban részletesen is elemzünk.



0.3. ábra:
A teljesítmények eloszlása az emelt szintű vizsgán 2008 és 2012 között
(Érettségi 2012 május-június, 2012)

A teljes vizsgán a nappali tagozatos képzésben a legjobb teljesítményt a gimnáziumi érettségizők hozzák. Az emelt szintű vizsgánál jó eredmények mutatkoznak, hiszen ezt a vizsgát többnyire azok a tanulók választják, akik német szakra kívánnak felvételizni.

0.12. táblázat: Német nyelvi érettségi átlagok eloszlása emelt szinten %-ban vizsgázói rétegek szerint

Vizsgaév	Gimnázium	Szakközépiskola	Felnőttoktatás	Tanulói jv. kívül
2010	74,46%	68,65%	74,69%	69,90%
2011	70,17%	62,43%	60,00%	68,95%
2012	70,32%	63,58%	64,23%	68,50%

A továbbiakban az emelt szintű érettségien az egyes vizsgarészek megoldottságát a középshinthez hasonló módon elemezzük. Először megvizsgáljuk az átlagteljesítményeket az egyes vizsgaéveken belül, majd a vizsgaévek közötti tendenciákra következtetünk. A fejezet további része Vígh (2010: 104–106. o.) munkája és további számítások alapján készült.

0.13. táblázat:

A német nyelvi emelt szintű vizsgák eredményei (százalékpontban megadott átlagok, zárójelben a szórások)
(Vígh, 2010: 105 és további számítások alapján)

Vizsgarész	Vizsgaév					
	2007	2008	2009	2010	2011	2012
Olvasott szöveg értése	78 (21)	82 (14)	89 (10)	68 (20)	69 (22)	74 (19)
Nyelvhelyesség	69 (18)	65 (19)	64 (17)	73 (20)	67 (19)	68 (17)
Hallott szöveg értése	68 (17)	74 (19)	79 (14)	69 (18)	53 (18)	53 (20)
Íráskészség	65 (26)	74 (21)	71 (23)	75 (20)	72 (23)	70 (24)
Beszédképesség	69 (26)	73 (23)	74 (23)	74 (24)	73 (25)	74 (22)

Az olvasott szöveg értése vizsgarész 2007 és 2009 között volt a legkönnyebb, az emelt szinten érettségizők 2009-ben 90% körüli átlagteljesítményt értek el, a szórás és a relatív szórás (11,29%) alacsony, így ez a feladatsor tesztelméleti szempontból nem megfelelő, a reliabilitás feltételezhetően alacsony. Ebben a három évben növekszik a vizsgázók átlagteljesítménye, és ezzel együtt csökken a szórás: a feladatlap mindhárom évben túl könnyű az érettségizőknek. Ugyanakkor 2010-ben és 2011-ben a vizsgázók jóval alacsonyabb átlageredményt értek el, amely azt mutatja, hogy ez a vizsgarész a korábbi évekhez képest nehezebb, 2012-ben viszont némiképp könnyebb volt.

Hasonló tendencia figyelhető meg a másik receptív készségnél, a hallott szöveg értése feladatsornál: 2007 és 2009 között némiképp növekedett az átlagteljesítmény, míg 2010-ben csökkent, az utolsó két évben pedig a korábbi évekhez képest jóval nehezebb volt ez a feladatsor. Mindez azt mutatja, hogy lényeges lenne megvizsgálni, hogy az emelt szintű feladatok hogyan viszonyulnak a vizsgakövetelményekben rögzített készségszintekhez és a vizsgázók nyelvtudás-szintjéhez, illetve mi jellemzi a felkészítési folyamatot.

A nyelvhelyességteszt a 2010. évet kivéve hasonló megoldottságú. Ez általában nehezebben teljesíthető a vizsgázóknak, viszont viszonylag egyenletes az átlagteljesítmény.

A két produktív készséget mérő vizsgarészben elért teljesítmények 70%-pont körüliek. Ugyanakkor ezekben a vizsgarészekben a többihez képest heterogénebb a populáció, jelentős különbségek lehetnek a vizsgázók között különböző háttérváltozók szerint.

5. A kutatás módszerei

5.1. A kutatás céljai, kérdései és hipotézisei

A kutatás fő célja a 2012. évi május-júniusi vizsgaidőszakban megoldott közép- és emelt szintű német nyelvi érettségi feladatsorainak, feladatainak és értékelési eljárásainak mérszemetodikai elemzése, működésük értékelése, valamint javaslatok megfogalmazása a vizsga továbbfejlesztésére vonatkozóan. A német nyelv érettségét mindkét szinten azonos alapelvek alapján kívánjuk vizsgálni, és az egyes vizsgarészek/vizsgarész-összetevők tesztszintű elemzése mellett két további célterületet jelölünk ki: külön elemezzük az itemalapú receptív készségeket (olvasott szöveg értése, hallott szöveg értése) és a nyelv-helyességet mérő vizsgarészeket/vizsgarész-összetevőket, illetve a produktív készségek közül az íráskészséget mérő feladatokat, itemeket és értékelési eljárásokat.

Az itemalapú feladatsorok minősége szempontjából célunk meghatározni a tesztek, a feladatok és az itemek nehézségi szintjét. Célunk elemezni, hogy a feladatok megfelelő nehézségűek-e a vizsgázók számára, az itemnehézségi indexek mennyire fedik le a vizsgázók nyelvi készségszintjei által meghatározott intervallumot. A feladatfejlesztés szempontjából arra keresünk választ, hogy melyik itemet és esetleg feladatot milyen okból kellene elhagyni a vizsgált tesztek közül és ezeknek milyen következményei vannak/lehetnek a vizsgázók teljesítményére. A vizsgálat célja tehát a feladatsorok, a feladatok és az itemek működésének részletes megismerése, a nehézségi szintek meghatározása, az abból levonható következtetések és a vizsga továbbfejlesztésére vonatkozó javaslatok megfogalmazása.

Az íráskészség vizsgarész/vizsgarész-összetevő külön elemzésének alapvető célja, hogy a teszt, a feladatok és főként az értékelési rendszer megfelelőségét elemezzük. Célunk a vizsgázók íráskészség-szintjének, az értékelési szempontok jellemzőinek és az értékelési skálák szerkezetének vizsgálata. Az értékelési szempontoknál meghatározzuk azok magyarázóerejét, a közöttük lévő összefüggéseket, a nehézségi szintet. Az érettségizők íráskészségével kapcsolatban azt kívánjuk elemezni, hogy a feladatok és értékelési eljárások milyen mértékben képesek a vizsgázók teljesítményét különböző képesség-szintekbe különíteni. Az értékelési skálák működése szempontjából azt elemezzük, hogy az értékelők minden szintleírást azonos mértékben használnak-e. A vizsgálat célja tehát olyan minőségi mutatókat jellemző paraméterek mentén elemezni ezt a vizsgarészt, amelyekkel a különböző évekből származó eredmények összehasonlíthatók, és e vizsgarész/vizsgarész-összetevő mérési-értékelési rendszerében elvégzett változások hatásai vizsgálhatók. Így elősegíthetjük azt a folyamatot, amelyben az elért eredmények alapján adható visszajelzés a vizsgafejlesztés számára.

A kutatás céljaihoz illeszkedve, az alábbi az alábbi kérdésekre keresünk válaszokat:

- I. A receptív készségek, a nyelvhelyesség és az íráskészség vizsgálata:
 1. Mennyire megbízhatóan mértek az egyes készségeket mérő feladatsorok?
 2. Milyen mértékben magyarázzák az egyes feladatok az adott vizsgarészben/vizsgarész-összetevőben elért összpontszámot?
- II. Az itemalapú feladatsorok vizsgálata közép- és emelt szinten:
 3. Milyen nehézségi szintbe sorolhatók az egyes feladatok és itemek?
 4. Mennyiben felelnek meg a feladatok itemnehézségi indexei a vizsgázók nyelvi készség szintjeinek?
 5. Mennyire megfelelően működnek az egyes vizsgarészekben/vizsgarész-összetevőkben használt zárt végű feladattípusok? Ha nem jól működnek, mi ennek az oka?
 6. Mennyire megfelelően működnek a nyílt végű feladattípusok? Mi az oka a nem megfelelő itemműködésnek ezekben a feladatokban? Hogyan működött ezeknél a feladatoknál az értékelési rendszer?
- III. Az íráskészség vizsgálata közép- és emelt szinten:
 7. Hogyan teljesítettek a tanulók az egyes feladatokban? Van-e jelentős különbség a két feladat megoldottságában? Van-e jelentős különbség a második feladatban megadott két téma megoldottsága között?
 8. Milyen kapcsolat van az értékelési szempontok között, ezek mennyiben határozzák meg az összpontszámot és milyen nehézségi szintekbe sorolhatók?
 9. Mennyire fedik le az értékelési skálapontok a vizsgázók íráskészség-szintjei alapján meghatározott intervallumot?
 10. Mennyire jól rendezettek az egyes értékelési szempontokhoz tartozó pontszámok az értékelési skálákban?

A kutatási kérdésekhez az alábbi hipotéziseket fogalmazzuk meg:

- I. A német nyelvi érettségi működésének jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján feltételezzük, hogy
 1. az egyes készségeket mérő feladatsorok megbízhatósága eltér, amely a mérendő konstruktum jellemzőitől és az alkalmazott feladattípusoktól függ;
 2. az egyes feladatok súlya a tesztekben eltér, amely a mért nyelvi szinttől, illetve az egyes feladatok minőségétől függ.
- II. A receptív készséget mérő feladatsorok jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján feltételezzük, hogy
 3. a feladatok itemnehézségi indexei nem fedik le teljesen a vizsgázók nyelvi készség szintjei által meghatározott intervallumot;
 4. a magasabb készségtartományból hiányoznak itemek;

5. a zárt végű feladatokban lévő összefüggő itemek, a nem megfelelő disztraktorok, a pontlevonás rontják az adott feladat és a teszt reliabilitását;
6. a nyílt végű feladatok működése függ a mért nyelvi szinttől és a mért gondolkodási művelettől.

III. Az íráskészség vizsgarészben/vizsgarész-összetevőben alkalmazott mérési-értékelési eljárások alapján feltételezzük, hogy

7. a vizsgázók teljesítménye eltér a két feladatban, de a választható témák megoldottságában nincs jelentős különbség;
8. az értékelési szempontok magyarázóereje és nehézsége eltér, amely függ attól, hogy az íráskészség mely dimenzióját mérik;
9. az értékelési skálapontok nem fedik le teljesen a vizsgázók íráskészség-szintjei által meghatározott intervallumot;
10. az egyes értékelési skálák működése eltér szempontonként és a skálákban definiált pontszámok alapján.

A vizsgán elért készségenkénti teljesítmények, illetve a receptív és a produktív készségeket mérő vizsgarészek/vizsgarész-összetevők elemzéséhez kapcsolódó kutatási kérdések vizsgálatában az Oktatási Hivatal megrendelésére az Expanzió Kft. végzett adatgyűjtést és statisztikai elemzéseket. A továbbiakban először az adatbázisokat jellemezzük, majd az adatelemzés módszereit részletezzük.

5.2. Az adatbázisok jellemzői

Az adatbázisok célja, hogy lehetővé tegyék a feladat- és itemszintű elemzéseket. A német nyelvi érettségénél a közép- és emelt szintű vizsgafeladatokhoz tartozó adatbázisok az olvasott és a hallott szöveg értéke, a nyelvhelyesség feladatsorok itemszintű adatait és az íráskészség vizsgarészben/vizsgarész-összetevőben elért részpontszámokat, valamint a vizsgázó nemét, a települést, az intézménytípust, illetve a vizsgázó nyelvi előkészítőn való részvételéről szóló adatait tartalmazzák. Ezek alapján információ szerezhető a feladatsorok minőségéről, a feladatok nehézségéről, az értékelési eljárások alkalmazásáról, a fejlesztés lehetőségeiről. Az adatbázisok problémája viszont, hogy nem teszik lehetővé a disztraktorelemzést, mert arról nem tartalmaznak információt, hogy a feleletválasztó feladatokban a tanuló melyik alternatívát jelölte be. Ezen kívül nem tartalmazzák a vizsgázó évfolyamát, illetve a vizsgafajtát. Míg az első probléma jelentősen befolyásolja a feladatok és feladattípusok beválásának vizsgálatát, addig a második kevésbé, hiszen a projektnek nem az volt az elsődleges célja, hogy a teszten nyújtott teljesítmény és különböző háttérváltozókkal való kapcsolatot elemezzük.

Az adatbázis középhalmozinten 992, emelt szinten 578 vizsgázó adatait tartalmazza. Ekkora mintaméret alapján ki lehet mutatni a lényeges tendenciákat, fontosabb összefüggéseket, de nem lehet ugyanolyan pontos becsléseket kapni, mint a teljes adatbázis – ez esetben nem kivitelezhető – elemzésével. Az Oktatási Hivatal által rendelkezésre bocsátott mintavételi eljárás-leírás szerint a cél az volt, hogy a minta a nappali tagozatos vizsgázók körére vonatkoztatva reprezentatív legyen régiók, képzési és településtípusok szerint. Vagyis az összteljesítmény alapján e háttérváltozók szerint történt a mintaillesztés. Ennek viszont az a következménye, hogy az egyes feladatlapok megoldottságában és az országos eredményekhez képest lehetnek eltérések. Az egyes vizsgarészek/vizsgarész-összetevők elemzését megelőzően nem volt mód a mintaillesztés elvégzésére. Az eredményekből – különösen középhalmozinten – nem lehet automatikusan következtetni a vizsgázói populáció teljesítményére, mert például a felnőttoktatásban résztvevőkre nem terjedt ki az adatfelvétel.

Az Expanzió Kft.-től az adatbázis eredményeiről kaptunk kimutatást, amely mindkét szinten nyolc részből áll:

1. Üresen hagyott feladatok⁸
2. Az összesített pontszámok eloszlása⁹
3. A részkérdések¹⁰ százalékos megoldottsága
4. A részkérdésenkénti (itemenkénti) pontszámok megoszlása és pontbiszeriális korrelációja
5. A feladatonkénti pontszámok megoszlása és pontbiszeriális korrelációja
6. A feladatok magyarázó ereje¹¹
7. A teszt¹² megbízhatósága (Cronbach-alfa)
8. A részkérdések nehézségének és a tanulók képességeinek megoszlása

A kimutatás mellett megkaptuk Excel-fájlban a kiszámított adatokat, az egyes receptív készségeket, és a nyelvhelyességet mérő vizsgarészekben/vizsgarész-összetevőkben lévő itemek karakterisztikus görbéit, illetve kérésünkre további elemzésként a feladatok megbízhatóságát jelölő Cronbach-alfa értékeket, az itemek nehézségi indexét és modellilleszkedését jelölő paramétereket. Mindezek mellett pedig az íráskészségben alkalmazott egyes értékelési szempontok megoldottságának a receptív készségeket és a nyelvhelyességet mérő vizsgarészek/vizsgarész-összetevők teljesítettségével való kapcsolatára vonatkozó adatokat is. Ezekkel az adatokkal ugyanakkor csak részben tudunk a megfogal-

⁸ Itt valójában az üresen hagyott itemekről van szó.

⁹ Itt nem a vizsgapontok, hanem a teszten elérhető feladatpontok alapján történt a pontszámok eloszlásának vizsgálata.

¹⁰ Itt valójában itemekről van szó.

¹¹ Itt valójában az itemek magyarázóerejéről van szó.

¹² A teszt megbízhatóságát vizsgarészenként számolták ki.

mazott kutatási kérdésekre választ adni. Ennek oka, hogy az egyes itemek paramétereit nem vizsgarészenként/ vizsgarész-összetevőnként, hanem a három itemalapú vizsgarészre/vizsgarész-összetevőre együtt számolták ki. Nem kaptuk meg továbbá az egyes vizsgarészekre/ vizsgarész-összetevőkre vonatkozóan külön az itemek nehézségének és a vizsgázók képességeinek eloszlását mutató ún. minta-itemtérképeket. Az íráskészség vizsgarész/ vizsgarész-összetevő esetében pedig nem állt rendelkezésre az értékelési szempontok megoldottsága, magyarázóereje, teljesítésének nehézsége, az értékelési skálapontok és a vizsgázók íráskészségének eloszlása. Bár a vizsgált minta az egyes értékelési szempontokhoz tartozó pontszámoknál elaprózódik, de e korlát ismeretében is lehet ezeket az elemzési szempontokat vizsgálni (I. Vígh, 2010; 2012). Mindezek alapján szükséges volt az egyes vizsgarészek/ vizsgarész-összetevők önálló másodelemzése is. Ennek elvégzése Vígh Tibor munkája, ezt az adott helyen VT rövidítéssel jelöljük.

5.3. Az adatelemzés módszerei

A receptív készségeket (olvasott szöveg értése, hallott szöveg értése), a nyelvhelyességet és az íráskészséget mérő feladatlapok minőségét (I. kutatási terület) azonos elvek szerint elemeztük. Mivel ezek különböző konstruktumokat mérnek, ezért vizsgarészenként/vizsgarész-összetevőnként vizsgáltuk az adott teszt működését. Ehhez kiszámítottuk a vizsgált teszt és a feladatok átlagát, szórását, megbízhatóságát és magyarázóerejét, illetve a teszten nyújtott teljesítmények eloszlását.

A teszt és a feladatok megbízhatóságát a Cronbach-alfa értékkel számítottuk, amely a reliabilitást alulról becsli. Képességmérő teszteknel 0,70 fölötti értéket tekintünk elfogadhatónak. Ennek értéke függ egyrészt az itemek számától, másrészt a szórástól, amely a csoport összetételét mutatja (I. Csapó, 2004; Csikos és B. Németh, 2002). Ezért feladat szinten azt vizsgáltuk, hogy az adott teszten belül egymáshoz képest milyen megbízhatósággal mérték az egyes feladatok.

A feladatok magyarázóereje azt adja meg, hogy az egyes feladatoknak mekkora a súlya az összpontszám alakulásában. Ezt lineáris regresszióanalízis módszerével állapítottuk meg és százalékban fejeztük ki.

A receptív készségeket mérő vizsgarészek/vizsgarész-összetevők elemzésekor (II. kutatási terület) a Rasch-modell (Rasch, 1960) segítségével az egyes itemek nehézségi indexét vizsgáltuk, amelyet a következő lépésben a mintába került diákok képességszintjével összekapcsolva közös képességskálán minta- és itemtérképen ábrázoltuk. E modellel az itemek nehézségét és modellilleszkedését is vizsgáltuk. Az egyes feladatokon belül az itemeknél a klasszikus tesztelemzési módszerek segítségével az itemek átlagát, szórását vagy differenciáló erejét, pontbiszeriális korrelációját és magyarázóerejét vizsgáltuk. A továbbiakban ezeket a paramétereiket jellemezzük.

A minta- itemtérkép segítségével középszinten azt elemeztük, hogy a feladatkészítők által megállapított nehézségi sorrend mennyire esik egybe a feladatok empirikus nehézségének növekedésével. Közép- és emelt szinten azt vizsgáltuk, hogy az itemek mennyire jól fedik le a vizsgázók nyelvi képességszintjei alapján meghatározott intervallumot. Ennek alapján megállapítható, hogy mely itemek túlságosan könnyűek, és melyek nagyon nehezek, amelyek már mutathatják, hogy melyiket célszerű elhagyni a tesztből.

Az itemek nehézségét a Rasch-modell segítségével számoltuk. Ez minden esetben az a képességszint, amely az item 50%-os valószínűséggel történő megoldásához szükséges. A negatív értékek az átlag alatti, a 0 körüli értékek az átlagos, a pozitív értékek az átlag feletti képességszinten 50%-os megoldási valószínűségű itemeket mutatják (Molnár, 2013).

Minden itemhez tartozik egy modellilleszkedés-paraméter, amelyet az elemzésben infit MNSQ-val jelölünk, és ez az item empirikus adatainak a Rasch-modellhez való illeszkedését mutatja. Az érték annál jobb, minél közelebb van 1-hez, ugyanis az item empirikus karakterisztikus görbéje ebben az esetben közel azonos meredekségű az elméleti görbével (Molnár, 2013: 93).

Az egyes itemek elemzésekor az átlaggal az itemek nehézségét jellemezhetjük. Ez 0 és 1 közé eső szám. Minél közelebb van az item nehézségi indexe 1-hez, annál többen adtak helyes választ, azaz annál könnyebb az adott item. Minél közelebb van az item nehézségi értéke 0-hoz, annál többen adtak rossz megoldást, azaz annál nehezebb az item (Csapó, 2004: 289). Általában 0,3 és 0,8 közötti értéket tekintünk megfelelőnek.

További mutatóként a szórás segítségével az itemek differenciáló erejét kaptuk meg. Ez megadja, hogy az item mennyire érzékeny a tanulók tudására, vagyis mennyire élesen tesz különbséget a különböző tudásszintű tanulók között (Csapó, 2004: 289).

A pontbiszeriális korreláció számításánál a Rasch-modellből adódó becsült képességpontokat használtuk. Ez azt mutatja, hogy az item azt méri-e, amit a teszt egésze, vagyis a teszttel azonos módon különíti-e el egymástól a különböző tudású tanulókat (vö. Csapó, 2004; Molnár, 2013). Az elemzéskor a dichotóm itemekre vonatkozóan csak az 1 pontszám pontbiszeriális korrelációját adjuk meg, mert a 0 pontszámnál a korreláció értéke ugyanannyi, iránya viszont ellentétes. Elemzésünkben azt várjuk, hogy ez az érték minél közelebb legyen 1-hez, de legalább 0,2 fölött legyen. Ha az 1 pontszám pontbiszeriális korrelációja ennél alacsonyabb, akkor ez azt jelenti, hogy az item nem azt méri, amit a teszt egésze, vagy nem jól méri azt, amit a teszt egésze. Ha az 1 pontszám pontbiszeriális korrelációja negatív, akkor az adott itemet tendenciaszerűen azok a vizsgázók oldották meg jól, akik a tesztet egészében rosszul. Ilyen esetben arra gyanakodhatunk, hogy az item nem működik, például a megfogalmazás félreérthető (Csapó, 2004).

Az itemek magyarázóereje (R^2) az adott item súlyát mutatja meg a tesztben. Ez egy 0 és 1 közötti érték. Minél nagyobb a magyarázó erő, annál inkább előrejelzi az adott itemre

kapott pontszám az adott vizsgarész/vizsgarész-összetevő pontszámát. Ez tehát megadja, hogy az adott item mennyire jól méri a vizsgált képességet.

Az íráskészség vizsgarész/vizsgarész-összetevő esetében (III. kutatási terület) először az egyes feladatokban az értékelési szempontok megoldottságát, magyarázóerejét, az egyes pontszámok eloszlását vizsgáltuk. A második feladatban azt is elemeztük, hogy a két téma mennyire tekinthető ekvivalensnek.

A 9. és 10. kutatási kérdés megválaszolásához a valószínűségi tesztelméleti eszközök közül Masters (1982) parciáliskredit-modelljét alkalmaztuk, mivel az egyes értékelési szempontokhoz tartozó skálák eltérő szerkezetűek. Először az adatokat közös skálára kalibráltuk a ConQuest programmal (Wu, Adams és Wilson, 1998) és az értékelési szempontok átlagos nehézségi indexét és modellilleszkedését vizsgáltuk. Ezek a paraméterek nagyon hasonlítanak a receptív készségeknél vizsgált itemparaméterekhez. A következő lépésben az értékelési skálapontok küszöbértékeit (vagy lépésparamétereit) a mintába került diákok íráskészség-szintjével kapcsoltuk össze és elemeztük. Ez a paraméter azt adja meg, hogy az adott értékelési szemponton belül milyen képességszint szükséges ahhoz, hogy a vizsgázó 50%-os valószínűséggel a magasabb pontszámot szerezzék meg. Ezt követően pedig a skálák karakterisztikus görbéit vizsgáltuk, melynek segítségével azt állapítottuk meg, hogy az értékelők hogyan használják az egyes pontszámokat, azaz, hogy minden pontszám és deskriptor azonos mértékben hasznosul-e.

Ezen kutatási célok, kérdések, hipotézisek és a hozzájuk rendelt elemzési módszerek alapján alakítottuk ki a tanulmány további részeit. A továbbiakban mindkét szintre vonatkozóan, a receptív készségeket és a nyelvhelyességet mérő vizsgarészekben/vizsgarész-összetevőkben belül végzünk feladat- és itemszintű elemzéseket (2. és 4. dokumentum), majd az íráskészség esetében megvizsgáljuk a mérési-értékelési eljárók minőségét (3. és 5. dokumentum). Ezen vizsgálattal a vizsgafejlesztési folyamatra vonatkozóan adunk visszacsatolást és továbbfejlesztési javaslatokat (6. dokumentum).

Irodalom

- 100/1997. (VI. 13.) Kormányrendelet az érettségi vizsga vizsgaszabályzatának kiadásáról (<http://net.jogtar.hu/jr/gen/getdoc.cgi?docid=99700100.kor>, megtekintés: 2012.11.25.)
- Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgakövetelmények.
http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv vk.pdf,
- Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgaleírás.
http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv vl.pdf (letöltve: 2013. május 20.)
- CSAPÓ BENŐ (2004): Tudásszintmérő tesztek. In: Falus Iván (szerk.): *Bevezetés a pedagógiai kutatás módszereibe*. Budapest: PSZMP–Keraban Kiadó, **277–316**.
- CSÍKOS CSABA – B. NÉMETH MÁRIA (2002): A tesztekkel mérhető tudás. In: Csapó Benő (szerk.): *Iskolai tudás*. Budapest: Osiris Kiadó, **91–122**.
- EINHORN ÁGNES (1998): Az 1997. évi német felvételi vizsga elemzése. In: Einhorn Ágnes (szerk.): *Vizsgatárgyak, vizsgamodellek I. Német nyelv*. Budapest: Országos Közoktatási Intézet, **125–152**. (=Mérés, értékelés, vizsga 5.)
- EINHORN ÁGNES (2006): A német nyelvi vizsgafeladatok fejlesztése – a tartalmi érvényesség és az értékelhetőség ellentmondásai. In: Horváth Zsuzsanna – Lukács Judit (szerk.): *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Budapest: Országos Közoktatási Intézet, **153–171**.
<http://www.ofi.hu/tudastar/erettsegi/uj-erettsegi/nemet-nyelvi-erettsegi> (letöltve: 2013. május 14.)
- EINHORN ÁGNES (2007a): A 2005. évi érettségi vizsga eredményeinek elemzése: Német nyelv. In: Einhorn Ágnes (szerk.): *A 2005-ös érettségi vizsga eredményeinek elemzése*. Budapest: Oktatókutató és Fejlesztő Intézet.
<http://www.ofi.hu/tudastar/erettsegi/2005-os-erettsegi/2005-evi-erettsegi-090617-2> (letöltve: 2013. május 24.)
- EINHORN ÁGNES (2007b): Az idegen nyelvi érettségi reformja. In: Vágó Irén (szerk.): *Fókuszban a nyelvoktatás*. Budapest: Oktatókutató és Fejlesztő Intézet, **73–105**.

- EINHORN ÁGNES (2008a): *A 2006. évi érettségi vizsga eredményeinek elemzése: Német nyelv*. Budapest: Oktatókutató és Fejlesztő Intézet.
<http://www.ofi.hu/tudastar/erettsegi/2006-os-erettsegi/2006-evi-erettsegi-090617-2> (letöltve: 2013. május 24.)
- EINHORN ÁGNES (2008b): Hogy vizsgázott az új érettségi németből? In: Bánkúti Zsuzsa/Lukács Judit (szerk.): *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés*. Budapest: Oktatókutató és Fejlesztő Intézet, **207–237**.
- EK, J.A. VAN és TRIM, J.L.M. (1991): *Threshold Level*. Brussel: Council of Europe.
- Érettségi 2012. május-június* (2012) Budapest: Oktatási Hivatal.
http://www.oktatas.hu/pub_bin/dload/kozoktatasi/erettsegi/prezentaciok/gyorsadatok_2012_maj_jun.pdf (letöltve: 2013. május 24.)
- HORVÁTH ZSUZSANNA – LUKÁCS JUDIT (2006): A kétszintű érettségi vizsga. In: Horváth Zsuzsanna – Lukács Judit (szerk.): *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Budapest: Országos Közoktatási Intézet, **11–38**.
<http://www.ofi.hu/tudastar/erettsegi/uj-erettsegi/keszszintu-erettsegi> (letöltve: 2013. május 24.)
- JÓZSA KRISZTIÁN – NIKOLOV MARIANNE (2005): Az angol és német nyelvi készségek fejlettségét befolyásoló tényezők. *Magyar Pedagógia* 105/3, **307–337**.
- Közös Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés* (2002): Európa Tanács Közoktatási Bizottsága. Közoktatási Bizottság. Élő Nyelvek Osztálya. Strasbourg. Piliscsaba: Pedagógus-továbbképzési Módszertani és Információs Központ Kht.
- MASTERS, G. N. (1982): A Rasch model for partial credit scoring. *Psychometrika* 47/2, **149-174**.
- MOLNÁR GYÖNGYVÉR (2013): A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában. Budapest: Gondolat Kiadó.
- NIKOLOV MARIANNE – PÉRCSEICH RICHÁRD – SZABÓ GÁBOR (2000): A puding próbája: Alapszintű angol feladatsorok bemérésének tapasztalatai. *Modern Nyelvoktatás* 6/4, **3–28**.
- PETNEKI KATALIN (2007): *Az idegen nyelvek oktatása Magyarországon az ezredfordulón*. Szeged: JATEPressz.
- RASCH, G. (1960): *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

- VÍGH TIBOR (2010): Az idegen nyelvi érettségi működése és hatása a tanulói teljesítmények és a tanári nézetek tükrében. PhD-értekezés. Szeged: Szegedi Tudományegyetem, Bölcsészettudományi Kar, Neveléstudományi Doktori Iskola.
www.edu.u-szeged.hu/phd/downloads/vigh_ertekezés.pdf (letöltve: 2013. május 24.)
- VÍGH TIBOR (2012): Qualitätskriterien in der Messung und Bewertung von Schreibfertigkeit: Die Ergebnisse einer Analyse mit dem Partial-Credit Modell. In: Arntz, R., Krings, H. P. és Kühn, B. (szerk.): *Autonomie und Assessment. Erträge des 3. Bremer Symposiums zum autonomen Fremdsprachenlernen*. Bochum: AKS-Verlag, **71–83**.
- VÍGH TIBOR (2013): A középszintű angol és német nyelvi érettségien elért teljesítmények változásai 2007-2012 között. *Modern Nyelvoktatás*, 19/1–2, **17–35**.
- WU, M., ADAMS, R. J. és WILSON, M. R. (1998): *ACER ConQuest. Generalised Item Response Modelling Software*. Australia: ACER Press.

VÍGH TIBOR és PETNEKI KATALIN

**A 2012. MÁJUSI NÉMET NYELV
KÖZÉPSZINTŰ VIZSGA
ELSŐ HÁROM FELADATSORÁNAK ELEMZÉSE**

A középszintű német nyelvi érettségi vizsga első három vizsgarész-összetevőjének elemzési célja az I. és a II. kutatási terület vizsgálata (vö. BEVEZETŐ/5.1. fejezet). Ennek megfelelően először áttekintést adunk a vizsga általános jellemzőiről, a 2012. évi eredményekről, majd bemutatjuk az első három vizsgarész-összetevő felépítését a vizsgaleírás és a feladatlap, illetve a javítókulcs alapján. Ezt követi készségenként a feladatok és az azokban előforduló itemek részletes elemzése. Végül összefoglaljuk az elemzés eredményeit.

1. A 2012. májusi vizsga általános jellemzői

Középszinten az írásbeli vizsga négy részből áll: olvasott szövegértés, nyelvhelyesség, hallott szövegértés és íráskészség. A középszintű feladatok pontszáma és itemszáma nem mindig egyezik, ezért az itemek után kapható feladatpontok vizsgapontszámmá alakításához átszámítási kulcs szükséges. A középszintű feladatsor pontszámai a következők:

0.1. táblázat:

Áttekintés az írásbeli vizsga itemeiről és a teljesítményre adható vizsgapontokról

Vizsgarész-összetevő	Itemek száma	Vizsgapont
1. Olvasott szöveg értése	33	33
2. Nyelvhelyesség	27	18
3. Hallott szöveg értése	21	33
4. Íráskészség	-	33
Összesen:		117

Legtöbb adatot az olvasott szövegértés, a nyelvhelyesség és a hallott szövegértés vizsgarész-összetevőhöz sikerült begyűjteni. E három vizsgarész-összetevő jellemzője, hogy a feladatokhoz itemek tartoznak, amiket vagy sikerül jól megoldani (1 pont), vagy nem (0 pont), így ezek statisztikai elemzése aránylag egyszerű.

A vizsga általános jellemzője, hogy sokféle feladattípust lehet alkalmazni. A lehetséges feladattípusokat a vizsgaleírás tartalmazza. Ezek közül a konkrét vizsga középszintű német nyelvi feladatsora a receptív készségek és a nyelvhelyesség vizsgarész-összetevő tekintetében az alábbi feladattípusokból áll:

0.2. táblázat:

Az írásbeli vizsga első három vizsgarész-összetevőjének feladattípusai

Olvasott szövegértés	Nyelvhelyesség	Hallott szövegértés
1. ál igaz-hamis (szelektálás)	1. feleletválasztó	1. kérdésekre válasz
2. kérdésekre válasz	2. feleletválasztó	2. igaz-hamis feladat
3. mondatok kiegészítése	3. szöveg kiegészítése megadott szavakkal	3. ál igaz-hamis (szelektálás)
4. igaz-hamis feladat	4. hozzárendelés	

A felsorolt feladattípusok nem csak egy bizonyos készséghez kötődnek, mind a három vizsgarész-összetevőben előfordulhatnak. Volt olyan feladattípus, ami csak egyetlen vizsgarész-összetevőben fordult elő. A nyelvhelyesség feladattípusai ezen a vizsgán csak ebben a vizsgarész-összetevőben szerepeltek.

0.3. táblázat:

Az írásbeli vizsga feladattípusai az első három vizsgarész-összetevőre elosztva

Feladattípus	Olvasás	Nyelvhelyesség	Hallás	Összesen
hozzárendelés		1		1
szöveg kiegészítése		1		1

kérdésre válasz	1		1	2
mondatok kiegészítése	1			1
igaz-hamis	1		1	2
ál igaz-hamis (szelektálás)	1		1	2
feleletválasztó		2		2

0.4. táblázat:

Áttekintés az itemek eloszlásáról a vizsgarész-összetevőkön, ill. feladatsoron belül

Vizsgarész-összetevő	Feladattípus	Itemszám	Összesen
Olvasott szövegértés	ál igaz-hamis (szelektálás)	9	33
	kérdésekre válasz	9	
	mondatok kiegészítése	8	
	igaz-hamis feladat	7	
Nyelvhelyesség	feleletválasztó	7	27
	feleletválasztó	7	
	szöveg kiegészítése megadott szavakkal	8	
	hozzárendelés	5	
Hallott szövegértés	kérdésekre válasz	7	21
	igaz-hamis feladat	8	
	megfelelő kijelentés kiválasztása (igaz-hamis)	6	

A középszintű vizsga további jellemzői, hogy a feladatok kb. 50%-ának az A2 szintet kell lefednie, a másik fele pedig a B1 szintet. Az egyes vizsgarész-összetevők feladatai tehát egy vagy két könnyebb feladattal kezdődnek, s nehezebbel záródnak. Az elemzés célja tehát az is, hogy megállapítsa, mennyire sikerült ezt az elvet a vizsgált feladatsorban teljesíteni.

A következőkben a vizsgarész-összetevőkön belüli feladatokat és azon belül az itemeket egyenként elemezzük. Az elemzést részben a megkapott bázisadatokból a megbízó által feldolgoztatott adatok, ábrák segítségével, részben önálló elemzés alapján¹³ mutatjuk be (v.ö. I. Bevezető).

¹³ Az önálló statisztikai elemzés Vigh Tibor munkája, amit VT rövidítéssel jelölünk a dokumentumban.

2. Készségenkénti (vizsgarész-összetevőnkénti) elemzés

2.1. Olvasott szöveg értése (I. vizsgarész-összetevő)

A vizsgarész-összetevő célja annak mérése, hogy a vizsgázó képes-e a mindennapi életben előforduló, különböző fajtájú autentikus szövegeket önállóan elolvasni, és az olvasási céloknak megfelelő mélységben megérteni. A vizsgafeladatok megoldásához a vizsgázónak képesnek kell lennie az értési céloknak megfelelő stratégiák alkalmazására (40/2002 OM rendelet, Vizsgaleírás: 31). A vizsgaleírás tartalmazza, hogy milyen olvasási stratégiák alkalmazása várható el az adott szinten, ezt milyen típusú szövegeken milyen típusú feladatokkal lehet mérni. A teszt feladattipológiailag kötetlen szerkezetű, azaz bármelyik felsorolt feladattípus szerepelhet benne. A feladatsor a szövegértés alábbi részkészségeit méri: globális, szelektív és részletes értés. A feladatsor 3–4 szövegből és 3–4 feladtból áll. Egy feladaton belül egy hosszabb szöveg helyett előfordulhat több rövidebb szöveg is (pl. apróhirdetések). A feladatokhoz felhasznált szövegek együttes terjedelme 900–1100 szó. A feladatsor 25–33 íteimből áll. Az elemzendő 1. vizsgarész-összetevő formailag megfelel az előírásoknak, mivel terjedelme az előírt határon belül van, 4 feladtból, összesen 33 íteimből áll, amire 33 vizsgapontot lehet kapni.

0.5. táblázat:

Az olvasott szöveg értése vizsgarész-összetevő áttekintése

Feladattípusok	Mérési cél	Szöveg	Szöveghossz ¹⁴
1. ál igaz-hamis (szelektálás)	szelektív értés	Karnevályok a világban	199 szó
2. kérdésekre válasz	részletes értés	Hotel a fán	234 szó
3. mondatok kiegészítése	részletes értés	Mesegyűjtőről	205 szó
4. igaz-hamis feladat	részletes értés	Az angol nyelv fontossága	415 szó
			össz.: 1.053 szó

A teszt egészének paramétereinek vizsgálatakor százalékpontban adjuk meg az eredményeket, hogy az eltérő itemszámú feladatok és a teszt átlaga és szórása összehasonlítható legyen (0.6. táblázat). A feladatsor a vizsgázók számára közepesen nehéz volt, a csoport összetétele heterogén volt. Ez tükröződik a megfelelő reliabilitásmutatóban is. A feladatok közül a legkönnyebb az első volt, majd ezt követi a negyedik, utána a második és végül a legnehezebb feladat a harmadik volt. Minden esetben az átlagok közötti különbség szignifikáns ($p < 0,001$). A két legnehezebb feladat megbízhatósága magas, hiszen

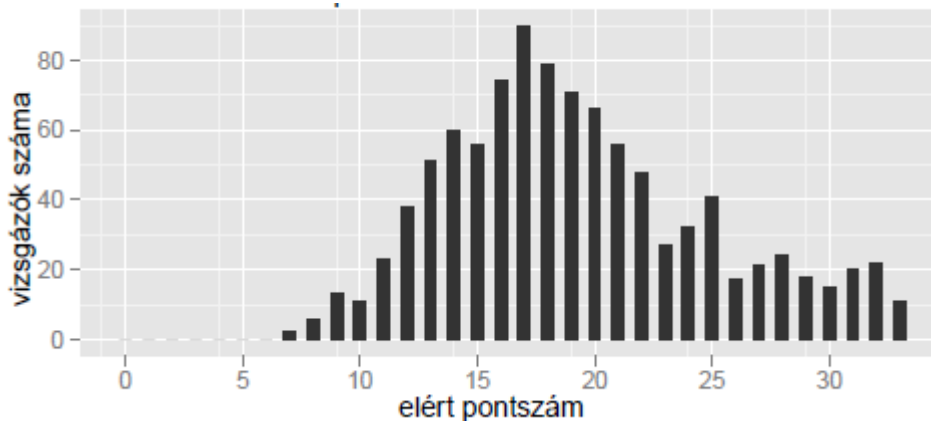
¹⁴ A szöveghosszat sem a feladatlap, sem a javítókulcs nem tartalmazza. Így a hosszúságot úgy állapítottuk meg, hogy a vizsgaszöveget átmásoltuk word fájlba, ami kimutatja a szavak számát a szövegben.

ezeknél heterogén a vizsgált minta összetétele, és ezeknek van a legnagyobb magyarázóereje.

0.6. táblázat:
Az olvasott szöveg értéke feladatsor jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Kategória	Feladat				Teszt
	1.	2.	3.	4.	
Cronbach-alfa	0,40 ¹⁵	0,75	0,74	0,30	0,82
Átlag (%p)	78,1	50,0	44,6	60,0	58,4
Szórás (%p)	16,0	27,1	26,6	20,9	17,1
Magyarázóerő (R ² %)	16,7	37,0	31,1	15,3	100

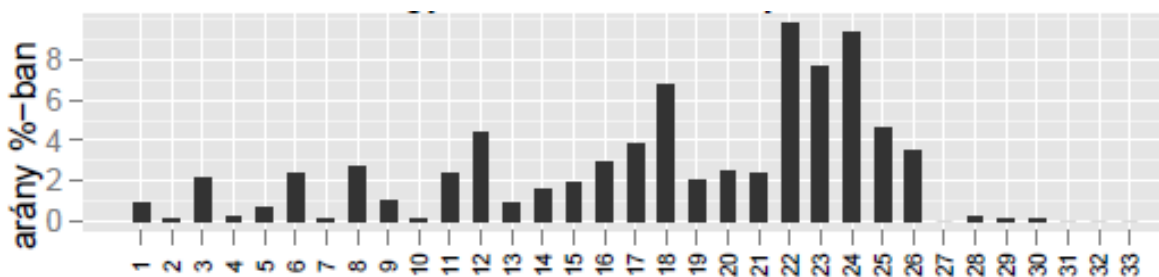
A 0.1. ábra azt mutatja, hogy az elérhető 33 feladatpontból milyen pontszámot értek el a vizsgázók a vizsgált 992 dolgozat esetében. A pontos számok az ábráról sajnos nem olvashatók le, de az látszik, hogy az eloszlás a magasabb teljesítmények felé, jobbra tolódik. A diákok többsége 17 pont (52%p) fölött teljesített.



0.1. ábra:
A teszten nyújtott pontszám abszolút gyakorisági eloszlása
(megbízó által készített grafikon)

Megbízónktól feldolgozott adatokat kaptunk, amelyek az itemeket vizsgarész-összetevőnként, illetve a három feladatlapon együttesen, nem pedig feladatszinten vizsgálják. A megbízótól kapott első adatsor a kihagyott itemek százalékos arányát mutatja az összes feladatra vonatkozóan. Ez alapján az látszik, hogy a könnyű 1. és 4. feladat itemeinél volt a legalacsonyabb a nem megválaszolt itemek aránya, a 2. és különösen a 3. feladat 22–24 iteme között fordult elő a legtöbb üresen hagyott item. Ez adódhat a feladattípusból (mindkettő nyílt végű feladat) és nehézségből. Ezt az itemanalízisnél részletesen tárgyaljuk.

¹⁵ A pirossal jelölt adatok a problémás értékekre vonatkoznak.



0.2. ábra:
Az üresen hagyott itemek relatív gyakorisági eloszlása
 (megbízó által készített grafikon)

A Rasch-modellel az egyes feladatokhoz tartozó itemek nehézségét és a mintába került diákok képességszintjét a következő minta- és itemtérképen ábrázoltuk, hogy össze tudjuk hasonlítani a diákok képességszintjét a feladatlapon szereplő itemek 50 százalékos valószínűséggel történő megoldásához szükséges képességszintekkel. Az ábráról leolvasható (0.1. ábra), hogy az adott item mennyivel könnyebb-nehezebb, illetve a vizsgált személy mennyivel magasabb-alacsonyabb képességű.

logit	személy	1.f	2.f	3.f	4.f	+item
4						
				XX		
				X		
				XXX		
3				XX		
				XXXXX		
				XX		
				XXXXX		
				XXXXXXX	19	
				XXXXXXX		
				XXX		
				XXXXXXX		
2				XXXXXXXXX		
				XXXXXXXXXX	22	
				XXXXXXXXXX	17	
				XXXXXXXXXX		24
				XXXXXXXXXX		
				XXXXXXXXXXXXXXXXXXXX	11	
				XXXXXXXXXXXXX		
				XXXXXXXXXXXXXXXXXXXX		
1				XXXXXXXXXXXXXXXXXXXX		26
				XXXXXXXXXXXXXXXXXXXX		26
				XXXXXXXXXXXXXXXXXXXX	16	33
				XXXXXXXXXXXXXXXXXXXX	15	
				XXXXXXXXXXXXXXXXXXXX	18	
				XXXXXXXXXXXXXXXXXXXX		27
				XXXXXXXXXXXXXXXXXXXX	12	
				XXXXXXXXXXXXXXXXXXXX		23
				XXXXXXXXXXXXXXXXXXXX		32
0				XXXXXXXXXXXXXXXXXXXX	8	13
				XXXXXXXXXXXXXXXXXXXX		29
				XXXXXXXXXXXXXXXXXXXX		30
				XXXXXXXXXXXXXXXXXXXX		
				XXXXXXXXXXXXXXXXXXXX	6	14
				XXXXXXXXXXXXXXXXXXXX		25
				XXXXXXXXXXXXXXXXXXXX	3	21
				XXXXXXXXXXXXXXXXXXXX		20
				XXXXXXXXXX		
				XXXXXXXXX		
-1				XXXXX	7	
				XXX		28
				XX	2	
				X	1	
					5	10
					9	
-2						
-3					4	

0.3. ábra:

Az olvasás vizsgarész-összetevő minta-ítemterképe (VT által készített ábra)

Az ábrán (0.3. ábra) minden egyes 'x' két tanulót reprezentál. Ugyanakkor az ábrázolt 'x'-eken kívül is található vizsgázók a képességskála magasabb, illetve alacsonyabb tartományában, csak kettőnél kevesebben vannak, így az ábra nem tartalmazza az adatokat.

A negatív számok átlag alatti képességet jelölnek, a 0 az átlagos, a pozitív számok az átlag feletti képességszintet mutatják. A vizsgált személyek képességében jelentős különbségek vannak, többségük az átlagos képességszint fölött helyezkedik el.

A minta-itemtérképen az egyes feladatokhoz tartozó itemek nehézségi szintjét látjuk a képességskála alapján. Megállapítható, hogy az 1. feladathoz tartozó mind a 9 item az átlagos képességszint alatt helyezkedik el, vagyis ezek az itemek jellemzőik alapján vélhetően az A2 szinten vannak. A 2. feladat itemnehézségében jelentős különbségek vannak, három item (10., 14., 13.) van a 0 logitérték alatt, a többi az e fölötti képességszintet méri. A 3. feladatban szintén található nagyon könnyű (20., 21., 25.) és nagyon nehéz itemek (24., 22., 29.), az itemek nehézségében jelentős különbségek vannak (a szórás 1,2 logit). A 4. feladat itemeinek nehézségében kisebbek a különbségek. Vagyis a teszt egy könnyebb feladattal kezdődik, de a 4. feladat nem a teszt legnehezebb feladata, ezért csak részben egyezik meg a vizsgaleírás által elvárt nehézségi sorrend az empirikus nehézség növekedésével.

A közös logitskálán ábrázolt feladatlapok minta- és itemtérképe alapján megállapítható, hogy az adott teszt itemeinek nehézsége mennyire felel meg a vizsgázók nyelvi szintjének. A mintához jól illesztett tesztnél a személyek képességparamétereit jelző 'x'-ek és az itemeket reprezentáló számok egymással párhuzamosan futnak. Ez a konkrét esetben többnyire teljesül, bár a magasabb képességszint méréséhez kevesebb item áll rendelkezésre.

A továbbiakban részletesen az egyes feladatokhoz tartozó itemeket elemezzük. Először a feladatokat jellemezzük, majd bemutatjuk az egyes itemek paramétereit és ezeket vizsgáljuk, végül a teljes feladatsorra vonatkozóan vonunk le következtetéseket.

1. Feladat: Karneval international (9 item)

Az internetes szöveg három város karneváli jellemzőit foglalja össze röviden, egyszerű nyelvezettel. Az európai, észak-amerikai és dél-amerikai városok karneváli szokásaiban közös vonások és eltérések is felfedezhetők. A szöveg hossza 199 szó. A feladat olyan kijelentéseket tartalmaz, amelyek egy vagy két városra igazak, s a vizsgázónak el kell döntenie, melyik kijelentés melyik városra igaz (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 4–5). Bár a szöveg tartalmaz néhány olyan szót, amelyek meghaladják nemcsak az A2, de a B1 szintet is (Eigentümlichkeiten, Verbände, einheimisch, toben), azonban nem gátolhatják a feladat megoldását. Maguk a kijelentések igen egyszerűek, de figyelmes olvasást kívánnak meg. A feladat nehézségi szintje nem lehet nehezebb, mint A2.

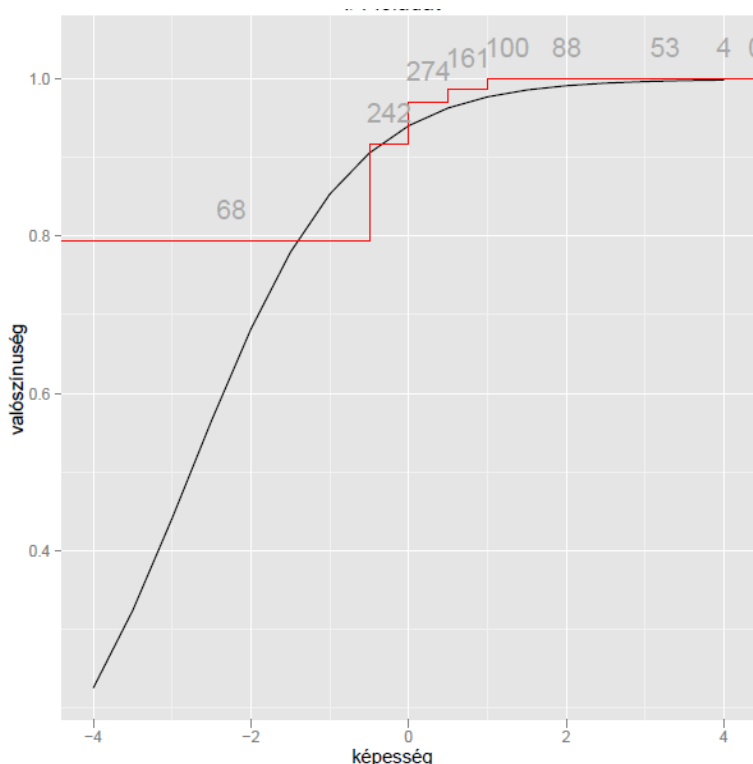
0.7. táblázat:
Az olvasás 1. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint ¹⁶
[1.]	0,9	0,83	0,38	0,25	0,06	-1,32	1,06	A2
[2.]	0,1	0,85	0,36	0,29	0,07	-1,27	1,02	A2
[3.]	2,1	0,71	0,45	0,30	0,07	-0,54	1,06	A2
[4.]	0,2	0,97	0,18	0,22	0,05	-2,87	0,97	A1/A2
[5.]	0,7	0,86	0,34	0,21	0,05	-1,58	1,05	A2
[6.]	2,3	0,67	0,47	0,29	0,07	-0,26	1,09	A2
[7.]	0,1	0,79	0,40	0,21	0,03	-0,95	1,11	A2
[8.]	2,7	0,58	0,49	0,35	0,11	0,17	1,08	A2
[9.]	1,0	0,88	0,33	0,22	0,05	-1,71	1,03	A2

A feladat valóban kifejezetten könnyűre sikerült. Az itemek nehézsége az átlagos képességszint alatt (logitérték=0) helyezkedik el. Az itemek a feladaton belül tehát könnyűek, vagyis magas az átlaguk, a diákok ezekre az itemekre 58 és 97%-ban válaszoltak helyesen. Mindegyik item pontbiszeriális korrelációja és modellilleszkedése elfogadható, egyik item kihagyása sem javítaná jelentősen a teszt megbízhatóságát. Ugyanakkor az itemek magyarázóereje nagyon alacsony, vagyis az átlagos képességszintnél lévő 8. item kivételével, súlytalanok a tesztben. Ezt a problémát már a feladatok magyarázóerőinél is láttuk. E feladat esetében pontlevonás jár abban az esetben, ha a vizsgázó az előre megadott 9 x helyett többet tesz. Ez a szankcionálás nem volt jellemző, a vizsgált mintában 13 tanulónál (1,3%) egy, 2 tanulónál (0,2%) két pont levonás történt. Összességében, könnyű feladatról van szó, amely a pontlevonás ellenére nem rontja jelentősen a teszt reliabilitását, de a többi feladathoz képest alacsonyabb mértékben járul hozzá az összpontszámhoz.

A legkönnyebb 97%-os teljesítettségű a 4. item volt a feladaton belül. A megbízó által készített 0.4. ábra is ezt támasztja alá. Az ábrán a vízszintes vonalak azt jelentik, hogy az adott tartományba eső képességszinttel rendelkező vizsgázók milyen arányban oldották meg az itemet helyesen. Ezen túl szürke számokkal jelöltük a különböző képességszinttartományokba eső vizsgázók számát.

¹⁶ A „Feltételezhető szintet” a Közös európai referenciakeret (2002) szintleírásai, a vizsgakövetelmények és vizsgaleírás, a nehézségi paraméterek, valamint saját tapasztalataink alapján határoztuk meg. Erre a későbbiekben már nem utalunk.



0.4. ábra:

Az olvasás 1. feladat 4. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

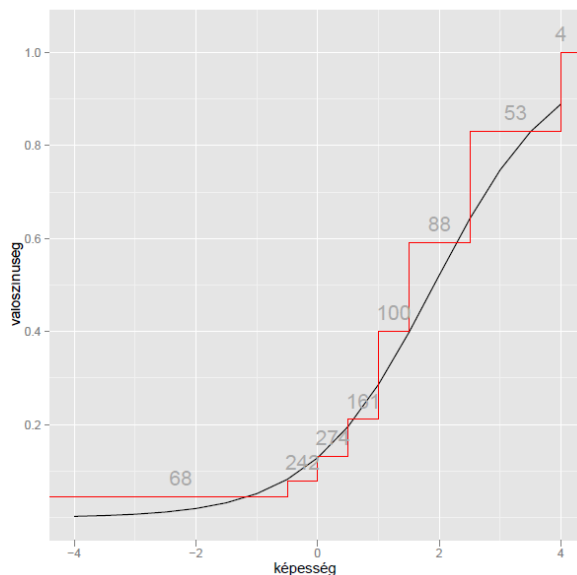
2. Feladat: Die skurrilsten Hotels: Baumhaus in Västerås (9 item)

A szöveg szintén egy internetes forrásból származik, és olyan szállodáról szól, amelyek kurióznak számítanak. A szöveg terjedelme nagyobb, mint az első feladaté, 234 szó. Természetesen a feladatot is el kell olvasni, de az már nem számít bele a szöveghosszba (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 6–7). Míg az első feladatnál csak jelölni kellett a megoldást egy x beírásával, itt válaszokat kellett adni tíz kérdésre, ahol az első kérdés a példa, ami mutatja, hogy milyen választ várnak el (rövid, nem kell teljes mondatot írni). A kérdések megválaszolásához azonban alapos olvasás szükséges, és az itemek némelyike már inkább a B1 szinten jár.

Az itemek átlaga és nehézségi paramétere alapján vélhetően 6 item van B1 szinten (0.8. táblázat). Ezekre az itemekre az jellemző, hogy a vizsgázóknak összefüggéseket kellett a szöveg alapján megadniuk, míg a könnyebb itemeknél (10., 13. és 14.) az adott információt kellett a szövegből visszakeresni és többnyire módosítás nélkül válaszként megadni. Ezeknek az itemeknek a magyarázóereje is alacsonyabb. Érdekes módon a 17. item megoldása szó szerint benne van a szövegben, de bonyolultabb szöveggörnyezetben, ezért itt ez volt a legnehezebb item, amelynek modellilleszkedése megfelelő, ahogyan a 0.5. ábra is látható.

0.8. táblázat:
Az olvasás 2. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[10.]	0,1	0,88	0,33	0,35	0,14	-1,56	0,94	A2
[11.]	2,3	0,34	0,47	0,62	0,40	1,46	0,83	B1
[12.]	4,4	0,55	0,50	0,54	0,28	0,38	0,90	B1
[13.]	0,9	0,61	0,49	0,47	0,20	0,09	0,92	A2
[14.]	1,6	0,67	0,47	0,46	0,18	-0,21	0,97	A2
[15.]	1,9	0,47	0,50	0,45	0,16	0,69	0,99	B1
[16.]	2,9	0,46	0,50	0,61	0,37	0,81	0,86	B1
[17.]	3,8	0,26	0,44	0,52	0,24	1,88	0,89	B1
[18.]	6,8	0,49	0,50	0,36	0,10	0,57	1,04	B1



0.5. ábra:
Az olvasás 2. feladat 17. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

A feladaton belül a mérni kívánt olvasási képességet a magyarázóerő alapján leginkább a 11. és a 16. item méri, mindegyik B1 szinten van. Ezen itemek jellemzője, hogy kapcsolatokat kellett a szövegben megállapítani, és a választ a szöveg megfelelő helyének megtalálását követően kellett megadni.

3. Feladat: Sie erzählt und sammelt Märchen (8 item)

A GEOlino-ról származó, 205 szó hosszúságú szöveg egy mesekutató nő munkásságáról szól. Ezt egy olyan feladat követi, amely lyukas mondatokból áll. A lyukakat csak úgy lehet kitölteni, ha a vizsgázó a szöveget elolvasta és megértette (Német nyelv, középzin-

tű írásbeli érettségi vizsga, feladatlap 2012: 8–9). A szöveg nehézségi szintje B1, bár témája nem kifejezetten a hétköznapi életről szól.

0.9. táblázat:

Az olvasás 3. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

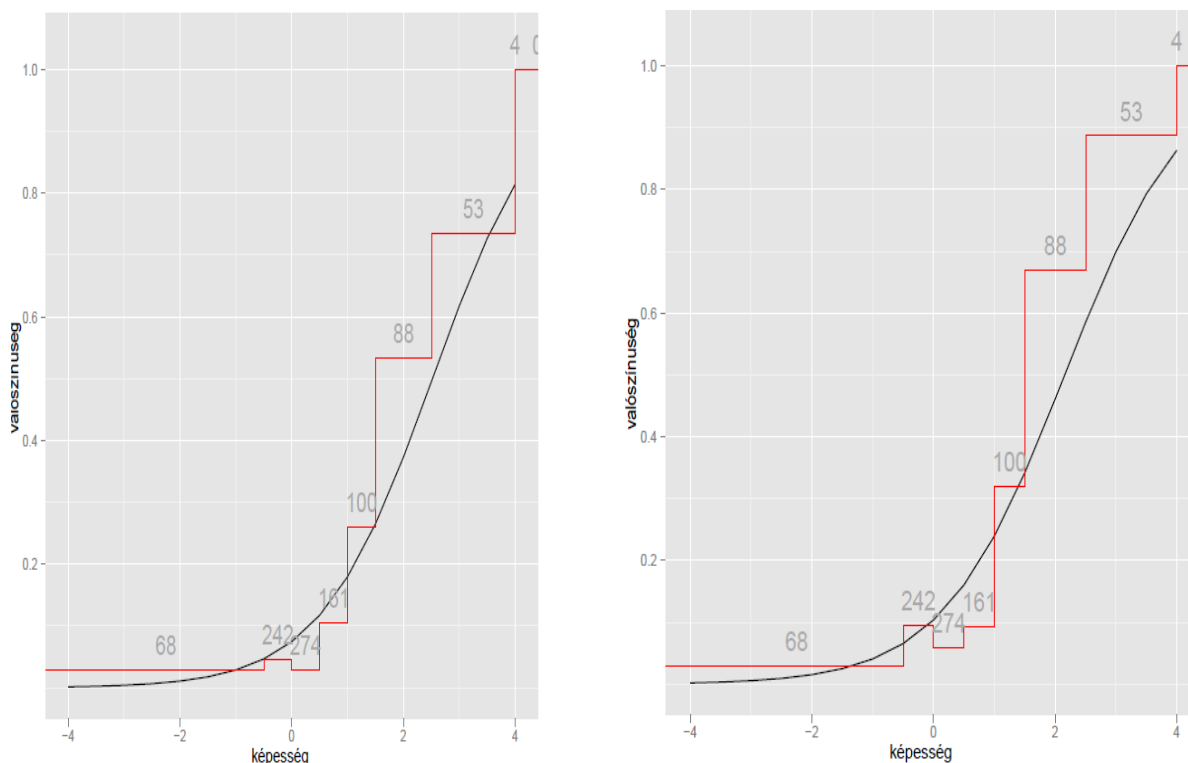
Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[19.]	2,0	0,18	0,38	0,54	0,27	2,49	0,82	B1/B2
[20.]	2,4	0,74	0,44	0,42	0,13	-0,62	0,94	A2
[21.]	2,3	0,73	0,44	0,43	0,14	-0,55	0,94	A2
[22.]	9,8	0,24	0,42	0,56	0,29	2,04	0,83	B1/B2
[23.]	7,7	0,55	0,50	0,47	0,20	0,31	0,97	B1
[24.]	9,4	0,27	0,44	0,54	0,27	1,77	0,88	B1
[25.]	4,6	0,69	0,46	0,38	0,13	-0,40	1,02	A2
[26.]	3,5	0,40	0,49	0,51	0,27	1,05	0,92	B1

Az itemek közül öt helyezkedik el az átlagos képességszint fölött. Ezek közül a 19. item volt nagyon nehéz (lásd. 0.6. ábra), amelyet a diákoknak csak 18%-a válaszolt meg helyesen. A nehézség okát csak a tanulói válaszok alapján lehetne megállapítani, feltételezhető azonban, hogy a diákok nem a javítókulcsban szereplő „Sammlerin”, „Forscherin” szavakat írták, hanem a szövegből származó „Märchen- und Sagenforscherin” kifejezést, amely a mondatba nem illik, és amelynek elfogadhatóságáról a javítókulcs így tájékoztatja a tanárokat: „... elfogadhatók azok a variánsok is, amelyek tartalmilag megegyeznek ezekkel. A megoldásokban előforduló nyelvi vagy helyesírási hibákat figyelmen kívül kell hagyni, amennyiben a válasz érthető” (Javítási-értékelési útmutató, 2012: 3.). Azt azonban nem tudjuk, hogy az értékelő tanárok ezt a javítási elvet hogyan alkalmazták az értékeléskor. Az item nehézsége alapján ugyanakkor azt feltételezzük, hogy nem vették figyelembe.

A teszt második legnehezebb iteme (22.) szintén ebben a feladatban van, ahol a „vor dem Aussterben bewahren”¹⁷ szerkezet megértése alapján kellett a választ módosítva beírni a mondatba. A 24. itemnél – bár nem kellett módosítani a szövegben szereplő szóalakon –, de itt következtetési művelettel kellett a diákoknak a feladatot megoldani. A 26. item nehézségét és alacsony megoldottságát az okozhatta, hogy ennél több ruhadarab (Seidenblusen, Rock, Hut) közül kellett a leginkább megfelelőt kiválasztani. Az A2 szintű itemekre főként az jellemző, hogy az adott szóalakon nem kellett változtatnia a diáknak. A 20. és 21. item ugyanarra a szöveghelyre vonatkozik, és minden paraméterben hasonlíthatnak egymásra, vagyis ezek az itemek egymással összefüggnek. Összességében ez volt

¹⁷ A kifejezés nem szerepel a „Profile deutsch” (Glaboniat és mtsai, 2006) katalógusban, valószínűleg meghaladja a B1 szintet.

a teszt legnehezebb feladata, az itemek nehézségét pedig az okozhatta, hogy a vizsgázóknak az adott szót a megfelelő mondatkörnyezetbe írva kellett megadnia.



0.6. ábra:

Az olvasás 3. feladat 19. és 21. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

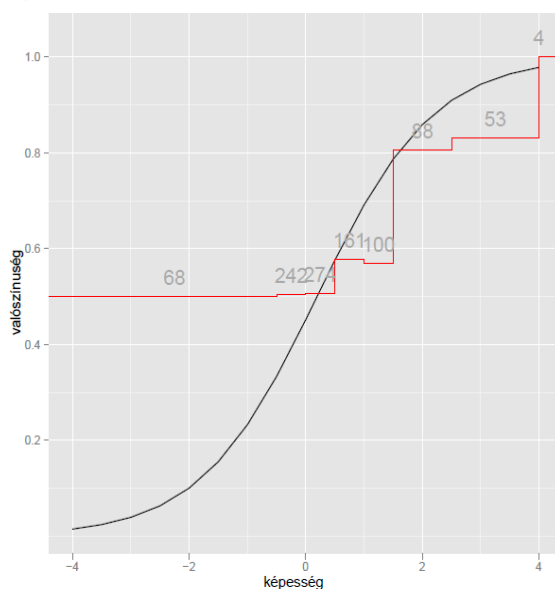
4. Feladat: Ohne Englisch geht im Job fast nichts mehr (7 item)

Az elolvasandó szöveg tipikus újságszöveg: taglalja, hogy milyen tendenciák figyelhetők meg a tanulandó nyelvek választásában, a választás okainál, a tanulási motivációkban. Ez a feladatsor leghosszabb szövege, 415 szóból áll. Ezt követi az igaz-hamis feladat (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 10–11). A feladattípus a könnyűek közé tartozik, hiszen 50%-os eséllyel lehet találgatni. De felületes olvasással könnyű elrontani. Minden biztonnyal azért választották a könnyebb feladattípust, hogy ellensúlyozzák a szöveg komolyabb terjedelmét és nehézségi fokát. A feladat szintjét B1 szintre tervezték, ugyanakkor a feladattípusból adódóan több item inkább A2 szintű. Ebben a feladatban a vizsgázók mindenhol adtak választ.

0.10. táblázat:
Az olvasás 4. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[27.]	0,0	0,52	0,50	0,19	0,02	0,51	1,20	B1
[28.]	0,2	0,81	0,39	0,27	0,06	-1,08	1,03	A2
[29.]	0,1	0,60	0,49	0,24	0,04	0,10	1,15	A2
[30.]	0,1	0,63	0,48	0,27	0,04	-0,06	1,12	A2
[31.]	0,0	0,70	0,46	0,33	0,09	-0,39	1,04	A2
[32.]	0,0	0,57	0,50	0,23	0,04	0,19	1,16	B1
[33.]	0,0	0,43	0,50	0,31	0,08	0,87	1,12	B1

Az itemek magyarázóereje rendkívül alacsony, amely problémát feladatszinten is láthattuk korábban (0.6. táblázat). Az itemek közül a 27., 29. és 32. item problematikus, mert ezek alacsony pontbiszeriális korrelációval és magyarázóerővel rendelkeznek, és nem megfelelő a modellilleszkedésük. A 27. itemnél feltehetően az a mondat zavarhatta meg a vizsgázókat, amely a kisebbségi álláspontot írja le, hogy a spanyol és a kínai lesz a következő világnyelv. Egy tesztfejlesztéskor indokolt ennek az itemnek a cseréje vagy módosítása. A 29. itemnek is alacsony a pontbiszeriális korrelációja a magyarázóereje. Itt vélhetően az a zavaró, hogy a vizsgázónak összehasonlítást kellett tennie. A 32. item esetében pedig feltehetően az item megfogalmazása nem volt kellőképpen pontos. Az ábrán példaként ennek az itemnek a karakterisztikus görbéjét választottuk (0.5. ábra), amely alapján jól látszik, hogy a tapasztalati és elméleti görbéje nem fedi egymást, vagyis itt a találgatásnak is nagy szerepe lehetett. Erre utal a magas infit paraméter is.



0.7. ábra:

Az olvasás 4. feladat 32. item elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

Összegezés: a középszintű vizsga 1. részének (olvasott szöveg értése) feladatsora megfelelő megbízhatósággal mért, többnyire megfelelt a vizsgázók képességszintjének, a feladatírói szándéknak megfelelően az első feladat volt a legkönnyebb. A 4. feladat viszont az elvártnál könnyebb volt, amely a zárt végű feladattípus jellegéből adódik. A teszten belül a 2. és a 3. feladat járult leginkább az összpontszámhoz, ezek megbízhatóan mértek, és ezekben a feladatokban volt a legtöbb

B1-es szintű és magas magyarázóerővel rendelkező item. A 3. feladat nehézségét leginkább az okozta, hogy az olvasott szövegben található válaszokat a diákoknak a megfelelő mondatkörnyezetbe integrálva kellett megadniuk. Ebben a két feladatban általában azok az itemek voltak A2 szintűek, amelyek információ visszakeresését igényelték, és a tanulóknak a választ a szöveg alapján kellett kimásolni. Azok voltak B1-es szintűek, amelyek következtetési művelettel oldhatók meg és/vagy a tanulóknak a válaszokat a mondatkörnyezetbe integrálva kellett megadniuk. A harmadik feladatban két item feltételezhetően már inkább B2 szinten van, amely egyrészt adódhat értékelési problémából, másrészt pedig feladatírói problémából, hiszen a megértendő szerkezet B1 szint fölött van, valamint a tanulóknak a válaszuk megadásakor módosítani kellett a szerkezeten.

2.2. Nyelvhelyesség (II. vizsgarész-összetevő)

A vizsgarész-összetevő célja annak mérése, hogy a vizsgázó rendelkezik-e azokkal a lexikai, grammatikai, szemantikai és pragmatikai ismeretekkel, amelyek képessé teszik a vizsgázót az önálló kommunikációra (40/2002 OM rendelet, Vizsgaleírás: 3). A nyelvhelyesség feladatai szövegalapúak, azaz a lexikai/nyelvtani jelenséget szövegkörnyezetben kell felismerni. Ez a vizsgarész-összetevő tehát nemcsak nyelvtani jelenségekre kérdez célzottan, hanem az alapvető szókincs ismeretére is. Az érettségi vizsgaszabályzat részletes követelmény-leírása táblázatban foglalja össze, hogy mely nyelvtani jelenségekre lehet célzottan rákérdezni: „A nyelvtani ismeretek alapvető szerepet játszanak mind a receptív, mind a produktív készségekben. A *receptív* készségek között felsorolt nyelvtani szerkezetek előfordulhatnak az olvasott vagy hallott szöveg értését mérő feladatokban, de ezekre a nyelvhelyességi rész nem kérdez rá. A *produktív* ismeretek közé sorolt nyelvtani szerkezetekre lehet feladatot készíteni a Nyelvhelyesség részben” (40/2002 OM rendelet, Vizsgakövetelmények: 14). A táblázat megkülönbözteti a közép- és emelt szintet, s felsorolja, hogy a megnevezett nyelvtani fogalmakat receptív vagy produktív szinten kell-e ismernie a tanulónak. A középszinten belül azonban nem tesz különbséget, hogy az adott item, ill. feladat az A2 vagy a B1 szintnek felel-e meg. Ezért itemenként kell megnézni, hogy azok melyik nyelvi szinten várhatóak el. A feladatírás egyik szempontja középszinten az, hogy az adott vizsgarész-összetevő első fele (egy vagy két feladat) inkább könnyebb, azaz A2 szintű legyen, míg a másik fele nehezebb, B1 szintű. Mivel ennek megállapítása csak a vizsgakövetelményekre és leírásra támaszkodva nem lehetséges, ezért segítségül kell hívni a „Profile deutsch” (Glaboniat és mtsai, 2006) c. programot, melynek segítségével részletesen megvizsgálható, hogy az itemek melyik szinten várhatóak el. Ennek tükrében érdemes megvizsgálni az itemek és feladatok megoldottságát is.

0.11. táblázat:

A nyelvhelyesség feladatsor jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

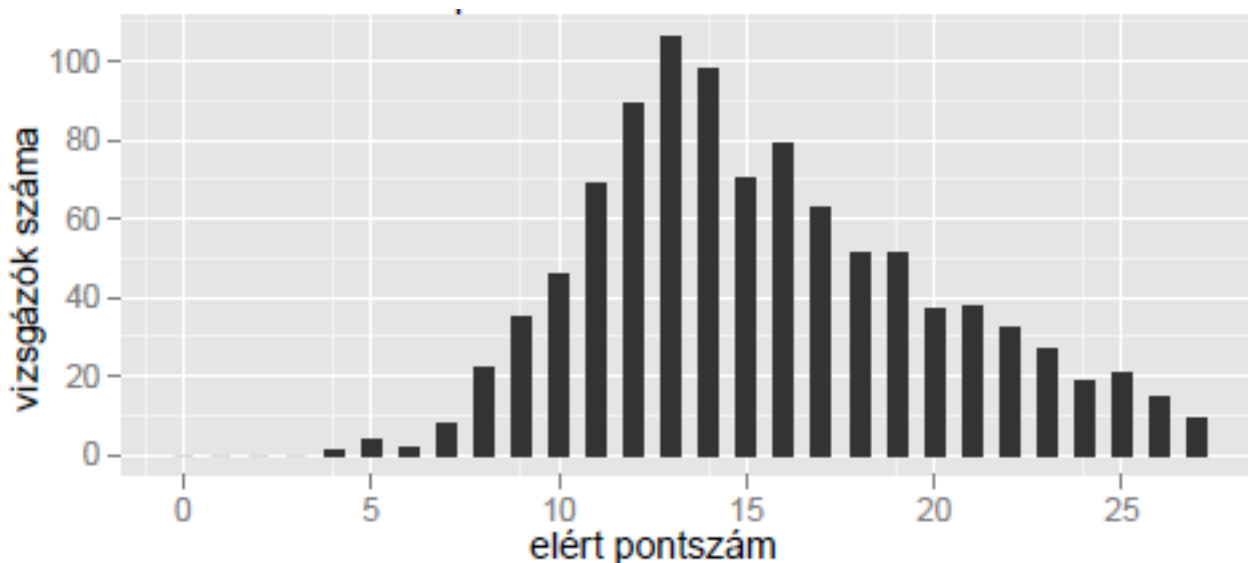
Kategória	Feladat				Teszt
	1.	2.	3.	4.	
Cronbach-alfa	0,41	0,36	0,55	0,74	0,77
Átlag (%p)	82,8	42,1	54,7	47,4	57,4
Szórás (%p)	16,6	21,3	22,8	34,4	17,0
Magyarázóerő (R ² %)	15,8	23,1	31,4	29,7	100

A 2. vizsgarész-összetevő feladatsora ezen vizsgakövetelmények elvárásait – úgy tűnik – nem egészen tudta megvalósítani. Az olvasáshoz viszonyított alacsonyabb, de még elfogadható Cronbach-alfa érték csak részben magyarázható az alacsonyabb itemszámmal, ebben az esetben a szórás is alacsonyabb. A feltűnően könnyű 1. feladat után csak a 3. feladatot sikerült valamivel több, mint a minta felének megoldani, a 2. és 4.

feladat megoldottsága azonban 50 % alatt maradt. Minden esetben az átlagok közötti különbség szignifikáns ($p < 0,001$).

A négy feladat közül a legalacsonyabb magyarázóerővel az első és a második feladat rendelkezik, amelyek szórása is alacsonyabb, vagyis a teszten belül az utolsó két feladat mért megbízhatóbban.

A megbízó elkészítette a 992 dolgozatban elért feladatpontok összesítését. Az ábráról leolvasható (0.8. ábra), hogy a vizsgázók többsége 13–14 pontot ért el (52% pont), és a teljesítmények jobbra, a magasabb teljesítménysávba tolnak.



0.8. ábra:
A teszten nyújtott pontszám abszolút gyakorisági eloszlása
(megbízó által készített grafikon)

A skála 27 pontra terjed ki, ami az itemekre adható feladatpontokat jelenti, ám ezt át kell számítani vizsgapontra, amennyiben a végén összesített eredményt szeretnénk kapni. Ezeket a vizsgapontokat azonban az adatbázisban nem rögzítették. A Javítási és értékelési útmutató (7. o.) tartalmazza az átszámítási táblázatot:

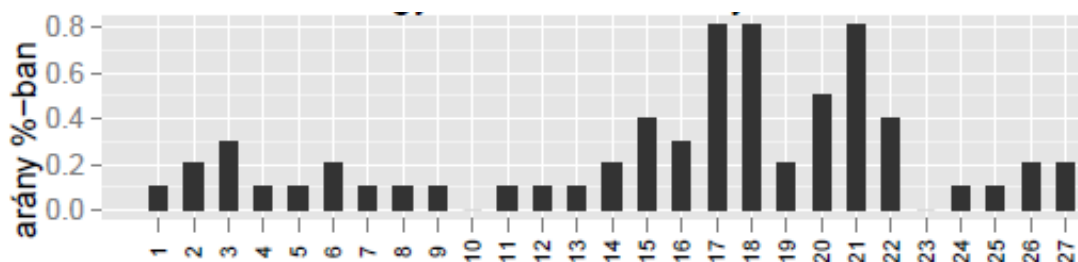
0.12. táblázat:

A javítókulcsban szereplő átszámítási táblázat

Feladatpont	Vizsgapont
27	18
26	17
25	17
24	16
23	15
22	15
21	14
20	13
19	13
18	12
17	11
16	11
15	10
14	9

Feladatpont	Vizsgapont
13	9
12	8
11	7
10	7
9	6
8	5
7	5
6	4
5	3
4	3
3	2
2	1
1	1

A 0.9. ábra azt mutatja, hogy ebben a vizsgarész-összetevőben milyen arányban hagytak üresen itemeket a mintavétel eredményeként kiválasztott vizsgázók. Ennek alapján megállapítható, hogy az itemek többségére válaszoltak a vizsgázók, még a nyílt végű [3. feladatban](#) is csak a vizsgázók 0,8%-ánál (8–10 tanuló) hiányzik a válasz.



0.9. ábra:

Az üresen hagyott itemek relatív gyakorisági eloszlása
(megbízó által készített grafikon)

Az itemekről a legjobb áttekintést a minta-itemtérkép adja (0.10. ábra), ami a mintában résztvevő diákok képességszintjét és a négy feladatban szereplő itemek 50 százalékos valószínűséggel történő megoldásához szükséges képességszinteket mutatja. Az ábrán minden 'x' két vizsgázót jelöl.

A minta-itemtérképről leolvashatjuk, hogy az első feladat, ami mindösszesen 7 itemből áll, túl könnyű volt, az összes item a 0 logitérték alatt szerepel. Ezen belül is három item (1., 3., 5.) akár el is hagyható lenne, hiszen igen messze került az átlagtól, nem befolyásolja a feladat megbízhatóságát. A 2. feladat, aminek még szintén relatív könnyűnek kellett volna lennie, jóval nehezebb lett. Itt a legkönnyebb item (11.) pont a 0 logittal esik egy vonalra, a többi mind föllette van, és ebben a feladatban található az egész feladatsor legnehezebb iteme is (8.). A 3. és 4. feladat itemei már közelebb kerültek egymáshoz és a mérni szánt képességszintekhez, bár a 4. feladat itemei is mind a 0 érték fölött helyezkednek el. Összességében a feladatok itemei a három legkönnyebb kivételével többnyire megfelelően fedik le a vizsgázók képességszintje alapján meghatározott intervallumot.

logit	személy	1.f	2.f	3.f	4.f	+item
4						
3						
2						
1						
0						
-1						
-2						
-3						

0.10. ábra:
A nyelvhelyesség vizsgarész-összetevő minta-itemtérképe
(VT által készített ábra)

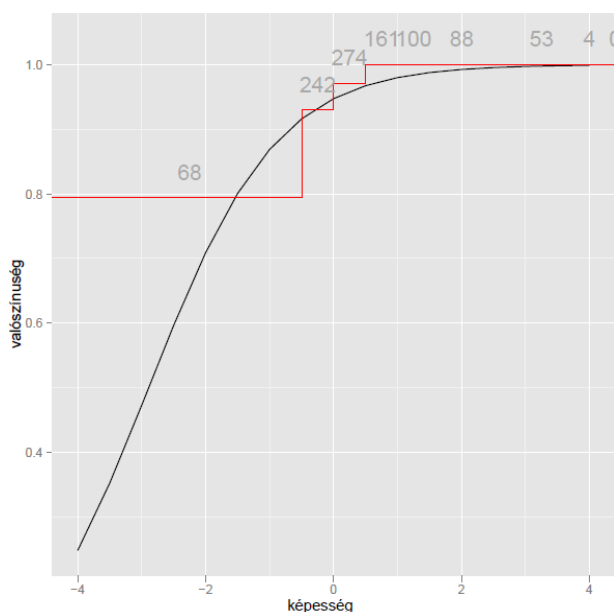
1. Feladat: Brieffreundschaften (7 item)

Az első feladat két rövid bemutatkozó szövegre épül, amelyben a megszólaló levelezőtársat keres (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 14). Ebből adódóan a két szöveg mind tartalmában, mind nyelvi megformálásában igen egyszerű. A két szöveghez egy feladat társul, melynek típusa feleletválasztó. Ez a feladattípus a könnyűek közé tartozik, mivel 25%-os eséllyel lehet találgatni. Ám annál nehezebb jó feladatot készíteni.

0.13. táblázat:
A nyelvhelyesség 1. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[1.]	0,1	0,94	0,23	0,22	0,04	-2,57	0,97	A1
[2.]	0,2	0,71	0,45	0,30	0,07	-0,57	1,04	A1
[3.]	0,3	0,93	0,26	0,29	0,08	-2,23	0,93	A1
[4.]	0,1	0,72	0,45	0,26	0,04	-0,59	1,08	A2 (rec.) B1 (prod)
[5.]	0,1	0,96	0,19	0,23	0,05	-3,05	0,92	A1
[6.]	0,2	0,81	0,40	0,29	0,05	-1,13	1,00	A1
[7.]	0,1	0,76	0,43	0,39	0,14	-0,86	0,99	A2

A 0.13. táblázat alapján megállapítható, hogy a feladat egésze alatta van a középszintű vizsga követelményeinek, hiszen a 7 itemből 5 az A1 szinten van. Ezt alátámasztják az itemnehézség adatai, amelyek mind negatív tartományban vannak, s ez arra utal, hogy átlag alatti képességszinten 50%-os megoldási valószínűségű itemek találhatóak a feladatban. Az adatokból az is látható, hogy az első 7 item megoldottsága magasabb, mint a többié. Ezen belül is az 5. item megoldottsága a legmagasabb (lásd. II.11. ábra). Amire az item célzottan rákérdez, azt a kezdő nyelvtanulók legkésőbb a harmadik órájukon megtanulják tehát, semmilyen kihívást nem jelent. Az itemek az utolsó kivételével súlytalanok a tesztben, ezek nagyon alacsony magyarázóerővel rendelkeznek.



0.11. ábra:

A 2. vizsgarész 1. feladat 5. itemének elméleti és tapasztalati megoldási valószínűségét jelölő

2. Feladat: Sibirien-Reise im Hundeschlitten (7 item)

A második feladat is feleletválasztó. A szöveg, amire a feladat épül, egy rövid és egyszerű hír a GEOlino-ból (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap

2012: 15). Ez a feladat már egyértelműen nehezebb. A 7 itemből háromnál lehet egyértelműen az A2 szintet beazonosítani, két item ennél könnyebb, A1 szintet képvisel, míg a másik kettő azért problémás, mert a vizsgakövetelményben középszinten ugyan követelményként megjelenik, ez azonban nem egyezik a Közös Európai Referenciakeretre épülő „Profile deutsch” listájával. Ez a két item biztosan a nehezebbek közé tartozik.

0.14. táblázat:

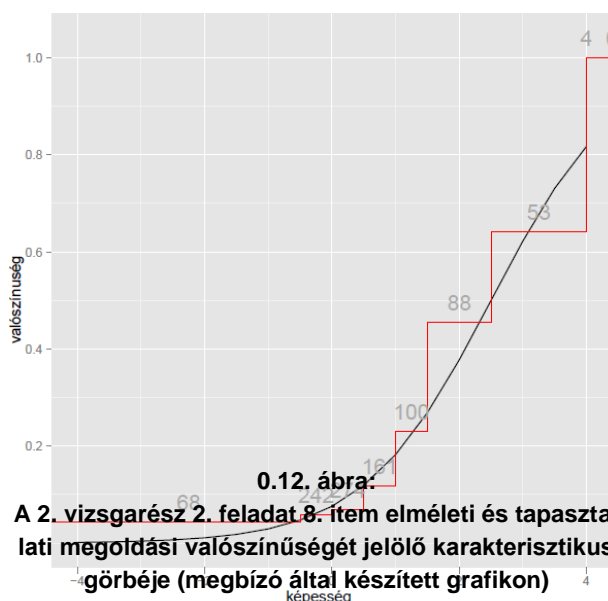
**A nyelvhelyesség 2. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)**

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[8.]	0,1	0,16	0,37	0,44	0,19	2,39	0,93	A2
[9.]	0,1	0,45	0,50	0,45	0,14	0,70	1,01	B1/B2
[10.]	0,0	0,49	0,50	0,30	0,05	0,52	1,11	A1
[11.]	0,1	0,64	0,48	0,16	0,01	-0,17	1,18	B1/B2
[12.]	0,1	0,57	0,50	0,37	0,10	0,13	1,04	A1
[13.]	0,1	0,40	0,49	0,34	0,10	0,98	1,08	A2
[14.]	0,2	0,26	0,44	0,24	0,04	1,72	1,15	A2

A 0.14. táblázatban a 11. item azért lóg ki az itemek sorából, mert a pontbiszerialis korrelációja alacsonyabb 0,2-nél, azaz az item nem azt méri, amit a teszt egésze, vagy nem jól méri azt. Ráadásul ennek az itemnek a teszten belül a legalacsonyabb magyarázóereje. Ez az item melléknévragozást vizsgál, ami köztudottan nehezen megy a tanulók számára. Ennek a modellilleszkedése sem megfelelő, a 14. itemhez hasonlóan. Ami a megoldottságot illeti, rögtön a feladat első iteme (8.) tűnik ki alacsony értékével, amelynek ugyanakkor modellilleszkedése megfelelő (lásd. II.12. ábra). Ha összevetjük a táblázat adatait a Közös Európai Referenciakeret szintleírásával, akkor furcsának tűnhet, hogy a 8. item volt az egész feladatsor legnehezebb iteme, ami a Referenciakeret szerint A2 szinten van. Ennek magyarázatát csak akkor tudnánk megadni, ha tudni lehetne, hogy a helyes megoldás helyett mit választottak a mintában részt vevő vizsgázók. Itt egyértelműen hiányzott a disztraktorelemzés lehetősége. De elképzelhető, hogy a c) disztraktor volt a zavaró, ami az idegen (francia) névből vezethető le. Ha ugyanis a diák nem figyelte a nyelvtani nemre, akkor könnyen félrevezethette a „Nicolas” név. Ennek tanulsága pedig az, hogy nem szabad olyan itemet készíteni, ami tulajdonnévre kérdez, ill. azzal kell egyeztetni.

3. Feladat: Warum Vegetarier kein Fleisch essen (8 item)

A szöveg jellege egyszerű ismeretterjesztő a vegetarianizmus kialakulásának okairól. A szöveg egyben maga a feladat is, ugyanis



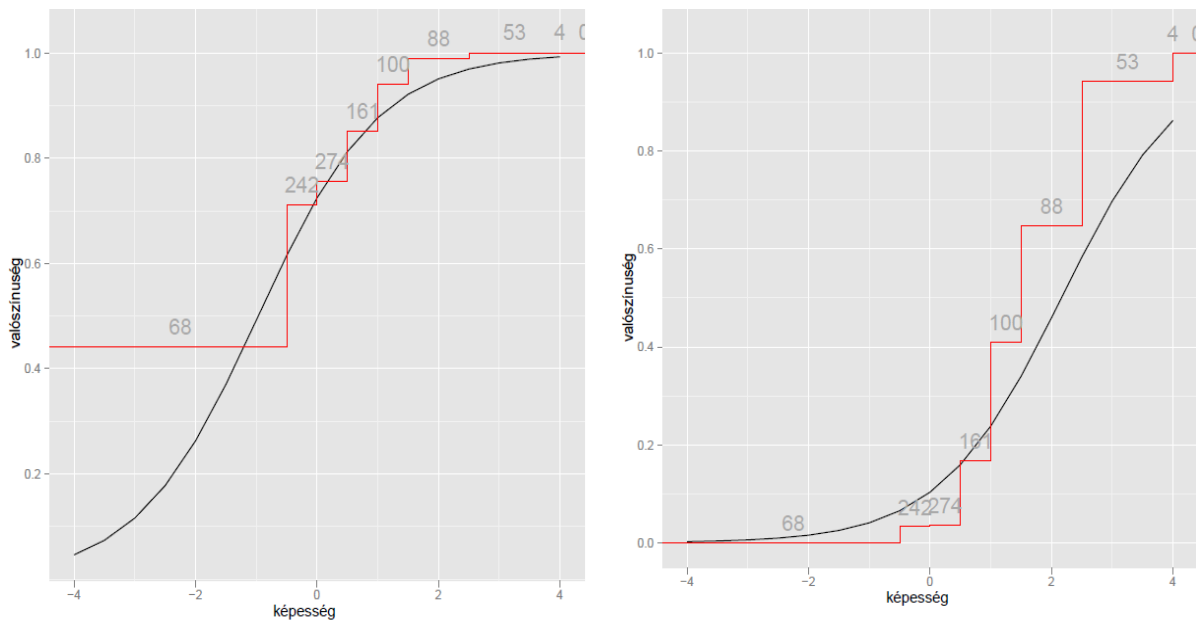
lyukas szövegről van szó (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 16). A feladat tartalmazza a lyukakból hiányzó szavakat. Ezek szófajlag mind igék. A feladat mérési célja az, hogy a vizsgázó képes-e az alany-állítmány megfelelő egyeztetésére, a szövegkörnyezetben használandó igeidő és igemód felismerésére és használatára. Az itemek nehézségi foka itt is meglehetősen vegyes. A feladatot azért lehet a B1 szintre helyezni, mert a 3 valóban B1 szintű item mellett az a tény, hogy a szövegbe ágyazottan kell megoldani a feladatot, már komplexebb, mint a feleletválasztás. A négy feladat közül ebben volt az üresen hagyott itemek legmagasabb aránya, bár ez a 17. és 18. itemnél mindössze 8 diákot jelentett.

0.15. táblázat:

A nyelvhelyesség 3. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[15.]	0,4	0,51	0,50	0,21	0,02	0,42	1,19	A2
[16.]	0,3	0,80	0,40	0,36	0,10	-1,07	0,96	A1
[17.]	0,8	0,20	0,40	0,62	0,39	2,06	0,82	B1
[18.]	0,8	0,47	0,50	0,34	0,09	0,59	1,09	B1
[19.]	0,2	0,79	0,41	0,35	0,09	-1,01	0,96	A1
[20.]	0,5	0,48	0,50	0,46	0,17	0,59	0,97	A1
[21.]	0,8	0,53	0,50	0,46	0,17	0,30	0,96	A1
[22.]	0,4	0,63	0,48	0,33	0,08	-0,14	1,06	B1

Ebben a feladatban is rögtön az első itemnél (15.) látható probléma, mivel az infit MNSQ mutató szerint ez kevésbé különbözteti el a különböző képességszintű tanulókat, valamint ennek alacsony a pontbiszeriális korrelációja és a magyarázóereje. A legkönnyebb item (16.) egy alapvető ige ragozatlan alakja volt, ugyanabban az alakban kellett beírni, mint ahogy az igét megadták. Nem csoda, hogy ezt az itemet 80%-ban jól oldották meg. Ezen kívül volt még egy magas, 79% megoldottságú item (19.) is, ami egy igen alapvető kifejezés felismerése volt. A másik végletet a 17. item mutatja, amit csak 20%-ban sikerült megoldani. Itt egy múlt idejű alakot kellett képezni egy elvált igekötős igéből. Ugyanakkor ennek az itemnek a legmagasabb a magyarázóereje, vagyis ez méri a legjobban a vizsgált konstruktumot. A 0.13. ábra e két item karakterisztikus görbét mutatja. A többi item megoldottsága 50% körül mozgott.



0.13. ábra:

A 2. vizsgarész-összetevő 3. feladat 16. (legkönnyebb) és a 17. (legnehezebb) itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

4. Feladat: Auf die sanfte Tour (5 item)

A rövid szöveg egy anya és fia közti esetet mesél el egyszerű köznyelven (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 17). A szövegben összetett mondatok vannak. A hozzárendelés feladat abban áll, hogy a szövegből kivett tagmondatok típusát felismerve azokat a megfelelő helyre tegye a vizsgázó. Ehhez fel kell ismernie, hogy a) a hiányzó rész főmondat-e vagy mellékmondat, b) ha mellékmondat, akkor milyen típusú: mellérendelt vagy alárendelt (ezt a szórendből lehet megállapítani), és c) mi a funkciója (ezt a tartalomtól és a kötőszókból lehet megállapítani). A mérési cél tehát a feladaton belül egységes.

0.16. táblázat:

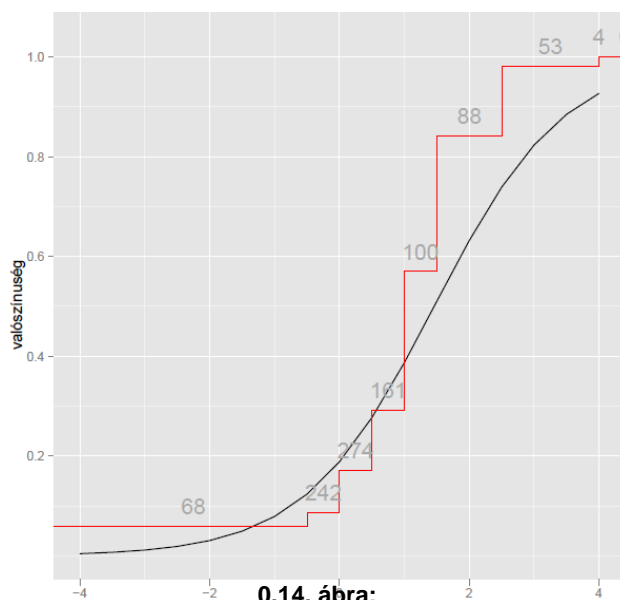
A nyelvhelyesség 4. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[23.]	0,0	0,55	0,50	0,55	0,28	0,24	0,91	A2
[24.]	0,1	0,32	0,47	0,63	0,36	1,39	0,82	A2
[25.]	0,1	0,54	0,50	0,52	0,22	0,28	0,92	A2
[26.]	0,2	0,44	0,50	0,55	0,29	0,75	0,88	A2
[27.]	0,2	0,54	0,50	0,53	0,24	0,30	0,90	A1

Az itemek nehézségi szintje az előző feladathoz képest könnyebbnek tűnik, az egyes mondattípusok – egy kivételével – az A2 szintet jelentik. Az egy kivétel pedig még ennél is egyszerűbb, A1 szinten van. Az egyetlen B1 nehézségű mondat maga a példa (Finalsatz), és a disztraktor is A1 szintet képvisel. A feladat nehézsége inkább abból adódik, hogy a mellékmondatok megtalálásával az egész szöveg koherenciáját kell helyreállítani.

A 0.16. táblázatból az derül ki, hogy a mindössze öt itemet tartalmazó feladat tűnik a legjobb mérőfeladatnak az egész feladatsorból. A diákok teljesítménye itt a legkiegyensúlyozottabb. A 24. item kivételével a többi 50% körül sikerült megoldani. Bár nem a legnehezebb feladat volt, de a csoport heterogén összetételéhez viszonyítva kellő nehézségű, nagyobb szórással. Nincsenek benne igazán kiugró eltérések, így ez a feladat volt a legmegbízhatóbb, az ebben található itemek magas magyarázóerővel rendelkeznek.

Összegzés: A nyelvhelyesség vizsgarész-összetevő 3. és 4. feladata megbízhatóbb volt az 1. és 2. feladatnál. Az 1. feladat messze alatta maradt a vizsgázók képességszintjének, a mérési célt is elvétette, a teszen belül minden item súlytalan volt. Az 1. és 2. feladat azonos feladattípus volt, mégis nagy volt a szintbeli különbség. A feladat megbízhatósága tehát ebben az esetben nem a teszt feladattípusától függ, hanem az abban alkalmazott itemektől. Középszinten érvényesülni kellene a fokozatosság elvének, azaz az első két feladatnak könnyebbnek, míg a második kettőnek nehezebbnek kellett volna lennie. Ebben a feladatsorban azonban a 2. feladat lett a legnehezebb. Az első két esetben a feladattípus igen gyakori a vizsgafeladatok között, többnyire bevált, így a problémák okát csak akkor tudtuk volna feltárni, ha kiderült volna, hogy mely disztraktorokat választották a vizsgázók. A feleletválasztós feladatoknál ugyanis a rosszul megválasztott disztraktorok is könnyen félreviszik a vizsgázót.



0.14. ábra: A 2. vizsgarész 4. feladat 24. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

2.3. Hallott szöveg értése (III. vizsgarész-összetevő)

A vizsga célja annak mérése, hogy a vizsgázó képes-e megérteni az adott nyelvterületen általánosan elfogadott nyelvhasználatról nem, vagy csak kissé eltérő anyanyelvi beszédet az értési céloknak megfelelően. A vizsgafeladatok megoldásához a vizsgázónak képesnek kell lennie az értési céloknak megfelelő stratégiák alkalmazására. A teszt feladattípológiai-

lag kötetlen szerkezetű, azaz bármelyik, a vizsgaleírásban felsorolt feladattípus szerepelhet benne. A feladatsor a szövegértés alábbi részkészségeit méri: globális, szelektív és részletes értés. A feladatsor 2–3 szövegből és 2–3 feladatból áll. Egy feladaton belül egy hosszabb szöveg helyett előfordulhat több rövidebb szöveg is. A feladatokhoz felhasznált szövegek együttes terjedelme 5–8 perc. A feladatsor 18–25 itemből áll (40/2002 OM rendelet, Vizsgaleírás: 5). Ezeket a kritériumokat a vizsgált feladatsor teljesíti: 3 feladatból és 21 itemből áll, terjedelme összesen 6:32 perc.

0.17. táblázat:
Áttekintés a 3. vizsgarész-összetevő felépítéséről

Feladattípus / Szöveg hossz percben ¹⁸	Itemszám	Összes item:	Vizsgapont
kérdésekre válasz / 1:35 perc	7	21	33
igaz-hamis feladat / 2:17 perc	8		
ál igaz-hamis (szelektálás) / 2:40 perc	6		

A feladatsor formailag mindenben megfelel az elvárásoknak, ám alapvető jellemzőit tekintve úgy tűnik, hogy a 3. feladat nem tölti be megfelelően funkcióját. A Cronbach-alfa érték a mínusz tartományban van, így ez nem is értelmezhető. A magyarázóerő szempontjából is ez a feladat a leggyengébb. A teljes teszt reliabilitása 0,71-re növelhető, amennyiben ezt a feladatot törölnénk. Ez az érték a képességmérő teszteknel az elfogadható tartományba kerülne. Az első feladat ugyanakkor megbízhatóan mér, a második feladat Cronbach-alfa értéke alacsonyabb, amely a kisebb szórásnak is köszönhető.

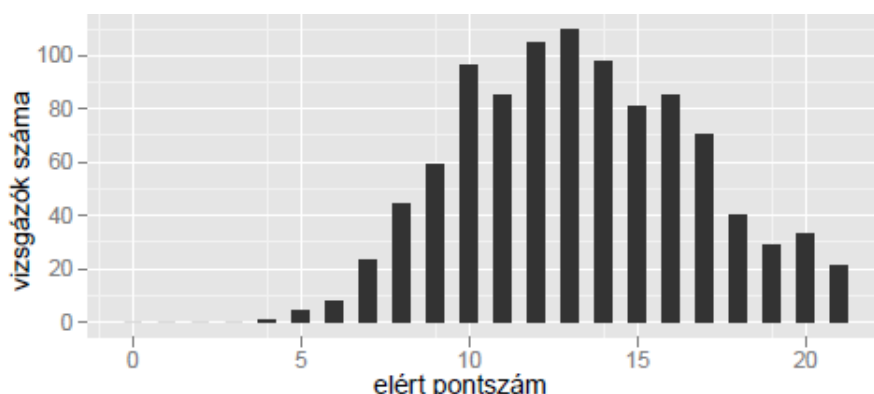
0.18. táblázat:
A hallott szöveg értéke feladatsor jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Kategória	Feladat			Teszt
	1.	2.	3.	
Cronbach-alfa	0,69	0,46	-0,32	0,68
Átlag (%p)	64,7	61,0	65,1	63,4
Szórás (%p)	26,0	21,6	16,9	16,7
Magyarázóerő (R ² %)	42,7	39,5	17,8	100

Az átlagos megoldottság mindhárom feladat esetében 60% fölött van. A legkönnyebb a harmadik feladat volt, amelyhez hasonló átlageredményt értek el a vizsgázók az első feladatban ($t=0,52$; $p>0,05$). Ezeknél szignifikánsan ($p<0,001$) alacsonyabb átlageredményt értek el a második feladatban, amely így a feladatsor legnehezebb feladata volt.

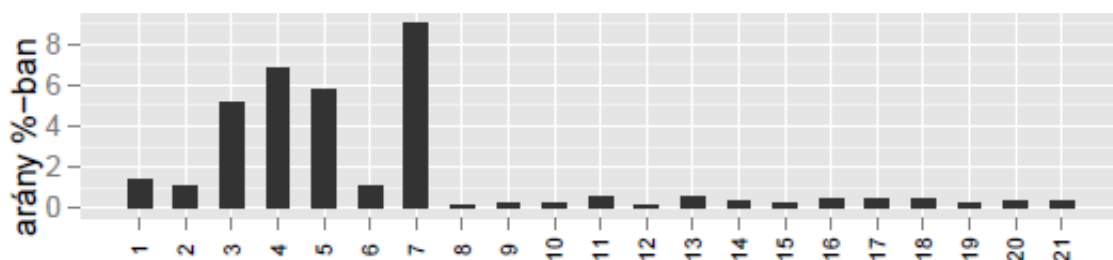
A megbízótól kapott adatsor csak a feladatpontokat összesítette, ami nem azonos a 3. vizsgarész-összetevőre kapható vizsgapontokkal. Az eloszlás (0.15. ábra) a magasabb teljesítmények felé, jobbra tolódik. A diákok többsége 12 pont (57%p) fölött teljesített.

¹⁸ Mivel az internetre feltöltött hanganyag összefüggően 30 percet tesz ki, nincs track-ekre bontva, így csak a lejátszáskor mért időt tudjuk megadni.



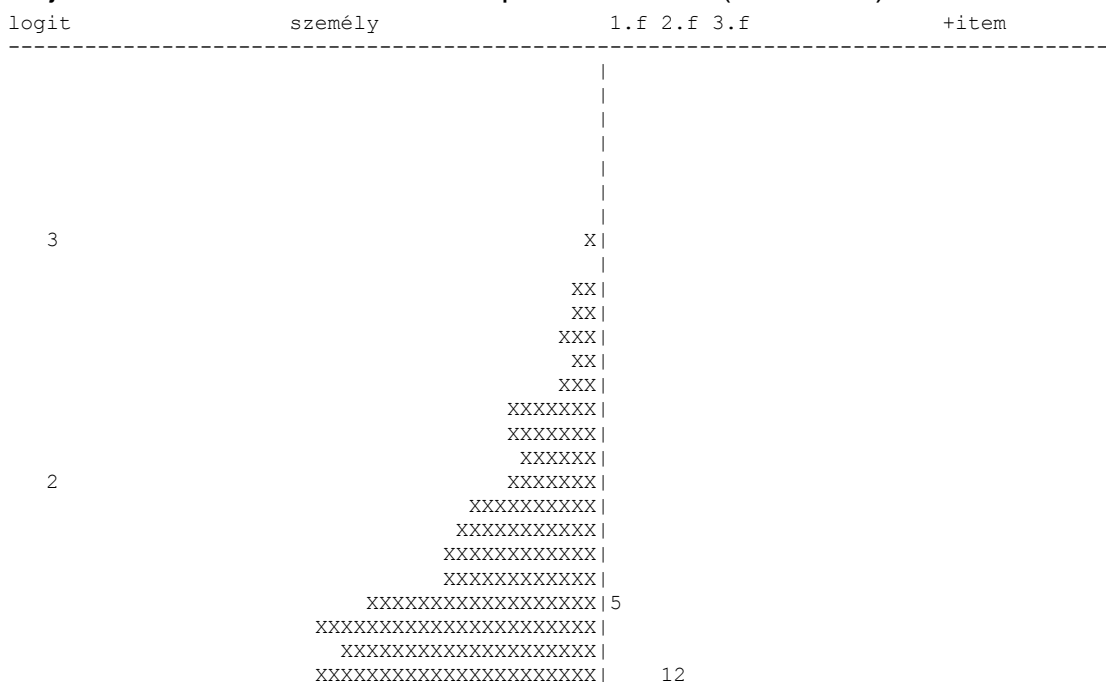
0.15. ábra:
 A teszten nyújtott pontszám abszolút gyakorisági eloszlása
 (megbízó által készített grafikon)

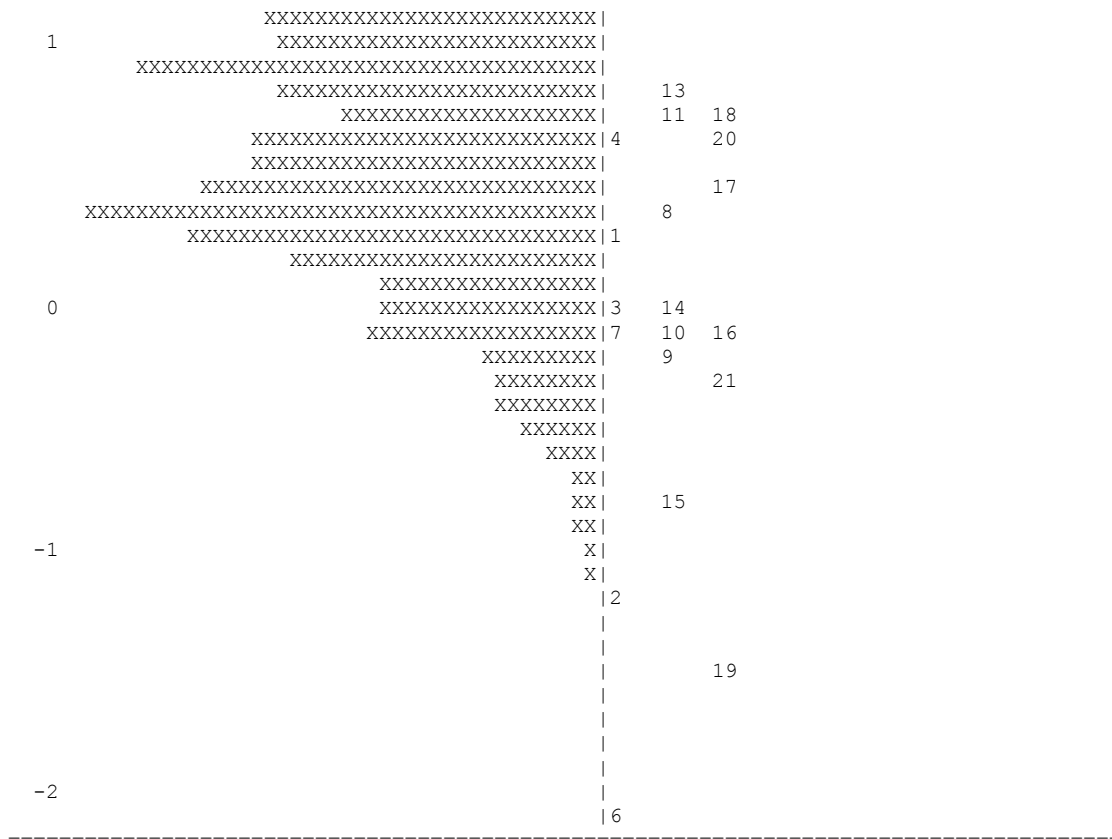
A megbízó által készített, és a 3. vizsgarész-összetevőben üresen hagyott itemek sorát mutató ábrán szembeötlő, hogy az 1–7 itemet átfogó első feladat okozhatott komolyabb problémát a vizsgázóknak. Ennek háttérében a nyílt végű feladattípus áll.



0.16. ábra:
 Az üresen hagyott itemek relatív gyakorisági eloszlása
 (megbízó által készített grafikon)

Az egyes feladatokhoz tartozó itemek nehézségét és a mintába került diákok képességszintjét az alábbi minta- és itemtérképen ábrázoltuk (0.17. ábra).





0.17. ábra:

A hallott szöveg értése feladatsor minta-itemtérképe
(VT által készített ábra) Minden 'x' két vizsgázót jelöl.

A térképről leolvasható, hogy az adott item mennyivel könnyebb vagy nehezebb, illetve a vizsgált személy mennyivel magasabb vagy alacsonyabb képességű. A vizsgált személyek képességében ugyan különbségek vannak, de többségük az átlagos képességszint fölött helyezkedik el. Bár a hallott szövegértést az olvasásnál nehezebbnek tartják, úgy tűnik, hogy ezen a vizsgán a hallott szövegértés vizsgarész-összetevőben a tanulók képességszintje magasabb volt, mint ugyanazon tanulók olvasott szövegértése. Ezt a két vizsgarész-összetevőnél elért átlagos megoldottság alapján is alá tudjuk támasztani, hiszen a hallott szövegértésnél a diákok jobban teljesítettek ($t=11,55$; $p<0,001$).

Ebben a vizsgarész-összetevőben is érvényesülnie kellene a fokozatosság elvének, tehát az 1. feladatnak aránylag könnyebbnek kellett volna lennie. De itemszinten azt látjuk, hogy az 1. feladat tartalmazza a feladatsor legkönnyebb és a legnehezebb itemét is. A többi feladat is tartalmaz mind az átlagos képességszint alatti, mind afölötti, 50%-os megoldási valószínűséggel rendelkező itemeket. Az itemek nehézsége többnyire megfelel a tanulók képességszintjének, bár az 1-es logitérték fölötti képességszint mérésére kevés item van. Ellenben három item (2., 6., 19.) túl könnyű volt a vizsgázók számára, amelyek így akár el is hagyhatók a tesztből.

1. Feladat: Unsere Direktorin (7 item)

Az első szövegben egy tanuló interjút készít egy tanárnővel. A feladat az, hogy a tanárnő válaszai alapján röviden válaszoljanak írásban a feladatlap kérdéseire (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 23). Az első feladatnak kelle-ne elméletileg a legkönnyebbnek lennie. A vizsgaleírás szerint a szövegnek az alábbi tulajdonságokkal kell rendelkeznie:

- hétköznapi nyelven hangzik el, szinte kizárólag gyakran használt szavakat és nyelvi szerkezeteket tartalmaz,
- lassú, de még normál beszédtempójú,
- a standard kiejtés(ek)nek megfelelő,
- tisztán artikulált,
- tartalmilag kifejezetten egyszerű (40/2002 OM rendelet, vizsgaleírás: 4).

Az alkalmazott szöveg stúdióban készült felvétel, a beszélők érezhetően olvasták a szöveget. Bár a két beszélő anyanyelvű, beszédük mégsem kelti az autentikus szöveg benyomását. A beszédtempó lassúsága, túlzott artikuláltsága mesterségesen hat. Mindazonáltal a gyenge nyelvtanulónak minden bizonnyal könnyebb egy ilyen, valószínűleg még A1 szintre is alkalmas szöveget megérteni. Valószínűleg ezt kívánták kompenzálni egy nehezebb feladattípussal. A három feladat közül ez ugyanis az egyetlen, ahol a vizsgázónak a célnyelven le kell írni valamit. Az ilyen jellegű feladatok mindig nehezebbek. Feltételezhető, hogy a nehezebb feladattípus kioltotta a szövegből fakadó könnyűséget, s így a szöveg és a feladat nem voltak összhangban.

0.19. táblázat:

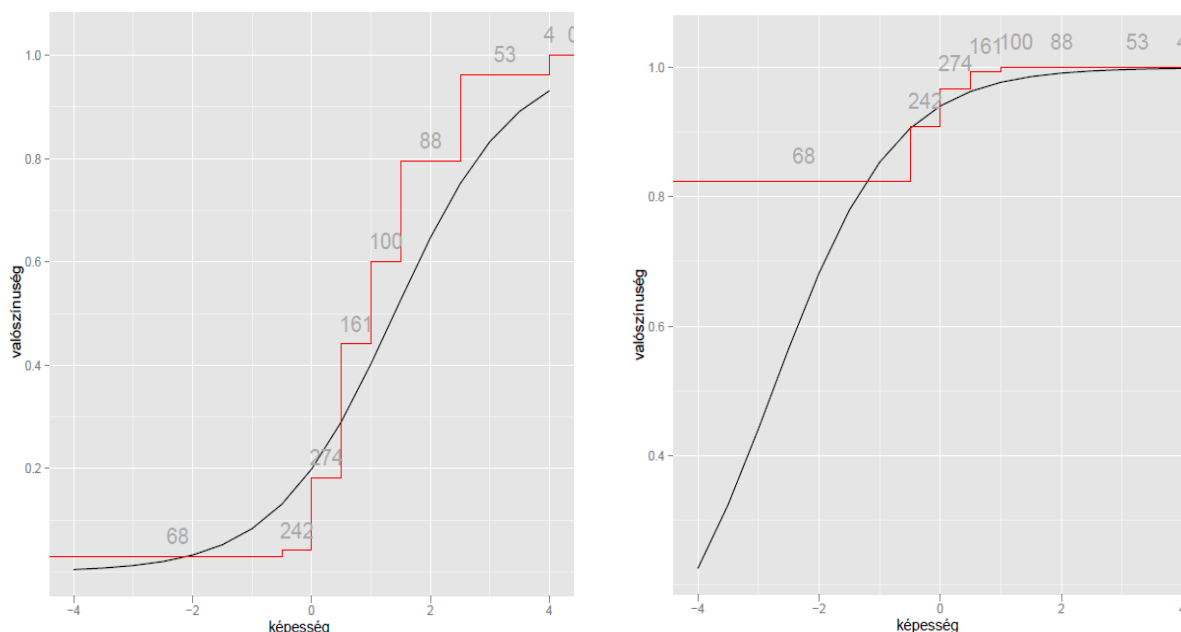
A hallott szöveg értéke 1. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[1.]	1,4	0,66	0,47	0,53	0,28	0,22	0,91	B1
[2.]	1,1	0,86	0,34	0,28	0,05	-1,21	1,00	A1
[3.]	5,1	0,67	0,47	0,45	0,13	-0,04	0,96	A2
[4.]	6,9	0,55	0,50	0,57	0,30	0,65	0,88	B1
[5.]	5,7	0,37	0,48	0,59	0,39	1,53	0,87	B1
[6.]	1,1	0,97	0,17	0,25	0,05	-2,79	0,96	A1
[7.]	9,1	0,71	0,45	0,55	0,32	-0,15	0,91	A2

A 2. és 6. item pontbiszerialis korrelációja a legalacsonyabb (0.19. táblázat), bár értéke 0,2 fölött van, de a magyarázóerejük nagyon alacsony, vagyis ezek az itemek nem jól mérik azt, amit a teszt egésze. Erre utal a 0,86 és 0,97 megoldottsági átlag is, és az itemnehézségi mutató negatív értéke. Érdekes, hogy a 2. item jobban sikerült, mint a 3. item, jóllehet, a két ütem összefügg. Oka valószínűleg az, hogy összesen két egyszerű informá-

ciót kellett leírni a három elhangzott információból, de ezek nagyon gyorsan követték egymást, így amíg az elsőt még sikerült leírni, a második már elhangzott és csak kevesebben tudták azt leírni.

A legkönnyebb item (6.) szórása igen alacsony, kevésbé differenciál. A legnehezebb, 5. item nehézségmutatója magas. Ebben az itemben a vizsgázónak olyan kifejezést kellett leírnia, ami ugyan elhangzott, ám nem abban a formában, ahogy a kérdésre adott választ hallhatta. Ez tanulságos lehet a feladatíróknak: ha már jegyzetelnie kell a vizsgázónak ebben a vizsgarész-összetevőben, ügyelni kell arra, hogy jellegében ne térjen el az item a többitől. A magyarázóerő alapján a hallott szöveg értését leginkább a B1 szintre sorolható itemek (1., 4., 5.), valamint a 7. item méri. Ezek jellemzője, hogy egy-egy szóval, az adott információ megértésével kellett a vizsgázóknak a választ megadniuk.



0.18. ábra:
3. vizsgarész-összetevő 5. (legnehezebb), 6. (legkönnyebb) itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített ábra)

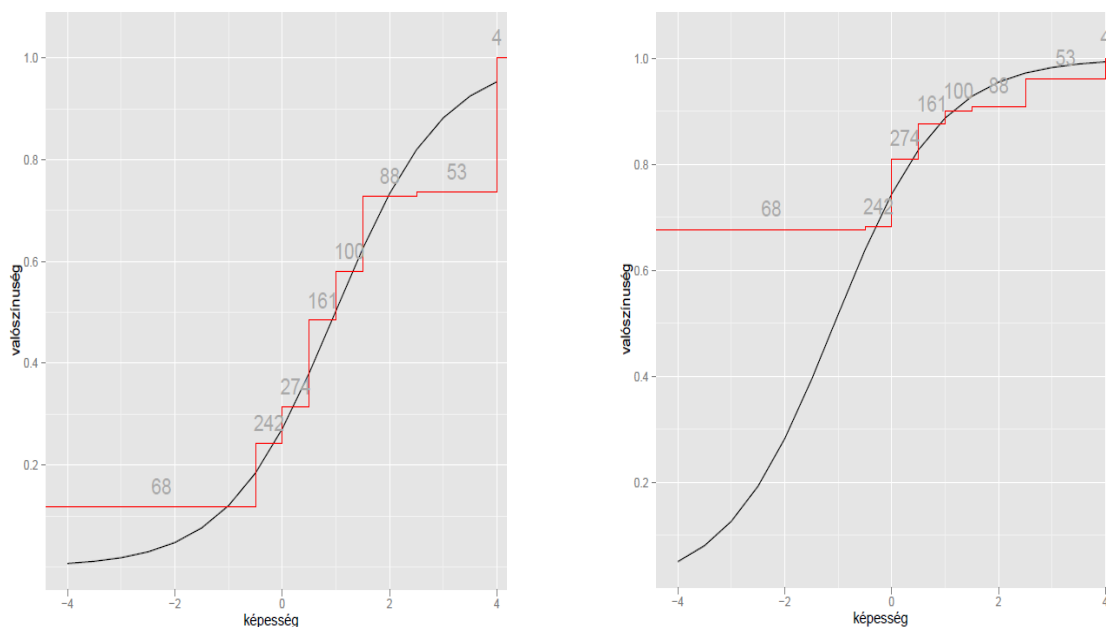
2. Feladat: Frank ist Vater geworden (8 item)

A vizsgázó egy monológot hall, amelyben egy édesapa elmeséli lányai születésének rövid történetét (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 24). A mesélő anyanyelvű, de a szöveget érezhetően felolvassa. A felvétel nem autentikus, stúdióban készült. A beszéd megfelel az A2 szint elvárásainak: hétköznapi téma, egyszerű kifejezések, normál beszédtempó, standard kiejtés, tiszta artikuláció. Az A2 szintű szöveghez könnyű, ú.n. igaz-hamis feladattípus tartozik, amit 50% eséllyel akár találgatni is lehet. A 8 itemen belül megegyezik az igaz és a hamis állítások száma.

0.20. táblázat:
A hallott szöveg értése 2. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
[8.]	0,1	0,61	0,49	0,52	0,15	0,34	0,94	B1
[9.]	0,2	0,74	0,44	0,32	0,05	-0,27	1,02	A2
[10.]	0,2	0,70	0,46	0,37	0,07	-0,12	1,00	A2
[11.]	0,5	0,50	0,50	0,39	0,08	0,78	1,00	B1
[12.]	0,1	0,43	0,50	0,39	0,15	1,23	1,01	B1
[13.]	0,5	0,50	0,50	0,34	0,05	0,80	1,03	B1
[14.]	0,3	0,67	0,47	0,31	0,03	-0,02	1,04	A2
[15.]	0,2	0,82	0,39	0,31	0,05	-0,80	1,01	A2

A feladat az utolsó item kivételével nem mutat különös eltéréseket, a pontbiszerialis korrelációja megfelelő. Ugyanakkor magyarázóerejük eléggé alacsony. A 12. item volt a legnehezebb. A feladattípusnak jellemzője, hogy a hamis állítások könnyen megzavarhatják a vizsgázót. Itt is ez az eset állhat fenn: először elhangzik egy állítás, majd annak az ellenkezője. Ha a pozitív állításra „ráugrik” a tanuló, s nem figyeli tovább a szöveget, könnyű eltéveszteni. Az utolsó item (15.) pedig valószínűleg azért lett könnyű, mert erről hosszabban mesél a beszélő, szinte szájbarágósan. A 0.19. ábra e két item karakterisztikus görbéjét összehasonlítva mutatja be.



0.19. ábra:

A 3. vizsgarész-összetevő 12. (legnehezebb) és 15. (legkönnyebb) itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített ábra)

3. Feladat: Italiener und Deutsche (6 item)

Az utolsó szöveg már autentikus, internetes rádióból származó hanganyag. Ezért – bár időtartamában csak fél perccel hosszabb az előző hanganyagnál – a szöveghez lényegesen hosszabb transzkripció tartozik. Ez azt jelenti, hogy nincs szó könnyített beszédtempóról. B1 szinten elvárható, hogy hétköznapi témájú, standard kiejtéssel sugárzott rövid rádióműsort értsen a vizsgázó. A szöveghez 12 kijelentés tartozik (Német nyelv, közép-szintű írásbeli érettségi vizsga, feladatlap 2012: 25). A vizsgázónak el kell döntenie, hogy ezek közül valójában mely kijelentésekről beszélgettek a rádióműsorban. Összesen 6 kijelentés felel meg ennek, de ehhez még 6 másik kijelentés is tartozik, így az előző feladathoz hasonlóan itt is fele-fele arányban voltak az igaz és hamis kijelentések, ám a hamis kijelentéseket nem értékeli a vizsga, nem jár értük feladatpont.

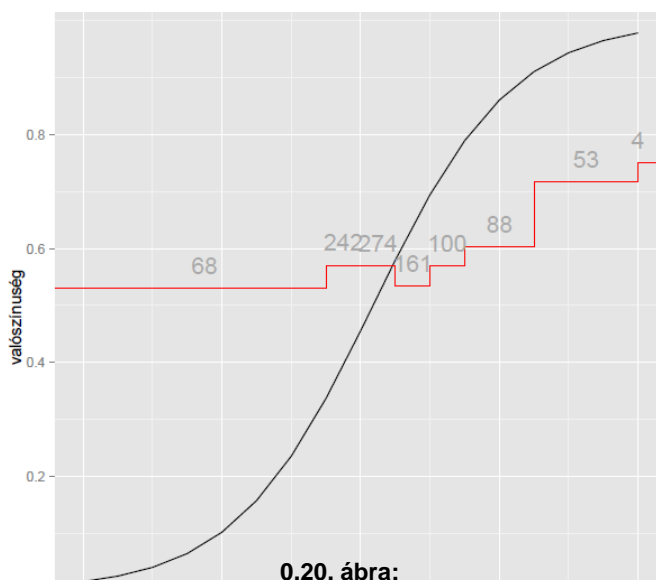
0.21. táblázat:

**A hallott szöveg értése 3. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)**

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ
[16.]	0,4	0,71	0,45	0,34	0,08	-0,15	1,01
[17.]	0,4	0,57	0,50	0,08	0,00	0,45	1,17
[18.]	0,4	0,53	0,50	0,28	0,09	0,73	1,08
[19.]	0,2	0,89	0,31	0,18	0,02	-1,50	1,02
[20.]	0,3	0,54	0,50	0,19	0,01	0,63	1,11
[21.]	0,3	0,74	0,44	0,25	0,06	-0,30	1,06

A feladat javítási-értékelési útmutatója azt az utasítást adja (2012: 8), hogy ha a vizsgázó 6-nál több mondatot jelölt be, akkor amennyivel többet jelölt, annyi pontot kell levonni megszerzett pontjaiból. 12 tanulónál (1,2%) fordult elő a pontlevonás. Közülük 9-en egy, és 1-1 vizsgázónál 2, 3 és 5 pont levonás történt. Ez utóbbi esetben a diák valószínűleg egy kivétellel minden mondatot beikszelt.

Már a feladatsort bemutató 0.18. táblázatban szereplő Cronbach-alfa érték is utal arra, hogy baj van ennek a feladatnak a reliabilitásával. Ennek háttérében az áll, hogy a vizsgázó akkor kap 0 pontot, ha az adott helyre, ahova kellett volna, nem tett x-et. Ezért az üresen hagyott itemek aránya ebben a feladatban nem értelmezhető. Az adatrögzítés sajátossága, hogy arról nem szerzünk információt, hogy a tanuló a helytelen válasz esetén hova tett ikset. Feladattipológiailag szabályosan tehát ennek a feladatnak 12 pontosnak kellett volna lennie, amelyben a vizsgázó jelöli, hogy melyik mondat helyes, melyik



0.20. ábra:

A 3. vizsgarész 3. feladat 17. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

hamis. Azzal, hogy csak a helyes válaszokat kell megjelölni, a feladatmegoldó tevékenység mesterségessé válik, hiszen a vizsgázónak arra is kell figyelnie, hogy hány elemet jelölt meg válaszként, így tehát az itemek nagy valószínűséggel összefüggnek. E tényezők mind hozzájárulhattak ahhoz, hogy ez a feladat ilyen formában működésképtelen. Ezért nem adtuk meg a 0.21. táblázatban a feltételezhető nyelvi szintet. Három itemnél külön problémák láthatók. A 17., 19. és 20. item pontbiszeralis korrelációja és magyarázóereje igen alacsony, a 19. item nehézségi mutatója az átlag alatti képességtartományban van. A 17. itemnek rossz az illeszkedése, amit az item elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje is jól mutat.

Összegzés: A hallott szövegértés feladatsora számos problémát tartalmaz. Az 1. feladat igen könnyű szövegre épült, ehhez egy nehezebb, nyílt végű feladattípus kapcsolódott. Maga a szöveg kérdés-feleletből állt, és a feladat is az volt, hogy kérdésekre kellett rövid választ írni. De a kérdések megfogalmazása más volt, mint az elhangzó szövegben. Valószínűleg ez a feladattípus egy nehezebb szöveggel B1 vagy B2 szinten jobban működik. A 2. feladatban se a szöveg, se a feladattípus nem volt nehéz. A 3. feladattal viszont alapvető problémák vannak. A relatív hosszabb szöveghez csak nagyon kevés item párosul, s azok nem mérnek megbízhatóan. Ez a feladattípus a tankönyvekben sem fordul elő, és a gyakorlókönyvek sem tartalmazzák. A vizsga bevezetésének előkészítésekor a minta-feladatsorok egyikébe se került bele. Valószínűleg ez a feladattípus inkább gyakorlásra való, s nem mérésre. Helyette célszerű lenne 3 elemből álló feleletválasztó feladattípust alkalmazni. Így megnőne az itemszám, és a megbízhatóságon is javítani lehetne.

3. A három, itemalapú vizsgarész-összetevő elemzésének összefoglalása

Kutatási kérdéseink (I. Bevezetés, 5. A kutatás módszerei) közül, amelyek egyrészt az első három vizsgarész-összetevő megbízhatóságára és a feladatok súlyára vonatkoztak, másrészt az itemalapú feladatok működésére, elemzésünkkel az alábbiakra kaptunk választ:

1. Mennyire megbízhatóan mérték az egyes készségeket mérő feladatsorok?
Mindegyik feladatsorról elmondható, hogy többnyire megbízhatóan mérték, bár a Cronbach-alfa értékek csökkenő tendenciát mutatnak: az olvasás eredménye a legjobb, és a hallás a leggyengébb, ez már az elvárható 0,70 alatt van.

0.22. táblázat:

A három feladatsor összesített jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Kategória	Olvasott szöveg értése	Nyelvhelyesség	Hallott szöveg értése
Cronbach-alfa	0,82	0,77	0,68
Átlag (%p)	58,4	57,4	63,4
Szórás (%p)	17,1	17,0	16,7

2. Milyen mértékben magyarázzák az egyes feladatok az adott vizsgarész-összetevőben elért összpontszámot?
A feladatok magyarázóereje a vizsgarész-összetevőkben eltért egymástól. A legalacsonyabb magyarázóerővel az 1. vizsgarész-összetevő 1. és 4. feladata; a 2. vizsgarész-összetevő 1. feladata és a 3. vizsgarész-összetevő 3. feladata rendelkezett. Ezekre a feladatokra általában az volt jellemző, hogy vagy a feladattípussal volt komoly probléma (lásd. hallott szöveg értése 3. feladat) vagy az adott feladaton belüli itemek voltak túlságosan könnyűek és ezáltal súlytalanok az adott vizsgarész-összetevőben.
3. Milyen nehézségi szintbe sorolhatók az egyes feladatok és itemek?
Azt tapasztaltuk mind a számítások, mind a minta-itemtérképek alapján (0.3. ábra., 0.10. ábra., 0.17. ábra), hogy a feladatok csak részben fedik le a vizsgakövetelményekben meghatározott nyelvi szintet. Bár a feladatoknak kb. fele-fele arányban kellene lefedni az A2 és a B1 nyelvi szintet, a feladatok inkább az A2 tartományban voltak. A teszten nyújtott pontszám abszolút gyakorisági eloszlását mutató ábrák (0.1. ábra., 0.8. ábra., 0.15. ábra) is inkább jobbra, a magasabb teljesítmények felé tolódtak.
4. Mennyiben felelnek meg a feladatok itemnehézségi indexei a vizsgázók nyelvi készségszintjeinek?

Az itemnehézségi indexek nem fedik le teljesen a mintában elemzett vizsgázói készségszinteket, ezek többsége alatta maradt a tanulók képességeinek. Ugyanakkor figyelembe kell vennünk azt is, hogy az adatelemzés kizárólag a nappali tagozatosok körére vonatkozott. Az érettségi eredményei azt mutatják, hogy a felnőttoktatásban végzetek jóval alacsonyabb nyelvtudással rendelkeznek a nappalisoknál. Vagyis feltételezhető, hogy ha a teljes vizsgázói populációra vonatkozóan reprezentatív a mintavétel, akkor jobban működnek a feladatsorok ebből a szempontból.

A feleletválasztó feladatoknál a legfőbb probléma az itemek (időnként extrémén) eltérő nehézségi szintje volt az adott feladaton belül. Egy feladaton belül nagyon eltérő nyelvi szinteket találtunk, a nehézségi index több esetben is tág határok között mozgott (pl. a hallás 1. feladata -2,79-től 1,53 értékig).

5. Mennyire megfelelően működnek az egyes vizsgarész-összetevőkben használt zárt végű feladattípusok? Ha nem jól működnek, mi ennek az oka?

A feladattípusok – egy kivétellel – általában jól működtek, de a feladaton belüli itemek nem mindig. A feladatok nehézségi szintje a vizsgaleírásnak nagyrészt megfelelt, bár inkább a könnyebb (A2) tartományba tartozott. Mivel mindegyik feladat autentikus szövegre épül, már a szöveg kiválasztásánál eldőlhet, sikerül-e benne megfelelően mérő itemelemeket találni. Megfigyelhető volt, hogy a szöveg egyszerű mivoltát egy nehezebb feladattípussal, a nehezebb szöveget pedig könnyebb feladattípussal próbálták kompenzálni. Ez nem helyes taktika, mert veszélyezteti a feladat megbízhatóságát. Könnyebb feladattípuson a zárt végű feladatokat értjük, ahol a vizsgázónak nem kell beírni célnyelvi kifejezést, csak a megfelelő választ aláhúzással, beikszeléssel jelölni (pl. igaz-hamis). Az igaz-hamis feladattípus itemenként 50%-os találgatási esélyt kínál, ezért könnyebbnek bizonyult a tanulók képességszintjénél. A feladatokon belül előforduló összefüggő itemek, a nem megfelelő disztraktorok és a pontlevonás rontják az adott feladat és az egész teszt reliabilitását.

Az egyetlen rosszul működő feladattípus az ún. ál igaz-hamis (szelektálás) feladat volt az 1. vizsgarész-összetevőben (olvasott szövegértés) és a 3. vizsgarész-összetevőben (hallott szövegértés). Itt a relatív hosszú szöveghez kevés számú item párosult. A megadott kijelentésről fel kellett ismerni, hogy miről volt szó a szövegben, és miről nem. De ami nem odaillő kijelentés, azt hiába ismeri fel a vizsgázó, nem jár a felismerésért pont. Viszont ha többet jelöl be, mint amennyi lehetséges, büntető pont jár érte, ami nincs összhangban a mérő funkcióval.

6. Mennyire megfelelően működnek a nyílt végű feladattípusok? Mi az oka a nem megfelelő itemműködésnek ezekben a feladatokban? Hogyan működött ezeknél a feladatoknál az értékelési rendszer?

A nyílt végű feladatoknál középszinten a kérdésre rövid válasz fordult elő az 1. és 3. vizsgarész-összetevőben. Ott, ahol valami célnyelvi kifejezést kell leírni, jobban

meg kell gondolni, miképp kérdez a feladat. Nem szerencsés, ha ugyanazon nehézségi szinten hol egyszerű információt kell beírni, hol a szövegből következtetve megadni a választ – ez utóbbi ugyanis jóval nehezebb. Itt is veszélyes, ha az itemek összefüggnek egymással. Az elemzés során nem tudtuk megállapítani, miképp működött az értékelési rendszer, mert nem tudhatjuk, milyen válaszokat adtak a vizsgázók, mennyire vették figyelembe az értékelő tanárok az értékelési útmutatóban megfogalmazott lehetséges változatokat.

A kutatási kérdésekhez megfogalmazott hipotézisek közül az alábbiak igazolódtak be: A német nyelvi érettségi működésének jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján helyesen feltételeztük, és elemzésünk nyomán beigazolódtott, hogy

1. az egyes készségeket mérő feladatsorok megbízhatósága eltér, amely a mérendő konstruktum jellemzőitől és az alkalmazott feladattípusoktól függ;
2. az egyes feladatok súlya a tesztekben eltér, amely a mért nyelvi szinttől, illetve az egyes feladatok minőségétől függ.

A receptív készségeket és a nyelvhelyességet mérő feladatsorok jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján helyesen feltételeztük, és elemzésünk nyomán beigazolódtott, hogy

3. a feladatok itemnehézségi indexei nem fedik le teljesen a vizsgázók nyelvi készségszintjei által meghatározott intervallumot;
4. a magasabb készségtartományból hiányoznak itemek;
5. a zárt végű feladatokban lévő összefüggő itemek, a nem megfelelő disztraktorok, a pontlevonás rontják az adott feladat és a teszt reliabilitását;
6. a nyílt végű feladatok működése függ a mért nyelvi szinttől és a mért gondolkodási művelettől.

Irodalom

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgakövetelmények.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv_vk.pdf,

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgaleírás.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv_vl.pdf (letöltve: 2013. május 20.)

Glaboniat, Manuela et al. (2006): *Profile deutsch*. Lernzielbestimmungen, Kannbeschreibungen und kommunikative Mittel für die Niveaustufen A1, A2, B1, B2, C1 und C2 des "Gemeinsamen europäischen Referenzrahmens für Sprachen". Buch mit CD-ROM. Klett-Langenscheidt Verlag.

Közös Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés (2002): Európa Tanács Közoktatási Bizottsága. Közoktatási Bizottság. Élő Nyelvek Osztálya. Strasbourg. Piliscsaba: Pedagógus-továbbképzési Módszertani és Információs Központ Kht.

Német nyelv. Középszintű írásbeli vizsga. 2012. május 11. (2012). Budapest: Nemzeti Erőforrás Minisztérium.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/kozep/k_nemet_12maj_fl.pdf (letöltve 2013. május 20.)

Német nyelv. Középszintű írásbeli vizsga. Javítási-értékelési útmutató. 2012. május 11. (2012). május 11. Budapest: Nemzeti Erőforrás Minisztérium.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/kozep/k_nemet_12maj_ut.pdf (letöltve 2013. május 20.)

Német nyelv. Középszintű írásbeli vizsga. Hanganyag. 2012. május 11. (2012). Budapest: Nemzeti Erőforrás Minisztérium. [k_nemet_12maj_fl.mp3](#)

http://www.oktatas.hu/kozneveles/erettsegi/feladatsorok/kozepszint_2012tavasz (letöltve 2013. május 20.)

VÍGH TIBOR, PETNEKI KATALIN

**A 2012. MÁJUSI NÉMET NYELV
KÖZÉPSZINTŰ
ÍRÁSKÉSZSÉG FÜZET/VIZSGARÉSZ (ÖSSZETEVŐ)
ELEMZÉSE**

Az íráskészség vizsgarész-összetevő elemzésének célja az I. és a III. kutatási terület vizsgálata (vö. 5.1. fejezet). Ennek megfelelően először áttekintést adunk a 4. vizsgarész-összetevő felépítéséről a vizsgaleírás és a feladatlap, illetve a javítókulcs alapján, majd a teszt legfőbb alapstatisztikai jellemzőit tárgyaljuk. Ezt követően megvizsgáljuk az egyes feladatok működését, végül az értékelési rendszer jellemzőit elemezzük.

1. Az íráskészség vizsgarész-összetevő felépítése

A 4. vizsgarész-összetevő két részből áll. Az első feladat a Közös Európai Referenciakeret (2002) A2 szintleírása szerinti elemi írásbeli produkciót vár, míg a második feladat a B1 szint önálló nyelvhasználói kritériumai szerint épül fel. A két feladat egymástól független. Értékelésük a központi javítási értékelési útmutatóban közölt analitikus skálák alapján történik (40/2002 OM rendelet, Vizsgaleírás: 6).

Az első feladat az A2 szint készségleírásának felel meg (40/2002 OM rendelet, Vizsgaleírás: 6). Ebben a vizsgázó verbális és/vagy vizuális segédanyagra reagálva 50–80 szavas közlést hoz létre (pl. üdvözlőlap, e-mail). A2 szinten a vizsgázónak képesnek kell lennie arra, hogy

- a személyével, illetve közvetlen környezetével kapcsolatos témákról írjon,
- az adott témáról egyszerű kifejezéseket használva írjon,
- egyszerű mondatokból álló rövid, összefüggő szöveget írjon,
- a szöveg mondatai között lévő alapvető logikai kapcsolatokat egyszerű nyelvi eszközökkel (pl. a leggyakrabban használt kötőszavakkal) tudja kifejezni,
- néhány egyszerű szerkezetet és alapvető helyesírási szabályt úgy használjon, hogy a hibák ellenére érthető legyen a mondanivaló.

Az első feladat teljesítését mindössze két kritérium szerint kell értékelni. Mindegyik kritérium alapján külön kell pontot adni attól függően, hogy mely képességleírás illik a legjobban az adott írásműre (Német nyelv, középszintű írásbeli érettségi vizsga, javítási és értékelési útmutató 2012: 15). A pontozás egyben súlyoz is. Így ha a vizsgázó kap is 1–2 pontot a dolgozatára, nem fogja vele elérni az eredményes vizsgarész-összetevő teljesítéséhez szükséges határt. Erre a feladatra maximálisan 10 pont adható.

0.1. táblázat:

Áttekintés az 1. feladat értékelési skálájáról

Kritérium	5 pont	4 pont	3 pont	2 pont	1 pont	0 pont
Tartalom	eléri az A2 szintet			nem éri el az A2 szintet		
Nyelvhelyesség	eléri az A2 szintet			nem éri el az A2 szintet		

Az 1. feladat megfelel a vizsgaleírás elvárásainak: a vizsgázónak egy e-mail-re kell válaszolnia. Az e-mail-ben egy ismerős kér tanácsot mobiltelefon vásárlásával kapcsolatban (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 31). A rövid és egyszerű nyelvezetű inputszöveg sok támpontot ad arra vonatkozóan, mire kell reagálnia a vizsgázónak. A téma egyszerű, és az input is ad nyelvi segítséget, hiszen hasonló szókincs kell a válaszhoz is, mint ami a kérdésekben volt.

A 4. vizsgarész-összetevő 2. feladata a B1 szintet célozza meg. A vizsgázónak olyan szövegeket kell létrehoznia (40/2002 OM rendelet, Vizsgaleírás: 6), amelyek

- meghatározott kommunikációs szándékkal jönnek létre,
- az olvasó számára világosak, érthetőek és alkalmasak a kommunikációs cél elérésére,
- szövegfajtája meghatározott,
- tematikusan megfelelnek ezen korosztály életpaszatának és általános érdeklődésének,
- megírásához nincs szükség az érettségi vizsga általános műveltségi szintjét meghaladó ismeretekre.

A vizsgázónak egy inputszövegre kell levélben reagálnia, meghatározott szempontok alapján. A feladat leírja a kommunikációs szituációt, amiben a szövegnek el kell készülnie. Ehhez tartozik az autentikus forrásból származó inputszöveg, ami lehet egy levél, egy hirdetés vagy egy hír. Ezt követi a feladat leírása és az öt irányító szempont, amit a levélnek tartalmaznia kell. A végén felhívják a vizsgázó figyelmét a levél formai követelményeinek betartására. Mivel a vizsgázónak itt már jóval hosszabb szöveget kell létrehoznia (100–120 szó), s nem biztos, hogy a vizsgázó egy adott témáról tud ennyit írni, ezért lehetőség van két téma közül választani. De mindkét témához ugyanaz a feladat tartozik, s egyforma a kritériumrendszer is, amely alapján értékelik.

A 2. feladatot öt értékelési szempont és az első feladatnál összetettebb értékelési skála alapján kell értékelni (40/2002 OM rendelet, Vizsgaleírás: 15). A kritériumok pontonkénti leírása – az 1. feladathoz hasonlóan – a Közös Európai Referenciakeret (2002) szintleírásához igazodik, ám itt a viszonyítási pont már a B1 szint elérése.

0.2. táblázat:

Áttekintés a 4. vizsgarész-összetevő 2. feladatának értékelési szempontjairól

Értékelési szempont	Maximális pontszám
Tartalom	5
Formai jegyek és hangnem	3
Szövegalkotás	5
Szókincs, kifejezésmód	5
Nyelvtan, helyesírás	5
Összesen:	23

Új az értékelésben, hogy a legfőbb kritérium a kommunikációs szándék megvalósításának sikeressége. A Javítási és értékelési útmutatóban ezt olvashatjuk: „Amennyiben a dolgozat a Tartalom szempont alapján 0 pontos, akkor a többi értékelési szempont pontszáma is 0 pont.” (Német nyelv, középszintű írásbeli érettségi vizsga, javítási és értékelési útmutató 2012: 15). Ezt a lenullázási szabályt a vizsgáztató tanárok többsége (a mintába került dolgozatok 98%-ában) megfelelően alkalmazta. 11 esetben fordult elő, hogy más értékelési szempontnál elért 0 pont esetében a teljes dolgozatra is 0 pontot adtak.

0.3. táblázat:
Áttekintés a 2. feladat értékelési skálájáról

Kritérium	5 pont	4 pont	3 pont	2 pont	1 pont	0 pont	
Tartalom							
Formai jegyek, hangnem							
Szövegalkotás	eléri		nem éri				
Szókincs, kifejezés	a B1		el a B1				
Nyelvtan, helyesírás	szintet		szintet				

2. Az íráskészség vizsgarész-összetevő alapstatisztikai jellemzői

Az íráskészség teszt a teljes mintán 51,6%-os teljesítettségű volt, azaz ez a készség volt a vizsgázóknak a legnehezebb a többihez képest (vö. 0.8. táblázat). A 0.4. táblázat a két feladat paramétereit tartalmazza, amelyben az átlagot és a szórást százalékpontokban adjuk meg, mert a maximális pontszám eltér, és célunk a két feladat nehézségének összehasonlítása.

0.4. táblázat:

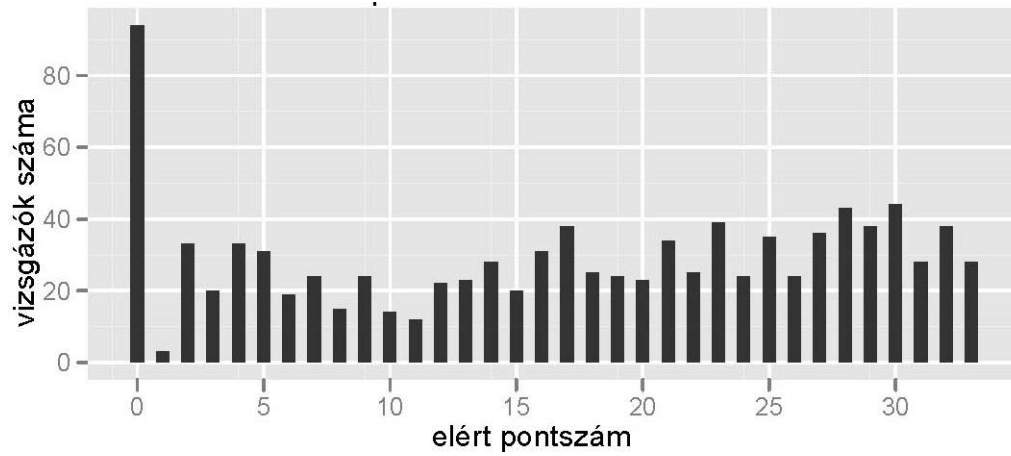
Az íráskészség feladatsor jellemzői (VT számításai alapján¹⁹)

Kategória	Feladat		Teszt
	1.	2.	
Cronbach-alfa	0,90	0,96	0,95
Átlag (%p)	59,5	48,1	51,6
Szórás (%p)	32,1	34,6	32,0
Magyarázóerő (R^2 %)	26,2	73,8	100

Eredményeink szerint az értékelési skálák használata magas megbízhatóságot eredményez. Az első feladatot az A2, a második feladatot a B1 szint mérésére készítették, ennek megfelelően az első feladatban a vizsgázók szignifikánsan jobb átlagteljesítményt értek el ($t=14,9$; $p<0,001$), mint a másodikban. A két feladat átlaga között 11,4%-nyi különbség van, ami abból a szempontból tűnik kevésnek, hogy az első feladat A2 szintű nyelvtudással is megoldható. Ennél tehát lényegesen magasabb is lehetne e feladat teljesítettsége. A két feladat közötti korreláció erős ($r=0,74$; $p<0,001$). A második feladat magyarázóereje majdnem háromszorosa az első feladaténak. Ennek oka a két feladat eltérő pontértéke.

A 0.1. ábra a teszten elért pontszámok eloszlását mutatja. Ebben feltűnő, hogy a vizsgázók közül 99-en (10%) értek el 0 pontot. Ennek oka egyrészt a két feladatban alkalmazott lenullázási szabály, másrészt, hogy a diákok közül 29-en (2,9%) a második feladatot nem töltötték ki. 1 pontot elvileg nem lehet elérni ebben a vizsgarész-összetevőben, a többi pontszám eloszlása nagyon hasonló, amely azt is jelzi, hogy jelentős különbségek vannak a tanulók íráskészség-szintjében, ahogy ezt az előző táblázat is mutatta.

¹⁹ VT rövidítéssel jelöltük Vigh Tibor számításait.



0.1. ábra:

A teszten nyújtott pontszám abszolút gyakorisági eloszlása (megbízó által készített grafikon)

3. Feladatszintű elemzés

A szükséges elemzések elvégzéséhez további másodelemzéseket végeztünk a megbízótól kapott elemzéseken felül.

Az első feladatban a tartalom, szöveghossz szempontnál a vizsgázók jobban teljesítettek (átlag: 3,4; szórás: 1,8), mint a nyelvhelyességnél (átlag: 2,6; szórás: 1,6). A két értékelési szempont teljesítettségében tehát szignifikáns különbség van ($t=25,8$; $p<0,001$) és szoros kapcsolat van közöttük ($r=0,83$; $p<0,001$). A tartalom szempont az első feladat varianciáját 53,2%-ban magyarázza, míg a második szempont 46,7%-ban. Ez összhangban van a vizsgakövetelményekben megfogalmazott elvárásokkal, hiszen a kommunikatív cél megvalósulását mérő szempontnak van némiképp nagyobb súlya e feladatra kapott pontszám alakulásában.



0.2. ábra:
Az 1. feladat megoldottsága szempontok szerint
(a megbízó által készített grafikon)

Megvizsgáltuk, hogy a 992 dolgozatból a tanulók hány százaléka érte el az A2 szintet, és hány százalékuk nem. Az első értékelési szempontnál a vizsgázók 27,5%-a, a másodikonál pedig 47,7% teljesített az A2 szint alatt. Ez összhangban van a korábbi tapasztalatokkal, amely azt mutatja, hogy a vizsgázóknak leginkább a nyelvileg megfelelő szöveg produkálása okoz problémát.

0.5. táblázat:
Az 1. feladat értékelési szempontjainál elért pontszámok abszolút és relatív gyakorisága (VT számításai alapján)²⁰

Pontszám	Tartalom és szöveghossz	Nyelvhelyesség
----------	-------------------------	----------------

²⁰ Kék színnel jelöltük azon vizsgázók számát és arányát, akik nem érik el az A2 szintet, barnával pedig azokat, akik elérik ezt a szintet a III/1. táblázatnak megfelelően.

	N	%	N	%
0	128	12,9	131	13,2
1	60	6,0	119	12,0
2	85	8,6	223	22,5
3	137	13,8	209	21,1
4	184	18,5	171	17,2
5	396	39,9	136	13,7
A2 szint alatt	273	27,5	473	47,7

A második feladatban a vizsgázók két téma közül választhattak. A 4. vizsgarész-összetevő első levelének témája szállásfoglalás kérése egy osztrák ifjúsági szálláson egy hirdetés alapján. A másik levélben egy tánciskolának kell írni és a kurzusokra való jelentkezés feltételeiről kell érdeklődni (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 31–33). A vizsgázók mindkét témát közel azonos arányban választották: az elsőt a diákok közül 451-en, a másodikat 512-en, 29-en nem oldották meg egyik feladatot sem. A választhatóság fontos alapelv, hiszen elsősorban nem a vizsgázók adott témáról megszerezhető háttértudását mérjük, hanem azt, hogy írásban sikeresen teljesítik-e a megadott kommunikációs célt. Ugyanakkor nagyon lényeges, hogy a feladatok megoldottságában ne legyen különbség. A továbbiakban ezt vizsgáljuk.

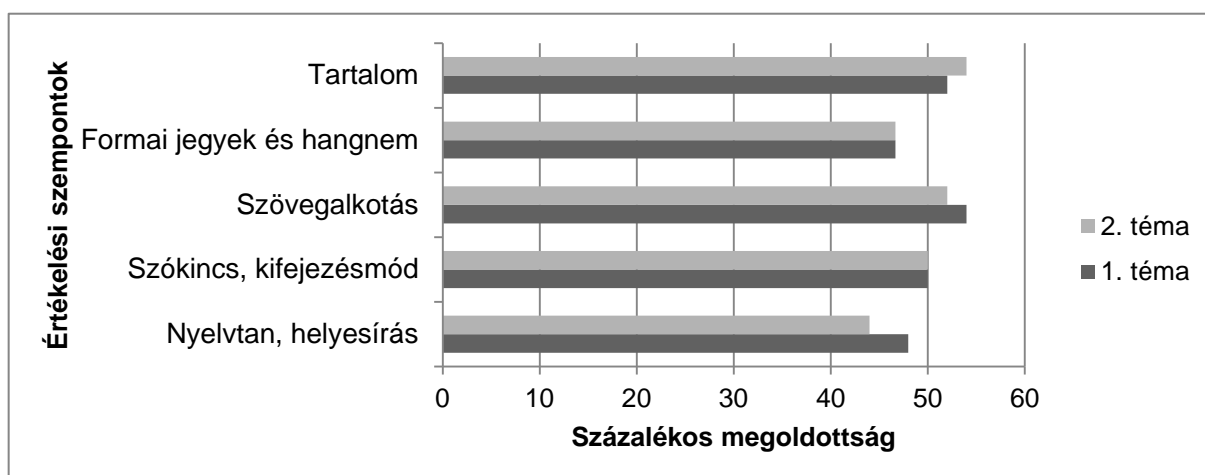
A második feladatban lévő két téma paraméterei közül a két téma átlagában nincs szignifikáns különbség ($d=0,3$; $p>0,05$). Ugyanez jellemző az összes értékelési szempont-ra is. Ez látható az egyes értékelési szempontok százalékos megoldottsága alapján is. Ezért a két témát nehézségében ekvivalensnek tekinthetjük, amely megfelel a vizsgaleírásban közölt elvárásoknak.

0.6. táblázat:

A 2. feladat értékelési szempontjainak átlaga, szórása témánként (VT számításai alapján)

Értékelési szempont	1. téma (N=451)		2. téma (N=512)		Két téma együtt (N=963)	
	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Tartalom	2,6	1,9	2,7	1,8	2,6	1,9
Formai jegyek és hangnem	1,4	1,3	1,4	1,2	1,4	1,2
Szövegalkotás	2,7	2,0	2,6	1,8	2,6	1,9
Szókincs, kifejezésmód	2,5	1,9	2,5	1,7	2,5	1,8
Nyelvtan, helyesírás	2,4	1,8	2,2	1,5	2,3	1,6
2. feladat	11,5	8,3	11,3	7,4	11,1	8,0

Az öt értékelési szempont megoldottságában szignifikáns különbség van ($\chi^2 = 213$; $p < 0,001$). A vizsgázók a tartalom és szövegalkotás szempontnál teljesítettek a legjobban (ezek megoldottságában nincs szignifikáns különbség: $t=0,8$; $p>0,05$). Ezekről szignifikánsan ($p<0,001$) alacsonyabb a szókincs, kifejezésmód értékelési szempont megoldottsága. A harmadik legnehezebben teljesíthető kritérium a formai jegyek és hangnem, amelyhez képest a nyelvtan, helyesírás szempontot a vizsgázók hasonló arányban teljesítették.



0.3. ábra:

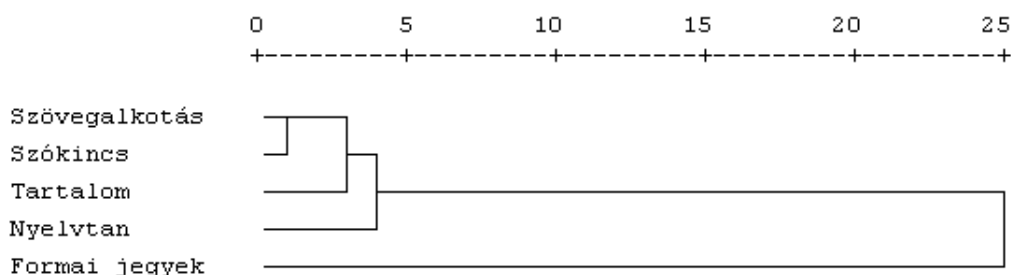
A 2. feladat megoldottsága szempontok szerint a két téma szerint (VT által készített grafikon)

Megvizsgáltuk, hogy az egyes értékelési szempontoknak mekkora a súlya az összpontszám alakulásában. Ennek eredményeit a 0.6. táblázat tartalmazza. Megvizsgáltuk továbbá klaszteranalízissel azt is, hogy milyen kapcsolat van az egyes értékelési szempontok között (0.4. ábra).

0.7. táblázat:

A 2. feladathoz tartozó értékelési szempontok magyarázóereje témánként és a teljes teszten
(VT számításai alapján)

Értékelési szempont	1. téma	2. téma	Teljes teszt
Tartalom	23,4	23,7	23,6
Formai jegyek és hangnem	11,9	12,4	12,2
Szövegalkotás	26,1	24,1	24,9
Szókincs, kifejezőmód	19,4	21,6	20,6
Nyelvtan, helyesírás	19,0	18,1	18,6
Megmagyarázott variancia (%)	100	100	100



0.4. ábra:

A 2. feladathoz tartozó értékelési szempontok klaszteranalízise
(VT által készített ábra)

Eredményeink szerint a tartalom és a szövegalkotás szempontnak van a legnagyobb magyarázóereje. Ezeket követi a két nyelvi szempont, a szókincs, kifejezőmód és a nyelvtan, helyesírás. A legkisebb magyarázóereje a formai jegyek és hangnem szempontnak van. E szempontnak a többihez képest alacsonyabb az összpontszámmal való korrelációja ($r=0,79$; $p<0,001$), és ez kapcsolódik a legtávolabbról a többi értékelési szempont-hoz (0.4. ábra). Az eredmények összhangban vannak a vizsgakövetelményekkel, hiszen a leglényegesebb, hogy a szerző a kommunikációs célnak megfelelő, koherens szöveget hozzon létre, és ez tükröződjön az egyes szempontoknak a tesztben lévő súlyában.

A 0.8. táblázat bemutatja témánként az egyes pontszámokat elérők számát és arányát. Ez alapján megállapítható, hogy kb. a vizsgázók fele teljesített a B1 szint alatt. A 0 pontot elérők száma és aránya elsősorban a tartalom szempont esetében releváns, hiszen ez alapján a többi szemponté is 0 pont. Az adatfelvétel következményeként elemzésünkben arról nem szerzünk információt, hogy mi jellemzi azokat az írásműveket, amelyekben a vizsgázók eltérnek a kommunikációs céltól. Erre nem tudunk válaszolni, hiszen tanulói dolgozatok nem állnak rendelkezésünkre, pedig ezek alapján az egyes irányító szempontok működését és a feladat esetleges problémáit is vizsgálhattuk volna.

0.8. táblázat:

A 2. feladat értékelési szempontjainál elért pontszámok abszolút és

relatív gyakorisága (zárójelben a százalékos arányok láthatók,
N1=451 és N2=512; VT számításai alapján)²¹

Pontszám	Tartalom		Szövegalkotás		Szókincs, kifejezőmód		Nyelvtan, helyesírás	
	1. téma	2. téma	1. téma	2. téma	1. téma	2. téma	1. téma	2. téma
0	122 (27)	102 (20)	123 (27)	111 (22)	122 (27)	104 (20)	121 (27)	104 (20)
1	30 (7)	49 (10)	23 (5)	44 (9)	23 (5)	49 (10)	30 (7)	66 (13)
2	61 (14)	75 (15)	43 (10)	68 (13)	53 (12)	88 (17)	63 (14)	112 (22)
3	54 (12)	78 (15)	70 (16)	96 (19)	88 (20)	96 (19)	96 (21)	116 (23)
4	80 (18)	100 (20)	77 (17)	89 (17)	79 (18)	108 (21)	86 (19)	86 (17)
5	104 (23)	108 (21)	115 (26)	104 (20)	86 (19)	67 (13)	55 (12)	28 (6)
B1 szint alatt	439 (46)		412 (43)		439 (46)		496 (51)	

Az alábbi táblázatból az értékelési skála egyik sajátossága is látható, hiszen az azonos sávba tartozó pontszámoknál (2–1 és 4–3) következetesen a magasabb pontszámot adják a vizsgáztatók, különösen a 2–1 pontoknál. Ez a jelenség felveti azt a kérdést, hogy célszerű-e a 4 szintleírásból álló értékelési skálát 6 pontra bontani. Ezt az értékelési rendszer jellemzőinek vizsgálatakor részletesebben is elemezzük.

²¹ Kék színnel jelöltük azon vizsgázók számát és arányát, akik nem érik el a B1 szintet, barnával pedig azokat, akik elérik ezt a szintet a 0.3. táblázatnak megfelelően.

4. Az értékelési rendszer elemzése

A továbbiakban az értékelési rendszer megfelelőségét vizsgáljuk és célunk annak megállapítása, hogy (1) milyen nehézségi szintekbe sorolhatók az egyes értékelési szempontok; (2) mennyire fedik le az értékelési skálapontok a vizsgázók íráskészség-szintjei által meghatározott intervallumot; (3) mennyire megfelelően működnek az értékelési skálák. Ebben a részben tehát a III. kutatási terület 8–10. kérdését elemezzük.

4.1. Az értékelési szempontok jellemzői és teljesítésének nehézsége

A parciáliskredit-modellel (Masters 1982) vizsgált adatbázis eredményeként a 0.9. táblázat az értékelési szempontok nehézségi indexét és modellilleszkedését mutatja a két feladatban. A táblázatban lévő logitérték a vizsgázók íráskészség-szintjének és az értékelési szempontokhoz tartozó nehézségi paraméterek közös skálájának egységét jelöli, amely intervallumskálán helyezkedik el. A középérték logittal az adott értékelési szempont átlagos nehézségét határoztuk meg a vizsgázók íráskészsége alapján. Egy nehezebb, magas középérték logittal rendelkező értékelési szempont esetében például a tanulók alacsonyabb pontértéket érnek el.

0.9. táblázat:

Az értékelési szempontok nehézségi indexe és modellilleszkedése (VT számításai alapján)

Feladat	Értékelési szempont	Középérték logitok	Infit MNSQ
1.	Tartalom és szöveg hossz	-1,25	1,7
	Nyelvhelyesség	-0,08	1,3
2.	Tartalom	0,02	0,9
	Formai jegyek és hangneme	0,45	1,6
	Szövegalkotás	0,01	0,7
	Szókincs, kifejezés mód	0,21	0,6
	Nyelvtan, helyesírás	0,64	0,8

Az első feladathoz tartozó értékelési szempontok teljesítésének nehézsége a vizsgázók átlagos készségszintje (0 logitérték) alatt helyezkedik el, a második feladat értékelési szempontjainak átlagos nehézsége az átlagos szintnél, valamint a formai jegyek és hangneme, illetve a nyelvtan, helyesírás szempont az átlag fölötti szinten van. Mindkét feladatban a leghigorúbban a nyelvi komponenseket értékelték. Az azonos szempontok közötti távolság a tartalomnál 1,3; míg a nyelvhelyességnél 0,7 logit, így a két feladat feltehetően különböző nyelvi szintet mér a vizsgakövetelmény elvárásainak megfelelően. A második feladatban az egyes értékelési szempontok megoldottságának megfelelően a vizsgázók

számára a tartalom és szövegalkotás szempont teljesítése bizonyult könnyebbnek. Ezeket követi a szókincs, kifejezésmód értékelési szempont átlagos nehézsége. A legnehezebben teljesíthető a nyelvtan, a helyesírás illetőleg a formai jegyek és hangnem szempontja. Ez összhangban van korábbi vizsgálatunkkal (Vígh, 2010: 128).

A 0.9. táblázatban szereplő infit MNSQ paraméterek megmutatják, hogy az adott értékelési szempont mennyire illeszkedik a többi közé. A megfelelően illeszkedő szempont értéke közel van 1-hez. Az elfogadható értékek sávja függ az elemszámtól, ezért nem tudunk egy előre meghatározott intervallumot kijelölni, hanem a vizsgálat céljainak függvényében kell mérlegelni (Molnár 2013: 93). A kb. 1000 fős mintánál 0,8 és 1,2 között határozhatjuk meg ezt a sávot.

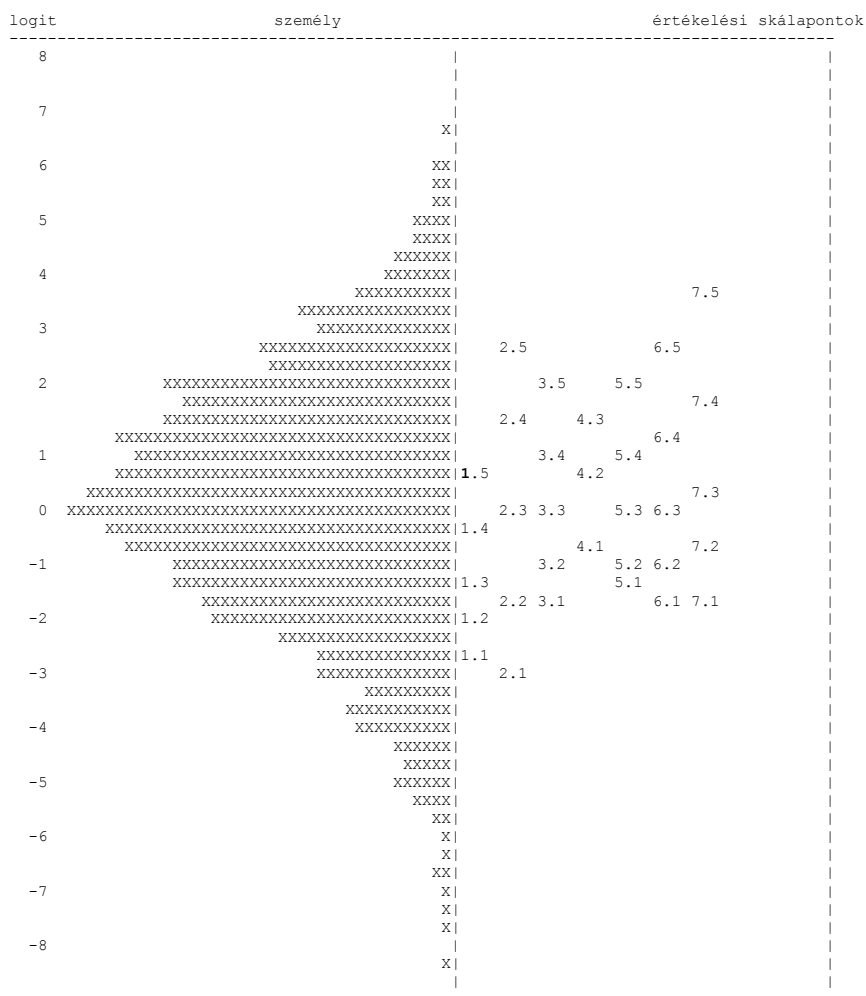
Az első feladat két szempontja, valamint a formai jegyek és hangnem nem illeszkedik, mivel az infit paraméter érték 1,2 fölött van (vö. 0.9. táblázat), ami azt jelenti, hogy e szempontokat többdimenzionalitás jellemzi, mást is mérnek. A jelenség értelmezésében érdemes a két feladatot önállóan is megvizsgálni. A két feladat elkülönített vizsgálata alapján az első feladathoz tartozó szempontok modellilleszkedése megfelelő. Ez azt jelzi, hogy az első feladat önállóan jól működik, a két feladat együtt viszont kevésbé, amely arra utal, hogy az első feladat az íráskészség más dimenzióját méri. Ennek oka, hogy a két feladatban az elérendő kommunikációs cél eltérő, a feladatok egymástól nemcsak nehézségükben, hanem az értékelési szempontok működésében is elkülönülnek. A formai jegyek és hangnem szempont az önálló elemzés során sem illeszkedik a többi szempont közé.

Két értékelési szempontnál utalnak az adatok túllilleszkedésre (az infit paraméter értékek 0,8 alatt vannak). Az önálló elemzés során is 0,7 körüli értéket kaptunk a szövegalkotás és a szókincs, kifejezésmód szempontnál (0.9. táblázat). Ez szintén összhangban van korábbi elemzéseinkkel (Vígh, 2010: 128–129). Mindkét esetben a diszkriminációs index magas (0,95). Ez arra utalhat, hogy az értékelési skála redundáns módon működik: egy bizonyos íráskészség-szintig alacsony lehet annak a valószínűsége, hogy a vizsgázók egy adott pontszámot érnek el, majd hirtelen megnő és szinte minden vizsgázónál a magasabb pontszám elérése a legvalószínűbb. Ennél a két szempontnál az értékelési skála nagy valószínűséggel nem tud különbséget tenni a magasabb szinten teljesítő diákoknál. A továbbiakban ezt a feltételezést empirikus adatokkal is megpróbáljuk alátámasztani.

4.2. A vizsgázók íráskészség-szintje és az értékelési szempontok nehézsége

A parciáliskredit-modell segítségével közös skálán tudjuk ábrázolni a vizsgázók íráskészség-szintjét és az értékelési skálapontok küszöbértékét. A küszöbértékek (vagy lépésparaméterek, Thurstone-küszöbértékek) megadják, hogy milyen képességszint szükséges ahhoz, hogy valaki 50%-os valószínűséggel érjen el egy

adott pontszámot, így ez az adott értékelési szempontozhoz tartozó átlagos nehézségi index értelmezésében játszik szerepet. A továbbiakban ezek segítségével az értékelési szempontrendszernek a vizsgázók íráskészség-szintjéhez való illeszkedést vizsgáljuk, amelyről a 0.5. ábra nyújt áttekintést.



0.5. ábra:
Az íráskészség vizsgarész-összetevő minta- és skálapont térképe
(VT által készített ábra)²² Minden 'x' két vizsgázót jelöl.

A 0.5. ábra első oszlopában az íráskészség feladatokon elért eredményeket egy közös skálán, ún. logitskálán helyezi el a ConQuest program (Wu, Adams és Wilson, 1998). A második oszlop a vizsgázók íráskészség-paramétereit tartalmazza, a harmadik az egyes értékelési szempontokhoz tartozó küszöbértékeket jeleníti meg. A pont előtti számjegy az értékelési szempontot, a pont utáni pedig a skálapontok küszöbértékét jelenti. Például a 3.2. a tartalom szempont 2-es pontszám elérésének küszöbértékét mutatja, amelynek 50%-os valószínűségi szintje a -1-es logitértéknél helyezkedik el. A továbbiakban először

²² A számok közül az első az alábbi értékelési szempontokat jelöli: 1=tartalom és szövegösszesség; 2=nyelvhasználat; 3=tartalom; 4=formai jegyek és hangneme; 5=szövegalkotás; 6=szóincs, kifejezőmód; 7=nyelvtan, helyesírás; a pont utáni második szám az adott értékelési szemponton belüli pontszám küszöbértékét.

a minta íráskészség-paramétereinek eloszlását, majd az értékelési skálapontok jellemzőit, végül a kettőt együtt elemezzük.

A mintatérképen minden 'x' két tanulót reprezentál. A logitskálán megadott negatív számok átlag alatti képességszintet, a 0 az átlagos, míg a pozitív számok az átlag feletti szintet jelölik. A térképről leolvashatók a tanulók íráskészség-paramétereinek jellemzői. A mintatérkép alapján megállapítható, hogy nagy különbség van a vizsgázók íráskészség-szintje között, a logitskála 15 egységnyi, így indokoltnak tűnik a vizsgán a két nyelvi szint mérése. A minta közel 50%-a az átlagos készségi szint felett helyezkedik el.

Az értékelési skálapontok akkor működnek az elvártnak megfelelően, ha az íráskészség szintjének növekedésével együtt arányosan nő a pontszám elérésének valószínűsége. A konkrét esetben az értékelési szempontokhoz tartozó küszöbértékek az íráskészség-szint növekedésével kivétel nélkül emelkednek. Minden értékelési szempont megfelelően működik: minél magasabb tudásszinttel rendelkeznek a tanulók, annál valószínűbb, hogy magasabb pontszámot érnek el.

Az első feladathoz tartozó értékelési skálapontok többsége (1. és 2. szempont) az átlagos készségi szint alatt helyezkedik el. Ugyanakkor a nyelvhelyességnél a 4. és 5. pont közötti távolság aránytalanul nagy, ha azt is figyelembe vesszük, hogy a feladatot az A2 szintre készítették. Egy korábbi, kismintás vizsgálat tapasztalatai azt mutatták (Einhorn, 2008), hogy ezt a szempontot a tanárok túlságosan szigorúan értékelik, nem differenciálnak az A2 nyelvi szintnek megfelelően: jellemző, hogy minden nyelvi hibát jelölnek és értékelnek a szövegben. Feltételezhetően e hibás gyakorlat okozhatja a nyelvhelyesség szempont nem megfelelő működését.

A második feladatban – a három pontszámból álló formai jegy és hangnem kivételével – a többi szempontnál az 1, 2 pontszám elérése az átlagos íráskészség-szint alatt jellemző, a 3 pontszám elérése pedig éppen a 0 logitértékre esik. Vagyis az e fölötti képességszintet a vizsgakövetelmények és vizsgaleírás alapján vélhetően a B1 szintet jelöli, ahogyan a 0.3. táblázatban láthatjuk. A kommunikatív komponenseket mérő szempontoknál a készségi szint növekedésénél viszonylag arányosan emelkedik az adott pontszám küszöbértéke. A szövegalkotás és a szókincs, kifejezőmód esetében az 1–2 pont közötti távolság kicsi (0,5 és 0,6) a 2–3 pont közötti távolsághoz (0,7 és 0,9) hasonlóan, a 3–4 pont közötti távolság a szövegalkotásnál 1, a szókincs, kifejezőmódnál 1,2 logitegység; a 4–5 pont között pedig 1,4 és 1,6. A nyelvhelyesség esetében az 1–2 és a 2–3 közötti távolság közel azonos (0,7 és 1,1 logitegység), majd növekszik: a 3–4 közötti távolság 1,6 logitegység, a 4–5 között pedig 1,9, így ennél a szempontnál az empirikus adatok nem követik a javítási-értékelési útmutatóban megadott skála szerkezetét. Az adatok arra utalnak, hogy ezt a szempontot a tanárok feltehetően túl szigorúan értékelik, elképzelhető, hogy főként e szempont mentén ítélik meg a vizsgázók íráskészség-szintjét. Ezt a

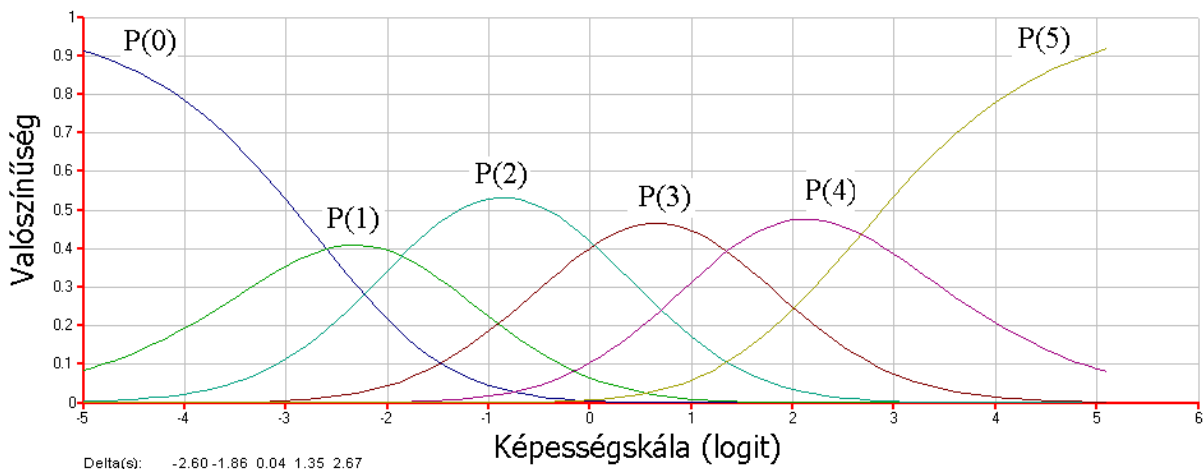
feltételezést azzal igazolhatjuk, hogy az ehhez a skálához tartozó pontszámok küszöbértékei szinte teljes mértékben lefedik a logitskálát.

A minta- és skálapont-térkép alapján megállapítható, hogy a 3-as logitértékig többnyire megfelelően fedik le az egyes értékelési szempontokhoz tartozó pontszámok küszöbértékei a vizsgázók íráskészség-szintje alapján meghatározott intervallumot.

4.3. Az értékelési skálák működése

Az előzőekben az egyes értékelési szempontok és skálapontok nehézségét, valamint a vizsgázók íráskészség-szintjét elemeztük, amelyek alapján számos megválaszolatlan kérdés maradt. Ezek közül a legfontosabb, hogy mi jellemzi az értékelési skálák működését, azaz a skálahasználat során minden deskriptor azonos mértékben hasznosul-e, vagy vannak-e olyan skálák, amelyeket a vizsgázatok nem használnak teljesen.

E kérdések megválaszolásához a parciáliskredit-moddal felrajzolható karakterisztikus görbéket vizsgáljuk. Ezek különböző íráskészség-szintek mellett annak valószínűségét mutatják, hogy a vizsgázó teljesítményét az adott értékkategoriába sorolják az értékelők. A továbbiakban két karakterisztikus görbét elemzünk, amelyek tipikusnak mondhatók. Ezek közül az első a nyelvtan, helyesírás értékelési szempontokhoz tartozó értékelési skála.



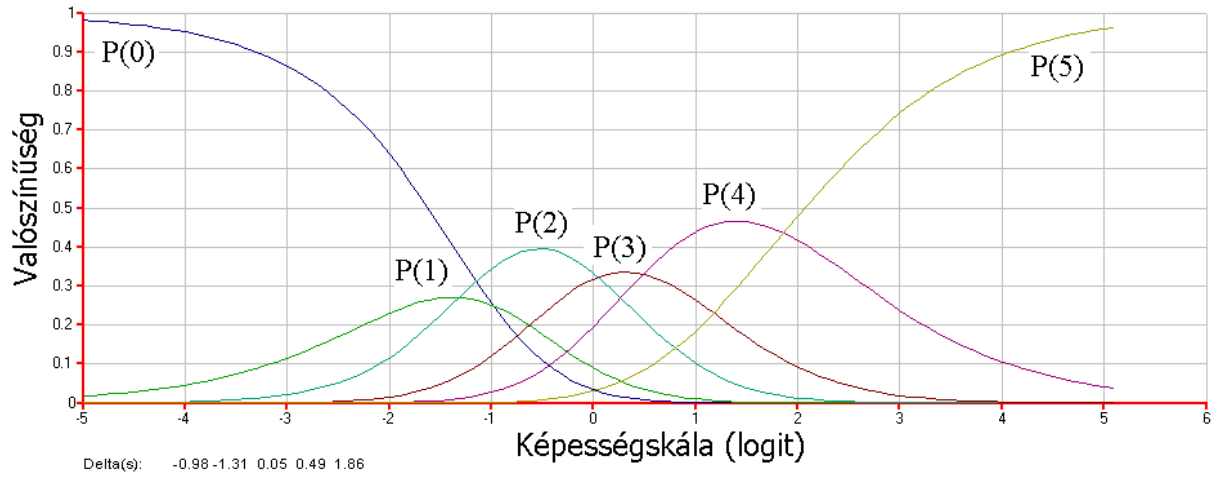
0.6. ábra:
Az első feladat nyelvhelyesség értékelési skálájának karakterisztikus görbéje (VT által készített ábra)

A 0.6. ábra nyelvtan, helyesírás szempontokhoz tartozó értékelési skála használatának eredményét mutatja. A legfelső, P(0) jelű görbe a 0 pontba tartozás valószínűségét jelöli, amely a készségi szint növekedésével folyamatosan csökken. Az 1 pontszám valószínűségi görbéje [P(1)] azt jelöli, hogy a készségi szint növekedésével hogyan növekszik, majd csökken az 1-es pontszámra való tartozás valószínűsége. Ugyanezt látjuk a többi pont-

számhoz tartozó görbénél is. Az egyes görbék metszéspontját a $\delta_1, \delta_2 \dots \delta_5$ paraméterek jelölik. A δ_1 paraméter (-2,6 logit) például azt a készségszintet jelöli, ahol azonos annak valószínűsége, hogy a vizsgázók 0 és 1 pontot érnek el. Az értékelési skálák működése szempontjából lényeges, hogy a delta értékek a képességskálán arányosan növekedjenek. Esetünkben az öt δ érték az elvártnak megfelelően emelkedik a készségszint növekedésével. Ugyanakkor a paraméterek közül a δ_1 és δ_2 egymáshoz közelebb van, mint a δ_2 és δ_3 , amely arra utal, hogy kevés azon diákok száma, akik 1 pontot kapnak és meghatározóvá válik az a vizsgázói csoport, amelyben 2, 3, 4 és 5 pontot érnek el. Ugyanakkor utalni kell arra is, hogy a δ értéke attól függ, hogy az egyes pontszám-kategóriában hány tanuló van, és éppen ebben a kategóriában vannak kevesebben és a mintában kizárólag a nappali tagozatos vizsgázók szerepelnek (vö. 5.2. fejezet).

A többi szempontnál hasonló skálaműködés látható. A karakterisztikus görbék azt mutatják, hogy a δ_1 és δ_2 értékek közel azonosak és előfordul, hogy fel is cserélődnek, majd a δ_3, δ_4 és δ_5 paraméterek többnyire azonos távolságra vannak. Ez azt jelzi, hogy az 1–2 pontba tartozás valószínűsége nagyon alacsony és meghatározóvá válnak azok az érettségizők, akik 2–5 pont közötti eredményeket érnek el, azaz a 6 pontszámból álló skála a nappali tagozatosok körében 5-fokozatúként működik. Ez arra utal, hogy az 1–2 pontos deskriptorra a vizsgált mintában nincs szükség, mert kevésbé jellemző a mintába került vizsgázókra, ezért az értékelők nem alkalmazzák ezeket a skálapontokat.

A tartalom értékelési szempontoz tartozó skála a második feladatban következetlenül működött (0.7. ábra). Jellemző az egyes paraméterek felcserélődése: $\delta_2 < \delta_1 < \delta_3 < \delta_4 < \delta_5$, ahol a δ_1 és δ_2 értéke hasonló és a δ_3 közel van a δ_4 -hez. Ez a skálaműködés feltehetően azért alakult így, mert az 1-es és a 3-as pontszámba való tartozás valószínűsége nagyon alacsony; ezek a görbék nagyon laposak. A 6-fokozatú skála 4-fokozatúként működik, vagyis mind a négy deskriptort használják az értékelők, csak az egyes pontszámok valószínűsége nem azonos. Feltehető az a vizsgáztatói gyakorlat, hogy az azonos deskriptorokat leíró 1–2 és 3–4 pont esetén a magasabbat adják a vizsgázónak.



0.7. ábra:
A második feladat tartalom értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)

5. Az íráskészség vizsgarész-összetevő elemzésének összefoglalása

Az íráskészség vizsgarész-összetevőhöz a bevezetőben (5.1. fejezet) két célterülethez (I. és III.) kapcsolódóan fogalmaztunk meg kutatási kérdéseket, amelyek közül az egyik a vizsgarész-összetevő megbízhatóságára (1. kérdés) és a feladatok súlyára (2. kérdés), a másik pedig az ebben a vizsgarész-összetevőben alkalmazott mérési-értékelési eljárásokra vonatkozott (7–10. kérdés).

A továbbiakban a kutatási kérdések számozásának megfelelően megadjuk először a vizsgált kérdést, majd pedig a választ.

1. Mennyire megbízhatóan mérték az egyes készségeket mérő feladatsorok?

Eredményeink szerint a vizsgarész-összetevőben szereplő két feladat mindegyike magas megbízhatósággal mért, vagyis kijelenthetjük, hogy az értékelési skálák használata megfelelő reliabilitást eredményezett mindkét feladatban.

2. Milyen mértékben magyarázzák az egyes feladatok az adott vizsgarész-összetevőben elért összpontszámot?

Az első feladat súlya kisebb az összpontszám alakításában, a másodiké nagyobb (0.4. táblázat). Ez abból adódik, hogy az első feladatnak alacsonyabb a pontértéke.

3. Hogyan teljesítettek a tanulók az egyes feladatokban? Van-e jelentős különbség a két feladat megoldottságában? Van-e jelentős különbség a második feladatban megadott két téma megoldottsága között?

Az első feladat a megoldottság és az értékelési szempontok nehézsége alapján a vizsgakövetelménynek megfelelően az A2, a második a B1 szinten helyezkedett el. E készség továbbra is jelentős kihívást jelent a vizsgázók számára, hiszen kb. a vizsgázók fele teljesített a B1 szint alatt.

A második feladatban szereplő választhatóság az elvártan megfelelően érvényesült. A mintába került vizsgázók közel azonos arányban választották mindkét témát, amelyek nehézsége sem összességében, sem az értékelési szempontok szerint nem különbözött jelentősen.

4. Milyen kapcsolat van az értékelési szempontok között, ezek mennyiben határozzák meg az összpontszámot és milyen nehézségi szintekbe sorolhatók?

A tanárok mindkét feladatban legszigorúbban a nyelvtan, helyesírás szempontot értékelték. Megállapítottuk, hogy a két feladat az íráskészség más dimenzióit méri. A második feladatban a formai jegyek és hangnem szempontnak alacsony a magyarázóereje, nem megfelelő a modellilleszkedése, jelentősen elkülönül a többi szemponttól. A vizsgafejlesztés folyamatában e szempontnak egyrészt az volt a funkciója, hogy segítse a vizsgázót abban, hogy a formai jegyek ismeretével

elérhessen néhány pontot, másrészt a vizsgakoncepció kialakításakor még a levél formai jegyeinek jelentős szerepe volt a német nyelv oktatásában. Azóta azonban a kézi levélírást felváltotta az e-mail, amelyben ezek a formai szabályszerűségek kevésbé érvényesülnek. Az eredmények tükrében javasoljuk e szempont törlését vagy esetleg az értékelési skála módosítását.

5. Mennyire fedik le az értékelési skálapontok a vizsgázók íráskészség-szintjei alapján meghatározott intervallumot?

Az értékelési skálapontok küszöbértékei az íráskészség-szint növekedésével többnyire arányosan emelkedtek és nagyrészt lefedték a vizsgázók készség szintje alapján meghatározott intervallumot. Ugyanakkor azt tapasztaltuk, hogy a tanárok vélhetően nem differenciálnak a nyelvtan, helyesírás szempontnál a mért nyelvi szinteknek megfelelően a hibák között, hiszen eredményeink arra utalnak, hogy ezt a szempontot a tanárok túl szigorúan értékelik és kiemelten kezelik. Ezen megállapítás pontosabb alátámasztásához ugyanakkor elengedhetetlen egyrészt az értékelők szigorúságának vizsgálata oly módon, hogy az is rögzítésre kerül, hogy egy vizsgáztató hány tanulói dolgozatot értékelt, és kellő elemszám esetében a többaspektusú modell (Linacre, 1994) segítségével történik az elemzés. Másrészt pedig kvalitatív módon az egyes tanulói dolgozatok vizsgálata is szükséges a tanári értékelés jellemzőinek elemzéséhez.

6. Mennyire jól rendezettek az egyes értékelési szempontokhoz tartozó pontszámok az értékelési skálákban?

Vizsgálatunkban az értékelési skálahasználatot karakterisztikus görbékkel jellemeztük, amelynek minőségi mutatói a delta-paraméterek. Ezek felcserélődése és egymástól való távolságuk jellemzik az értékelési skálák redundáns működését, valamint azt, hogy az értékelési skála bizonyos pontszámait az értékelők nem használják (például a tartalom szempontnál). Eredményeink szerint több értékelési skálánál az 1–2 pontot leíró deskriptorokkal volt probléma. Ugyanakkor figyelembe kell vennünk azt is, hogy az adatelemzés kizárólag a nappali tagozatosok körére vonatkozott. Az érettségi eredményei azt mutatják, hogy a felnőttoktatásban végzetek jóval alacsonyabb nyelvtudással rendelkeznek a nappalisoknál. Vagyis feltételezhető, hogy ha a teljes vizsgázói populációra vonatkozóan reprezentatív a mintavétel, akkor az értékelési skálák jobban működnek. Ezért ezek módosítására jelen elemzésünk alapján nem lehet javaslatot tenni, hiszen egyrészt szükséges kiterjeszteni a vizsgálatot más vizsgázói rétegekre is, másrészt elengedhetetlen a fenti paraméterek mentén több vizsgaév eredményeinek részletes elemzése és nyomon követése.

A kutatási kérdésekhez megfogalmazott hipotézisek közül az alábbiak igazolódtak be: A német nyelvi érettségi működésének jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján feltételeztük és elemzésünk során bebizonyosodott, hogy

1. az íráskészség vizsgarész-összetevő magas megbízhatósággal mér, amely a mérendő konstruktum jellemzőitől függ: tehát az értékelési skálák használata magas reliabilitást eredményez, valamint heterogén a csoport összetétele;
2. a két feladat súlya a tesztben eltér, amely ebben az esetben a mért nyelvi szinttől, az egyes feladatok minőségétől, valamint az értékelési rendszer jellemzőitől függ.

Az értékelési rendszer megfelelőségét korábbi kutatásunkban (Vígh, 2010) kialakított minőségi mutatók mentén elemeztük a parciáliskredit-modellel. Ebben a 2006. évi középszintű német nyelvi érettségi elemzése (Vígh, 2010: 127–131) alapján megfogalmazott kutatási eredmények alapján feltételeztük és jelenlegi elemzésünk alapján bebizonyosodott, hogy

7. a vizsgázók teljesítménye eltér a két feladatban, de a választható témák megoldottságában nincs jelentős különbség;
8. az értékelési szempontok magyarázóereje és nehézsége eltér, amely függ attól, hogy az íráskészség mely dimenzióját mérik;
9. az értékelési skálapontok többnyire lefedik a vizsgázók íráskészség-szintjei által meghatározott intervallumot;
10. az egyes értékelési skálák működése eltér szempontonként és a skálákban definiált pontszámok alapján.

Irodalom

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgakövetelmények.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv_vk.pdf (letöltve: 2013. május 20.)

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgaleírás.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv_vl.pdf (letöltve: 2013. május 20.)

EINHORN ÁGNES (2008): *A 2006. évi érettségi vizsga eredményeinek elemzése: Német nyelv*. Budapest: Oktatáskutató és Fejlesztő Intézet.

<http://www.ofi.hu/tudastar/erettsegi/2006-os-erettsegi/2006-évi-erettsegi-090617-2> (letöltve: 2013. május 24.)

Közös *Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés* (2002): Európa Tanács Közoktatási Bizottsága. Közoktatási Bizottság. Élő Nyelvek Osztálya. Strasbourg. Piliscsaba: Pedagógus-továbbképzési Módszertani és Információs Központ Kht.

LINACRE, J. M. (1994): *Many-faced Rasch Measurement*. Chicago: MESA press.

MASTERS, G. N. (1982): A Rasch model for partial credit scoring. *Psychometrika* 47/2, 149-174.

MOLNÁR GYÖNGYVÉR (2013): *A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában*. Budapest: Gondolat Kiadó.

Német nyelv. Középszintű írásbeli érettségi vizsga. 2012. május 11. (2012). Budapest: Nemzeti Erőforrás Minisztérium.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/kozep/k_nemet_12maj_fl.pdf (letöltve: 2013. június 27.)

Német nyelv. Középszintű írásbeli érettségi vizsga. Javítási és értékelési útmutató. 2012. május 11. (2012). Budapest: Nemzeti Erőforrás Minisztérium.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/kozep/k_nemet_12maj_ut.pdf (letöltve: 2013. június 27.)

VÍGH TIBOR (2010): *Az idegen nyelvi érettségi működése és hatása a tanulói teljesítmények és a tanári nézetek tükrében*. PhD-értekezés. Szeged: Szegedi Tudományegyetem, Bölcsészettudományi Kar, Neveléstudományi Doktori Iskola.

www.edu.u-szeged.hu/phd/downloads/vigh_ertekezés.pdf (letöltve: 2013. május 24.)

WU, M., ADAMS, R. J. és WILSON, M. R. (1998): *ACER ConQuest. Generalised Item Response Modelling Software*. Australia: ACER Press.

PETNEKI KATALIN, VÍGH TIBOR

**A 2012. MÁJUSI NÉMET NYELV
EMELT SZINTŰ VIZSGA
ELSŐ HÁROM FELADATSORÁNAK ELEMZÉSE**

Az emelt szintű német nyelvi érettségi vizsga első három vizsgarészének elemzési célja az I. és a II. kutatási terület vizsgálata (vö. I. Bevezető/4.1. fejezet²³). Ennek megfelelően először áttekintést adunk a vizsga általános jellemzőiről, a 2012. évi eredményekről, majd bemutatjuk az első három vizsgarész felépítését a vizsgaleírás és a feladatlap, illetve a javítókulcs alapján. Ezt követi készségenként a feladatok és az azokban előforduló itemek részletes elemzése. Végül összefoglaljuk az elemzés eredményeit.

²³ V.ö. nemet_dok_I.pdf dokumentum

1. A 2012. májusi vizsga általános jellemzői

Emelt szinten az írásbeli vizsga – a középszinthez hasonlóan – négy részből áll: olvasott szövegértés, nyelvhelyesség, hallott szövegértés és íráskészség. Az emelt szintű feladatokban elérhető vizsgapontok száma és itemszáma többnyire nem egyezik, így a 2012. évi vizsgában sem, ezért az itemek után nyert feladatpontokat át kell alakítani, amihez átszámítási kulcs szükséges.

0.1. táblázat:
Áttekintés a vizsgarészek feladatpontjairól (itemek) és vizsgapontjairól

Vizsgarész	Itemek száma	Vizsgapont
5. Olvasott szöveg értése	27	30
6. Nyelvhelyesség	45	30
7. Hallott szöveg értése	22	30
4. Íráskészség	-	30
(Szóbeli	-	30)
Összesen:		150

Jelen tanulmány a szóbeli vizsgarésszel nem foglalkozik, mivel ehhez nem állnak rendelkezésre adatok. Legtöbb adatot az olvasott szövegértés, a nyelvhelyesség és a hallott szövegértés vizsgarészhez sikerült begyűjteni. E három vizsgarész jellemzője, hogy a feladatokhoz itemek tartoznak, amiket vagy sikerül jól megoldani (1 pont), vagy nem (0 pont), így ezek statisztikai elemzése aránylag egyszerű.

A vizsga általános jellemzője, hogy sokféle feladattípust lehet alkalmazni. A lehetséges feladattípusokat a vizsgaleírás tartalmazza. Ezek közül a konkrét vizsga emelt szintű német nyelv feladatsora a két receptív készség és a nyelvhelyesség vizsgarész tekintetében az alábbi feladattípusokból áll:

0.2. táblázat:**Áttekintés a 2012. májusi emelt szintű vizsga első három vizsgarészének feladattípusairól**

Olvasott szövegértés	Nyelvhelyesség	Hallott szövegértés
5. sorba rendezés	5. szöveg kiegészítése	4. ál igaz-hamis (szelektálás)
6. hozzárendelés	6. szöveg kiegészítése	5. mondatok kiegészítése
7. igaz-hamis	7. hibazonosítás	
8. hozzárendelés	8. szöveg kiegészítése	

A feladattípusok nem csak egy bizonyos készséghez kötődnek, hanem mind a három vizsgarészben előfordulhatnak. A feladatsor e téren nem mutat túl nagy változatosságot.

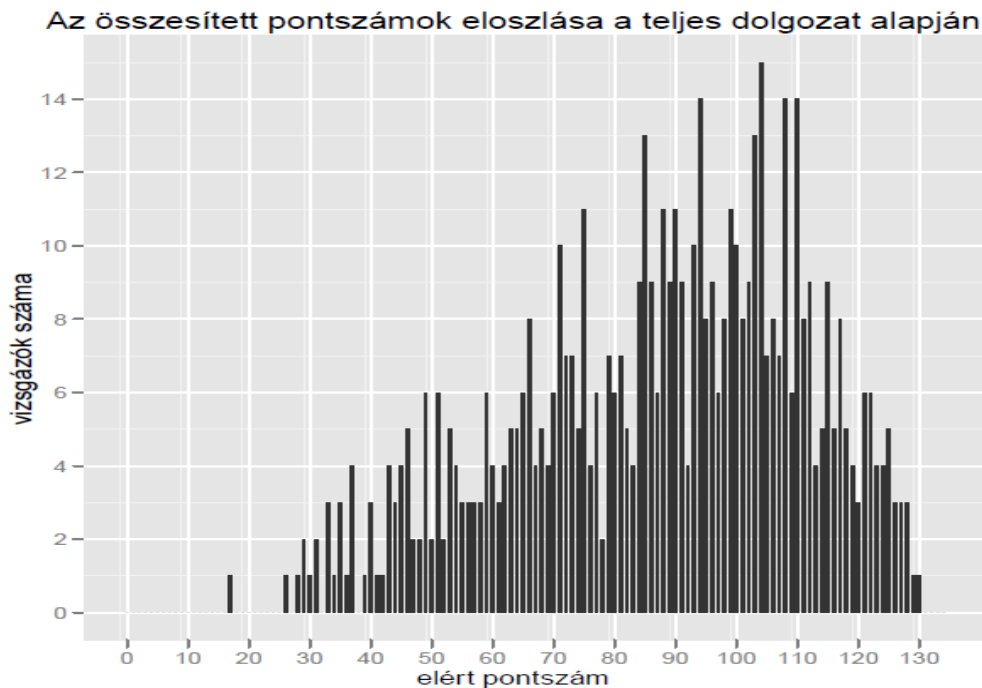
0.3. táblázat:**Áttekintés a 2012. májusi emelt szintű vizsga feladattípusainak előfordulásáról az egyes vizsgarészekben**

Feladattípus	Olvasás	Nyelvhelyesség	Hallás	Összesen
sorba rendezés	1			1
hozzárendelés	2			2
igaz-hamis	1			1
ál igaz-hamis (szelektálás)			1	1
szöveg kiegészítése: megadott szavak helyének megtalálása és beillesztése		3		3
mondatok kiegészítése			1	1
hibafelismerés		1		1

0.4. táblázat:**Az itemek eloszlása feladattípusonként az egyes vizsgarészekben, illetve feladatsoron belül**

Vizsgarész	Feladattípus	Itemszám	Összesen
Olvasott szövegértés	sorba rendezés	8	27
	hozzárendelés	6	
	igaz-hamis	7	
	hozzárendelés	6	
Nyelvhelyesség	szöveg kiegészítése	8	45
	szöveg kiegészítése	10	
	hibazonosítás	15	
	szöveg kiegészítése	12	
Hallott szövegértés	ál igaz-hamis (szelektálás)	9	22
	mondatok kiegészítése	13	

Emelt szintre 578 dolgozat alapján készült adatgyűjtés és –feldolgozás. A megbízó által készített adatfeldolgozásban kaptunk kész ábrákat, így a teljes vizsgán elért pontszámokról is (0.1. ábra).



0.1. ábra:

**Az összesített pontszámok eloszlása a teljes dolgozat alapján
(a megbízó által készített ábra – 10 ponttal elcsúsztatva,
tehát az ábrán 10 pont egyenlő nullával, 130 pont pedig 120 pontnak felel meg –
informatikai megjelenítési probléma következtében)**

A következőkben a feladatokat és azon belül az itemeket egyenként elemezzük. Az elemzést részben a megkapott bázisadatokból a megbízó által feldolgozott adatok, ábrák segítségével, részben önálló elemzés alapján mutatjuk be.

2. Készségenkénti (vizsgarészenkénti) elemzés

2.1. Olvasott szöveg értése (I. vizsgarész)

„A vizsga célja annak mérése, hogy a vizsgázó képes-e a mindennapi életben előforduló, különböző fajtájú autentikus szövegeket önállóan elolvasni, és az olvasási céloknak megfelelő mélységben megérteni. A vizsgafeladatok megoldásához a vizsgázónak képesnek kell lennie az értési céloknak megfelelő stratégiák alkalmazására is. (...) A közép- és az emelt szint közötti különbség az adott téma megközelítési módjában, a szövegek terjedelmében, tartalmi és nyelvi komplexitásában, valamint a feladatok eltérő nehézségi fokában jelenik meg” (40/2002 OM rendelet, Vizsgaleírás: 9). A vizsgaleírás azt is részletezi, hogy pontosan miben áll ez a magasabb szint. Itt ugyanis már nem elegendő tényyszerű információkat kiszűrni, hanem absztraktabb témákat is tudni kell követni, véleményeket megkülönböztetni, stb.

A feladatsor 3–4 szövegből és 3–4 feladatból áll. Egy feladaton belül egy hosszabb szöveg helyett előfordulhat több rövidebb szöveg is. A feladatokhoz felhasznált szövegek együttes terjedelme 1300-1500 szó. A feladatsor 25-30 itemből áll. A 2012. májusi feladatsor 4 szöveget és 4 feladatot tartalmazott, amire összesen 27 feladatpontot lehetett kapni. A szövegek terjedelme csak minimálisan haladta meg az előírt maximális összterjedelmet.

0.5. táblázat:

Áttekintés az 1 vizsgarész (olvasott szövegértés) feladatsoráról

Feladattípusok	Mérési cél	Szöveg	Szöveghossz ²⁴
sorba rendezés	részletes értés	A távcső története	265 szó
hozzárendelés	részletes értés	A piercing szerepe	426 szó
igaz-hamis	szelektív értés	A helyes életmód szenzora	332 szó
hozzárendelés	részletes értés	A víz alatti világról	554 szó
		össz.:	1577 szó

Az 1. vizsgarész 4 feladatból, összesen 27 itemből áll, amire 30 vizsgapontot lehet kapni. A teszt egészének paramétereinek vizsgálatkor százalékpontban adjuk meg az eredményeket, hogy az eltérő itemszámú feladatok és a teszt átlaga és szórása összehasonlítható legyen. A feladatsor jellemzőit mutató 0.6. táblázat alapján megállapítható, hogy a teszt egésze magas megbízhatósággal mért. Ugyan a feladatsort átlagosan 60% felett sikerült megoldani, de az is látszik, hogy a 3. feladat esetében komoly problémák vannak: A reliabilitást mérő számnak magasabbnak kellene lennie, a szórás értéke is itt a leg-

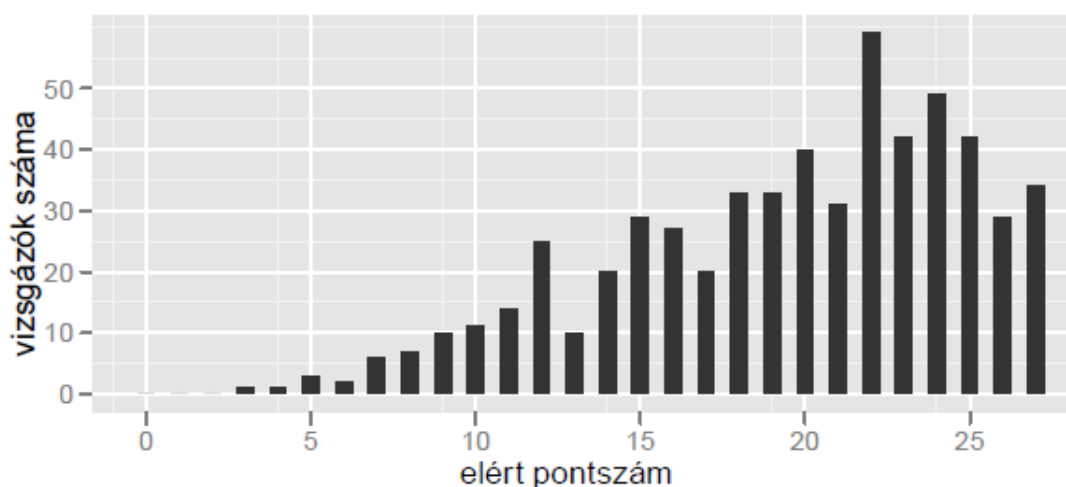
²⁴ A szöveghosszat sem a feladatlap, sem a javítókulcs nem tartalmazza. Így a hosszúságot úgy állapítottuk meg, hogy a letölthető vizsgaszöveget átmásoltuk word fájlba, ami kimutatja a szavak számát a szövegben.

alacsonyabb, ugyanúgy, mint a magyarázó erőé, tehát a feladat súlya az összpontszám alakításában csekély. Ez a feladat az elsővel együtt a legkönnyebb volt a tesztben. Ezek-től szignifikánsan nehezebb volt a negyedik feladat ($p < 0,001$), míg az egész teszt legnehezebb feladata a 2. volt.

0.6. táblázat:
Az olvasott szöveg értéke feladatsor jellemzői
(a megbízótól kapott adatok és VT²⁵ számításai alapján)

Kategória	Feladat				Teszt
	1.	2.	3.	4.	
Cronbach-alfa	0,84	0,65	0,37 ²⁶	0,75	0,86
Átlag (%p)	76,4	62,5	76,8	70,0	72,0
Szórás (%p)	28,9	27,7	18,3	29,0	19,7
Magyarázóerő (R ² %)	36,5	22,7	14,9	25,9	100

A 0.2. ábra azt mutatja, hogy az elérhető 27 feladatpontból milyen pontszámot értek el a vizsgázók a vizsgált 578 dolgozat esetében. A pontos számok az ábráról nem olvashatók le, de az igen, hogy sokan értek el 20 pontnál többet. A teljesítmények jobbra, a magasabb teljesítménysávba tolódnak. A statisztikai elemzésből kiderült az is, hogy ebben a vizsgarészben senki sem hagyott üresen itemet.



0.2. ábra:
A teszten nyújtott pontszám abszolút gyakorisági eloszlása
(a megbízó által készített grafikon)

²⁵ VT rövidítéssel jelöltük Vigh Tibor számításait.

²⁶ A pirossal jelölt adatok a problémás értékekre vonatkoznak.

A Rasch-modellel az egyes feladatokhoz tartozó itemek nehézségét és a mintába került diákok képességszintjét a következő minta- és itemtérképen ábrázoljuk (0.3. ábra). Így össze tudjuk hasonlítani a diákok képességszintjét a feladatlapon szereplő itemek 50 százalékos valószínűséggel történő megoldásához szükséges képességszintekkel. Mivel a Rasch-modellel elemző szoftverek az adatokat intervallumskálára transzformálják, a térképről leolvasható, hogy az adott item mennyivel könnyebb vagy mennyivel nehezebb, illetve a vizsgált személy mennyivel magasabb-alacsonyabb képességű. Mivel emelt szinten a mintában részt vevő vizsgázók száma viszonylagosan kevés, ezért itt – szemben a középszinttel – minden 'x' csak egy vizsgázót jelöl. (A középszintű minta-itemtérképeken²⁷ 'x'=2 vizsgázó.) Ez alapján a vizsgázók kb. 70%-a teljesített az átlagos (logitérték=0) képességszint felett.

²⁷ V.ö. nemet_dok_II.pdf dokumentum

logit	személy	1.f	2.f	3.f	4.f	+item
6						
5		X				
		X				
		XX				
		XX				
4		XX				
		XXXXXX				
		XXXXXXXXXX				
		XXXXXX				
		XXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
3		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
2		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX	11 13			
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX			22	
		XXXXXXXXXXXXXXXXXX				
1		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX	12	16 21		
		XXXXXXXXXXXXXXXXXX				
		XXXXXXXXXXXXXXXXXX			23 26	
		XXXXXXXXXXXXXXXXXX			24	
		XXXXXXXXXXXXXXXXXX	6 7	9		
		XXXXXXXXXXXXXXXXXX				
0		XXXXXXXXXXXXXXXXXX	5 8			
		XXXXXXXXXXXXXXXXXX	4	14		
		XXXXXXXXXXXXXXXXXX			17	
		XXXXXXXXXXXXXXXXXX	1 2 3	10		
		XXXXXX				
		XXXXX			18	25
-1		XXXX				
		XXX				
		XXX				
		XXX			15	27
		X				
		X				
-2					20	
-3						

0.3. ábra:
Az olvasott szöveg értéke feladatsor minta-itemtérképe
(VT által készített ábra)

A minta-itemtérképen az egyes feladatokhoz tartozó itemek nehézségi szintjét látjuk a képességskála alapján. Az itemek eloszlása azt mutatja, hogy leginkább az 1. feladat

felelt meg a mintában részt vevő diákok átlagos képességeinek, vagyis ez a feladat vélhetőleg a B2 szintet képviseli, míg a 3. feladat itemeinek többsége a negatív tartományba került, ami alatta van a vizsgált mintában résztvevők képességszintjének. A 2. feladat kissé nehezebb, csupán két item (14., 16.) van a 0 logit-érték alatt, a többi fölött. Ebben a feladatban volt a feladatsor két legnehezebb iteme (11., 13.). A 4. feladat igen eltérő nehézségű itemeket tartalmazott. Az itemeknek kb. a fele a 0 érték fölött, míg a másik fele alatta volt. Ebből 7 item került az átlagos 0 érték közelébe. Az emelt szintű feladatoknak elméletileg azonos szintet, a Közös Európai Referenciakeret (2002) B2 szintjét kellene mérnie, de úgy tűnik, hogy ez nem teljes egészében valósult meg. Az itemek nehézsége nem fedi le a vizsgázók képességszintjei alapján meghatározott intervallumot.

A továbbiakban részletesen az egyes feladatokhoz tartozó itemeket elemezzük. Először a feladatokat jellemezzük, majd bemutatjuk az egyes itemek paramétereit és ezeket vizsgáljuk, végül a teljes feladatsorra vonatkozóan vonunk le következtetéseket.

1. Feladat: Das Teleskop oder Fernrohr (8 item)

A szöveg a távcső felfedezésének és fejlődésének rövid története, jellegében tehát népszerű ismeretterjesztő szövegnek tekinthető (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 4). A szöveget azonban nem egyben olvashatják a vizsgázók, hanem 9, kb. egyenlő részre vágva és összekeverve. A feladat abban áll, hogy helyre kell állítani az eredeti szöveget a szétvágott részek helyes sorrendbe rakásával. Ez intenzív olvasást kíván. A vizsgázó annyi segítséget kap, hogy a feladatlap megadja a szöveget kezdő részletet, így összesen 8 részt kell a megfelelő helyre rakni. Feladatpontot a megfelelő kapcsolódási pontokra lehet adni. A feladat érdekessége, hogy az itemek erős összefüggése ellenére ez a feladat magas megbízhatósággal rendelkezik (v.ö. 0.6. táblázat).

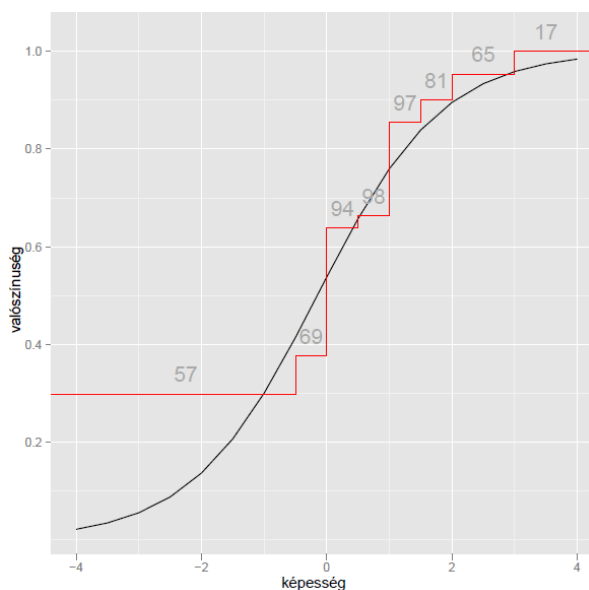
0.7. táblázat:

**Az olvasás 1. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)**

Itemszám	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ
[1.]	0,83	0,38	0,66	0,34	-0,63	0,76
[2.]	0,82	0,38	0,54	0,24	-0,59	0,89
[3.]	0,81	0,39	0,62	0,31	-0,54	0,84
[4.]	0,77	0,42	0,60	0,26	-0,22	0,86
[5.]	0,74	0,44	0,56	0,21	-0,01	0,91
[6.]	0,68	0,47	0,58	0,25	0,36	0,92
[7.]	0,70	0,46	0,54	0,23	0,27	0,97
[8.]	0,76	0,43	0,56	0,26	-0,13	0,89

A feladat nehézségi indexe – két item kivételével – mínusz tartományban van, így ez az átlagosnál könnyebbnek tekinthető. Mivel az itemek összefüggnek, ezért többségük hasonló képességszintet mér, itt a legkisebb a különbség az itemek nehézségében, amely

a feladattípusból adódik. Az átlag feletti képességszinten 50%-os megoldási valószínűséggel rendelkező itemek megoldottsága alacsonyabb (70% körüli). Minden item pontbiszeriális korrelációja és magyarázóereje magas. Az Infit MNSQ az item empirikus adatainak a Rasch-modellhez való illeszkedését mutatja. Az érték annál jobb, minél közelebb van 1-hez. Ezt az értéket leginkább az 5. 6. és 7. item közelíti meg a legjobban. Ez az item karakterisztikus görbéjén is látható, mert ebben az esetben közel azonos meredekségű az elméleti görbével. Ugyanakkor a minta elemszámát figyelembe véve minden item modellilleszkedése elfogadható.



0.4. ábra:
Az olvasás 1. feladat 7. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (a megbízó által készített grafikon)

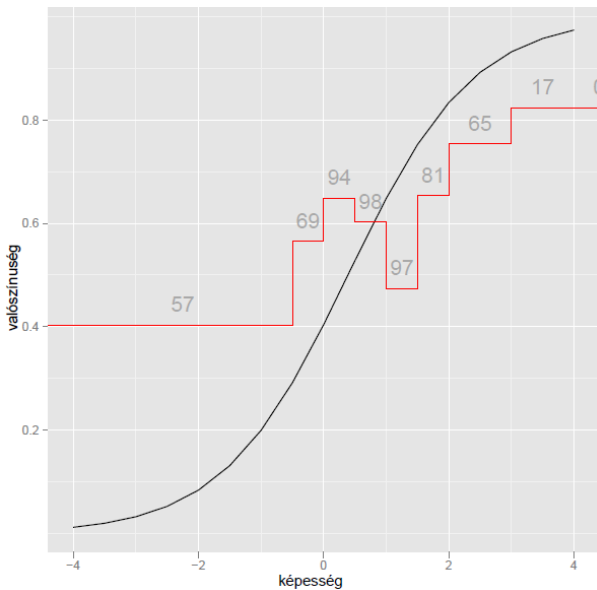
2. Feladat: Die meisten Mädchen wollen sich verschönern (6 item)

A szöveg a Spiegelonline-ról származó interjú egy pedagógussal, akivel arról beszélget az újságíró, hogy vajon miért népszerű a fiatal lányok körében a piercing. A szövegtípusból adódóan a szöveg kérdésekből és válaszokból áll. A feladat helyes sorrendben tartalmazza a válaszokat, de a hozzá tartozó kérdések össze vannak keverve, ráadásul eggyel több kérdés van, mint amennyire szükség van (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 6-7). A feladat abban áll, hogy meg kell találni, melyik kérdés illik az adott válaszhoz, és így kell helyreállítani az eredeti szöveget. A meglehetősen hosszú szöveghez mindössze 6 item kapcsolódik.

0.8. táblázat:
Az olvasás 2. feladat itemeinek jellemzői (a megbízótól kapott adatok és VT számításai alapján)

Itemszám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ
----------	-------	--------	---	----------------	----------	------------

[9.]	0,70	0,46	0,51	0,19	0,23	0,95
[10.]	0,82	0,38	0,59	0,28	-0,61	0,82
[11.]	0,43	0,50	0,46	0,14	1,69	1,05
[12.]	0,60	0,49	0,25	0,02	0,84	1,29
[13.]	0,42	0,49	0,37	0,07	1,77	1,13
[14.]	0,78	0,42	0,50	0,24	-0,24	0,96



A nehézségi index alapján két itemet (10., 14.) találunk mínusz tartományban, kettőt (11., 13.) pedig messze az 1 érték felett. A 12. item pontbiszerialis korrelációja és magyarázóereje alacsony, azaz az item nem azt méri, amit a teszt egésze, vagy nem jól méri azt, amit a teszt egésze. Az Infit MNSQ értéke meghaladja az 1 értéket. Ha az érték nagyobb, mint 1, akkor laposabb az item empirikus görbéje, mint a modellált görbe, azaz nem különbözteti el kellően a különböző képességszintű tanulókat. E feladat minőségének elemzéséhez nagyon hiányzik, hogy az adatbázisban kizárólag az itemek pontértéke szerepel, de az nem, hogy az egyes

itemeknél hibázáskor melyik alternatívát választották a vizsgázók. A feladat sajátossága,

0.5. ábra:

Az olvasás 2. feladat 12. itemének elméleti és tapasztalati megoldási valószínűségét jelölő

hogy ha egy helyen rontanak, akkor az a többi item működését is befolyásolja. Feltehetően ez jellemzi a 11., 12., 13. itemet. Például a 11. itemnél a D, a 13. itemnél a C válaszlehetőség jelenthetett problémát, amely befolyásolhatja a 12. item működését.

3. Feladat: Sensor macht gesündere Lebensführung sichtbar (7 item)

Az ismeretterjesztő szöveg egy olyan eszköztől számol be, ami visszajelzi használatának, hogy javított-e életstílusán. A szöveghez kijelentések tartoznak, amelyekről el kell dönteni, hogy azok a szöveg alapján igazak-e vagy hamisak (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 8-9). Ez a feladattípus a könnyűnek mondott feladattípusok közé tartozik, 50% találgatási eséllyel jár, ám kérdés, mennyire sikerült úgy megfogalmazni a kijelentéseket, hogy azok a mérési célnak megfeleljenek.

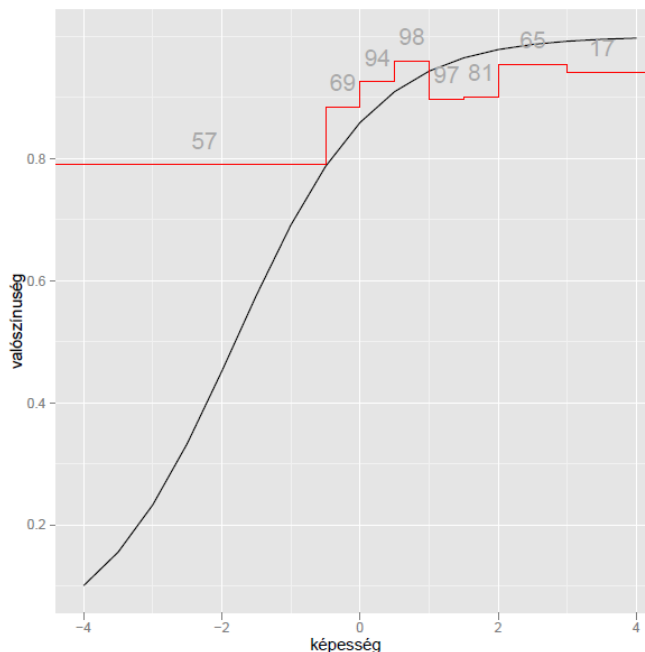
0.9. táblázat:

A 3. feladat itemeinek jellemzői (a megbízótól kapott adatok és VT számításai alapján)

Itemszám	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R^2	Nehézség	Infit MNSQ
----------	-------	--------	---	-------	----------	------------

[15.]	0,91	0,29	0,14	0,01	-1,48	1,15
[16.]	0,62	0,49	0,29	0,07	0,73	1,26
[17.]	0,79	0,41	0,37	0,10	-0,33	1,09
[18.]	0,86	0,35	0,45	0,19	-0,91	0,97
[19.]	0,67	0,47	0,46	0,22	0,44	1,03
[20.]	0,94	0,24	0,21	0,04	-2,02	1,07
[21.]	0,60	0,49	0,08	0,00	0,81	1,48

Ebben a feladatban a 15., 16. és 21. item modellilleszkedése nem megfelelő, a 16. és 21. item esetében vélhetőleg a találgatásnak is szerepe volt. A 15. és 21. itemnek pedig rendkívül alacsony a pontbiszerialis korrelációja és magyarázóereje, tehát nem jól mér. A 15. és 20. item nagyon könnyű volt, ezek nem érik el a B2 szintet. Feltételezhető oka ennek az lehet, hogy a kijelentések igazságtartalmát a szöveg megértése nélkül is sejteni lehetett. Itt tehát valószínűleg nem sikerült jól megfogalmazni a szöveghez rendelt igaz-hamis kijelentéseket. A feladat egészét tekintve, hétből csak három, a 17., 18. és 19. item megfogalmazása volt megfelelő.



0.6. ábra:
Az olvasás 3. feladat (legtöbb problémát mutató) 15. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

4. Feladat: Tauchen: Unterwasserwelt mit Abstand betrachten (6 item)

Az utolsó, egyben leghosszabb szöveg arról szól, hogy aki nyaraláskor a tengerekben búvárkodik, hogyan óvhatja a vizek élővilágát (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 10–11). A hosszú szövegben helyenként hiányzik 1-1 mondat, összesen (a példát leszámítva) hat. Viszont hét mondat közül kell kiválasztani, hogy a szövegösszefüggés alapján melyik mondat illik a legjobban az adott helyre. Ez nagyon alapos olvasást kíván.

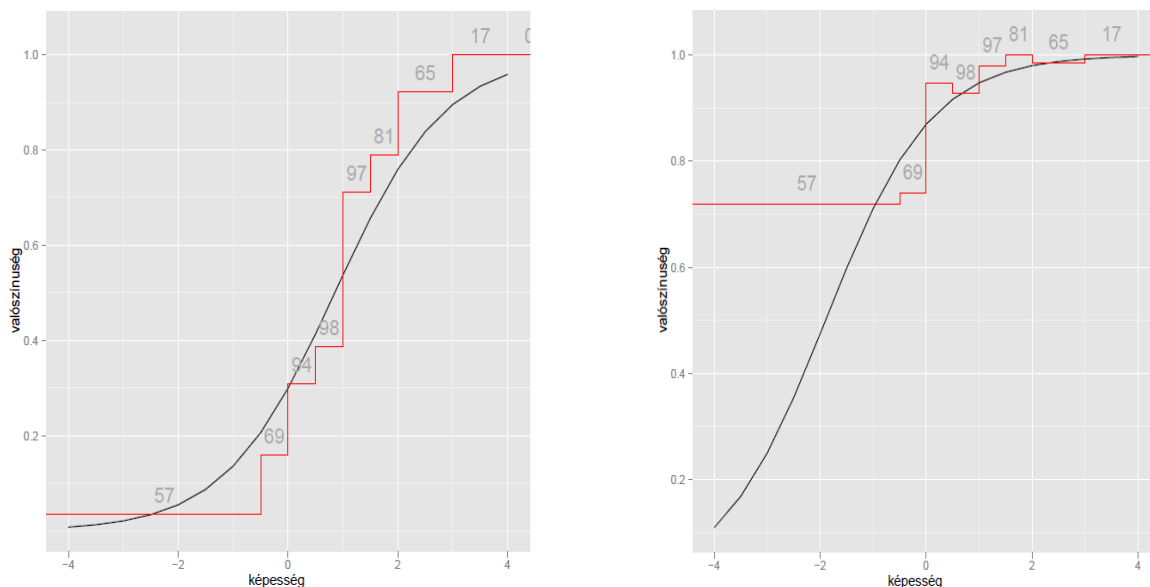
0.10. táblázat:

Az olvasás 4. feladat itemeinek jellemzői (a megbízótól kapott adatok és VT számításai alapján)

Itemszám	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ
[22.]	0,50	0,50	0,61	0,35	1,34	0,87
[23.]	0,63	0,48	0,59	0,32	0,66	0,91
[24.]	0,66	0,47	0,53	0,26	0,48	0,94
[25.]	0,86	0,35	0,43	0,17	-0,95	0,98

[26.]	0,63	0,48	0,59	0,32	0,64	0,89
[27.]	0,92	0,28	0,37	0,11	-1,59	1,00

Ebben a feladatban a 22. item volt a legnehezebb és az utolsó, a 27. item a legkönnyebb. Ez utóbbira a legvalószínűbb magyarázat az lehet, hogy a feladat végére csak a disztraktor és egy mondat maradt, a disztraktor pedig nagyon eltért attól, ami abba a szövegrészbe kell, tehát könnyen megkülönböztethető volt.



0.7. ábra: Az olvasás 4. feladat legnehezebb, 22. itemének és legkönnyebb, 27. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

Összegzés: Az 1. vizsgarész feladatsorában a legjobb mutatókkal az 1. feladat rendelkezik, de a 2. és 4. feladat is megfelel az elvárásoknak. Ez a három feladat lényegében az adott szöveg valamilyen formában történő rekonstrukciója, így a feladatíró tulajdonképpen nem tesz hozzá semmit, csak a rekonstrukció módját szabja meg. A kiugróan problémás 3. feladat viszont az igaz-hamis feladattípust képviselte, ahol a kijelentéseket a feladatírók készítik. Feltehető, hogy ez a típus emelt szinten nem igazán mér, hiszen reliabilitása alacsony. Célszerű lenne helyette más feladattípust alkalmazni.

2.2. Nyelvhelyesség (II. vizsgarész)

A 2. vizsgarész célja annak a mérése, hogy a középszintű érettségi elementáris nyelvhasználójával szemben képes-e változatos és bonyolultabb nyelvi szerkezeteket és lexikai kifejezéseket felismerni és megfelelő formában használni. Ennek során komplexebb autentikus szövegek komplexebb szerkezeteinek felismerését és használatát várja el a vizsga. Mivel azonban az itt használható feladattípusok zártak, nehéz a nyelvhelyességet valóban B2 szinten mérni. A vizsgakövetelmények nyelvhelyességi

elemeket felsoroló táblázata (40/2002 OM rendelet, Vizsgakövetelmények: 15–19) csak nagyon kevés olyan elemet sorol fel, amit nem kell már a középszinten is tudni. Ez a Plusquamperfekt használata aktív és passzív formában, a módbeli segédigék összetett múlt idejű alakjai, az „es” általános névmás használata, a határozatlan névmás ragozott alakjának ismerete, a vonatkozó névmás birtokos esete, valamint a birtokos esettel járó előjárók és a dupla kötőszavak ismerete és használata. Mondattanból csak az óhajtó mondat és az „als ob” kötőszóval járó mellékmondatok jelentik az emelt szintet – ezekre a középszinten nem lehet rákérdezni. Igen nehéz tehát olyan feladatsort összeállítani, ami ráadásul autentikus szövegekre épül, hogy elsősorban az előbb említett, B2 szintű nyelvtanra kérdezzen. A legtöbb jelenség ugyanis már a középszinten követelmény. Így csak a nehezebb szöveggörnyezet emelheti magasabb szintre a feladatokat. A feladatsor 4–5 feladatból, 30–45 itemből áll, a feladatokhoz felhasznált szövegek együttes terjedelme 700–900 szó. A konkrét vizsgán nem vizsgáljuk a terjedelmet, mert formázási okokból ebben a vizsgarészben különösen nehéz utólag meghatározni a pontos szószámot.

0.11. táblázat:
Áttekintés a 2. vizsgarész (nyelvhelyesség) feladatsoráról

Feladattípus	Itemszám	Összesen
szöveg kiegészítése	8	45
szöveg kiegészítése	10	
hibaazonosítás	15	
szöveg kiegészítése	12	

A vizsgarész négy, zárt típusú feladatból áll, amelyek jellemzőit a 0.12. táblázatban foglaltuk össze:

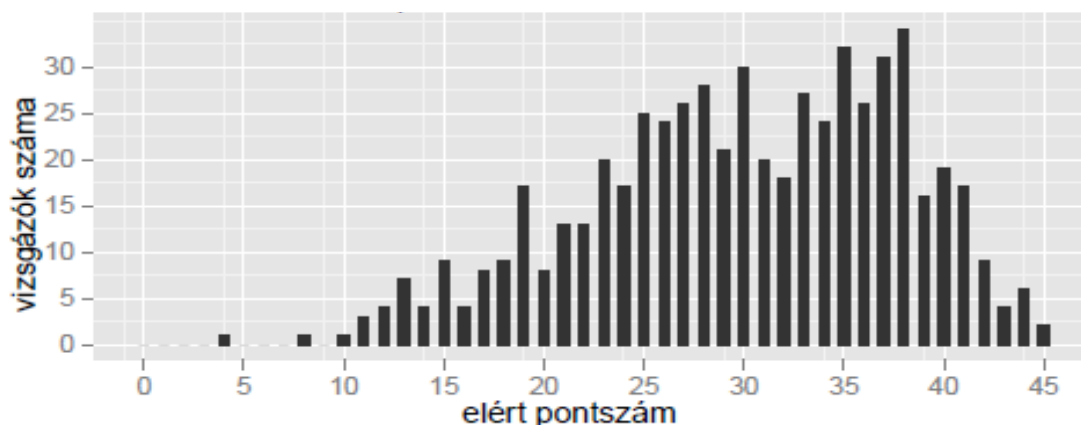
0.12. táblázat:

A nyelvhelyesség feladatsor jellemzői (a megbízótól kapott adatok és VT számításai alapján)

Kategória	Feladat				Teszt
	1.	2.	3.	4.	
Cronbach-alfa	0,52	0,46	0,69	0,80	0,88
Átlag (%p)	56,0	75,0	68,1	64,2	66,5
Szórás (%p)	21,4	16,1	19,1	25,5	17,3
Magyarázóerő (R ² %)	17,5	15,0	32,0	35,4	100

A feladatsor egésze – köszönhetően a magas itemszámnak – megfelelő megbízhatósággal mért, egyes elemei közül a legtöbb problémát az amúgy legjobb átlaggal megoldott 2. feladat mutatja. Ennek a feladatnak a leggyengébb a reliabilitás-mutatója, szórása alacsony, és gyenge a magyarázóereje is. Így a feladatok és itemek elemzésénél megpróbáljuk ennek okát kideríteni.

A megbízónktól feldolgozott adatokat kaptunk, amelyek az itemeket nem feladatonként, hanem együttesen ábrázolják.

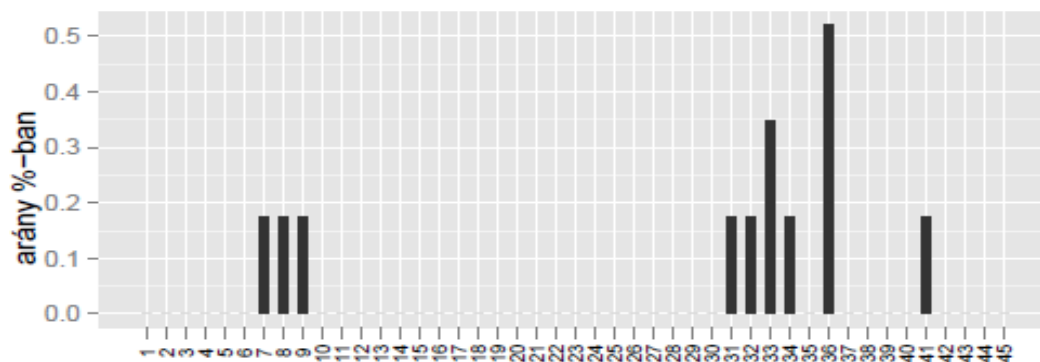


0.8. ábra:

A teszten nyújtott pontszám abszolút gyakorisági eloszlása (megbízó által készített grafikon)

Az 1. vizsgarészhez hasonlóan a 38. ábrán a feladatpontok összesítéséről van szó. Az összesen adható 45 feladatpontból az 578 diák közül a legtöbben 25-38 pont között teljesítettek. A teljesítmények eloszlása itt is a magasabb teljesítmények felé, jobbra tolódik.

A megbízótól kapott egy másik adatsor a kihagyott itemek százalékos arányát mutatja az összes feladatra vonatkozóan. A 39. ábrából az derül ki, hogy némely itemnél (7–9, 31–34, 36 és 41) aránylag sokan nem írtak be semmit az itemhez. Ezekre az itemekre a részletes elemzésnél visszatérünk.



0.9. ábra:
Az üresen hagyott ítemek aránya (a megbízó által készített grafikon)

A következő minta-itemtérkép a mintában résztvevő diákok képességszintjét és a négy feladatban szereplő ítemek 50 százalékos valószínűséggel történő megoldásához szükséges képességszinteket mutatja (0.10. ábra).

logit	személy	1.f	2.f	3.f	4.f	+item
5						
4						
						X
						XX
						X
						XXX
						XXXXXXXX
3						XXXX 7
						XXXXX
						XXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX
					24	XXXXXXXXXXXX
2						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX 8
				16		XXXXXXXXXXXX 5
						XXXXXXXXXXXX
				13		XXXXXXXXXXXX
						XXXXXXXXXXXX 2
1						XXXXXXXXXXXX
						XXXXXXXXXXXX
				31		XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX
				9		XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX 1
						XXXXXXXXXXXX
0						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX 4
						XXXXXXXXXXXX
						XXXXXXXXXXXX
						XXXXXXXXXXXX 3
				14		XXXXXXXXXXXX
						XXXXXXXXXXXX
				11		XXXXXXXXXXXX
						XXXXXXXXXXXX
				15		XXXXXXXXXXXX
-1						XXX
						XXXX
						XXX 6
				17 18		21 30 33
						X
				12		26
						X
						28
-2						45
-3						
						10

0.10. ábra:
A nyelvhelyesség feladatsor minta-itemtérképe
(VT által készített ábra) Minden 'x' egy vizsgázót jelöl.

A minta-itemtérkép alapján úgy látszik, hogy a 4. feladat fedi le aránylag legjobban a tanulók képességszintjét, bár ebben is volt egy item (45.), ami eléggé könnyű volt, hisz a negatív tartományba került. A tanulók képességszintje többségében a 0 logitérték fölött helyezkedik el, míg az itemek kb. fele ez alatt található. Az 1. feladat 7. iteme és a 2. feladat 10. iteme elképesztően kilóg a sorból, csak ellenkező előjellel: a 7. item a legmagasabb, míg a 10. item a legalacsonyabb logitértékkel rendelkezik. Ezután a négy feladat itemeinek feladatokra lebontott részletes elemzése következik.

1. Feladat: Knöpfungsortnung (8 item)

Az autentikus szövegre épülő feladat célja a névelő és az ahhoz hasonló vonatkozó névmás használata megfelelően ragozott formában. A feladatban egy hiányos szövegbe a megadott szavakat megfelelően ragozott formában kell beírni (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 15). A feladat a szavak nyelvtani kategóriáját nem adja meg, így ezek lehetnek névelők, de ugyanúgy lehetnek mutató vagy vonatkozó névmások, ill. kötőszavak. Ennek megfelelően változik az a nyelvi szint is, ahol ezeket produktívan kell tudni alkalmazni, s ezért céltartan rá lehet kérdezni. Ez egyben azt is jelenti, hogy az itemek eltérő nyelvi szintet mutatnak.

0.13. táblázat:

A nyelvhelyesség 1. feladat egyes itemeinek jellemzői a KER szintleírása alapján

Item	Mérési cél	Szint ²⁸
1.	lexika: az adott főnév névelője alanyesetben	A1
2.	'anhand' előljáró vonzata	B2
3.	lexika: az adott főnév névelője alanyesetben	A1
4.	lexika: az adott főnév névelője alanyesetben	B1
5.	lexika: kifejezés	B1
6.	vonatkozó névmás alanyesetben, kötőszó funkcióban	A2
7.	vonatkozó névmás részesesetben, kötőszó funkcióban	B1/B2
8.	lexika: az adott főnév névelője tárgyesetben	A2

Az itemek fele az önálló nyelvhasználat szintjén van, de ebből többségében B1 (3 item) szinten, s csak egy, esetleg két item (2. és 7.) van, ami az emelt szint, azaz a B2 elvárásait jeleníti meg. Az itemek másik fele A1 és A2 szintű. Ezek azok az itemek, ahol névelőre kérdez a feladat, s a szint attól függ, hogy mennyire tartozik az adott főnév, aminek a névelőjére kérdez az item, az elementáris szóhasználatba. Ezek után az adatelemzéssel vetjük össze az itemek jellemzőit.

0.14. táblázat:

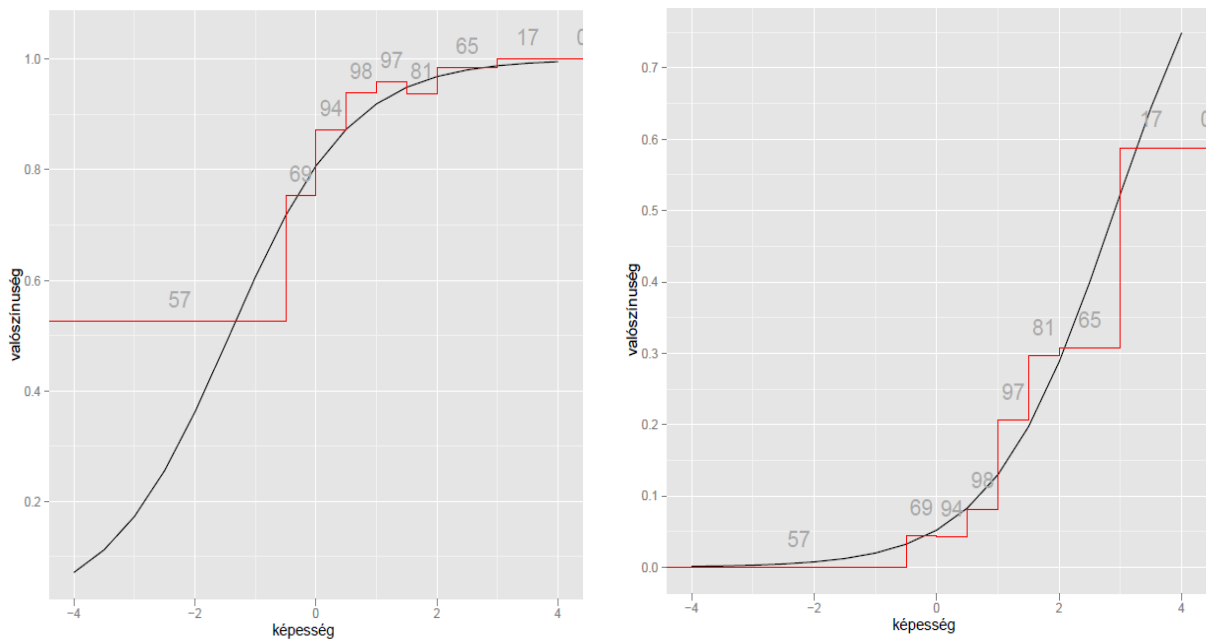
A nyelvhelyesség 1. feladat itemeinek jellemzői (a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott	Átlag	Szórás	1 pontszám pontbiszerialis	R ²	Nehézség	Infit MNSQ
-----------	----------------	-------	--------	----------------------------	----------------	----------	------------

²⁸ „Profile deutsch” c. software (Glaboniat és mtsai, 2006) alapján meghatározva – mindegyik kifejezés ill. szerkezet megtalálható a katalógusban.

	itemek (%)			korrelációja			
[1.]	0	0,65	0,48	0,33	0,09	0,30	1,09
[2.]	0	0,48	0,50	0,57	0,33	1,12	0,86
[3.]	0	0,78	0,41	0,36	0,11	-0,48	1,03
[4.]	0	0,72	0,45	0,27	0,06	-0,09	1,14
[5.]	0	0,42	0,49	0,36	0,11	1,42	1,09
[6.]	0	0,88	0,33	0,39	0,13	-1,27	0,94
[7.]	0,2	0,15	0,36	0,36	0,12	3,07	0,99
[8.]	0,2	0,41	0,49	0,46	0,20	1,51	0,97

A minta-itemtérkép már felhívta a figyelmet a 7. itemre, ami a feladatban a legnehezebbnek bizonyult. Az előtte lévő legkönnyebb, 6. item érdekes módon ugyanarra a nyelvtani szerkezetre épül (vonatkozó névmás kötőszói funkcióban), de alanyesetben – igaz passzív szerkezetben – míg a 7. itemben a kötőszóként használt vonatkozó névmást részes esetben kellett tenni. Ez azt mutatja, hogy önmagában egy nyelvtani struktúra nem feltétlenül határozza meg egy item nehézségét. Ezt az itemet többen üresen hagyták. Az 1. és 4. item magyarázóereje alacsony, akár el is hagyhatók.



0.11. ábra:

A 2. vizsgarész 1. feladat 6. (legkönnyebb) és 7. (legnehezebb) itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (a megbízó által készített grafikon)

2. Feladat: Lotto-Glück in Kanada (10 item)

A második feladat is egy lyukas szöveget tartalmaz, amibe megadott szavakból kell kiválasztani a mondatba az egyedüli odailő szót (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 16). A megadott szavak mind azonos szófajhoz tartoznak:

előjáró szavak. Így itt nem kell ezeket semmilyen módon megváltoztatni (pl. ragozni), de tudni kell pl. a velük járó ragozásokat, és persze a mondatban alkotott funkciójukat fel kell tudni ismerni. Önmagukban ezek nehézségét nehéz meghatározni, hiszen különböző funkciókban fordulnak elő a szövegben: társhatározó, időhatározó, igevonzat, melléknév vonzata, kifejezés része, stb. Ezek többsége elvárható már a középszinten is, a nehézséget a komplex szöveggörnyezet adja, ahol a megfelelő szó kiválasztása nem mindig könnyű.

0.15. táblázat:

A nyelvhelyesség 2. feladat egyes itemeinek jellemzői a KER (2002) szintleírása alapján

Item	Mérési cél	Szint ²⁹
9.	ige vonzata	A2
10.	társhatározó	A1
11.	időhatározó	A1
12.	melléknév vonzata	A1
13.	melléknév vonzata	B1
14.	ige vonzata	B1
15.	kifejezés	A2
16.	ige vonzata	B1
17.	ige vonzata	A1
18.	helyhatározó	A1

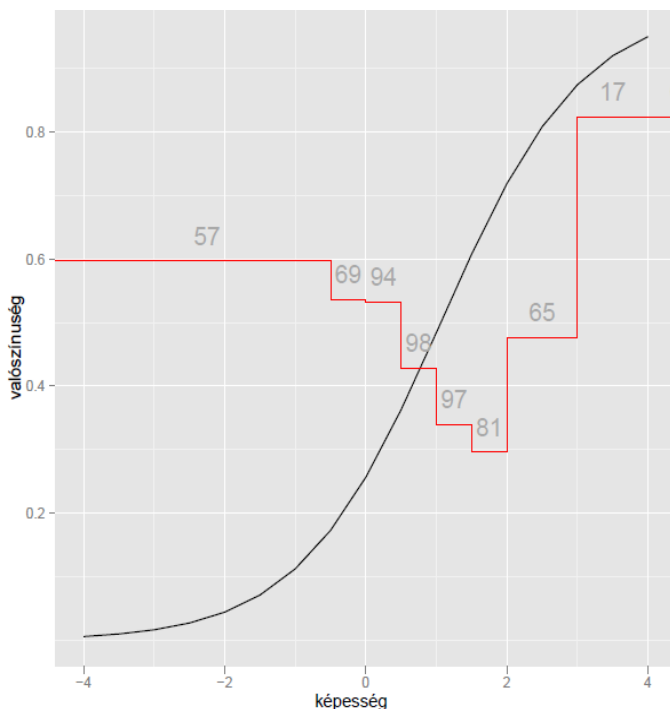
A feladat nehézségi szintje ennek alapján egyáltalán nincs a B2 szinten. A megoldás azon múlik, hogy a vizsgázó mennyire ismeri e szavak különböző funkcióját, de a konkrét itemekben alapvető helyzetben vannak. Meg kell jegyezni továbbá, hogy az utolsó item elméletileg nem mér: Mire a vizsgázó az utolsó itemhez ér, már el kellett használnia az összes megadott szót (néhányat többször is), és csak egyetlen egy maradt, ami még nem fordult elő: az utolsó item megoldása. Így kérdéses, hogy ez az item használható-e. A teljes feladatsor áttekintő adattáblázata is mutatta már, hogy ezzel a feladattal komolyabb problémák vannak. A minta-itemtérkép (0.10. ábra) szerint a legtöbb item ebben a feladatban az átlagos képességszint alatt van.

²⁹ „Profile deutsch” c. software (Glaboniat és mtsai, 2006) alapján meghatározva – nem mindegyik kifejezés ill. szerkezet található meg a katalógusban.

0.16. táblázat:
A nyelvhelyesség 2. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbizeriális korrelációja	R ²	Nehézség	Infit MNSQ
[9.]	0,2	0,60	0,49	0,34	0,12	0,56	1,10
[10.]	0	0,99	0,10	0,17	0,03	-4,01	0,98
[11.]	0	0,80	0,40	0,44	0,18	-0,61	0,96
[12.]	0	0,90	0,30	0,29	0,08	-1,53	0,98
[13.]	0	0,46	0,50	-0,06	0,01	1,24	1,51
[14.]	0	0,80	0,40	0,45	0,17	-0,59	0,93
[15.]	0	0,83	0,38	0,26	0,07	-0,80	1,07
[16.]	0	0,40	0,49	0,38	0,15	1,52	1,04
[17.]	0	0,88	0,33	0,35	0,10	-1,27	0,98
[18.]	0	0,87	0,34	0,49	0,20	-1,16	0,90

Adataink is azt mutatják, hogy a 2. feladat számos problémát rejt. A 10. item, amely már a minta-itemtérképről is csaknem „lelógott”, rendkívül könnyű volt a diákoknak, magyarázóereje és a pontbizeriális korrelációja is igen alacsony. A helyzet az, hogy a kb. egy oldalas szövegben nem nagyon lehetett emelt szintre való itemeket létrehozni – valószínűleg jobb lett volna ez a feladat középszinten. De a többi item is tartalmaz problémát. Nagyon könnyűnek bizonyult még a 12., 17. és a 18. item – ez utób-
 biról említettük, hogy ez technikailag tulajdonképpen nem mér. A legnehezebb a 16. item volt, majd ezt követi a 13. item, ami további furcsaságokat is rejt. Rendkívül alacsony a pontbizeriális korrelációja és a magyarázóereje. Ha pedig megnézzük ennek az itemnek a megoldási valószínűséget jelölő görbét, akkor kiderül, hogy a gyengék jobban oldották meg ezt az itemet, mint a jobb képességűek. Ennek magyarázatát csak akkor tudnánk kideríteni, ha lehetne tudni, hogy milyen más megoldásokat írtak be a tanulók. Összességében megállapítható, hogy ez a feladat alkalmatlan az emelt szintű nyelvhe-

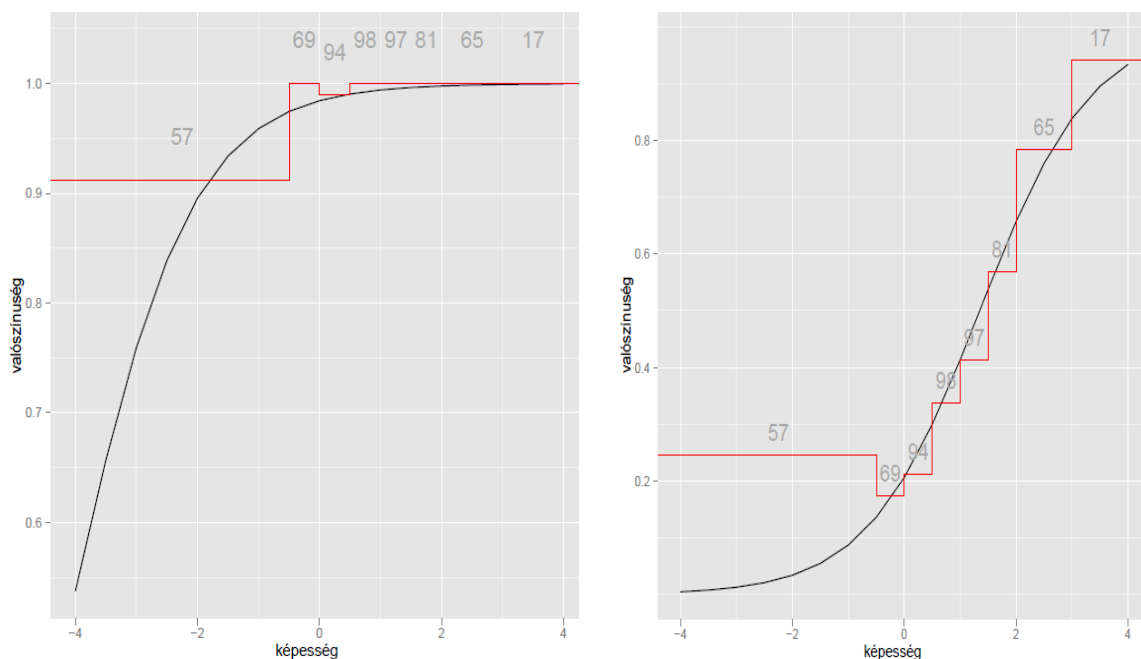


0.12. ábra:

2. vizsgarész 2. feladat 13. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (a megbízó által készített grafikon)

szintre való itemeket létrehozni – valószínűleg jobb lett volna ez a feladat középszinten. De a többi item is tartalmaz problémát. Nagyon könnyűnek bizonyult még a 12., 17. és a 18. item – ez utób-
 biról említettük, hogy ez technikailag tulajdonképpen nem mér. A legnehezebb a 16. item volt, majd ezt követi a 13. item, ami további furcsaságokat is rejt. Rendkívül alacsony a pontbizeriális korrelációja és a magyarázóereje. Ha pedig megnézzük ennek az itemnek a megoldási valószínűséget jelölő görbét, akkor kiderül, hogy a gyengék jobban oldották meg ezt az itemet, mint a jobb képességűek. Ennek magyarázatát csak akkor tudnánk kideríteni, ha lehetne tudni, hogy milyen más megoldásokat írtak be a tanulók. Összességében megállapítható, hogy ez a feladat alkalmatlan az emelt szintű nyelvhe-

lyesség mérésére. Ennek oka valószínűleg nem a feladattípusban keresendő, hanem a mérni kívánt nyelvhelyességi kategória (mérési cél) megválasztásában.



0.13. ábra:

A 2. vizsgarész 2. feladat 10. (legkönnyebb) és 16. (legnehezebb) itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (a megbízó által készített grafikon)

3. Feladat: Spaß und mehr gibt's nur bei der Jugendfeuerwehr (JFW) (15 item)

A szöveg egy, a fiatalok szabadidő eltöltésével kapcsolatos lehetőségről számol be, témájában és szókincsében sem különösebben nehéz (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 17). A szöveg azonban nyelvi hibákat tartalmaz. Ez problematikussá teszi a hibaaazonosításnak nevezett feladattípust. Meg kell jegyezni, hogy ilyen típusú feladatok egyáltalán nem fordulnak elő tankönyvekben, de nem nagyon találunk ilyent az új érettségire célzottan felkészítő segédanyagokban sem³⁰. Mégis hogy kerülhetett be ez a feladattípus a vizsgába? Az ok abban keresendő, hogy a vizsga bevezetésekor egyedül a nyelvszakok írták elő bemeneti követelményként az emelt szintű érettségit, ámde elvárták, hogy a jelentkezők e vizsga keretében adjanak számot célnyelvi grammatikai ismereteikről és nyelvtani tudatosságukról is. Erre próbált megoldást találni ez a feladattípus. 2012-től azonban kezd tágulni azon szakterületek köre, ahol díjazzák az emelt szintű érettségi vizsgát. Nekik csak a magasabb szintű nyelvi kompetenciáról kell(ene) tanúbizonyságot tenni, így kérdéses, hogy milyen jogosultsága van a továbbiakban ennek a feladattípusnak.

³⁰ PI. Lázárné és Panthóné (2004) kötete nem tartalmaz ilyen feladatot, viszont Sominé (2009) könyve igen.

A feladat mérési célja az, hogy felismeri-e a vizsgázó, hogy az adott mondat (ill. az adott sorban található szövegrész) helyes-e vagy sem. Mivel így nehéz meghatározni a grammatikai kompetencia pontos kategóriáját, ezért a szintbesorolás is nehéz. Így nem tudjuk az egyes itemeket a KER szintjeivel egyeztetni.

0.17. táblázat:

**A nyelvhelyesség 3. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)**

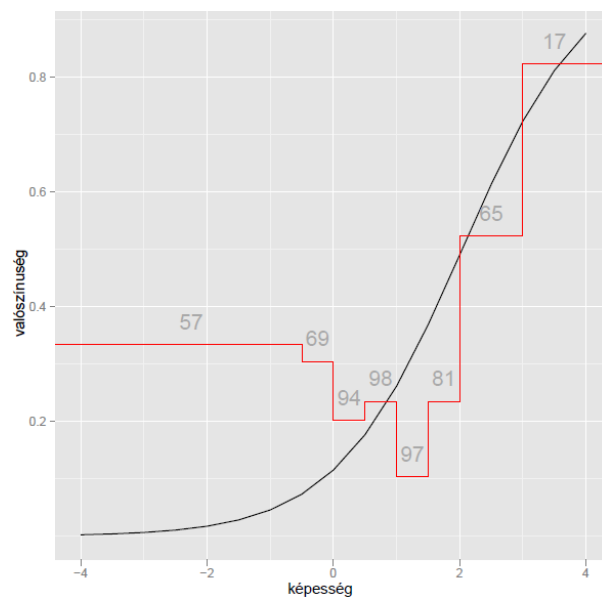
Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszériális korrelációja	R ²	Nehézség	Infit MNSQ
[19.]	0	0,83	0,38	0,31	0,09	-0,83	1,02
[20.]	0	0,76	0,43	0,29	0,04	-0,35	1,09
[21.]	0	0,87	0,33	0,19	0,03	-1,25	1,07
[22.]	0	0,38	0,49	0,53	0,30	1,66	0,91
[23.]	0	0,71	0,46	0,54	0,26	-0,02	0,89
[24.]	0	0,28	0,45	0,08	0,01	2,21	1,24
[25.]	0	0,48	0,50	0,60	0,34	1,15	0,87
[26.]	0	0,90	0,30	0,30	0,06	-1,55	0,99
[27.]	0	0,71	0,45	0,46	0,19	-0,05	0,96
[28.]	0	0,91	0,29	0,35	0,10	-1,62	0,93
[29.]	0	0,50	0,50	0,33	0,07	1,06	1,14
[30.]	0	0,87	0,34	0,30	0,07	-1,20	1,05
[31.]	0,2	0,53	0,50	0,45	0,16	0,88	1,00
[32.]	0,2	0,80	0,40	0,37	0,13	-0,60	0,99
[33.]	0,3	0,87	0,33	0,32	0,09	-1,22	0,98

Ebben a feladatban – a 2. feladathoz hasonlóan – az utolsó item sajnos szintén nem mér. A feladatban jelölni kell mind a helyes sorokat, mind a hibásakat – a hibásoknál a hibát kell az adott helyre kiírni. Mindegyik döntésre jár feladatpont. A feladat utasítása pedig pontosan megadja, hogy a szövegben hány helyes mondat található (itt 6 helyes volt). Mivel a 32. itemig már 5 helyes mondatos item volt, nyilvánvaló, hogy az utolsó sorba kerül a helyes mondat 6. jelölése.

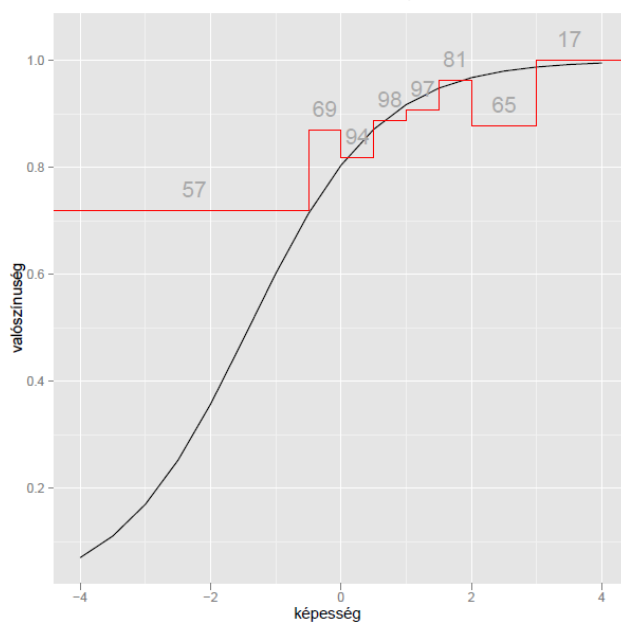
A minta-itemtérképen (0.10. ábra) az volt látható, hogy ebben a feladatban az itemek elhelyezkedése tág határokat fog át a tanulók képességszintjén, de az átlagos tartományban kevés item helyezkedik el. Ezzel magyarázható, hogy az itemek egy része könnyű (21., 26., 28., 30 és 33), más része nehéz volt (22., 24.). Nehéz megmondani, hogy mi volt az oka annak, hogy a 24. item volt ebben a feladatban a legnehezebb, mivel ez egy helyes mondat (ill. hibátlan sor) volt, ezt kellett felismerni. Itt ismerni kellett volna a tanulói megoldásokat ahhoz, hogy megtudjuk, mi okozhatta a nehézséget. Ráadásul az item megoldási valószínűsége jelölő görbéjén látszik (IV.14. ábra), hogy a jobb képességű tanulók egy csoportja rosszabb eredményt ért el, mint a gyengék. Ám ennek az itemnek alacsony a pontbiszerialis korrelációja és a magyarázó ereje is, vagyis nem azt méri, amit kellene. Érdekes, hogy ugyanez a helyzet a könnyűnek nevezhető 21. itemnél is, ahol ugyancsak a hibátlan mondatot kellett felismerni és jelölni. De gyenge a magyarázó ereje a 20. itemnek is, ahol szintén nem volt hiba az adott sorban. A helyes sort mutató ✓ jelet hatszor kellett kitenni: a 20., 21., 24., 26., 30. és 33. itemnél. Ezek mindegyikénél volt kisebb-nagyobb probléma. Ugyanakkor azok az itemek, amelyekben a vizsgázónak a helytelen alakot kellett megtalálni (19., 23., 25., 27., 29., 31., 32. item) jobban működtek. Ezért el kell gondolkodni azon, hogy ez a feladattípus megfelelő-e a B2 szintű nyelvhelyességi képesség mérésére.

4. Feladat: Mensch lernt tatsächlich im Schlaf (12 item)

Az utolsó feladat egy lyukas szöveget tartalmaz. Témája a népszerű tudományos ismeretterjesztő kategóriába tartozik, időnként eléggé speciális, ámde egy emelt szinten érettségizőnek felismerhető szókinccsel (pl. Hippocampus). A hiányzó szavakat a feladat a lyukkal azonos sorban megadja,



0.14. ábra:
A 2. vizsgarész 3. feladat 24. itemének elméleti és tapasztalati megoldási valószínűsége jelölő karakterisztikus görbéje (megbízó által készített grafikon)



0.15. ábra:
A 2. vizsgarész 3. feladat 21. itemének elméleti és tapasztalati megoldási valószínűsége jelölő karakterisztikus görbéje (megbízó által készített grafikon)

ám egy másik szófaji alakban (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 18-19). Az alapvető mérési cél tehát a szóképzés alapos ismerete. De nem elég felismerni, hogy pl. a megadott főnévből igét kell-e képezni vagy melléknévet, hanem azt a mondatban a megfelelően ragozott formában kell beírni. A Profile deutsch c. kiadvány (Glaboniat és mtsai, 2006) ugyan részletesen tárgyalja a szóképzést, de nem rendel hozzá szinteket, azaz hogy a különböző szóképzési lehetőségeket mely szinten kell alkalmazni tudni. A 40/2002 OM rendelet Vizsgakövetelmények c. dokumentuma e fogalomkörben csak annyit ír, hogy „a szóképzés alapvető szabályait” már középszinten ismerni kell (18. o.). Ezért ennél a feladatnál sem tudjuk az egyes itemeket a KER (2002) szintjeivel egyeztetni. Természetesen azért lehet vizsgálni az egyes itemek szintjét, pl. a passzív igealakok (37.) vagy a jelen idejű melléknévi igenév jelzői alkalmazása (42.) minden bizonyosan B2 szintet jelent.

0.18. táblázat:

A nyelvhelyesség 4. feladat itemeinél elvárt nyelvtani/lexikai ismeretek

Item	Mérési cél
34.	melléknévből főnév
35.	főnévből melléknév, jelzői funkcióban (melléknévragozás)
36.	képzett melléknévből ige
37.	főnévből ige, passzív igealak
38.	igéből főnév
39.	melléknévből főnév
40.	melléknévből főnév
41.	főnévből ige, passzív igealak
42.	igéből jelen idejű melléknévi igenév
43.	főnévből ige
44.	főnévből melléknév, jelzői funkcióban (melléknévragozás)
45.	képzett főnévből ige

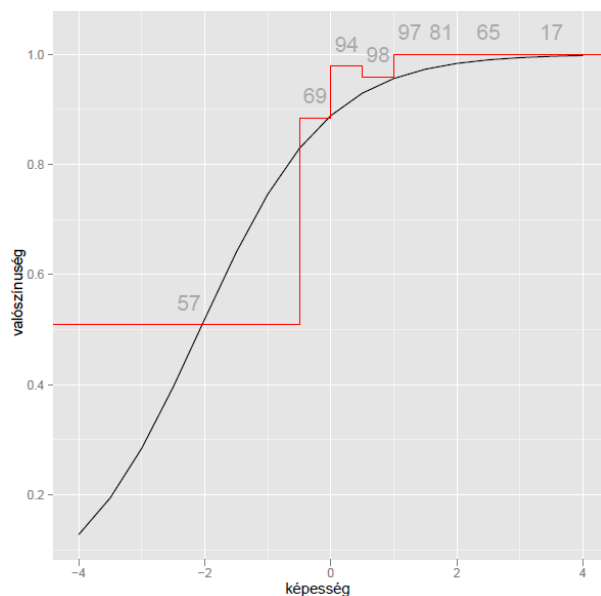
A következőkben megvizsgáljuk, hogy a fenti nyelvi jelenségekre készült itemek miképp működtek a gyakorlatban.

0.19. táblázat:
A nyelvhelyesség 4. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszériális korrelációja	R ²	Nehézség	Infit MNSQ
[34.]	0,2	0,47	0,50	0,52	0,25	1,19	0,94
[35.]	0	0,73	0,44	0,50	0,20	-0,16	0,93
[36.]	0,5	0,62	0,48	0,66	0,40	0,44	0,78
[37.]	0	0,76	0,43	0,52	0,24	-0,30	0,90
[38.]	0	0,54	0,50	0,54	0,26	0,84	0,92
[39.]	0	0,74	0,44	0,34	0,09	-0,24	1,07
[40.]	0	0,41	0,49	0,49	0,21	1,50	0,95
[41.]	0,2	0,68	0,47	0,53	0,26	0,14	0,92
[42.]	0	0,57	0,50	0,48	0,23	0,70	0,97
[43.]	0	0,50	0,50	0,51	0,21	1,03	0,93
[44.]	0	0,78	0,41	0,53	0,26	-0,45	0,88
[45.]	0	0,93	0,25	0,42	0,18	-1,91	0,88

Az egész feladatsorban ez a feladat volt aránylag a legkiegyensúlyozottabb. Az itemek nehézségi szintje között nem voltak olyan kiugró különbségek, mint a korábbi feladatok esetében. Ugyanakkor az utolsó item (45.) teljesítménye 93% volt, magyarázó ereje is elég alacsony. A 39. item is hasonló tulajdonságokat mutat. Mindkét esetre jellemző, hogy a tanulóknak nem a megadott szóból kellett egy másik szófajban egy szót képezni, hanem egy már képzett szóból kellett fölismernie azt a szót, amiből a megadott szót képezték. Azaz visszafelé kellett gondolkodni, ez okozhatta, hogy ezek az itemek túl könnyűek lettek.

Összegzés: A 2. vizsgarész feladatai közül a 3. és 4. feladat volt aránylag a legmegbízhatóbb mérőfeladat, bár ezeken belül is volt néhány problémás item. Az első két feladat reliabilitása már jóval gyengébb volt, közülük is a 2. feladat volt a leggyengébb. Itt változtatás nélkül kellett a megadott szavakból a megfelelőt kiválasztani. Ez minden bizonnyal alatta van az emelt szint követelményének, így nem méri jól a vizsgázók nyelvhelyességi kompetenciáját. A korábbi, 2006. évi



0.16. ábra:
A 2. vizsgarész 4. feladat 45. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

feladatelemzésből még az derült ki, hogy „mindkét szinten a nyelvhelyesség vizsgarész/vizsgarész-összetevő minősége és szintje volt a legmegfelelőbb” (Einhorn, 2008: 219). Úgy látszik, hogy hat év alatt elkopott az első évek minőségi feladatkészítése. De az említett tanulmány megemlíti azt is, hogy „2006-ban már érezhető, hogy bizonyos – korábban nehéznek gondolt – feladattípusok megoldása” (Einhorn, 2008: 219) nem jelent valódi kihívást. Ezért érdemes lenne ezt az egész vizsgarészt az emelt szinten újra alaposan átgondolni és nehézségi szintjét felülvizsgálni.

2.3. Hallott szöveg értése (III. vizsgarész)³¹

Az olvasott szövegértéshez hasonlóan e vizsgarész mérési célja is az értési céloknak megfelelő stratégia alkalmazásával a hallott szöveg megértése. De míg a középszinten egyszerű, hétköznapi szövegek lényegének a megértése a cél, emelt szinten komplexebb a feladat:

A vizsgázó legyen képes a szöveg

- gondolatmenetét részleteiben is követni,
- megértésén túl a szövegkörnyezetből következtetni az egyes beszélők álláspontjára,
- megértésén túl a szövegkörnyezetből következtetni a beszélők érzelmeire és egymáshoz való viszonyára (40/2002 OM rendelet, Vizsgakövetelmények: 4, Vizsgaleírás: 11).

A feladatsor általában 2–3 szövegből és 2–3 feladatból áll. Egy feladaton belül egy hosszabb szöveg helyett előfordulhat több rövidebb szöveg is. A feladatokhoz felhasznált szövegek együttes terjedelme 7–10 perc. A feladatsor 20–30 itemből áll (40/2002 OM rendelet, Vizsgaleírás: 12). A vizsgált feladatsor összesen 2 szövegből és két feladatból áll, ami összesen 22 itemet tartalmaz, összterjedelme 8:43 perc. Mindkét feladat internetről származó, autentikus felvétel, standard beszédtempóban és kiejtésben.

³¹ Hanganyag letöltve: http://193.225.13.214/erettsegi2012/emeltszintu_erettsegi_2012.html

0.20. táblázat: Áttekintés a hallott szöveg értése feladatsor szerkezetéről

Feladattípus / Szöveghossz percben ³²	Item-szám	Összes item	Vizsgapont
megfelelő kijelentés kiválasztása (igaz-hamis) / 4:08 perc	9	22	30
mondatok kiegészítése a hallott szöveg alapján / 4:35 perc	13		

A feladatsor formailag megfelel az elvárásoknak, az adatelemzés alapján azonban jellemzőit tekintve az első feladat nem tölti be mérő funkcióját, mert rendkívül alacsony a Cronbach-alfa értéke, a szórása és a magyarázó ereje is gyenge. A két feladat nehézségi szintje nincs összhangban, hiszen az első jóval könnyebb a másodiknál. Ugyanakkor a 2. feladat megbízhatósága megfelelő, gyakorlatilag ez teszi elfogadható reliabilitásúvá ezt a vizsgarészt.

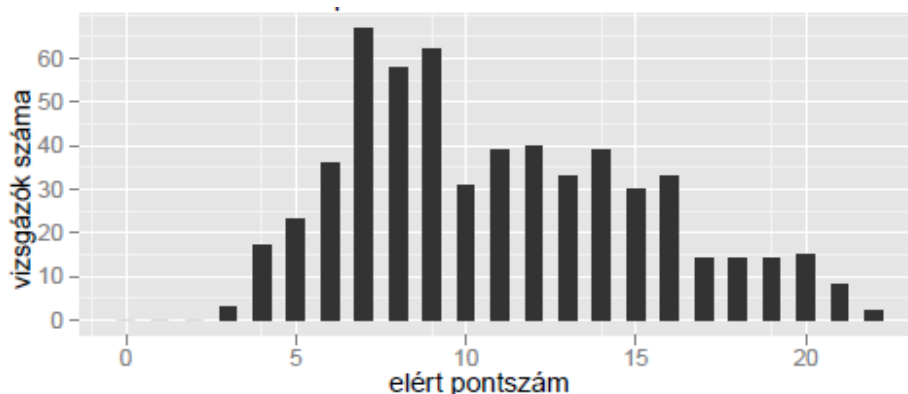
0.21. táblázat:

**A hallott szöveg értése feladatsor jellemzői
(a megbízótól kapott adatok és VT számításai alapján)**

Kategória	Feladat		Teszt
	1.	2.	
Cronbach-alfa	0,04	0,83	0,80
Átlag (%p)	72,5	34,2	49,9
Szórás (%p)	14,7	26,4	19,7
Magyarázóerő (R ² %)	23,3	76,7	100

A 0.17. ábra a két feladattal elért összpontszámot mutatja a feladatpontok alapján. Míg az első két vizsgarészben az értékek jobbra tolódtak, azaz magasabb pontszámot értek el a vizsgázók (v.ö. 0.2. ábra és 0.8. ábra), addig a hallott szövegértés esetében balra tolódás figyelhető meg, azaz kevesebb pontot szerzett a mintában résztvevők többsége.

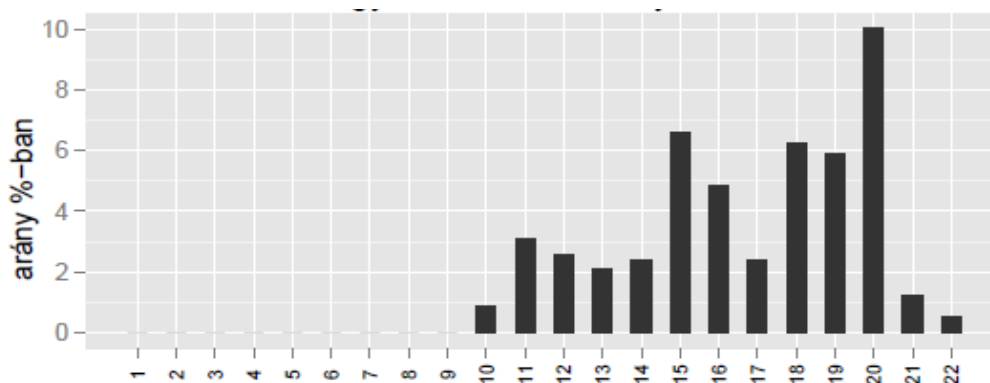
³² Mivel a terjedelem nincs megadva sehol, az internetre feltöltött hanganyag összefüggően 30 percet tesz ki, nincs trackekre bontva, így csak a lejátszáskor mért időt tudjuk megadni.



0.17. ábra:

A teszten nyújtott pontszám abszolút gyakorisági eloszlása
(megbízó által készített grafikon)

A 0.18. ábrából az derül ki, hogy a 10. itemtől kezdődő második feladatban sokan nem tudtak beírni semmit se, amely egyértelműen a nyílt végű feladattípusnak köszönhető. Az első feladatban az üresen hagyott itemek aránya nem is értelmezhető, hiszen ott már a feladat maga olyan volt, hogy nem volt szabad minden kockát bejelölni. Ezt csak úgy lehetne megállapítani, ha azt néznénk, hogy a jelölt a 16 válaszlehetőség közül melyiket ikszelte be és melyiket nem.



0.18. ábra:

Az üresen hagyott itemek aránya
(a megbízó által készített grafikon)

Az egyes feladatokhoz tartozó itemek nehézségét és a mintába került diákok képességi szintjét a következő minta- és itemtérképen ábrázoljuk (0.19. ábra). A térképről leolvasható, hogy az adott item mennyivel könnyebb vagy nehezebb, illetve a vizsgált személy mennyivel magasabb vagy alacsonyabb képességű. A vizsgált személyekben ugyan különbségek vannak, de többségük az átlagos képességi szint fölött helyezkedik el.

logit	személy	1.f	2.f	+item
4				
3		X		
		X		
		X		
		XX		
		XXX		
		XX		
		XX		
		XXXXXXX		
		XXXXXXXXX		
2		XXXXXX	14	
		XXXXXXX	18	
		XXXXXXXXXXX		
		XXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX	15 20	
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX	10 12	
1		XXXXXXXXXXXXXXXX	19	
		XXXXXXXXXXXXXXXX	16	
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX	11 17 21	
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX		
0		XXXXXXXXXXXXXXXX	22	
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX	1 6	
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX	13	
		XXXXXXXXXXXXXXXX		
		XXXXXXXXXXXXXXXX	9	
-1		XXXXXXXXXXXXXXXX	7	
		XXXXXXXXXXXXXXXX	8	
		XXXXXXXXXXXXXXXX	3	
		XXXXXXXXXXXX		
		XXXXXXXXXX		
		XXXXXXXXXX		
		XXX	5	
-2		XX	2	
		X		
		X		
		X		
		X	4	
		X		
-3				
-4				

0.19. ábra:
A hallott szöveg értése feladatsor minta-itemtérképe
(VT által készített ábra) Minden 'x' egy vizsgázót jelöl.

A két feladat nagyon eltérő nehézsége az itemekben tükröződik, hiszen míg az első feladat itemeinek nehézsége kivétel nélkül mind a 0 logitérték alatt helyezkedik el, addig

a 2. feladatban csak egyetlen item van alatta, egy azonos szinten, a többi mind föllette helyezkedik el. Ezzel viszont a két feladat kiegészíti egymást és többnyire lefedik a vizsgázók képességszintjei alapján meghatározott intervallumot.

1. Feladat: Kinder- und Jugendchor Juventus Vocalis (9 item)

Egy interjút hallunk egy kórusvezető tanárnővel és néhány kórustaggal, közte összekötő szöveggel (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 23). A szöveget végig háttérzaj kíséri, hol erősebben, hol gyengébben. Az információk elég nagy közökkel követik egymást. A feladattípus az igaz-hamis feladattípushoz hasonlít, de mégse egészen az. A vizsgázó feladata az, hogy döntse el, melyik kijelentésről hallott a rádióműsorban, s melyikről nem. A feladat meghatározása kiköti azt is, hogy összesen hány olyan kijelentés lehet, ami elhangzott a szövegben. Ha a vizsgázó ennél több ikszet ír be a kockákba, azok mindegyikéért le kell vonni egy pontot.

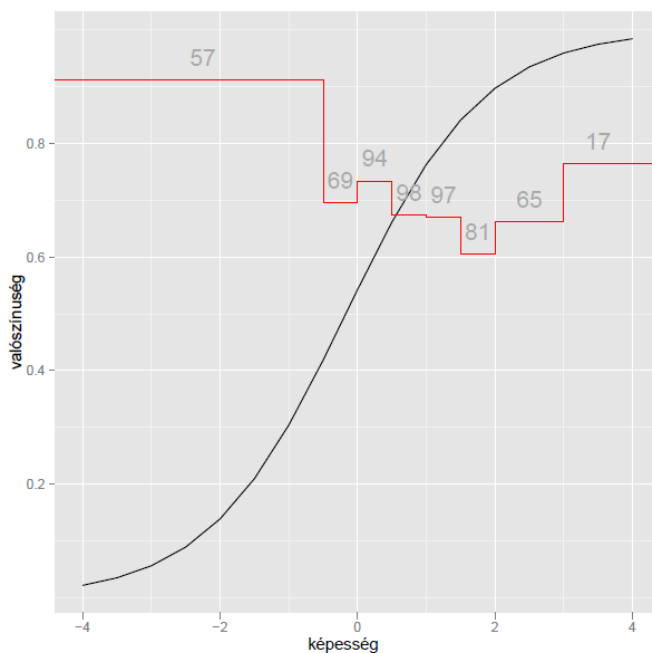
Már a feladatsort bemutató 0.21. táblázatban szereplő Cronbach-alfa érték is utal arra, hogy baj van ennek a feladatnak a reliabilitásával. Ennek hátterében az áll, hogy a vizsgázó akkor kap 0 pontot, ha az adott helyre nem tett x-et. Az adatrögzítés sajátossága, hogy arról nem szerünk információt, hogy a tanuló a helytelen válasz esetén hova tett ikszet. Feladattipológiai szabályosan tehát ennek a feladatnak 16 pontosnak kellett volna lennie, amelyben a vizsgázó jelöli, hogy melyik mondat helyes, melyik hamis. Azzal, hogy csak a helyes válaszokat kell megjelölni, a feladatmegoldó tevékenység mesterségessé válik, hiszen a vizsgázónak arra is kell figyelnie, hány elemet jelölt meg válaszként, így tehát az itemek nagy valószínűséggel összefüggnek. E tényezők mind hozzájárulhattak ahhoz, hogy ez a feladat ilyen formában működésképtelen. Ezt az itemszintű adatok is alátámasztják.

0.22. táblázat: A 3. vizsgarész 1. feladat itemeinek jellemzői (a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Átlag	Szórás	1 pontszám pontbiszerialis korrelációja	R ²	Nehézség	Infit MNSQ
[1.]	0,56	0,50	0,28	0,04	-0,27	1,19
[2.]	0,85	0,36	0,27	0,06	-1,92	1,07
[3.]	0,76	0,43	0,31	0,08	-1,29	1,07
[4.]	0,91	0,29	0,24	0,06	-2,56	1,05
[5.]	0,83	0,38	0,33	0,12	-1,80	1,03
[6.]	0,59	0,49	0,42	0,12	-0,34	1,06
[7.]	0,70	0,46	-0,10	0,02	-0,98	1,48
[8.]	0,71	0,45	0,24	0,05	-1,05	1,14
[9.]	0,69	0,46	0,33	0,08	-0,77	1,10

Az itemek nehézségi mutatója minden esetben a mínusz tartományban helyezkedik el. Ezen belül is kiugróan könnyű a 2., 4. és 5. item. Alacsony a magyarázó ereje az 1. itemnek, valamint a 7. itemnek, ahol a pontbiszerialis korreláció értéke elfogadhatatlan,

hiszen az 1 pont pontbiszeriális korrelációjának a pozitív tartományban kellene lennie. Ez azt jelenti, hogy a 0 pontértéknek van szignifikáns kapcsolata a teszt egészén nyújtott teljesítménnyel. Mindezek mellett rossz az item illeszkedése is, amit a 0.20. ábra is igazol. Itt a gyengébb képességűek jobban oldották meg az itemet, mint a jobb képességűek. A többi (3., 6., 8. és 9.) item magyarázóereje 10% körül van, amely azt mutatja, hogy ezek súlytalanok a tesztben.



0.20. ábra:

A 3. vizsgarész 1. feladat 7. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

Érdeemes lenne ezt a feladattípust újragondolni, mert – bár ez volt a könnyebb feladat – minden bizonnyal nem megfelelően mér. Veszélyes, hogy csaknem ugyanannyi hamis kijelentés van a feladatban, mint ahány helyes megoldás, de a hamis kijelentések felismerését nem értékeli a feladat. Ráadásul könnyen becsúszhat olyan kijelentés a feladatba, ami csapdás (pl. 5. item). Ezzel a feladattípussal problémák voltak a középszintű feladatsorban is.

2. Feladat: Sprachreisen für die Generation 50plus (13 item)

Az internetes rádióból származó riport idős emberek célországban folytatott nyelvtanfolyamairól szól. A szöveg tartalmát lyukas mondatokban foglalja össze a feladat, ezeket kell értelemszerűen kiegészíteni a hallott szöveg alapján (Német nyelv, emelt szintű írásbeli érettségi vizsga, feladatlap 2012: 24-25). A feladaton belül kétszer fordul elő, hogy összefüggenek az itemek: a 10. és 11., valamint a 19. és 20. item. Itt a felsorolásból kellett két információt lejegyezni ugyanahhoz a tartalomhoz. Mivel itt nem csak egy íkszet kell beírni, hanem szavakat, kifejezéseket, ez nyilvánvalóan fokozott kihívást jelent a vizsgá-

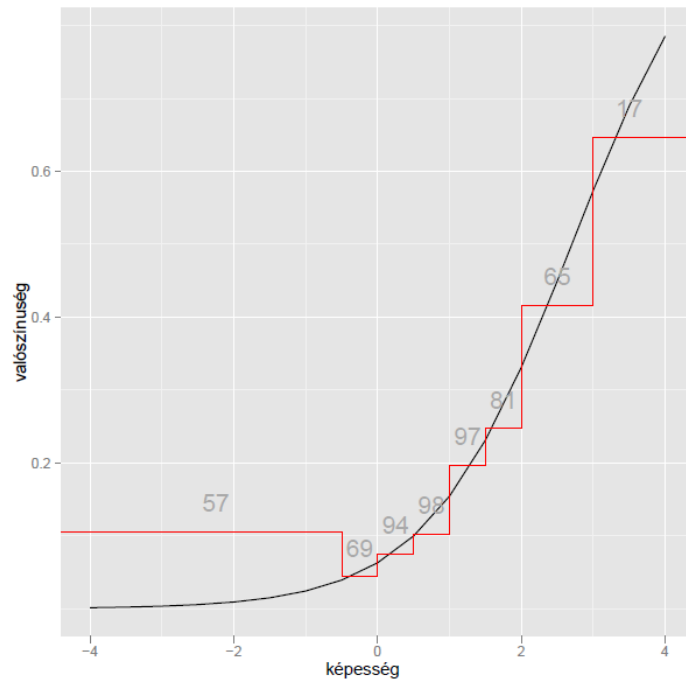
zóknak. Ez megmutatkozik abban, hogy a 13. és 22. item kivételével az összes item a minta-itemtérképen a 0 logit-érték fölött található.

0.23. táblázat:

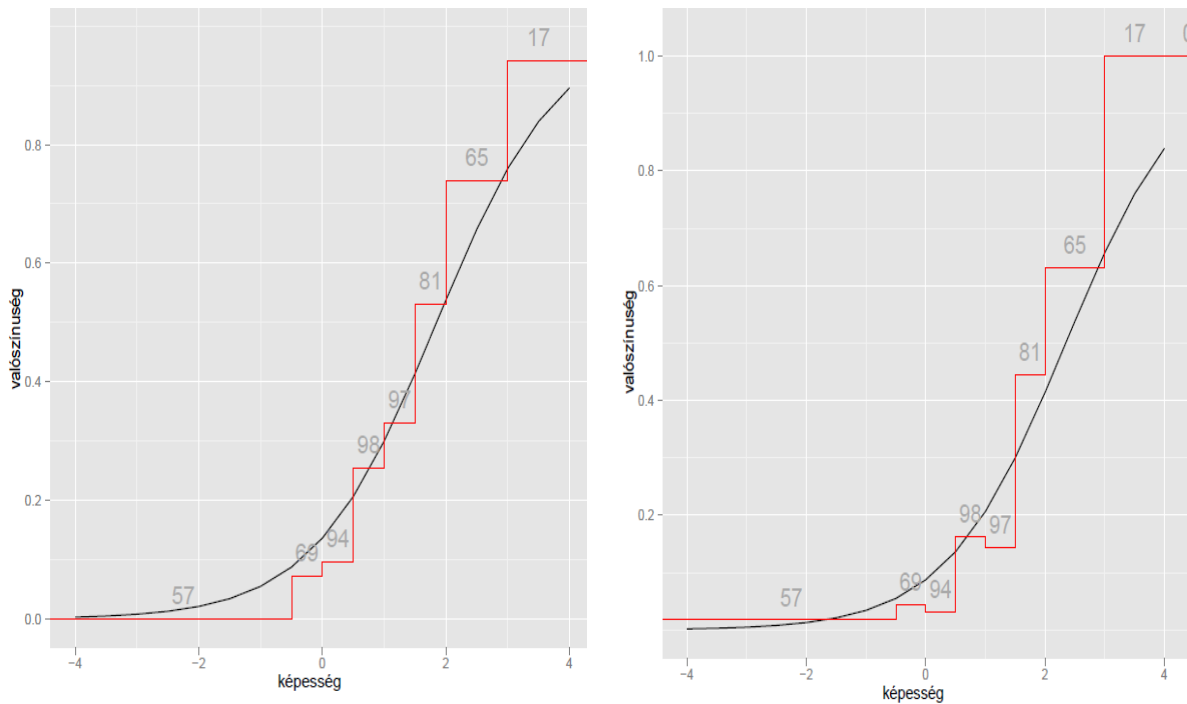
A 3. vizsgarész 2. feladat itemeinek jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Item-szám	Üresen hagyott itemek (%)	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ
[10.]	0,9	0,30	0,46	0,49	0,18	1,15	0,95
[11.]	3,1	0,44	0,50	0,62	0,30	0,38	0,85
[12.]	2,6	0,29	0,45	0,61	0,31	1,13	0,82
[13.]	2,1	0,64	0,48	0,57	0,34	-0,60	0,88
[14.]	2,4	0,19	0,39	0,41	0,09	1,87	1,02
[15.]	6,6	0,25	0,44	0,53	0,18	1,36	0,93
[16.]	4,8	0,35	0,48	0,60	0,23	0,85	0,87
[17.]	2,4	0,44	0,50	0,56	0,26	0,36	0,92
[18.]	6,2	0,19	0,39	0,49	0,19	1,83	0,92
[19.]	5,9	0,32	0,47	0,63	0,27	0,96	0,84
[20.]	10,0	0,24	0,43	0,62	0,25	1,41	0,82
[21.]	1,2	0,43	0,50	0,60	0,26	0,35	0,90
[22.]	0,5	0,53	0,50	0,51	0,22	-0,04	0,96

Az egész írásbeli vizsgán ebben a feladatban volt a legtöbb üresen hagyott item. A legnehezebb a 14., 18. és 20. item volt, de a többi item is – a 13. és 22. kivételével – a 0 érték feletti tartományban van, teljesítettségi átlaguk alacsony. Az itemek eloszlása a hangzó szövegben meglehetősen egyenetlen. Nem csak az összefüggő itemek vannak túl közel egymáshoz, hanem pl. a 13. és 14. is. És mivel túl hamar hangzik el a 13. item után a 14. item információja, ezért ez az item természetesen nehéz, hisz még az előző itemet írja a vizsgázó, miközben a következő item információja elhangzik. Hasonló okra vezethető vissza a 20. item nehézsége és a magas arányban üresen hagyott itemek aránya. Ez az item összefügg a 19. itemmel, és az információk, amelyekből kettőt kellett leírni, túl gyorsan követik egymást. A 18. item nehézségére viszont csak akkor tudnánk magyarázatot adni, ha tudnánk, hogy oda mit írtak be a vizsgázók, s azt a javító tanárok hogyan értékelték.



0.21. ábra:
A 3. vizsgarész 2. feladat 14. itemének elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)



0.22. ábra:
A 3. vizsgarész 2. feladat két összefüggő itemének, a 19. és 20. itemnek elméleti és tapasztalati megoldási valószínűségét jelölő karakterisztikus görbéje (megbízó által készített grafikon)

Összegzés: A 3. vizsgarész (hallott szövegértés) mindössze két feladatból áll, és ez rizikóval jár együtt. Esetünkben az 1. feladat nem volt alkalmas a B2 szinten a vizsgázók képességeinek megbízható mérésére. Ez a probléma azonban nem a szinttől függ, hanem a feladattípustól, hisz az 1. feladat típusa a középszinten is rossz eredményeket mutatott. Így csupán a 2. feladatot lehet valóban mérő feladatnak tekinteni, ami nagyon kevés egy receptív készség méréséhez. Mivel mindössze két feladattal csak kevés itemet lehet létrehozni, az „itemszaporítás” érdekében belekerülnek a feladatokba egymással összefüggő itemek is, amelyek mérési céljai egymástól nem különíthetők el. Ez a 2. feladatban kétszer is előfordult. A hallott szövegértés mérésénél is tovább kellene fejleszteni a feladatkészítés eljárását, hogy a vizsgamodellnek jobban megfelelő és reliabilis feladatsorok jöjjenek létre.

3. A három, itembázisú vizsgarész összefoglalása

Kutatási kérdéseink (I. Bevezetés, 5. A kutatás módszerei) egyrészt az első három vizsgarész megbízhatóságára és a feladatok súlyára vonatkoztak, másrészt az itemalapú feladatok működésére. Elemzésünkkel az alábbi kérdésekre kaptunk választ:

1. Mennyire megbízhatóan mértek az egyes készségeket mérő feladatsorok?

Összességében a feladatsorok adataink szerint jól mértek, bár nem voltak problémáktól mentesek. Mindegyik feladatsor megfelelő Cronbach-alfa értékkel rendelkezett. Az összesítésből azonban nem derül ki, hogy pl. a 3. vizsgarész csak a második feladatának köszönhetően mért megbízhatóan. De az első és a második vizsgarész is tartalmazott problémás feladatot.

0.24. táblázat:

A három feladatsor összesített jellemzői
(a megbízótól kapott adatok és VT számításai alapján)

Kategória	Olvasott szövegértése	Nyelvhelyesség	Hallott szöveg értése
Cronbach-alfa	0,86	0,88	0,80
Átlag (%p)	72,0	66,5	49,9
Szórás (%p)	19,7	17,3	19,7

2. Milyen mértékben magyarázzák az egyes feladatok az adott vizsgarészben elért összpontszámot?

A feladatok magyarázóereje a vizsgarészekben eltért egymástól. A legalacsonyabb magyarázóerővel az 1. vizsgarész 3. feladata, a 2. vizsgarész 2. feladata és a 3. vizsgarész 1. feladata rendelkezett. Ezekre a feladatokra általában az volt jellemző, hogy vagy a feladattípussal volt komoly probléma (v.ö. hallott szöveg értése 1. feladat), vagy az adott feladaton belüli itemek voltak túlságosan könnyűek és ezáltal súlytalanok az adott vizsgarészben.

3. Milyen nehézségi szintbe sorolhatók az egyes feladatok és itemek?

Azt tapasztaltuk mind a számítások, mind a minta-itemtérképek (0.3. ábra, 0.10. ábra, 0.19. ábra) alapján, hogy a feladatok csak részben fedik le a vizsgakövetelményekben meghatározott nyelvi szintet. Igazán csak a 3. vizsgarész 2. feladata fedett le magasabb képességszintet, így erről a feladatról jelenthető ki, hogy valószínűleg a B2 szinten volt. A 2. vizsgarészben csak egy-egy item közelítette meg a B2 szintjét, többségük könnyebb volt. Az olvasott szövegértés itemei a -2 és 2 közötti logitérték között mozogtak, többségükben az átlagos képességszint körül.

4. Mennyiben felelnek meg a feladatok itemnehézségi indexei a vizsgázók nyelvi készségszintjeinek?

Az itemnehézségi indexek nem fedik le teljesen a mintában elemzett vizsgázói készség szinteket, ezek többsége alatta maradt a tanulók képességeinek. Ez mindenféleképp felveti azt a kérdést, hogy valóban B2 szintet mérnek-e az emelt szintű feladatok. Ha a feladatok B2 szintűek, akkor indokolt lehet a vizsgakövetelmények módosítása és a feladatok nehézségi szintjének emelése C1 szint közelébe. Amennyiben a feladatok a B2 szint alatti szintet mérnek, akkor szükséges a vizsgafeladatokat a követelményekben megadott szinthez igazítani. Ezek alapján lényeges következtetésünk, hogy az emelt szintű feladatsornál is szükséges megvizsgálni a KER (2002) szintjeihez történő illeszkedést.

5. Mennyire megfelelően működnek az egyes vizsgarészekben használt zárt végű feladattípusok? Ha nem jól működnek, mi ennek az oka?

A feladattípusok – egy kivételével – általában jól működtek, de a feladaton belüli itemek nem mindig. A feladatok nehézségi szintje a vizsgaleírásnak nagyrészt megfelelt, bár inkább könnyebb volt a B2 szintnél. Az 1. vizsgarész 3. feladata, a 2. vizsgarész 2. feladata és a 3. vizsgarész 1. feladata a megkapott adatok részletes elemzése szerint nem volt megbízható, nem jól mért. Az 1. és 2. vizsgarészben ez a jelenség annyiban magyarázható a feladattípussal, hogy mindkettő a kifejezetten könnyű feladattípusok közé tartozik. Hiába épült a feladat egy nehezebb szövegre, ha az nincs összhangban a könnyű feladattípussal. Az 1. vizsgarész 1. feladata egy igaz-hamis típus volt, ahol itemenként 50%-os a találgatási esély, a 2. vizsgarész 2. feladata pedig egy lyukas szöveget tartalmazott, ahova a megadott szavakat változatlan alakban kellett a megfelelő helyre beírni. Ráadásul itt még a szöveg is meglehetősen könnyű volt, tehát nem felelt meg az emelt szint követelményeinek. E két esetben nem a feladattípus maga okozta a problémát, hanem az, hogy azt nem megfelelő módon alkalmazták.

Az egyetlen igazán rosszul működő feladattípus az ún. ál igaz-hamis (szelektálás) feladat volt a 3. vizsgarész 1. feladatában. Itt a relatív hosszú szöveghez kevés számú item párosult. A megadott kijelentésről fel kellett ismerni, hogy miről volt szó a szövegben, és miről nem. De ami nem odaillő kijelentés, azt hiába ismeri fel a vizsgázó, nem jár a felismerésért pont. Viszont ha többet jelöl be, mint amennyi lehetséges, büntető pont jár érte, ami nincs összhangban a mérő funkcióval.

6. Mennyire megfelelően működnek a nyílt végű feladattípusok? Mi az oka a nem megfelelő itemműködésnek ezekben a feladatokban? Hogyan működött ezeknél a feladatoknál az értékelési rendszer?

Az emelt szint itemalapú vizsgarészeiben csupán egyetlen nyílt végű feladat fordult elő a 3. vizsgarész utolsó feladataként. Köztudott, hogy a nyílt végű feladatok nehezebbek, így ez volt a legnehezebb feladat, és ennek volt a legmagasabb magyarázó ereje. Problémák leginkább az összefüggő itemek esetében merültek

fel, ahol a beírandó információk túl gyorsan követték egymást, így az összefüggő itemek közül az elsőt többnyire sikerült jól megoldani, de a következőt már nem volt idő se meghallani, se leírni. Erre fel kell hívni a feladatírók figyelmét.

A kutatási kérdésekhez megfogalmazott hipotézisek közül az alábbiak igazolódtak be: A német nyelvi érettségi működésének jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján helyesen feltételeztük, és elemzésünk nyomán beigazolódtott, hogy

1. az egyes készségeket mérő feladatsorok megbízhatósága megfelelő volt, annak ellenére, hogy mindegyikben volt 1-1 problémás feladat. Ezért a hipotézisünk csak részben igazolódtott be.
2. az egyes feladatok súlya a tesztekben eltér, amely a mért nyelvi szinttől, illetve az egyes feladatok minőségétől függ.

A receptív készségeket és a nyelvhelyességet mérő feladatsorok jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján helyesen feltételeztük, és elemzésünk nyomán beigazolódtott, hogy

3. a feladatok itemnehézségi indexei nem fedik le teljesen a vizsgázók nyelvi készségszintjei által meghatározott intervallumot;
4. a magasabb készségtartományból hiányoznak itemek;
5. a zárt végű feladatokban lévő összefüggő itemek, a nem megfelelő disztraktorok, és a pontlevonás rontják az adott feladat és a teszt reliabilitását;
6. a nyílt végű feladatok működése függ a mért nyelvi szinttől és a mért gondolkodási művelettől. A 3. vizsgarész 2. feladatában az összefüggő itemek is befolyásolták e nyílt végű feladat működését.

Irodalom

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgakövetelmények.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv_vk.pdf,

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgaleírás.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv_vl.pdf (letöltve: 2013. május 20.)

EINHORN ÁGNES (2008): Hogy vizsgázott az új érettségi németből? In: Bánkuti Zsuzsa/Lukács Judit (szerk.): *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés*. Budapest: Oktatáskutató és Fejlesztő Intézet, **207–237**.

GLABONIAT, MANUELA et al. (2006): *Profile deutsch*. Lernzielbestimmungen, Kannbeschreibungen und kommunikative Mittel für die Niveaustufen A1, A2, B1, B2, C1 und C2 des "Gemeinsamen europäischen Referenzrahmens für Sprachen". Buch mit CD-ROM. Klett-Langenscheidt Verlag.

Közös Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés (2002): Európa Tanács Közoktatási Bizottsága. Közoktatási Bizottság. Élő Nyelvek Osztálya. Strasbourg. Piliscsaba: Pedagógus-továbbképzési Módszertani és Információs Központ Kht.

LÁZÁR GYÖRGYNÉ és PANTÓNÉ NASZÁLYI DÓRA (2004): *Készüljünk az új érettségire! Német feladatsorok. Emelt szint*. Budapest: Nemzeti Tankönyvkiadó.

Német nyelv. Emelt szintű írásbeli vizsga. 2012. május 11 (2012). Budapest: Nemzeti Erőforrás Minisztérium. Feladatlap.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/emelt/e_nemet_12maj_fl.pdf (letöltve: 2013. május 20.)

Német nyelv. Emelt szintű írásbeli vizsga. 2012. május 11. (2012) Budapest: Nemzeti Erőforrás Minisztérium. Javítási-értékelési útmutató.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/emelt/e_nemet_12maj_ut.pdf (letöltve: 2013. május 20.)

Német nyelv. Emelt szintű írásbeli vizsga. 2012. május 11. (2012) Budapest: Nemzeti Erőforrás Minisztérium. Hanganyag. [e_nemet_12maj_fl.mp3](#)

http://193.225.13.214/erettsegi2012/emeltszintu_erettsegi_2012.html (letöltve: 2013. május 20.)

SOMINÉ HREBIK OLGA (2009): *Érettségi Tréning – Német*. Budapest: Maxim Kiadó.

PETNEKI KATALIN, VÍGH TIBOR

**A 2012. MÁJUSI NÉMET NYELV
EMELT SZINTŰ
ÍRÁSKÉSZSÉG FÜZET/VIZSGARÉSZ (ÖSSZETEVŐ)
ELEMZÉSE**

Az emelt szintű íráskészség vizsgarész elemzésének célja az I. és a III. kutatási terület vizsgálata (vö. 5.1. fejezet). Ennek megfelelően először áttekintést adunk a 4. vizsgarész felépítéséről a vizsgaleírás és a feladatlap, illetve a javítókulcs alapján, majd a teszt legfőbb alapstatisztikai jellemzőit tárgyaljuk. Ezt követően megvizsgáljuk az egyes feladatok működését, végül az értékelési rendszer jellemzőit elemezzük.

1. Az íráskészség vizsgarész felépítése

A vizsgarész célja annak megállapítása, hogy a jelölt képes-e a B2 szintnek megfelelő minőségű szöveget létrehozni. A középszinthez hasonlóan az emelten is két feladatból áll az utolsó vizsgarész. Az ok azonban más: míg középszinten azért van két feladat, hogy a két nyelvi szintet (A2 és B1) mérni lehessen, itt mindkét feladat értékelése a B2 szint kritériumaihoz igazodik. Az analitikus értékelési skálák tehát úgy vannak felépítve (0.1. táblázat), hogy magasabb pontszámot csak az kaphat, akinek teljesítménye eléri a B2 szintet, azaz képes „álláspontját viszonylag árnyaltan, érvelését rendszerezetten kifejtteni, a nyelvi eszközök széles skálájának változatos alkalmazásával összefüggő, megfelelően tagolt, logikusan felépített szöveget létrehozni” (40/2002 OM rendelet, Vizsgaleírás: 12) az adott témáról. Az analitikus skála ezért a két feladatnál kizárólag a tartalmi kategóriánál tér el, ugyanis a két feladatban a kommunikációs célhoz más-más szövegtípust kell alkalmazni, és a tartalmi kritérium az ennek való megfelelést vizsgálja. A többi szempontból (szövegalkotás, szókincs, nyelvtan) a skála egységes.

0.1. táblázat: A 4. vizsgarész értékelési szempontjainak és pontszámainak szintezése

Kritérium	5 pont	4 pont	3 pont	2 pont	1 pont	0 pont
Tartalom	eléri a B2 szintet			nem éri el a B2 szintet		
Szövegalkotás	eléri a B2 szintet			nem éri el a B2 szintet		
Szókincs, kifejezés	eléri a B2 szintet			nem éri el a B2 szintet		
Nyelvtan, helyesírás	eléri a B2 szintet			nem éri el a B2 szintet		

Az első feladat hasonlít a középszint második feladatához: levelet kell írni input szövegre reagálva, megadott irányító szempontok szerint, az értékelési szempontok megnevezése pedig azonos. Azonban nemcsak a szöveg hossz más (120–150 szavas levelet kell írni), hanem többnyire komplexebb szöveget kell létrehozni, mint középszinten. Ennek megfelelően az értékelési skálákban van eltérés a két szint között

0.2. táblázat: Az emelt szint íráskészség 1. feladatának értékelési szempontjai és pontjai

Értékelési szempont	Maximális pontszám
Tartalom	5
Formai jegyek és hangnem	2
Szövegalkotás	3
Szókincs, kifejezés mód	5
Nyelvtan, helyesírás	5
Feladatpont	20
Vizsgapont	10

Az első feladat kiértékelésekor a vizsgaleírás figyelembe veszi azt a tényt, hogy egy levél megírása valójában inkább a B1 szinthez tartozik, ezért a kapott feladatpontokat a végén felezni kell, így kapjuk meg a vizsgapontokat. Ezzel biztosítják, hogy a két feladat

aránya az értékelésben 1/3–2/3 legyen. A levél formai jegyeinek betartásával tehát nem lehet sok pontot szerezni, a nyelvhasználat minőségére pedig szigorúbb pontozás jár, mint középszinten. Amennyiben a dolgozat a tartalom, szövegalkotás, szókincs, kifejezésmód vagy a nyelvtan, helyesírás értékelési szempont alapján 0 pontos, akkor a többi szempont alapján is 0 pont (Német nyelv, emelt szintű írásbeli érettségi vizsga, javítási és értékelési útmutató 2012: 13). Ezt a lenullázási szabályt a vizsgált dolgozatok közül csupán 3 esetben (0,5%-ban) hagyták figyelmen kívül az értékelők.

A második feladatban a vizsgázónak saját véleményét kell kifejtenie 200–250 szavas szövegben egy adott témáról. Támpontot itt is egy inputszöveg nyújt, és a feladat négy irányító szempontot tartalmaz. Szövegfajtának többnyire olvasói levél szerepel, mert ez felel meg leginkább a véleménykifejtés szituációjának. A feladat a B2 szint leírásának felel meg, így a kritériumok alapján adható pontok alapján egyből a vizsgapontokat kapjuk meg. Amennyiben a dolgozat bármelyik értékelési szempont alapján 0 pontos, akkor a többi értékelési szempont pontszáma is 0 pont (Német nyelv, emelt szintű írásbeli érettségi vizsga, javítási és értékelési útmutató 2012: 19). Ezt a vizsgált mintában csupán 10 esetben (1,7%-ban) hagyták figyelmen kívül a javító tanárok.

0.3. táblázat:

Az emelt szint íráskészség 2. feladatának értékelési szempontjai és pontjai

Értékelési szempont	Maximális pontszám
Tartalom	5
Szövegalkotás	5
Szókincs, kifejezésmód	5
Nyelvtan, helyesírás	5
Összesen	20

Mivel a vizsgázónak a második feladatban már jóval hosszabb szöveget kell létrehoznia, s nem biztos, hogy az érettségiző egy adott témáról tud a megadott terjedelemben írni, ezért lehetőség van két téma közül választani. De mindkét témához ugyanaz a feladat tartozik, s egyforma a kritériumrendszer is, ami alapján értékelik.

2. Az íráskészség vizsgarész alapstatisztikai jellemzői

A 2012. évi emelt szintű érettségi íráskészség feladatlap két feladatot tartalmaz, az 1. feladat inputszövege egy reklám a stájerországi Styrassic Parkba. A vizsgázóknak egy félhivatalos levelet kell írni és a programok iránt részletesen érdeklődni. A 2. feladatban két téma közül választhattak az érettségizők: a lakókörnyezet megszépítésére irányuló verseny, ill. az önkéntes munka szerepe az ember életében (Német nyelv, középszintű írásbeli érettségi vizsga, feladatlap 2012: 31–35).

Az íráskészség vizsgarész a teljes mintán 66,5%-os teljesítettségű volt, azaz jobban sikerült, mint a hallott szöveg értése, de rosszabbul, mint az olvasott szöveg értése feladatsor, és a nyelvhelyesség teszt megoldottságától nem különbözött szignifikánsan (vö. 0.24. táblázat). Az 0.4. táblázat a két feladat paramétereit tartalmazza, amelyben az átlagot és a szórást százalékpontokban adjuk meg, hogy össze tudjuk hasonlítani a teljes teszt megoldottságával. Az első feladat eredményét a nyerspontok és nem az átszámított vizsgapontok alapján számítottuk, valamint a továbbiakban ezzel számoltunk.

0.4. táblázat: Az íráskészség feladatsor jellemzői (VT számításai alapján³³)

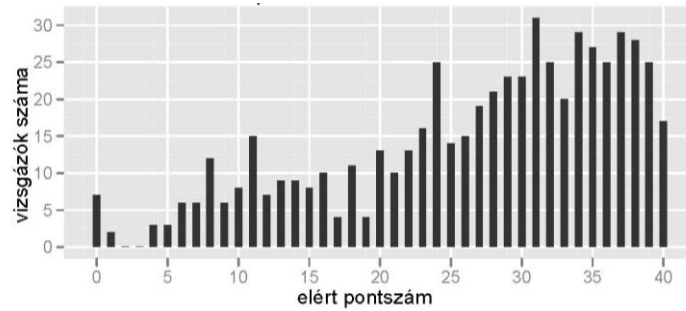
Kategória	Feladat		Teszt
	1.	2.	
Cronbach-alfa	0,85	0,96	0,92
Átlag (%p)	70,9	62,1	66,5
Szórás (%p)	22,4	32,6	25,2
Magyarázóerő (R ² %)	39,0	60,0	100

Eredményeink szerint az értékelési skálák használata magas megbízhatóságot eredményez. Ugyanakkor az első feladat megbízhatósága alacsonyabb, mint a második feladaté. Ennek oka a feladatok eltérő nehézsége és az értékelési szempontok különböző működése. A két feladatot azonos nyelvi szintre készítették, azonban az első feladatban a vizsgázók szignifikánsan jobb átlagteljesítményt értek el ($t=8,7$; $p<0,001$). Ennek oka az egyes értékelési szempontok nehézségében, az értékelési skálák és a feladatban megadott kommunikációs célok, valamint a létrehozandó szövegfajta különbségében keresendő. A két feladat közötti korreláció erős ($r=0,67$; $p<0,001$). A második feladat magyarázóereje másfélszerese az első feladaténak.

Az 0.1. ábra a teszten elért pontszámok eloszlását mutatja. 1–3 pontot nem lehet elérni ebben a vizsgarészben, amely a lenullázási szabályból adódik. Az ábra alapján megállapítható, hogy az eredmények a magasabb pontszámok felé, jobbra tolódnak, ugyanakkor

³³ VT rövidítéssel jelöltük Vigh Tibor számításait.

jelentős különbségek vannak a tanulók íráskészség-szintjében, ahogyan ezt az 0.4. táblázat is mutatta.



0.1. ábra:
A teszten nyújtott pontszám abszolút gyakorisági eloszlása
(megbízó által készített grafikon)

3. Feladatszintű elemzés

A megbízótól kapott adatok mellett további másodelemzéseket végeztünk, amelynek egyik eredménye az 0.5. táblázat, ami az első feladat alapstatisztikai jellemzőit tartalmazza.

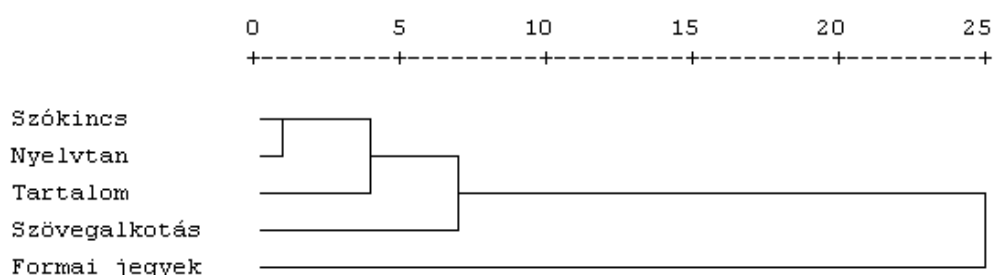
0.5. táblázat:

Az 1. feladat értékelési szempontjainak átlaga, szórása és megoldottsága
(a megbízótól kapott adatok és VT számításai alapján)

Értékelési szempont	Maximális pontszám	Átlag	Szórás	Megoldottság (%)	Magyarázó erő (%)
Tartalom	5	3,47	1,44	69,4	27,5
Formai jegyek és hangnem	2	1,03	0,81	51,5	9,7
Szövegalkotás	3	2,43	0,75	81,0	13,5
Szókincs, kifejezőmód	5	3,77	1,24	75,4	25,1
Nyelvtan, helyesírás	5	3,49	1,21	69,8	23,8

Az öt értékelési szempont megoldottságában szignifikáns különbség van ($\chi^2=337$; $p<0,001$). A vizsgázók legjobban a szövegalkotás szempontnál teljesítettek. Ettől szignifikánsan ($p<0,001$) alacsonyabb a szókincs, kifejezőmód megoldottsága, amelytől szignifikánsan ($p<0,001$) alacsonyabb eredményt értek el a nyelvtan, helyesírás és a tartalom szempontnál. A legnehezebben teljesíthető kritérium a formai jegyek és hangnem.

Az 0.5. táblázat azt is mutatja, hogy az egyes értékelési szempontoknak mekkora a súlya az első feladatra kapott összpontszám alakulásában. Eredményeink szerint a tartalom, illetve a két nyelvi szempont (a szókincs, kifejezőmód és a nyelvtan helyesírás) rendelkezik magas magyarázóerővel. Ezeket követi a szövegalkotás; a legkisebb magyarázóereje a formai jegyek és hangnem szempontnak van. E szempontnak a többihez képest alacsonyabb az összpontszámmal való korrelációja ($r=0,54$; $p<0,001$) és ez kapcsolódik a legtávolabbról a többi értékelési szemponthoz (0.2. ábra).



0.2. ábra:

Az 1. feladathoz tartozó értékelési szempontok klaszteranalízise (VT által készített ábra)

Az 0.6. táblázat bemutatja témánként az egyes pontszámok elérők számát és arányát. Ez alapján megállapítható, hogy a vizsgázók többsége elérte a B2 szintet, az arányuk a tartalom szempontnál alacsonyabb. A 0 pontot elérők száma és aránya azért azonos

a formai jegyek és hangnem kivételével, mert, ha ezek valamelyike 0 pontos, akkor a többi is. Ugyanakkor arról nem szerzünk információt, vajon mi jellemzi azokat az írásműveket, amelyekben a vizsgázók eltérnek a kommunikációs céltól. Erre nem tudunk válaszolni, hiszen tanulói dolgozatok sem állnak rendelkezésünkre, pedig ezek alapján az egyes írányító szempontok működését és a feladat esetleges problémáit is vizsgálhattuk volna.

0.6. táblázat:

Az 1. feladat értékelési szempontjainál elért pontszámok abszolút és relatív gyakorisága (VT számításai alapján)

Pontszám	Tartalom		Formai jegyek és hangnem		Szövegalkotás		Szókincs, kifejezőmód		Nyelvtan, helyesírás	
	N	%	N	%	N	%	N	%	N	%
0	15	2,6	179	31,0	13	2,2	14	2,4	14	2,4
1	36	6,2	201	34,8	52	9,0	13	2,2	27	4,7
2	143	24,7	198	34,3	186	32,2	66	11,4	63	10,9
3	38	6,6	---	---	327	56,6	111	19,2	153	26,5
4	158	27,3	---	---	---	---	169	29,2	198	34,3
5	188	32,5	---	---	---	---	205	35,5	123	21,3
B2 szint alatt	194	33,6	nem értelmezhető		65	11,2	93	16,0	104	18,00

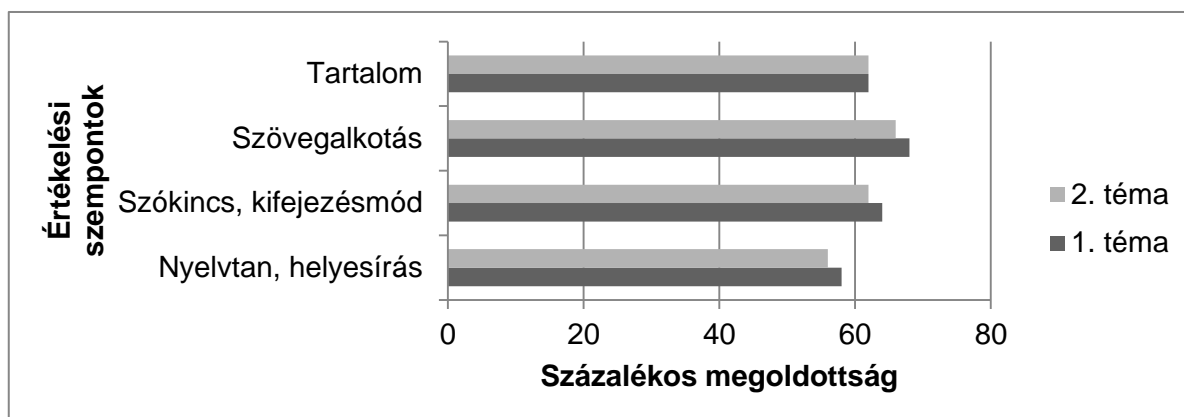
A második feladatban a vizsgázók két téma közül választhattak. Az érettségizők mindkét témát közel azonos arányban választották: az elsőt a diákok közül 276-an, a másodikat 302-en. A választhatóság fontos alapelv, hiszen elsősorban nem a vizsgázók adott témáról megszerzett háttértudását mérjük, hanem, hogy írásban sikeresen teljesítik-e a megadott kommunikációs célt. Ugyanakkor nagyon lényeges, hogy a feladatok megoldottságában ne legyen különbség. A továbbiakban ezt vizsgáljuk.

A második feladatban lévő két téma paramétereit közül az átlagban nincs szignifikáns különbség ($t=0,8$; $p>0,05$). Ugyanez jellemző az összes értékelési szempontra is (0.7. táblázat). Ez látható az egyes értékelési szempontok százalékos megoldottsága alapján is (0.3. ábra). Ezért a két témát nehézségében ekvivalensnek tekinthetjük, ami megfelel a vizsgaleírásban közölt elvárásoknak.

0.7. táblázat:
A 2. feladat értékelési szempontjainak átlaga, szórása témánként
(VT számításai alapján)

Értékelési szempont	1. téma (N=276)		2. téma (N=302)		Két téma együtt (N=578)	
	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Tartalom	3,1	1,8	3,1	1,9	3,1	1,8
Szövegalkotás	3,4	1,6	3,3	1,8	3,4	1,7
Szókincs, kifejezés mód	3,2	1,6	3,1	1,7	3,2	1,7
Nyelvtan, helyesírás	2,9	1,5	2,8	1,6	2,9	1,5
2. feladat	12,7	6,2	12,2	6,8	12,4	6,5

A négy értékelési szempont megoldottságában szignifikáns különbség van ($\chi^2=200$; $p<0,001$). A vizsgázók az első feladathoz hasonlóan a legjobban a szövegalkotás szempontnál teljesítettek. Ezekből szignifikánsan ($p<0,001$) alacsonyabb a szókincs, kifejezés mód és a tartalom megoldottsága. A harmadik legnehezebben teljesíthető kritérium a nyelvtan, helyesírás.

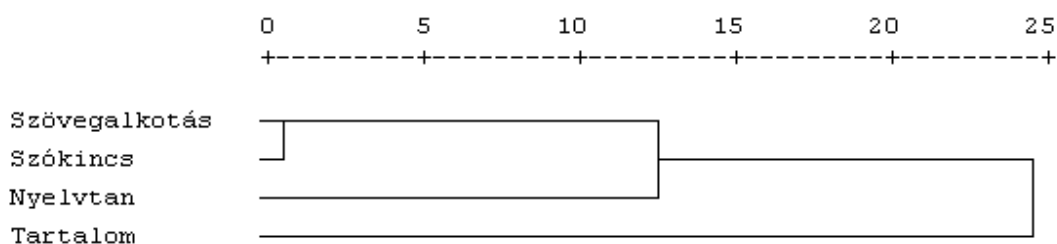


0.3. ábra:
A 2. feladat megoldottsága szempontok szerint a két téma alapján
(VT által készített ábra)

Megvizsgáltuk, hogy az egyes értékelési szempontoknak mekkora a súlya a 2. feladatra kapott összpontszám alakulásában. Ennek eredményeit az 0.8. táblázat tartalmazza. Megvizsgáltuk továbbá klaszteranalízissel azt is, milyen kapcsolat van az egyes értékelési szempontok között (0.4. ábra).

0.8. táblázat:
A 2. feladathoz tartozó értékelési szempontok magyarázóereje témánként és együtt
(VT számításai alapján)

Értékelési szempont	1. téma	2. téma	Két téma együtt
Tartalom	29,9	27,8	28,8
Szövegalkotás	20,4	23,8	22,1
Szókincs, kifejezés mód	25,0	25,3	25,2
Nyelvtan, helyesírás	24,1	22,9	23,6
Megmagyarázott variancia (%)	100	100	100



0.4. ábra:
A 2. feladathoz tartozó értékelési szempontok klaszteranalízise
(VT által készített ábra)

Eredményeink szerint mindkét témánál külön és együtt az értékelési szempontoknak hasonló a súlya a második feladatra kapott összpontszám alakításában, mindegyik közel van 25%-hoz, bár következetesen a kommunikációs cél teljesülését vizsgáló tartalom-szemponthoz nagyobb a magyarázóereje, és ez kapcsolódik legtávolabbról a többi értékelési szemponttal. Ennek háttérében az áll, hogy ez a szempont az íráskészség más dimenzióját méri. Az eredmények összhangban vannak a vizsgakövetelményekkel, hiszen lényeges, hogy a szerző a kommunikációs céljainak megfelelő, koherens és megfelelő nyelvi szintű szöveget hozzon létre. Ez tükröződik az egyes szempontok súlyában.

A 0.9. táblázat bemutatja témánként az egyes pontszámokat elérők számát és arányát. Ennek alapján megállapítható, hogy a tartalmi szempontnál a vizsgázók 39%-a, a többi szempontnál kb. 30%-a teljesített a B2 szint alatt. A 0 pontot elérők száma és aránya azért azonos, mert ha az értékelési szempontok valamelyike 0 pontos, akkor a többi is. A második témánál ez némiképp gyakrabban fordult elő, mint az első témában. Ugyanakkor arról nem szerzünk információt, hogy mi jellemzi azokat az írásműveket, amelyekben a vizsgázók eltérnek a kommunikációs céltól.

0.9. táblázat:

A 2. feladat értékelési szempontjainál elért pontszámok abszolút és relatív gyakorisága (zárójelben a százalékos arányok láthatók, N1=276 és N2=302; VT számításai alapján)³⁴

Pontszám	Tartalom		Szövegalkotás		Szókincs, kifejezőmód		Nyelvtan, helyesírás	
	1. téma	2. téma	1. téma	2. téma	1. téma	2. téma	1. téma	2. téma
0	36 (13)	53 (18)	33 (12)	51 (17)	33 (12)	49 (16)	33 (12)	50 (17)
1	23 (8)	16 (5)	8 (3)	10 (3)	11 (4)	14 (5)	14 (5)	22 (7)
2	47 (17)	52 (17)	18 (7)	16 (5)	32 (12)	26 (9)	34 (12)	27 (9)
3	26 (9)	18 (6)	56 (20)	49 (16)	60 (22)	59 (20)	87 (32)	87 (29)
4	54 (20)	59 (20)	67 (24)	72 (24)	70 (25)	76 (25)	73 (26)	78 (26)
5	90 (33)	104 (34)	94 (34)	104 (34)	70 (25)	78 (26)	35 (13)	38 (13)
B2 szint alatt	227 (39)		136 (24)		165 (29)		178 (31)	

A 0.9. táblázatból az értékelési skála egyik sajátossága is látható, hiszen az azonos sávba tartozó pontszámoknál (2-1 és 4-3) következetesen a magasabb pontszámot adják a vizsgáztatók. Ez a jelenség felveti azt a kérdést, hogy célszerű-e a 4 szintleírásból álló értékelési skálát 6 pontra bontani. Ezt az értékelési rendszer jellemzőinek vizsgálatakor részletesebben is elemezzük.

³⁴ Kék színnel jelöltük azon vizsgázók számát és arányát, akik nem érik el a B2 szintet, barnával pedig azokat, akik elérik ezt a szintet az V/1. táblázatnak megfelelően.

4. Az értékelési rendszer elemzése

A továbbiakban az értékelési rendszer megfelelőségét vizsgáljuk, és célunk annak megállapítása, hogy (1) milyen nehézségi szintekbe sorolhatók az egyes értékelési szempontok; (2) mennyire fedik le az értékelési skálapontok a vizsgázók íráskészség-szintjei által meghatározott intervallumot; (3) mennyire megfelelően működnek az értékelési skálák.

4.1. Az értékelési szempontok jellemzői és teljesítésének nehézsége

A parciáliskredit-moddal (Masters 1982) vizsgált adatbázis eredményeként az 0.10. táblázat az értékelési szempontok nehézségi indexét és modellilleszkedését mutatja a két feladat alapján. A táblázatban lévő logitérték a vizsgázók íráskészség-szintjének és az értékelési szempontokhoz tartozó nehézségi paraméterek közös skálájának egységét jelöli, amely intervallumskálán helyezkedik el. A középérték logittal az adott értékelési szempont átlagos nehézségét határoztuk meg a vizsgázók íráskészsége alapján. Egy nehezebb, magas középérték logittal rendelkező értékelési szempont esetében például a tanulók alacsonyabb pontértéket érnek el.

0.10. táblázat:
Az értékelési szempontok nehézségi indexe és modellilleszkedése
(VT számításai alapján)

Feladat	Értékelési szempont	Középérték logitok	Infit paraméter
1.	Tartalom	-0,35	1,6
	Formai jegyek és hangnem	0,96	1,6
	Szövegalkotás	-1,35	1,0
	Szókincs, kifejezőmód	-0,73	1,0
	Nyelvtan, helyesírás	-0,39	1,0
2.	Tartalom	0,46	1,0
	Szövegalkotás	0,20	0,7
	Szókincs, kifejezőmód	0,41	0,7
	Nyelvtan, helyesírás	0,79	0,7

Az első feladathoz tartozó értékelési szempontok teljesítésének nehézsége a formai jegyek és hangnem kivételével az átlagos íráskészség-szint (0 logitérték) alatt, a második feladat szempontjai az átlagos szintnél helyezkednek el. Ez is azt mutatja, hogy a két feladat eltérő nehézségű. Mindkét feladatban a legenyhébben értékelt szempont a szövegalkotás, míg a legszigorúbban az első feladatban a formai jegyek és hangnem, a másodikban pedig a nyelvtan, helyesírás. Ezek az eredmények teljesen összhangban vannak a 2006. évi német nyelv érettségi korábbi elemzésének tanulságaival (Vígh, 2010: 112). Az

írás-készség azonos komponenseihez tartozó értékelési szempontok szerinti képesség-szintek közötti különbség tartalomnál a legkisebb (0,8), közel azonos a két nyelvi komponensnél (a szókincs, kifejezésmódnál 1,1; a nyelvtan, helyesírásnál 1,2 logit) és a legnagyobb a szövegalkotásnál (1,6). Az itt megjelenő nagyobb különbség feltehetően a két feladatban megjelenő eltérő skálaszerkezettel magyarázható.

Az infit paraméterek azt adják meg, hogy az adott értékelési szempont mennyire illeszkedik a többi közé. A megfelelően illeszkedő szempont értéke közel van 1-hez. Az elfogadható értékek sávja függ az elemszámtól, ezért nem tudunk egy előre meghatározott intervallumot kijelölni, hanem a vizsgálat céljainak függvényében kell mérlegelni (Molnár 2013: 93). Egy kb. 1000 fős mintánál 0,8 és 1,2 között határozhatjuk meg ezt a sávot, az 578 fős minta esetében ezt 0,7 és 1,3 között határoztuk meg.

Az 0.10. táblázatban az első feladatnál két nem illeszkedő szempontot találunk, az infit paraméter értéke 1,3 fölött van, amely azt jelenti, hogy ezeket a szempontokat többdimenzionalitás jellemzi, a többihez képest mást is mérnek. E jelenség értelmezésében érdemes a két feladatot önállóan is megvizsgálni. A feladatok elkülönített elemzése ugyanazt mutatta: az első feladatban az első két szempont, a második feladat önálló elemzésekor pedig a tartalom nem illeszkedik. Ezeket az íráskészség kommunikatív komponenseiként határozhatjuk meg. Feltételezhetjük, hogy ezek a szempontok függenek az adott feladat megértésétől, a kommunikatív keret felismerésétől, ezért a vizsgázókból nagy arányban válhatnak ki következtelen válaszokat, de az is elképzelhető, hogy az értékelési skálák nem működnek megfelelően, vagy az értékelők nem használják következetesen az egyes skálapontokat. Ezek a megállapítások összhangban vannak a korábbi elemzésünkkel (Vígh, 2010; 2012).

4.2. A vizsgázók íráskészség-szintje és az értékelési szempontok nehézsége

A parciáliskredit-modell segítségével közös skálán tudjuk ábrázolni a vizsgázók íráskészség-szintjét és az értékelési skálapontok küszöbértékét. A küszöbértékek (vagy lépéssparaméterek, Thurstone-küszöbértékek) megadják, milyen képességszint szükséges ahhoz, hogy valaki 50%-os valószínűséggel érjen el egy adott pontszámot, így ez az adott értékelési szempontoz tartozó átlagos nehézségi index értelmezésében játszik szerepet.

Az 0.5. ábra első oszlopában az íráskészség feladatokon elért eredményeket ún. logitskálán helyezi el a ConQuest program (Wu, Adams és Wilson, 1998). A második oszlop a vizsgázók íráskészség-paramétereit tartalmazza, a harmadik az egyes értékelési szempontokhoz tartozó küszöbértékeket jeleníti meg. A pont előtti számjegy az értékelési szempontot, a pont utáni pedig a skálapontok küszöbértékét jelenti. Például a 2.1. a formai jegyek és hangnem szempont 1-es pontszám elérésének küszöbértékét mutatja, amely a 0 logitértéknél helyezkedik el. A továbbiakban először a minta íráskészség-paramétereinek eloszlását, majd az értékelési skálapontok jellemzőit, végül a kettőt együtt elemezzük.

A mintatérképen minden 'x' négy vizsgázót reprezentál, ugyanakkor az ábrázolt 'x'-eken kívül is vannak még vizsgázók a logitskála magasabb-alacsonyabb tartományában, csak négynél kevesebben vannak, így az ábra nem tartalmazza az adataikat. A logitskálán megadott negatív számok átlag alatti íráskészség-szintet, a 0 az átlagos szintet, míg a pozitív számok az átlag feletti szintet jelölik. A mintatérképről leolvashatók a vizsgázók íráskészség paramétereinek jellemzői. A vizsgázók íráskészsége tág határok között mozog, 6 logitegységnyi a skála. Az érettségizők többsége - a minta kb. 70%-a - helyezkedik el az átlagos készségszint felett.

Az értékelési skálapontok akkor működnek az elvártnak megfelelően, ha az íráskészség szintjének növekedésével együtt arányosan nő a pontszám elérésének valószínűsége. A konkrét esetben az értékelési szempontokhoz tartozó küszöbértékek az íráskészség-szint növekedésével kivétel nélkül emelkednek. Minden értékelési szempont megfelelően működik: minél magasabb tudásszinttel rendelkeznek a tanulók, annál valószínűbb, hogy magasabb pontszámot érnek el.

logit	személy	értékelési skálapontok								
5										
	X									
	X									
4	X									
	X									
	XXX									
	XXX									9.5
	XXX									
3	XX									
	XXXXXX					5.5				
	XXXXXX									
	XXXXXX									8.5
2	XXXXXXXX									
	XXXXXXXX	1.5	2.2		4.5		6.5	7.5		
	XXXXXXXX									9.4
	XXXXXXXX									
1	XXXXXXXX						6.4		8.4	
	XXXXXXXX	1.4		3.3		5.4	6.3	7.4		
	XXXXXXXX									
	XXXXXXXX	1.3			4.4					
0	XXXXXXXX		2.1						8.3	9.3
	XXXXXX							7.3		
	XXXX						6.2			9.2
	XXX							7.2	8.2	
	XXX				4.3	5.3	6.1	7.1	8.1	9.1
-1	XXXXX									
	XXX									
	XX			3.2						
	XX	1.2				5.2				
-2	X									
	X				4.2					
-3		1.1			4.1	5.1				
				3.1						
-4										

0.5. ábra:
Az íráskészség vizsgarész minta- és skálapont térképe
(VT által készített ábra)³⁵

³⁵ A számok közül az első az alábbi értékelési szempontokat jelöli: 1=tartalom; 2=formai jegyek és hangnem; 3=szövegalkotás; 4=szókincs, kifejezés mód; 5=nyelvtan, helyesírás; 6=tartalom 7=szövegalkotás; 8=szókincs, kifejezés mód; 9=nyelvtan, helyesírás; a pont utáni második szám az adott értékelési szemponton belüli pontszám küszöbértékét.

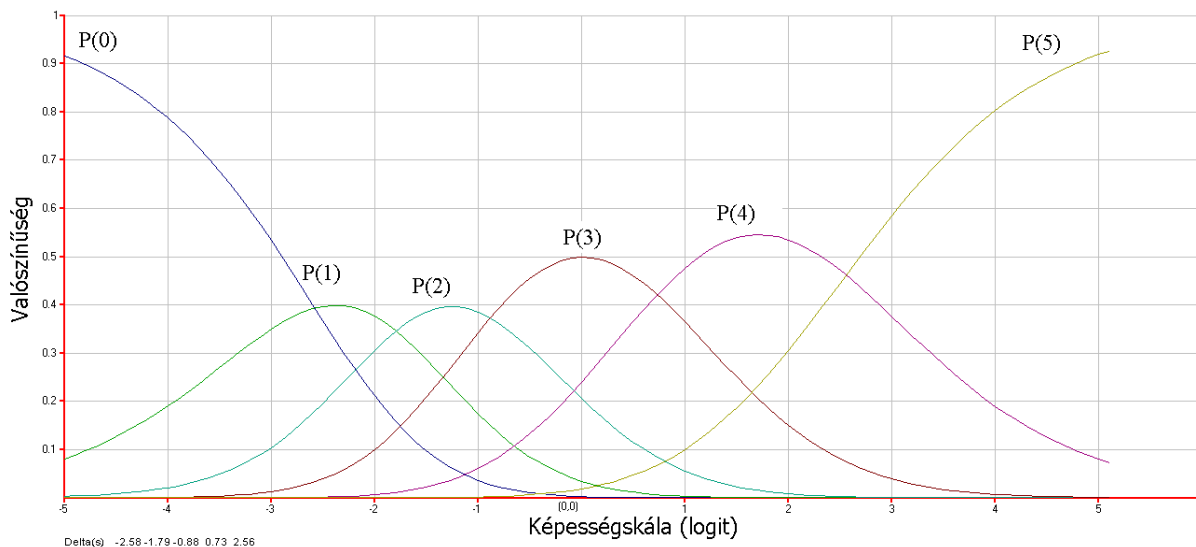
A két feladatban az egyes szempontokhoz tartozó pontszámok eltérően működtek. Az első feladatban (1, 3–5. szempont) jellemzően a küszöbértékek közül az 1 és a 2 az átlagos képességszint alatt helyezkedik el a -3 és -1 logitérték közötti tartományban és a többi az e fölötti készségszinten. A második feladatban (az V/5. ábrán 6–9. szempont) jellemzően a -1 és 3 logitérték közötti tartományban vannak a küszöbértékek. Ez abból adódik, hogy a két feladat nehézsége eltér egymástól.

A vizsgázók íráskészség-szintjét és az értékelési skálapontok küszöbértékét együtt vizsgálva az ábráról leolvashatjuk (0.5. ábra), hogy ezek mennyire jól fedik le a vizsgázók íráskészség-szintjei által meghatározott intervallumot. A mintához jól illesztett tesztnél a készségparamétereket jelző 'x'-ek és az értékelési skálapontokat jellemző küszöbértékek egymással párhuzamosan futnak. Az ábrán az 1-es logitértékig a küszöbértékek megfelelően fedik le a vizsgázók íráskészség-szintjét. Az e fölötti szintet, azaz a minta kb. 50%-át az értékelési skálákhoz tartozó maximális pontszámot leíró deskriptorok jellemzik, amelyek nem fedik le teljesen a vizsgázók íráskészség-szintjét. Az 1–2 pontot leíró deskriptor a minta 15%-ára jellemző. A minta többi részét a 3–4 pontot tartalmazó szintleírások jellemzik. Ezek az értékek arra utalnak, hogy az alacsonyabb pontszámot leíró deskriptorok kevésbé jellemzők az érettségizőkre, míg a magasabb íráskészség-szint méréséhez további differenciálás lenne indokolt az értékelési skálában.

4.3. Az értékelési skálák működése

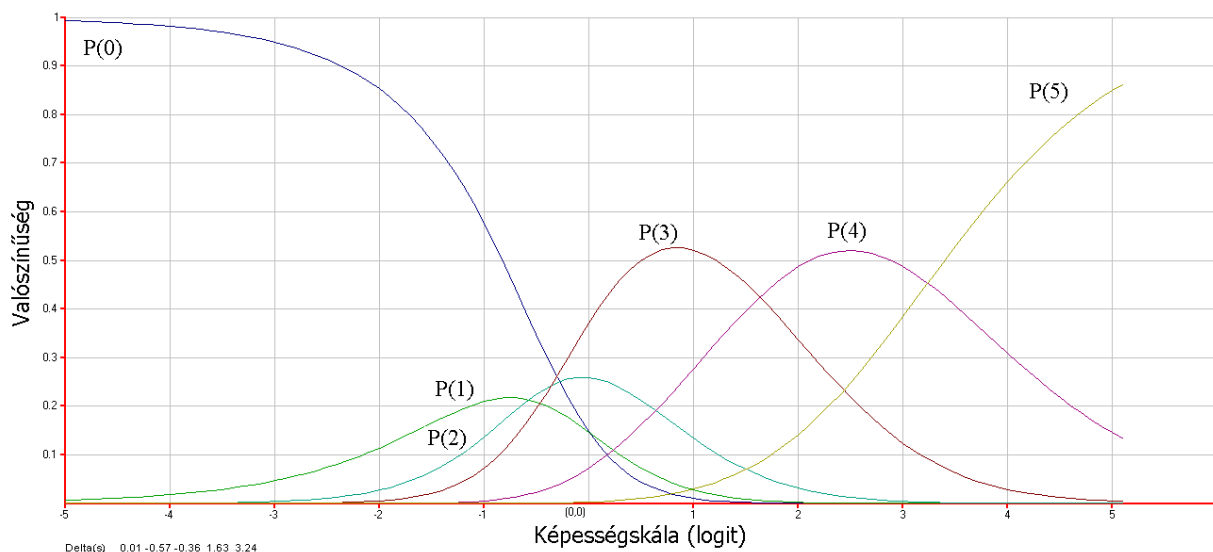
Az előzőekben az egyes értékelési szempontok és skálapontok nehézségét, valamint a vizsgázók íráskészség-szintjét elemeztük, amelyek alapján számos megválaszolatlan kérdés maradt. Ezek közül a legfontosabb, hogy mi jellemzi az értékelési skálák működését, azaz a skálahasználat során minden deskriptor azonos mértékben hasznosul-e, vagy vannak-e olyan skálák, amelyeket a vizsgáztatók nem használnak teljesen.

E kérdések megválaszolásához a parciáliskredit-moddellel felrajzolható karakterisztikus görbéket vizsgáljuk. Ezek különböző íráskészség-szintek mellett annak valószínűségét mutatják, hogy a vizsgázó teljesítményét az adott értékkategóriába sorolják az értékelők. A továbbiakban néhány értékelési szempontoz tartozó karakterisztikus görbét elemzünk, amelyek tipikusnak mondhatók. Ezek közül az első példa (0.6. ábra) a nyelvtan, helyesírás szempontoz tartozó értékelési skála használatának eredményét mutatja, amely többnyire megfelelően működik.



0.6. ábra:
Az első feladat nyelvhelyesség értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)

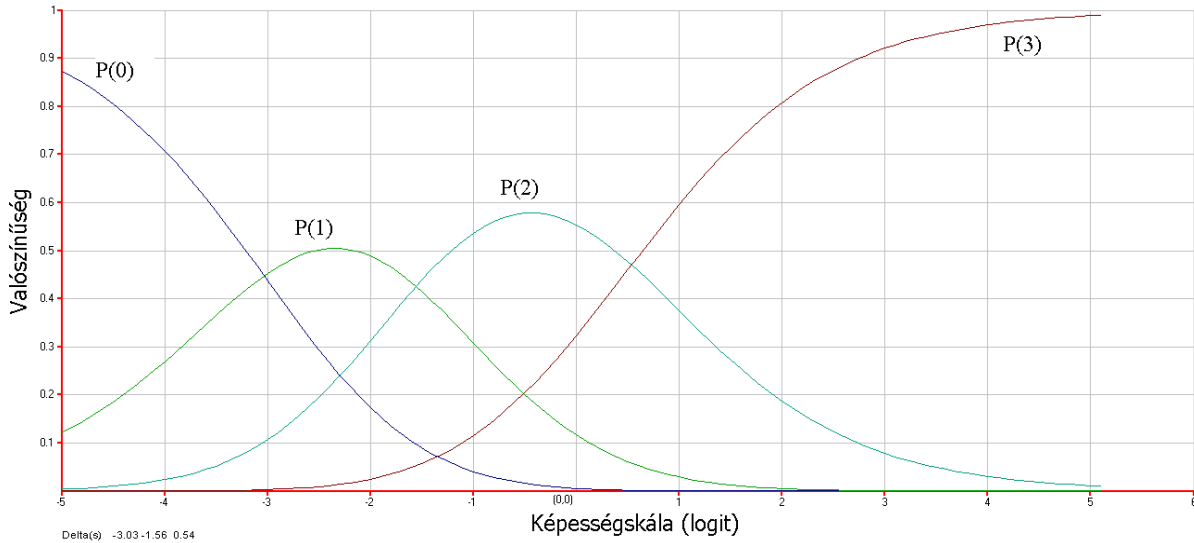
Az 0.6. ábrán a legfelső, P(0) jelű görbe a 0 pontba tartozás valószínűségét jelöli, amely azt mutatja, hogy a képességszint növekedésével hogyan csökken a 0 pontszám elérésének valószínűsége. Az 1 pontszám valószínűségi görbéje [P(1)] azt jelöli, hogy a képességszint növekedésével hogyan növekszik, majd csökken az 1-es pontszámra való tartozás valószínűsége. Ugyanezt látjuk a többi pontszámhoz tartozó görbénél is. Az egyes görbék metszéspontját a $\delta_1, \delta_2 \dots \delta_5$ paraméterek jelölik. A δ_1 paraméter például azt a képességszintet jelöli, ahol azonos annak valószínűsége, hogy a vizsgázók 0 és 1 pontot érnek el. Az értékelési skálák működése szempontjából lényeges, hogy a delta értékek a képességskálán arányosan növekedjenek. Esetünkben az öt δ érték (-2,58; -1,79; -0,88; 0,73; 2,56) az elvártnak megfelelően emelkedik a képességszint növekedésével. Ugyanakkor a paraméterek közül a δ_1, δ_2 és δ_3 egymáshoz közelebb van, mint δ_4 és δ_5 , amely arra utal, hogy kevés azon diákok száma, akik 1–2 pontot kapnak és meghatározóvá válik az a vizsgázói csoport, amelyben 3–4 és 5 pontot érnek el. Ugyanakkor utalni kell arra is, hogy a δ értékek attól függ, hogy az egyes pontszám-kategóriákban hány tanuló van, és éppen az alacsonyabb, a B2 alatti képességszintet jelölő kategóriákban vannak kevesebben (0.9. táblázat). Ehhez hasonló, de problémásabb működést mutat a második feladathoz ugyanennek az értékelési skálának a karakterisztikus görbéje (0.7. ábra).



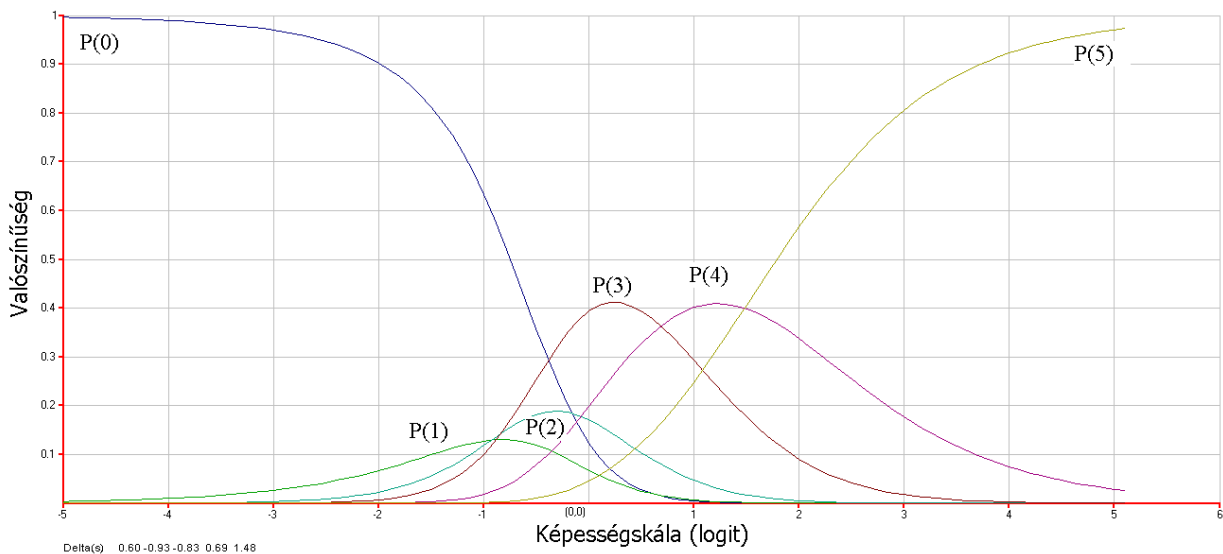
0.7. ábra:
A második feladat, nyelvtan helyesírás értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)

A karakterisztikus görbék azt mutatják, hogy a δ_1 és δ_2 felcserélődnek, majd a δ_3 , δ_4 és δ_5 paraméterek többnyire azonos távolságra vannak (0.7. ábra). A többi nyelvi szempontnál hasonló skálaműködés látható. Ez azt jelzi, hogy az 1–2 pontba tartozás valószínűsége nagyon alacsony, és meghatározóvá válnak azok az érettségizők, akik 3–5 pont közötti eredményeket érnek el, azaz a 6 pontszámból álló skála 4-fokozatúként működik. Ez arra utal, hogy az 1–2 pontos deskriptor megfogalmazása nem megfelelő, vagy nincs szükség rá, mert kevésbé jellemző a vizsgázókra, ezért az értékelők nem alkalmazzák ezeket a skálapontokat.

A szövegalkotás szemponthoz emelt szinten kétféle skáaelrendezés tartozik. Ugyanakkor a két karakterisztikus görbe hasonló működést mutat. Az első feladathoz tartozó skálán a δ értékek folyamatosan növekednek (0.8. ábra), a második feladathoz tartozó δ paraméterek a nyelvtan helyesírás értékelési skálájának karakterisztikus görbéjéhez hasonló képet mutatnak, szélsőséges a skálahasználat (0.9. ábra). Ez a skálaműködés azt jelzi, hogy a -0,5-ös logitértékig a 0 pontszám elérése, majd a 3 pontba tartozás valószínű. Az első feladathoz tartozó szövegalkotás szempont értékelési skálájához hasonlóan a második feladatban lévő 6-fokozatú skála 4-fokozatúként működik, így nem teljesül az a szándék, hogy a vizsgázói teljesítményeket minél pontosabban jellemezzük.



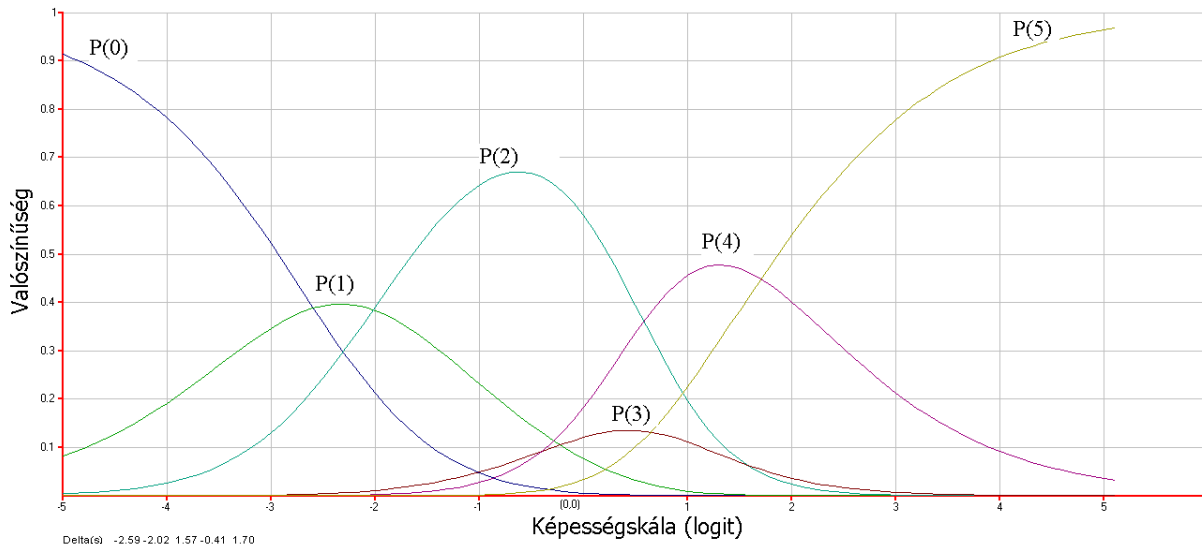
0.8. ábra:
Az első feladat szövegalkotás értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)



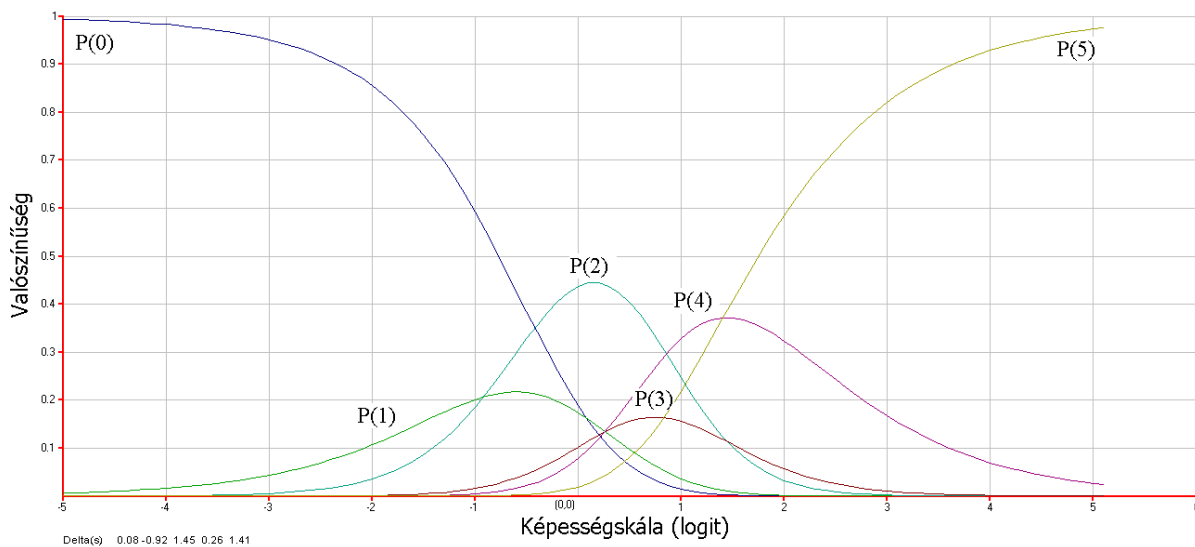
0.9. ábra:
A második feladat szövegalkotás értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)

A tartalom szempontoz tartozó skála mindkét feladatban hasonlóan következetlenül működött (0.10. ábra és 0.11. ábra). Jellemző mindkét esetben az egyes delta paraméterek felcserélődése: δ_1 és δ_2 értéke hasonló és a δ_4 közel van a δ_3 -hoz. Ez a skálaműködés feltehetően azért alakult így, mert az 1-es és a 3-as pontszámra való tartozás valószínűsége nagyon alacsony; ezek a görbék laposak. A 6-fokozatú skála 4-fokozatúként működik, de a szövegalkotás szemponttól eltérően többnyire mind a négy deskriptort használják

az értékelők, csak az egyes pontszámok valószínűsége nem azonos. Feltehető az a vizsgáztatói gyakorlat, hogy az azonos deskriptorokat leíró 1–2 és 3–4 pont esetén a magasabbat adják a vizsgázónak.

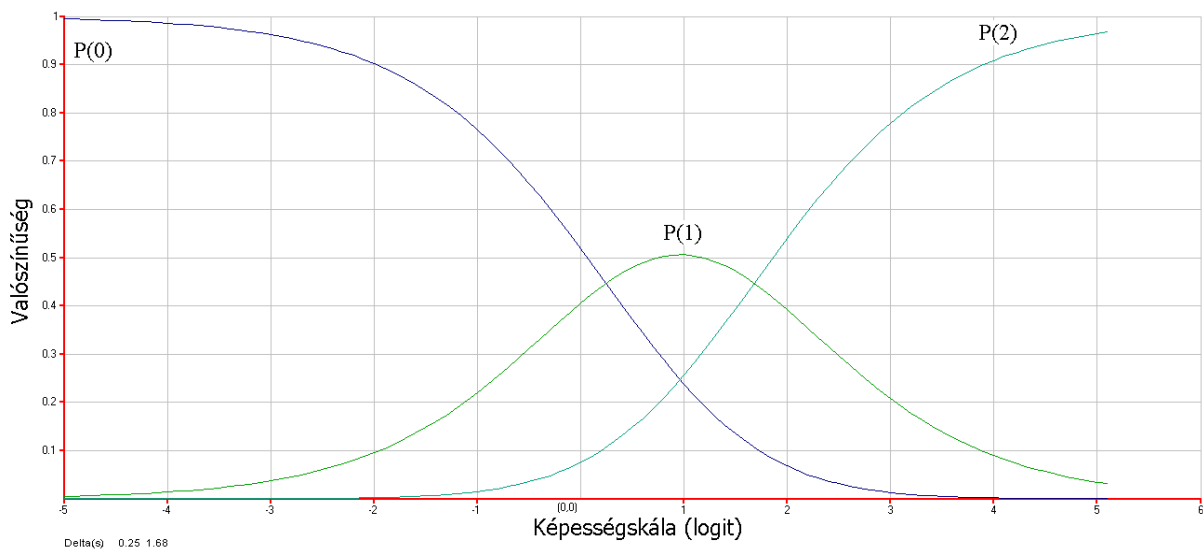


0.10. ábra:
Az első feladat tartalom értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)



0.11. ábra:
A második feladat tartalom értékelési skálájának karakterisztikus görbéje
(VT által készített ábra)

A formai jegyek és hangnem szempontnál egy dichotóm itemhez hasonló jelleggörbe rajzolódik ki. A 0,3 és 1,6 képességszint között nő meg annak a valószínűsége, hogy a vizsgázók inkább 1, mint 0 vagy 2 pontot érnek el ennél a szempontnál. A többi esetben inkább az jellemző, hogy a szöveg formai jegyei teljesen megfelelnek a szövegfajtának, hangneme a közlési szándéknak és a címzethez való viszonynak, vagy pedig nem felelnek meg. Feltehetően ez a skálaműködés az oka annak, hogy e szempont nem illeszkedik a többi közé, a legtávolabbról kapcsolódik a többi szemponthoz és a legalacsonyabb magyarázóerővel rendelkezik.



0.12. ábra:

Az első feladat formai jegyek és hangnem értékelési skálájának karakterisztikus görbéje (VT által készített ábra)

5. Az íráskészség vizsgarész elemzésének összefoglalása

Az íráskészség vizsgarészhez a bevezetőben (5.1. fejezet) két célterülethez (I. és III.) kapcsolódóan fogalmaztunk meg kutatási kérdéseket, amelyek közül az egyik a vizsgarész megbízhatóságára (1. kérdés) és a feladatok súlyára (2. kérdés), a másik pedig az ebben a vizsgarészben alkalmazott mérési-értékelési eljárásokra vonatkozott (7–10. kérdés).

A továbbiakban a kutatási kérdések számozásának megfelelően megadjuk először a vizsgált kérdést, majd pedig a választ.

1. Mennyire megbízhatóan mértek az egyes készségeket mérő feladatsorok?

Eredményeink szerint a vizsgarészben szereplő két feladat mindegyike megfelelő megbízhatósággal mért. Bár az első feladat Cronbach-alfa értéke némiképp alacsonyabb, mint a második feladaté. Ugyanakkor mindkét érték elfogadható, vagyis kijelenthetjük, hogy az értékelési skálák használata megfelelő reliabilitást eredményezett.

2. Milyen mértékben magyarázzák az egyes feladatok az adott vizsgarészben elért összpontszámot?

Az első feladat súlya kisebb az összpontszám alakításában, a másodiké nagyobb. Ez abból adódik, hogy az első feladat könnyebb volt a másodiknál, valamint az első feladatban a formai jegyek és hangnem értékelési szempont súlya kisebb a többinél.

3. Hogyan teljesítettek a tanulók az egyes feladatokban? Van-e jelentős különbség a két feladat megoldottságában? Van-e jelentős különbség a második feladatban megadott két téma megoldottsága között?

Az íráskészség vizsgarészhez tartozó két feladatot a B2 nyelvi szint méréséhez készítették, ugyanakkor az elvárásoktól eltérően, a megoldottság és az értékelési szempontok nehézsége alapján az első feladat könnyebb volt, mint a második, vagyis ezek nehézségi szintje eltért egymástól.

A második feladatban szereplő választhatóság jól érvényesült, közel azonos arányban választották a vizsgázók mindkét témát, amelyek nehézsége sem összességében, sem az értékelési szempontok szerint nem különbözött szignifikánsan.

4. Milyen kapcsolat van az értékelési szempontok között, ezek mennyiben határozzák meg az összpontszámot és milyen nehézségi szintekbe sorolhatók?

A tanárok mindkét feladatban legszigorúbban a nyelvtan, helyesírás szempontot értékelték. Megállapítottuk, hogy a két feladat az íráskészség más dimenzióit méri. A második feladatban a formai jegyek és hangnem szempontnak alacsony a ma-

gyarázóereje, nem megfelelő a modellilleszkedése, jelentősen elkülönül a többi szemponttól. A vizsgafejlesztés folyamatában e szempontnak egyrészt az volt a funkciója, hogy segítse a vizsgázót abban, hogy a formai jegyek ismeretével elérhessen néhány pontot, másrészt a vizsgakoncepció kialakításakor még a levél formai jegyeinek jelentős szerepe volt a német nyelvoktatásban. Azóta azonban a kézi levélírást felváltotta az e-mail, amelyben ezek a formai szabályszerűségek kevésbé érvényesülnek. Az eredmények tükrében javasoljuk e szempont törlését vagy esetleg az értékelési skála módosítását.

5. Mennyire fedik le az értékelési skálapontok a vizsgázók íráskészség-szintjei alapján meghatározott intervallumot?

Az értékelési skálapontok küszöbértékei a képességszint növekedésével többnyire arányosan emelkedtek, ugyanakkor azonosítottunk egy magasabb képességszintet, amelyet az egyes skálapontok küszöbértékei nem fednek le. A vizsgálatunk következtetései összhangban vannak korábbi elemzéseinkkel (Vígh, 2010; 2012). E két vizsgálat megerősíti, hogy az értékelési szempontokhoz tartozó alacsonyabb pontszámok inflálódtak, vagyis az alacsonyabb szintleírás kevésbé jellemző az érettségizőkre. Ennek oka, hogy a feladatokat és az értékelési rendszert a diákok a vizsgára való felkészülés során megismerik, ami jelentősen befolyásolja a teljesítményüket. Mindezek alátámasztják, hogy szükséges a vizsgafeladatok és az értékelési eljárások nehézségi szintjének a Közös Európai Referenciakerethez (2002) való illeszkedését vizsgálni és ezek alapján – szükség szerint – a vizsgakövetelményeket módosítani. Ehhez is alkalmazhatók az elemzésünkben használt valószínűségi tesztelméleti modellek.

6. Mennyire jól rendezettek az egyes értékelési szempontokhoz tartozó pontszámok az értékelési skálákban?

Az értékelési skálahasználatot karakterisztikus görbékkel jellemeztük, amelyek minőségi mutatói a delta-paraméterek. Ezek felcserélődése és egymástól való távolóságuk jellemzik az értékelési skálák redundáns működését (például a nyelvi szempontoknál és a szövegalkotásnál), valamint azt, hogy az értékelési skála bizonyos pontszámait az értékelők nem használják (például a tartalom szempontnál). Ezek az eredmények szintén összhangban vannak a 2006. évi német nyelv emelt szintű érettségi eredményeivel (Vígh, 2010; 2012). Elemzésünk azt mutatta, hogy

- a. a tartalom értékelési szempontnál a köztes deskriptorokban (4–3 és 2–1) az alacsonyabb pont nem tölti be a funkcióját, az értékelők ezt nem alkalmazzák. Vélhetően jobban működne ez az értékelési skála, ha – például az angolhoz hasonlóan – 0–3 pontos lenne.
- b. a második feladatban a szövegalkotás értékelési skálája az első feladatban jobban működött, mint a másodikban. Vagyis célszerű ezt az első feladathoz hasonlóvá alakítani.

- c. a szókincs, kifejezésmód és a nyelvtan, helyesírás szempontnál a 2–1 pontos deskriptor nem működik, vagyis itt az értékelési skála leírásán célszerű változtatni.
- d. jobban működik úgy az értékelési skála, ha egy szintleíráshoz csak egy pont tartozik.

A kutatási kérdésekhez megfogalmazott hipotézisek közül az alábbiak igazolódtak be: A német nyelvi érettségi működésének jellemzői, mérési-értékelési eljárásai és a korábbi kutatási eredmények alapján feltételeztük és elemzésünk során bebizonyosodott, hogy

1. az íráskészség vizsgarész magas megbízhatósággal mér, amely a mérendő konstruktum jellemzőitől függ: tehát az értékelési skálák használata magas reliabilitást eredményez, valamint heterogén a csoport összetétele;
2. a két feladat súlya a tesztben eltér, amely ebben az esetben abból adódik, hogy az azonos nyelvi szintre tervezett két feladat közül az első könnyebb volt a másodikonál.

Az értékelési rendszer megfelelőségét korábbi kutatásunkban (Vígh, 2010; 2012) kialakított minőségi mutatók mentén elemeztük a parciáliskredit-moddal. Ebben a 2006. évi emelt szintű német nyelvi érettségi elemzése (Vígh, 2012) alapján megfogalmazott kutatási eredmények alapján feltételeztük és jelenlegi elemzésünk alapján is bebizonyosodott, hogy

3. a vizsgázók teljesítménye eltér a két feladatban, de a választható témák megoldottságában nincs jelentős különbség;
4. az értékelési szempontok magyarázóereje és nehézsége eltér, amely függ attól, hogy az íráskészség mely dimenzióját mérik;
5. az értékelési skálapontok nem fedik le teljesen a vizsgázók íráskészség-szintjei által meghatározott intervallumot: az alacsonyabb pontszámok kevésbé jellemzők az érettségizőkre, ezért indokolt a további differenciálás;
6. az egyes értékelési skálák működése eltér szempontonként és a skálákban definiált pontszámok alapján, ezért célszerű az értékelési skálák felülvizsgálata.

Irodalom

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakovetelmenyek2012/nemet nyelv_vk.pdf,
http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakovetelmenyek2012/nemet nyelv_vl.pdf (letöltve: 2013. május 20.)

MASTERS, G. N. (1982): A Rasch model for partial credit scoring. *Psychometrika* 47/2, **149-174.**

Német nyelv. *Emelt szintű írásbeli érettségi vizsga. 2012. május 11.* (2012). Budapest: Nemzeti Erőforrás Minisztérium.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/emelt/e_nemet_12maj_fl.pdf (letöltve: 2013. június 27.)

Német nyelv. *Emelt szintű írásbeli érettségi vizsga. Javítási és értékelési útmutató. 2012. május 11.* (2012). Budapest: Nemzeti Erőforrás Minisztérium.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/emelt/e_nemet_12maj_ut.pdf (letöltve: 2013. június 27.)

Közös *Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés* (2002): Európa Tanács Közoktatási Bizottsága. Közoktatási Bizottság. Élő Nyelvek Osztálya. Strasbourg. Piliscsaba: Pedagógus-továbbképzési Módszertani és Információs Központ Kht.

MOLNÁR GYÖNGYVÉR (2013): *A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában.* Budapest: Gondolat Kiadó.

VÍGH TIBOR (2010): *Az idegen nyelvi érettségi működése és hatása a tanulói teljesítmények és a tanári nézetek tükrében.* PhD-értekezés. Szeged: Szegedi Tudományegyetem, Bölcsészettudományi Kar, Neveléstudományi Doktori Iskola.

www.edu.u-szeged.hu/phd/downloads/vigh_ertekezés.pdf (letöltve: 2013. május 24.)

VÍGH TIBOR (2012): Qualitätskriterien in der Messung und Bewertung von Schreibfertigkeit: Die Ergebnisse einer Analyse mit dem Partial-Credit Modell. In: Arntz, R., Krings, H. P. és Kühn, B. (szerk.): *Autonomie und Assessment. Erträge des 3. Bremer Symposiums zum autonomen Fremdsprachenlernen.* Bochum: AKS-Verlag, **71–83.**

WU, M., ADAMS, R. J. és WILSON, M. R. (1998): *ACER ConQuest. Generalised Item Response Modelling Software.* Australia: ACER Press.

PETNEKI KATALIN, VÍGH TIBOR

**A KÖZÉP- ÉS EMELT SZINTŰ NÉMET NYELV
ÉRETTSÉGI MÉRÉSI-ÉRTÉKELÉSI RENDSZERÉNEK
MÓDOSÍTÁSÁRA VONATKOZÓ JAVASLATOK**

1. Mi az, ami bevált, ami a vizsga továbbfejlesztésekor ne változzon

A német nyelvi érettségi az elmúlt nyolc évben általában véve megfelelően működött, a bevezetést követően az első három év eredményeit számos tanulmányban elemezték (Einhorn 2006, 2007, 2008a; Einhorn és Major 2006; Vígh, 2010, 2012). Az Oktatási Hivatal évről-évre közzétett érettségi gyorsjelentései is többnyire pozitív visszajelzéseket adtak a német nyelvi érettségi vizsgákról. Itt volt például a benyújtott felülvizsgálati kérelmek száma a legalacsonyabb.

A német nyelv érettségi vizsga középszinten lezáró típusú, emelt szinten lezáró és egyben tudásszintmérő vizsga. A lezáró jellegből adódóan a vizsgának azt kell mérnie, amit a Nemzeti Alaptanterv (2007, 2012) követelményrendszere előír. A készségek különböző szintjét az Európa Tanács dokumentuma, a Közös Európai Referenciakeret (2002) írja le, a Nemzeti Alaptanterv élő idegen nyelv része erre épül, ezért a vizsgának is azt kell mérnie, hogy a tanuló elérte-e az első, ill. második nyelvből a különböző nyelvi készségek terén az előírt nyelvi szintet vagy sem. Mivel a vizsga a kommunikatív készségeket céltan méri, elkerülhetetlen, hogy a tanórán e készségek fejlesztése álljon a középpontban. A vizsga szemlélete jelentős visszahatással bír a napi oktatási gyakorlatra, szükségessé teszi az áttérést a kommunikatív szemléletű nyelvoktatásra.

A nyelvi szintek mérésénél már a vizsgamodell kialakításakor is meglehetősen nagy problémát okozott, hogy a középszinten a vizsga két nyelvi szintet mér: A2 és B1 szintet. Ezért többször felmerült, hogy a középszint is – az emelt szinthez hasonlóan – csak egy szintet mérjen. Adataink azonban azt mutatták, hogy ennek még nem jött el az ideje. A minta-itemtérképek középszinten azt mutatják, hogy ugyan sokan már vélhetően a B1 szinten vannak, de még elég nagy számban vannak olyanok is, akik ezt nem tudnák teljesíteni. Az adatelemzésünk kizárólag a nappali tagozatosokra vonatkozott, a felnőttoktatásban érettségizőkre nem, és e vizsgázói réteg jellemzően alacsonyabb, A2 szintű nyelvtudással rendelkezik. Eredményeink szerint még a nappali tagozatosok körében is leginkább az íráskészség esetében lenne problémás, ha csak a magasabb szintet mérné a vizsga. Célul ki lehet tűzni, hogy ez a későbbiekben változzon, de ehhez folyamatosan kell monitorozni a képességszintek alakulását.

A vizsga bevezetését hosszú előkészítő munka előzte meg. Ennek eredményeképpen egy alapos és részletes vizsgadokumentáció jött létre, ami elősegítette a tanárok átállását az egységes szemléletre. Ez a dokumentáció erősítette a vizsga megbízható lebonyolítását. A részletesen kidolgozott Javítási és értékelési útmutató fontos támpont mind az értékelő tanárok, mind a vizsgázó diákok számára.

Javasoljuk, hogy

- a) az érettségi vizsga továbbra is a kommunikatív szemléletre épüljön, alapjaiban, szemléletében lényegesen ne változzon.
- b) egyelőre ne változzon a középszint belső kétszintűsége.
- c) továbbra is legyen részletesen kidolgozott vizsgadokumentáció. Az esetleges változtatások végrehajtásánál figyelni kell a vizsgamodell koncepciójára.

2. Mi az, amit a feladatokkal és feladattípusokkal kapcsolatban érdemes újragondolni

Ebben a részben először áttekintést adunk a közép- és emelt szintű receptív készségek (olvasott és hallott szöveg értése), valamint a nyelvhelyesség vizsgarészekben szereplő feladatok és itemek működéséről. Ezt követően mindkét vizsgaszinten a feladattípusok beválását elemezzük és értékeljük. Végül az íráskészség vizsgarészben alkalmazott feladatok működését vizsgáljuk, illetve ezek alapján teszünk javaslatokat a vizsga továbbfejlesztésére.

A 0.1. táblázatban azt összegeztük, hogy mely itemekkel volt az egyes vizsgarészekben probléma. Mellékletben kigyűjtöttük a II. és IV. dokumentumból azokat az itemeket, amelyek elhagyása javítaná az adott vizsgarész működését. Ezen itemek törlésének következményeit adatok hiányában részletesen nem vizsgáltuk (mivel nem álltak rendelkezésünkre az ilyen típusú korigált adatok). Ugyanakkor meg kell jegyeznünk, hogy ennek az utólagos korrekciónak a 2012. májusi vizsga eredményeire nem lenne hatása. Ráadásul a megadott itemek törlésével közép- és emelt szinten az olvasott és a hallott szöveg értése vizsgarész/vizsgarész-összetevő lecsökkentett itemszáma nem felelne meg a vizsgaleírásban megadott elvárásoknak (40/2002. OM rendelet). Ezért a 0.1. táblázatban lévő információkat úgy kell értelmezni, hogy ha mód lenne a vizsgafeladatok előzetes kipróbálására, akkor mely itemek pótlása javítaná az egyes vizsgarészek működését. Mindezek mellett, az egyes feladatok és feladattípusok beválásának vizsgálatához, illetve a javaslatunk alátámasztásához a táblázatban lévő összegzés kíván kiindulópontot nyújtani³⁶.

³⁶ Lásd még: Melléklet

0.1. táblázat:

A közép- és emelt szintű feladatsorokban szereplő problémás itemek sorszáma és száma

Probléma	Középszint	Emelt szint	Összesen
1. Túl könnyű, súlytalan itemek	Olvasás: 4., 5., 9., 10. Nyelv ³⁷ : 1., 3., 5. Hallás: 2., 6., 15.	Nyelvh.: 6., 10., 12., 17., 18., 20., 21., 26., 28., 30., 33.	21 item
2. Túl nehéz itemek	Olvasás: 19., 22. Nyelv.: 8.	Nyelvh.: 7., 24.	5 item
3. Nem megfelelő modellilleszkedésű, súlytalan itemek	Olvasás: 27., 29., 32. Nyelv.: 11., 14., 15.	Olvasás: 12., 15., 16., 20., 21., 27. Nyelv.: 13.	12 item
4. Elfogadhatatlan reliabilitású feladatban szereplő itemek	Hallás: 16–21.	Hallás: 1–9.	15 item
Összesen	25 item (31%)	29 item (31%)	53 item

Mindkét szinten az itemek 31 százalékával volt valamilyen probléma. Ezekből négy csoportot alakítottunk ki az adott itemben megjelenő elsődleges probléma szerint. (1) A problémás itemek 40%-a mindkét szinten abból adódott, hogy ezek túlságosan könnyűek voltak a vizsgázók számára és ezeknél magas volt a megoldottság, az itemek differenciáló ereje kisebb volt, kevésbé járultak hozzá az összpontszám alakításához, illetve a nehézségi paramétereik negatív értékűek voltak, vagyis az átlag alatti képességtartományban helyezkedtek el. (2) Összesen 5 olyan itemet találtunk, amelyek túl nehezek voltak a vizsgázók számára. A 0.1. táblázatban a középszinten lévő három item megoldottsága, jellemzői és nehézségi paramétere alapján inkább már a B2 szintet mérte; míg emelt szinten ezek az itemek nyílt végű feladatokban szerepeltek, és működésük ellenőrzéséhez szükség lett volna a tanulói válaszok elemzésére, amelyre adatok hiányában nem került sor. E csoportba tartozó itemek mind testelméleti, mind felhasználói szempontból is problémásak, hiszen a tanulói, a javítói és a szülői észrevételek ezekre az itemekre vonatkozóan érkezhettek. Ugyanakkor mindkét vizsgaszinten ezen itemek száma annyira alacsony, hogy ezek elhagyása a vizsgarészekből érdemben nem befolyásolná sem az egyes tanulók teljesítményének megállapítását, sem az adott vizsgarész megbízhatóságát. A 3. csoportba sorolt itemek modellilleszkedése nem volt megfelelő, vagyis ezek mást is mértek, mint a teszt egésze. Ezek az itemek egy kivételével (középszint, nyelvhelyesség 15. item) zárt végű feladatokban szerepeltek, tipikusan alternatív választást igénylő, ún. igaz-hamis feladatokban, kisebb részt hozzárendelés és feleletválasztás feladatokban. E dokumentum további részében ezeket a feladattípusokat részletesebben is elemezzük. (4) Két hallott szöveg értése feladatban volt elfogadhatatlan a Cronbach-alfa értéke (0.18. táblázat; 0.20. táblázat). Ezek törlése javítaná ugyan mindkét esetben a vizsgarész megbízhatóságát, de az alacsony itemszám miatt a feladat törlésének következményeként

³⁷ Nyelv. = Nyelvhelyesség

egyik vizsgarész sem felelne meg a vizsgaleírásban szereplő elvárásoknak (40/2002. OM rendelet), vagyis ezeket a feladatokat javítani kell, ezért a módosításnak az egyes tanulók teljesítményére és az adott vizsgarész teljesítménystruktúrára gyakorolt hatását kipróbálás hiányában nem tudjuk elemezni. Ezért szükséges e feladattípus javítása, amelyre e fejezet további részében teszünk javaslatot.

A 2012. májusi vizsga részletes elemzése megmutatta, hogy vannak feladattípusok, amelyek megbízhatóan működnek, de van néhány olyan típus is, ahol komolyabb problémát észleltünk. A továbbiakban először áttekintést adunk a vizsgált érettségiben felhasznált feladattípusokról (0.2. táblázat). Itt már nem differenciáltunk az egyes típusokon belül, mint az elemzésben, tehát pl. egy szöveg kiegészítése éppúgy lehet olvasott és hallott szöveg értését mérő szövegrekonstrukció, ahol a szövegből kivett mondatokat kell a koherencia helyreállításával a lyukba visszahelyezni, vagy lexikai, ill. nyelvtani kompetenciát mérő feladat.

0.2. táblázat:
Áttekintés a közép- és emelt szintű vizsgán előfordult feladattípusokról

Feladattípus	Közép olvasás	Emelt olvasás	Közép Nyelvh	Emelt Nyelvh	Közép hallás	Emelt hallás	Összesen
1. hozzárendelés		2	1	1			4
2. szövegkiegészítés			1	2			3
3. feleltválasztás			2				2
4. igaz-hamis	1	1			1		3
5. ál igaz-hamis (szeklektálás)	1				1	1	3
6. sorbarendezés		1					1
7. kérdésre válasz	1				1		2
8. mondatkiegészítés	1					1	2
9. hibaaazonosítás				1			1

1. A hozzárendelés feladattípus különböző módon működött, attól függően, hogy mit kellett mihez hozzárendelni. A legjobban működő hozzárendeléses feladat a középszint nyelvhelyesség (2. vizsgarész) 4. feladata volt. Itt a szövegben az összetett mondatok tagmondatait kellett megtalálni a kötőszavak megfelelő ismerete alapján. Az emelt szint 2. vizsgarészének a 2. feladatában egy autentikus szövegbe kellett beírni a megadott szavakat, de ezeken nem kellett változtatni, (ragozni). Ez az emelt szinten túl könnyűnek bizonyult. Itt tehát nem a feladattípus okozta a problémát, hanem a nyelvi szint meghatározása.

Az emelt szinten az olvasás 2. feladatában egy interjú válaszaihoz kellett a megfelelő kérdést hozzárendelni. A kérdések száma ugyan eggyel több volt, mint ahány válasz volt az interjúban, de végeredményképpen ez is egyfajta szövegrekonstrukció volt, amibe a feladatírók nem nagyon nyúlhattak bele. Ez a feladat a vizsgán jól működött. Az emelt olvasás utolsó feladata is szövegrekonstrukció, ahol a hiányzó mondatok helyét kellett

beazonosítani, de eggyel több mondat volt, mint amennyi hiányzott, ezért szerepel a hozzárendelés típusban. Ebben a feladatban az utolsó, a 27. item volt a legkönnyebb. Erre a legvalószínűbb magyarázat pedig az lehet, hogy a feladat végére csak a disztraktor és egy mondat maradt, a disztraktor pedig nagyon eltért attól, ami abba a szövegrészbe kell, tehát könnyen megkülönböztethető volt.

Kijelenthetjük, hogy ez a feladattípus megvalósítja a mérési célt, továbbiakban is érdemes alkalmazni mindkét szinten, de ügyelni kell a nyelvi szint megállapítására és a disztraktor megfogalmazására. Azt, hogy megfelelő nehézségű-e a feladat és jó-e a disztraktor, leginkább úgy lehet megállapítani, ha az éles alkalmazás előtt mód van a feladat kipróbálására, és a kipróbálás elemzésére, kiértékelésére és a feladat módosítására.

2. A szövegkiegészítés feladattípus az olvasott és hallott szövegértés mérésére nem fordult elő, hanem egyszer a középszint 2. vizsgarész-összetevőjében, és kétszer az emelt szint nyelvhelyesség vizsgarészében. A nyelvhelyesség feladatai mind szöveg alapúak. A középszint 2. vizsgarész-összetevő 3. feladatánál a megadott igéket kellett a megfelelő alakban beírni. Itt leginkább az egyes igékkel kapcsolatos ragozások eltérő nehézségi szintjei okoztak problémát. Erre a feladatkészítésnél jobban kell ügyelni. Az emelt szinten ez utóbbi feladatra hasonlított a nyelvhelyesség vizsgarész 4. feladata, csak itt nem igéket kellett ragozva beírni, hanem a megadott szavakat egy másik szófajban (és többnyire ragozva). Ez is a megbízhatóan működő feladatok közé tartozott. A harmadik szövegkiegészítés a 2. vizsgarész 1. feladata volt. Ebben az esetben a megadott szavakat kellett megfelelő ragozott alakban a lyukakba írni és ebben a feladatban volt e vizsgarész egyik legkönnyebb (6.) és legnehezebb (7.) iteme.

Kijelenthetjük, hogy bevált ez a feladattípus a nyelvhelyesség mérésére. Ugyanakkor ügyelni kell az egyes feladatokba kerülő itemek nyelvi szintjének meghatározására.

3. A középszint nyelvhelyesség vizsgarész-összetevőben kétszer is előforduló feleletválasztó feladatnál is az előbbiekhöz hasonló a probléma: nem a feladattípussal van baj, hanem a szint és a disztraktorok meghatározásával. Ennek is az lenne a megoldása, hogy előzőleg kipróbáljuk a feladatot, majd az eredményt elemezve javítjuk.

4. Az ún. igaz-hamis feladattípus mindkét szinten az olvasott és a hallott szöveg értéke vizsgarészben/vizsgarész-összetevőben fordult elő. Ennek a feladattípusnak alapvető problémája, hogy 50%-ban találgathat a vizsgázó. Elemzésünk igazolta, hogy ennél a típusnál szinte minden itemnél probléma van. A középszinten az 1. vizsgarész-összetevő utolsó feladatában alkalmazták ezt a típust. Mindegyik itemnek nagyon alacsony volt a magyarázóereje, és több itemnek sem megfelelő a modellilleszkedése. Kiugróan problémás volt emelt szinten az igaz-hamis típusú 3. feladat, ahol a kijelentéseket a feladatírók készítik. Feltehető, hogy ez a típus az emelt szinten nem igazán mér, hiszen reliabilitása alacsony. Célszerű lenne tehát emelt szinten helyette más feladattípust alkalmazni.

E feladattípus működése nagyban függ attól, hogy mennyire sikerül a feladathoz tartozó igaz és hamis kijelentéseket jól megfogalmazni. Ezt a feladatíró igazán nem tudja másképp ellenőrizni, csak ha lehetőség van a feladat kipróbálására és korrigálására.

5. A 3. vizsgarészben/vizsgarész-összetevőben (hallott szöveg értése) előfordult mindkét szinten olyan feladat, amelyben a vizsgázónak a megfelelő kijelentést kell kiválasztania. Ez tulajdonképpen egy ál igaz-hamis (szelektálás) feladat. Ez a feladattípus a hallott szöveg értése vizsgarészben/vizsgarész-összetevőben se középszinten, se emelt szinten nem működött. Középszinten a szöveghez 12 kijelentés tartozik. A vizsgázónak el kell döntenie, hogy ezek közül valójában mely kijelentésekről beszélgettek a rádióműsorban.

0. wo die Herder-Buchhandlung liegt.
 wie groß die Buchhandlung ist.
16. wer die Buchhandlung leitet.
 warum Monika die Deutschen mag.
 wie viele Deutsche Monika kennt.
17. was für eine Meinung die Deutschen von Italien haben.
18. was für eine Meinung die Italiener von den Deutschen haben.
19. wie die Italiener das deutsche Essen finden.
 wie Tatjana ihren Mann kennen gelernt hat.
 in welchem Jahr Tatjana nach Italien gezogen ist.
20. was man an Italien schön findet.
21. warum Tatjana nach Deutschland zurück möchte.
 wie die Arbeitsbedingungen in Deutschland sind.

0.1. ábra:

A középszint hallott szövegértés 3. feladatának megoldása
 (Német nyelv. Középszintű írásbeli vizsga. Javítási-értékelési útmutató, 2012: 8)

Összesen 6 kijelentés felel meg a feladatkiírásnak, de ehhez tartozik még 6 másik kijelentés is, így a valódi igaz-hamis feladatokhoz hasonlóan itt is kb. fele-fele arányban voltak az igaz és hamis kijelentések, ám a hamis kijelentéseket nem értékeli a vizsga, nem jár értük feladatpont. Emelt szinten is hasonló volt a feladat. Mindkét esetben a Cronbach-alfa érték arra utal, hogy ez a feladat nem megbízható, az itemparaméterek is igen gyenge értékeket mutatnak.

A középszinten az olvasásnál is előfordult egy ilyen feladat. Itt meg kellett találni, hogy melyik kijelentés melyik városra illik. A problémát az okozta, hogy a feladatban több hely volt a feladatlapon, mint ahány ikszet a vizsgázónak be kellett írni. A három városhoz összesen kilenc információt lehetett rendelni. A Javítási és értékelési útmutató szerinti megoldás nem értékeli azt a teljesítményt, hogy hova nem kell írni semmit se.

0.3. táblázat:

A középszint 1. vizsgarész-összetevő 1. feladatának megoldó kulcsa
 (Német nyelv. Középszintű írásbeli vizsga. Javítási-értékelési útmutató, 2012: 8)

Venedig	Rio de	New
---------	--------	-----

		Janeiro	Orleans
Dieser Karneval ist der berühmteste in Europa.	(0)X		
Die Leute ziehen zum Karneval interessante Kostüme oder Masken an.	(1)X	(3)X	
Die Masken dieses Karnevals sind Karikaturen der regionalen Besonderheiten.	(2)X		
Hier sind Tanzproduktionen typisch.		(4)X	
Dieser Karneval findet in einem bestimmten Teil der Stadt statt.			(7)X
Hauptattraktion des Karnevals ist ein typischer Tanz oder eine typische Musik.		(5)X	(8)X
Typisch für diesen Karneval ist der Umzug durch die Straßen.		(6)X	(9)X

Ezt a feladatot úgy lehetne javítani, hogy pl. elhagyni azokat az itemeket a hozzá tartozó kijelentéssel együtt, amelyek csak egy városra igazak. A többinél pedig jelölni kellene (pl. áthúzással) azokat a helyeket, amelyek nem illenek az adott városra a szöveg szerint.

Úgy gondoljuk, hogy ezt a feladattípust ki kellene hagyni a feladatok tárházából. Helyette jól működhetne egy három elemből álló feleletválasztó feladat, de ilyen a receptív készségeket mérő feladatsorokban egyszer sem szerepelt.

6. Sorbarendezés csak az emelt szintű olvasott szövegértés esetében fordult elő egyszer. A kronológiai sorrend alapján rekonstruálható feladattípus megbízható volt, ennek voltak a legjobb mutatói. Nehézsége okán jogos, hogy ez a típus csak az emelt szinten fordult elő. E feladattípus veszélye ugyanakkor, hogy az itemek összefüggnek és hasonló nehézségi szintűek.

7. Mindössze két feladatnál fordult elő, hogy a feladat abban állt, hogy kérdésekre kellett rövid választ adni, mindkettő középszinten. Az 1. vizsgarész-összetevő 2. feladata és a 3. vizsgarész-összetevő 1. feladata képviselte ezt a feladattípust. Az olvasás 2. feladatában adatelemzésünk szerint az itemek átlaga és nehézségi paramétere alapján vélhetően a kilenc itemből hat volt B1 szinten. Ezekre az itemekre az jellemző, hogy a vizsgázóknak összefüggéseket kellett a szöveg alapján megadniuk, míg a könnyebb itemeknél az adott információt kellett a szövegből visszakeresni és többnyire módosítás nélkül válaszként megadni. Ezeknek az itemeknek a magyarázóereje alacsonyabb volt. A hallás 1. feladata nagyon könnyű volt, egyszerű információkat kellett lejegyezni. Egy item volt nehezebb, ahol a válasz nem pont abban a formában hangzott el, ahogy azt le kellett írni. Ennél a feladatnál tehát nem mindegy, hogy egyszerű információt kell válaszként megadni, vagy a szövegből kell a válaszra következtetni, ami sokkal nehezebb.

8. A mondatkiegészítés feladattípusban az elolvasott vagy meghallgatott szöveg tartalmát nem összefüggő szövegben, hanem egy-egy különálló mondatban fogalmazták meg, és az ezekből hiányzó lényeges információt kell a szöveg megértése alapján a lyukba beírni. Ilyen volt a középszint olvasás 3. feladata és az emelt szint hallott szövegértés

2. feladata. A középszint olvasás 3. feladata volt annak a vizsgarész-összetevőnek a legnehezebb feladata, míg emelt szinten a hallott szövegértést jól mérte a feladat – jóllehet, problémás itemek mindegyikben előfordultak, amelyek főleg abból adódtak, hogy az adott mondatba beírandó szó esetén milyen mértékben kell módosítani a vizsgázónak a választ.

A 7. és a 8. feladattípusba tartozó feladatok és az itemek nehézségi szintjét a feladatíró kétféle szempontból tudja befolyásolni. Egyrészt aszerint, hogy milyen szövegértési műveletet mér az adott item (információ visszakeresést vagy következtetést). Másrészt pedig az adott item nehézségét alapvetően az is befolyásolja, hogy a kérdésre a választ mennyire kell a feltett kérdéshez igazítani a tanulónak (ki kell másolni a szövegből vagy nyelvileg kell rajta változtatni).

9. A hibazonosítás nevű feladattípus kizárólag az emelt szint nyelvhelyesség vizsgarészben fordul elő. A feladat mérési célja az, hogy felismeri-e a vizsgázó, hogy az adott mondat (ill. az adott sorban található szövegrész) helyes-e vagy sem. A feladatban jelölni kell mind a helyes sorokat, mind a hibásakat – a hibásoknál a hibát kell az adott helyre kiírni. Mindegyik döntésre jár feladatpont. A feladat utasítása pedig pontosan megadja, hogy a szövegben hány helyes mondat található. Adatelemzésünkben azt láttuk, hogy azoknak az itemeknek, amelyek a hibátlan részt jelölik, alacsony a pontbiszeralis korrelációja és a magyarázó ereje, így nem jól mérnek. Ugyanakkor azok az itemek, amelyekben a vizsgázónak a hibát kellett megtalálni, jobban működtek. A feladat meglehetősen életidegen, ilyen tudásra leginkább csak a bölcsész- ill. tanárszakra készülőknek lehet szüksége. Ezért el kell gondolkodni azon, hogy ez a feladattípus megfelelő-e a B2 szintű nyelvhelyességi képesség általános mérésére.

Végül, e rész zárásaként az íráskészség feladatok bevalását elemezzük. Mind közép-, mind emelt szinten azonos e vizsgarész/vizsgarész-összetevő felépítése: két feladatból áll, ahol a másodikban a vizsgázók két téma közül választhatnak. Mindkét feladat mindkét szinten eltérő nyelvi szinten volt. Ez a középszinten megfelel az elvárásoknak, míg emelt szinten nem, hiszen ott mindkét feladatnak azonos nyelvi szinten kellett volna lennie (40/2000 OM rendelet) azonban, az első feladat jóval könnyebb volt. Ugyanakkor a második feladatban a választhatóság mindkét szinten jól érvényesült, a feladatok nehézségük szerint ekvivalensnek tekinthetők. A feladatok minőségét, az irányító szempontok működését, esetleges problémáit a tanulói vizsgadolgozatok kvalitatív elemzésével lehetett volna elemezni, amelyre ugyanakkor az adatfelvétel nem terjedt ki. Középszinten mindkét feladatban és témánál magas volt azoknak az érettségizőknek a száma, akik a tartalom szempontnál 0 pontot értek el, vagyis az alapvető kommunikációs célt is eltévesztették. Elemzésünkéből adatok hiányában nem derült ki, hogy ez mennyiben köszönhető a feladat esetleges problémájának.

Javasoljuk, hogy

- a) a feladatírók ne feladatsorokat készítsenek, hanem külön-külön önálló feladatokat, mert így jobban lehet a feladat itemeinek minőségére figyelni,
- b) az itemek jobb minősége érdekében legyen lehetőség a feladatok kismintás kipróbálására, elemzésére és javítására,
- c) azok a feladattípusok, amelyek az elemzés során komoly problémákat mutatnak, kerüljenek ki a vizsga feladattípusai közül,
- d) legyen lehetőség a vizsgafeladatok folyamatos elemzésére és kiértékelésére (amibe be lehetne vonni pl. egyetemi műhelyeket),
- e) legyen lehetőség a zárt végű vizsgafeladatokban alkalmazott disztraktorok elemzésére,
- f) az íráskészség vizsgarészhez legyen lehetőség egy kismintán a vizsgafeladatok működésének elemzésére,
- g) az íráskészség vizsgarészben a második feladatban mindkét szinten maradjon meg a választhatóság alapelve, ugyanakkor elemezni kell, hogy a feladatok mennyiben tekinthetők ekvivalensnek,
- h) legyen lehetőség a szóbeli vizsgafeladatok minőségének vizsgálatára is.

3. Mi az, amit az értékelési rendszerben újra kellene gondolni

A zárt végű feladatok értékelése többnyire egyértelmű és egyszerű. Minden helyes megoldásra 1 pont jár, ha a megoldás nem egyezik meg a javítási és értékelési útmutató megoldásával, azaz hibás, akkor 0 pontos. Ám van ez alól kivétel, mint például a középszint 1. vizsgarész-összetevő 1 feladata (lásd előző pontban). E feladat esetében pontlevonás jár abban az esetben, ha a vizsgázó az előre megadott számú x helyett többet tesz. Ez a szankcionálás nem volt ugyan jellemző, a vizsgált mintában 13 tanulónál (1,3%) egy, 2 tanulónál (0,2%) két pont levonás történt. Hasonló probléma merül fel mindkét szint 3. vizsgarészében/vizsgarész-összetevőben használt ál-igaz-hamis feladatok esetében. A feladat javítási útmutatója ugyanis azt az utasítást adja, hogy ha a vizsgázó a megadottnál több mondatot jelölt be, akkor ahánnyal többet jelölt, annyi pontot kell levonni megszerzett pontjaiból. Ez az eljárás azért problémás, mert pontot csak a jó válaszokból lehet levonni. Így ez büntető funkciót kap, s a mérő funkciót rontja.

A nyílt végű, pl. kérdésekre rövid választ elváró feladatnál (középszint, 1. vizsgarész-összetevő, 3. feladat) vagy a lyukas mondat kiegészítése (emelt, 3. vizsgarész 2. feladat) volt sok jó, de sok hibás megoldás is, ahol az itemek kifejezetten nehéznek mutatkoztak. A nehézség okát csak a tanulói válaszok alapján lehetne megállapítani. Feltételezhető azonban, hogy a diákok nem a javítókulcsban szereplő szavakat írták, hanem a szövegből származó kifejezést, amely azonban az adott mondatba nem illik, és amelynek elfogadhatóságáról a javítókulcs így tájékoztatja a tanárokat: „... elfogadhatók azok a variánsok is, amelyek tartalmilag megegyeznek ezekkel. A megoldásokban előforduló nyelvi vagy helyesírási hibákat figyelmen kívül kell hagyni, amennyiben a válasz érthető” (Német nyelv. Középszintű írásbeli vizsga. Javítási-értékelési útmutató, 2012: 3). Azt azonban nem tudjuk, hogy az értékelő tanárok ezt hogyan alkalmazták az értékeléskor. Az item nehézsége alapján ugyanakkor azt feltételezzük, hogy ezt nem vették figyelembe.

Az íráskészség vizsgarész/vizsgarész-összetevő értékelési rendszerének megfelelőségét mindkét szinten külön elemeztük (III. és V. dokumentum). Az értékelési szempontok mindkét szinten általában véve megfelelően működtek és alkalmazásuk magas megbízhatóságot eredményezett. Megállapítottuk, hogy a vizsgáztatók a nyelvhelyességet mérő szempontokat szigorúan értékelték, középszinten jellemzően ezek alapján ítélték meg a tanulók írásműveit, az A2 szintet mérni kívánó első feladatban vélhetően nem differenciálnak a nyelvi szintnek megfelelően. A középszint második feladatában és az emelt szint első feladatában a formai jegyek és hangnem értékelési szempont nem működött: a többi értékelési szemponthoz képest mást mér, jelentősen elkülönül a többitől az elvárásokhoz képest, teljesítése nehezebb a vizsgázók számára.

Az értékelési skálák középszinten többnyire megfelelően működtek. Az alacsonyabb szintleírás volt problémás, de figyelembe kell vennünk, hogy a mintavétel kizárólag a nap-pali tagozatosok körére vonatkozott. Ha a felnőttoktatásban érettségizőkre is kiterjedt volna az adatfelvétel, akkor vélhetően a skálák is jobban működnek. Az emelt szintű íráskészséget mérő vizsgarészt illetően két vizsgaév (2006. és 2012.) eredményeiről áll rendelkezésünkre részletes információ. Ezekből hasonló következtetésekre jutunk: (1) a mintaitemtérképek alapján azonosítottunk egy magasabb képességszintet, amelyet az egyes értékelési skálapontok nem fednek le. (2) Általában véve az 1–2 és 0 pontot leíró deskriptorok kevésbé jellemzőek a vizsgázókra. (3) Elemzéseink a skálák redundáns működését is mutatják (pl. a nyelvi szempontoknál és a szövegalkotásnál), valamint azt, hogy a skála bizonyos pontszámait az értékelők nem használják (pl. a tartalomnál).

Az íráskészség vizsgarész további jellemzője, hogy ha a vizsgázó adott értékelési szempontnál 0 pontot ér el, akkor az összpontszám is nulla pont. Ez az ún. lenullázási szabály az élő idegen nyelvekre vonatkozóan jelentősen eltér³⁸. A középszint első feladatában kizárólag a német nyelv és a spanyol nyelv érettségire jellemző, hogy „amennyiben a dolgozat valamelyik szempont alapján 0 pontos, akkor a másik értékelési szempont pontszáma is 0 pont” (Német nyelv. Középszintű írásbeli érettségi vizsga. Javítási és értékelési útmutató, 2012: 14). A többi idegen nyelvi vizsgatárgy esetében pedig amennyiben a vizsgázó alapvető kommunikációs célját nem éri el, akkor 0 pontos a dolgozat, amely a második feladatra is érvényes. Ebben az esetben mindegyik idegen nyelvi érettségi vizsgán ugyanez a szabály érvényesül. Az emelt szint első és második feladatában az angol és az olasz érettségi kivételével – ahol szintén csak a 0 pontos tartalmi szempont esetében nullázzák le az összpontszámot –, az jellemző, hogy „amennyiben a dolgozat bármelyik értékelési szempont alapján 0 pontos, akkor a többi értékelési szempont pontszáma is 0 pont” (Német nyelv. Emelt szintű írásbeli érettségi vizsga. Javítási és értékelési útmutató, 2012: 19). Ez az eltérő gyakorlat azért problémás, mert így az azonos mérési céllal rendelkező íráskészség vizsgarész az idegen nyelvek között nem összehasonlítható. Épp ezért lenne szükséges ebből a szempontból egységesen kezelni az értékelési alapelveket.

Javasoljuk, hogy

- a. a zárt végű feladatoknál szűnjön meg a pontlevonás lehetősége, a feladatokat ennek értelmében kellene átalakítani.
- b. a nyílt végű feladatoknál legyen lehetőség a tanulói megoldások kismintás elemzésére is, mert csak így lehetne megtudni, hogy az értékelő tanárok helyesen használják-e az útmutatót, hogy milyen az értékelés minősége.

³⁸ A bekezdés további részét az angol, a német, a francia, az olasz és a spanyol érettségi íráskészség vizsgarészben alkalmazott értékelési eljárás elemzése alapján készítettük.

- c. az íráskészség vizsgarész értékelési szempontjai közül töröljék a formai jegyek és hangnem szempontot vagy dolgozzák át az értékelési skálát.
- d. az íráskészség vizsgarész értékelési skálái középszinten ne változzanak, emelt szinten viszont szükséges ezek felülvizsgálata.
- e. az írásbeli vizsgarészben alkalmazott lenullázási szabály legyen az élő idegen nyelvek esetében egységes. Amennyiben a vizsgázó nem valósítja meg az alapvető kommunikációs célt, akkor legyen kizárólag az összpontszám is 0 pont.
- f. a szóbeli vizsgálathoz készült értékelési útmutatóban szereplő szempontokat és skálákat, illetve azok alkalmazását is vizsgálják, mert csak így kaphatunk teljes képet a vizsga egészének működéséről és bevalásáról.

4. Mi az, amit az egész koncepción érdemes továbbgondolni és változtatni

A feladatkészítés koncepcióját szükséges módosítani a vizsgaminőség javítása érdekében. Elemzésünk igazolta és alátámasztotta, hogy szükséges folyamatosan önálló feladatokat és nem feladatsorokat fejleszteni, a létrejött feladatokat kismintán kipróbálni és javítani. Javasoljuk, hogy a 2005 és 2007 között zajlott feladatfejlesztési modellnek megfelelően készüljenek a vizsgafeladatok (lásd. Einhorn, 2008b: 227–231).

A vizsgafeladatok minőségének nyomon követése érdekében szükséges, hogy az összes vizsgázónak az egyes vizsgaidőszakban elért pontszámait tartalmazó publikus adatbázisai mellett legyen lehetőség a feladat- és itemszintű elemzésekre, mert csak ezzel garantálható a feladatsorok minőségének ellenőrzése, az értékelési eljárások felülvizsgálata és korrigálása. Ehhez javasoljuk a jelen projekt keretében kialakított eljárásrendnek megfelelően mindkét szinten egy kisebb, reprezentatív mintán a vizsgarészek/vizsgarész-összetevők feladatainak itemszintű rögzítését, elemzését, mert csak a vizsgaeredmények folyamatos ellenőrzése alapján érdemes változtatásokat elvégezni, és azok következményeit vizsgálni.

Időszerű lenne mindkét vizsgaszint követelményszintjének felülvizsgálata és a feladatsorok nehézségi szintjének empirikus vizsgálata, a Közös Európai Referenciakeret (2002) szintjeihez történő illesztése. Középszinten ez azért szükséges, mert a vizsga két nyelvi szintet, az A2–B1 szintet mér, míg az emelt szint a B2 nyelvi szintet. Ugyanakkor emelt szinten az látszik a minta-itemtérképek alapján, hogy minden készség esetében a vizsgarészekben alkalmazott itemek nehézségi szintje és az íráskészségben alkalmazott értékelési szempontrendszer alatta van az érettségizők készség szintjének. Ezért szükséges a feladatok szintjének hozzáillesztése a nemzetközi szintleírásokhoz, illetve a közép- és emelt szint között lévő tartalmi és mennyiségbeli különbségek empirikus úton történő meghatározása és a két szint egymáshoz illesztése. Ennek eredményeként hosszú távon egy feladatbank építhető ki, amely kipróbált és az egyes szintekhez kalibrált feladatokkal lehetővé tenné a teljesítmények összehasonlíthatóságát. Ugyanakkor ennek létrehozását megnehezíti, hogy a jelenlegi rendszerben a felhasznált aktuális érettségi vizsgafeladatok azonnal nyilvánosságra kerülnek. Szükséges tehát a szerkezeti változtatás, amelynek eredményeként azonban a teljesítményekben bekövetkező változások megbízhatóan nyomon követhetőek lehetnek.

Bár elemzésünk tantárgyspecifikus volt, szükséges lenne az érettségi vizsga továbbfejlesztése során közelebb hozni egymáshoz az élő idegen nyelvek vizsgáját, nagyobb összhangot létrehozni a feladattípusok, értékelési felfogások és eljárások között. Bár

a vizsgafejlesztés idején történtek kísérletek arra, hogy az egyforma követelményhez egyformán működő vizsga kapcsolódjon minden élő idegen nyelv esetében, ezt akkor nem sikerült végigvinni. Ismeretes, hogy több vizsgarészben is eltérnek pl. a német nyelv és angol nyelv vizsga feladattípusai és értékelési kritériumai. Ezért javasoljuk, hogy jöjjön létre egy munkabizottság, amelyik az élő idegen nyelvi érettségi témáját újra felvállalja és összhangját megteremti, vagy legalább az egységesítés útján előbbre viszi.

Irodalom

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgakövetelmények.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv vk.pdf, (letöltve: 2013. május 20.)

Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet. Német nyelv. Vizsgaleírás.

http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/vizsgakövetelmények2012/nemet nyelv vl.pdf (letöltve: 2013. május 20.)

EINHORN ÁGNES (2006): A német nyelvi vizsgafeladatok fejlesztése – a tartalmi érvényesség és az értékelhetőség ellentmondásai. In: Horváth Zsuzsanna és Lukács Judit (szerk.): *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Budapest: Országos Közoktatási Intézet, **153–171**.

<http://www.ofi.hu/tudastar/erettsegi/uj-erettsegi/nemet-nyelvi-erettsegi>

(letöltve: 2013. május 14.)

EINHORN ÁGNES (2007): A 2005. évi érettségi vizsga eredményeinek elemzése: Német nyelv. In: Einhorn Ágnes (szerk.): *A 2005-ös érettségi vizsga eredményeinek elemzése*. Budapest: Oktatókutató és Fejlesztő Intézet.

<http://www.ofi.hu/tudastar/erettsegi/2005-os-erettsegi/2005-évi-erettsegi-090617-2> (letöltve: 2013. május 24.)

EINHORN ÁGNES (2008a): *A 2006. évi érettségi vizsga eredményeinek elemzése: Német nyelv*. Budapest: Oktatókutató és Fejlesztő Intézet.

<http://www.ofi.hu/tudastar/erettsegi/2006-os-erettsegi/2006-évi-erettsegi-090617-2> (letöltve: 2013. május 24.)

EINHORN ÁGNES (2008b): Hogy vizsgázott az új érettségi németből? In: Bánkuti Zsuzsa/ Lukács Judit (szerk.): *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés*. Budapest: Oktatókutató és Fejlesztő Intézet, **207–237**.

EINHORN ÁGNES és MAJOR ÉVA (2006): Az idegen nyelvek – vizsgafejlesztés nemzetközi kontextusban. In: Horváth Zsuzsanna és Lukács Judit (szerk.): *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Budapest: Országos Közoktatási Intézet, **127–137**.

<http://www.ofi.hu/tudastar/erettsegi/uj-erettsegi/idegen-nyelvek> (letöltve: 2013. május 24.)

- GLABONIAT, MANUELA et al. (2006): *Profile deutsch*. Lernzielbestimmungen, Kannbeschreibungen und kommunikative Mittel für die Niveaustufen A1, A2, B1, B2, C1 und C2 des "Gemeinsamen europäischen Referenzrahmens für Sprachen". Buch mit CD-ROM. Klett-Langenscheidt Verlag.
- Közös *Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés* (2002): Európa Tanács Közoktatási Bizottsága. Közoktatási Bizottság. Élő Nyelvek Osztálya. Strasbourg. Piliscsaba: Pedagógus-továbbképzési Módszertani és Információs Központ Kht.
- Nemzeti alaptanterv* (2007). Budapest: Oktatási és Kulturális Minisztérium.
<http://www.nefmi.gov.hu/kozoktatas/tantervek/nemzeti-alaptanterv-nat> (letöltve: 2013. július 1.)
- Nemzeti alaptanterv* (2012). Budapest: Emberi Erőforrások Minisztériuma, Oktatásért Felelős Államtitkárság http://www.kormany.hu/download/c/c3/90000/MK_12_066_NAT.pdf (letöltve: 2013. július 1.)
- Német nyelv. Középszintű írásbeli vizsga. Javítási-értékelési útmutató. 2012. május 11* (2012). Budapest: Nemzeti Erőforrás Minisztérium.
http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/kozep/k_nemet_12maj_ut.pdf (letöltve 2013. május 20.)
- Német nyelv. Emelt szintű írásbeli érettségi vizsga. Javítási és értékelési útmutató. 2012. május 11.* (2012). Nemzeti Erőforrás Minisztérium, Budapest.
http://www.oktatas.hu/pub_bin/dload/kozoktatas/erettsegi/feladatok2012tavasz/emelt/e_nemet_12maj_ut.pdf (letöltve: 2013. június 27.)
- VÍGH TIBOR (2010): *Az idegen nyelvi érettségi működése és hatása a tanulói teljesítmények és a tanári nézetek tükrében*. PhD-értekezés. Szeged: Szegedi Tudományegyetem, Bölcsészettudományi Kar, Neveléstudományi Doktori Iskola.
www.edu.u-szeged.hu/phd/downloads/vigh_ertekezes.pdf (letöltve: 2013. május 24.)
- VÍGH TIBOR (2012): *Qualitätskriterien in der Messung und Bewertung von Schreibfertigkeit: Die Ergebnisse einer Analyse mit dem Partial-Credit Modell*. In: Arntz, R., Krings, H. P. és Kühn, B. (szerk.): *Autonomie und Assessment. Erträge des 3. Bremer Symposiums zum autonomen Fremdsprachenlernen*. Bochum: AKS-Verlag, **71–83**.

Melléklet

A táblázatokban pirossal jelöltük, hogy az adott itemnél mi a probléma. A „Feltételezhető szintet” az olvasott és a hallott szöveg értése feladatsorokban a Közös Európai Referenciakeret (2002) szintleírásai, a vizsgakövetelmények és vizsgaleírás, a nehézségi paraméterek, valamint saját tapasztalataink alapján határoztuk meg. A nyelvhelyesség vizsgarészben/ vizsgarész-összetevőben mindezek mellett a „Profile Deutsch” (Glaboniat és mtsai, 2006) szoftvert is felhasználtuk.

Feladat	Item-szám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
1.	[4.]	0,97	0,18	0,22	0,05	-2,87	0,97	A1/A2
	[5.]	0,86	0,34	0,21	0,05	-1,58	1,05	A2
	[9.]	0,88	0,33	0,22	0,05	-1,71	1,03	A2
2.	[10.]	0,88	0,33	0,35	0,14	-1,56	0,94	A2
3.	[19.]	0,18	0,38	0,54	0,27	2,49	0,82	B1/B2
	[22.]	0,24	0,42	0,56	0,29	2,04	0,83	B1/B2
4.	[27.]	0,52	0,50	0,19	0,02	0,51	1,20	B1
	[29.]	0,60	0,49	0,24	0,04	0,10	1,15	A2
	[32.]	0,57	0,50	0,23	0,04	0,19	1,16	B1

A középszintű olvasott szöveg értése feladatsor problémás itemeinek jellemzői

Feladat	Item-szám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
1.	[1.]	0,94	0,23	0,22	0,04	-2,57	0,97	A1
	[3.]	0,93	0,26	0,29	0,08	-2,23	0,93	A1
	[5.]	0,96	0,19	0,23	0,05	-3,05	0,92	A1
2.	[8.]	0,16	0,37	0,44	0,19	2,39	0,93	A2
	[11.]	0,64	0,48	0,16	0,01	-0,17	1,18	B1/B2
	[14.]	0,26	0,44	0,24	0,04	1,72	1,15	A2
3.	[15.]	0,51	0,50	0,21	0,02	0,42	1,19	A2

A középszintű nyelvhelyesség feladatsor problémás itemeinek jellemzői

Feladat	Item-szám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ	Feltételezhető szint
1.	[2.]	0,86	0,34	0,28	0,05	-1,21	1,00	A1
	[6.]	0,97	0,17	0,25	0,05	-2,79	0,96	A1
2.	[15.]	0,82	0,39	0,31	0,05	-0,80	1,01	A2
3.	[16.]	0,71	0,45	0,34	0,08	-0,15	1,01	--
	[17.]	0,57	0,50	0,08	0,00	0,45	1,17	--
	[18.]	0,53	0,50	0,28	0,09	0,73	1,08	--
	[19.]	0,89	0,31	0,18	0,02	-1,50	1,02	--
	[20.]	0,54	0,50	0,19	0,01	0,63	1,11	--
	[21.]	0,74	0,44	0,25	0,06	-0,30	1,06	--

A középszintű hallott szöveg értése feladatsor problémás itemeinek jellemzői

Feladat	Itemszám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ
2.	[12.]	0,60	0,49	0,25	0,02	0,84	1,29
3.	[15.]	0,91	0,29	0,14	0,01	-1,48	1,15
	[16.]	0,62	0,49	0,29	0,07	0,73	1,26
	[20.]	0,94	0,24	0,21	0,04	-2,02	1,07
	[21.]	0,60	0,49	0,08	0,00	0,81	1,48
4.	[27.]	0,92	0,28	0,37	0,11	-1,59	1,00

Az emelt szintű olvasott szöveg értése feladatsor problémás itemei

Feladat	Itemszám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ
1.	[6.]	0,88	0,33	0,39	0,13	-1,27	0,94
	[7.]	0,15	0,36	0,36	0,12	3,07	0,99
2.	[10.]	0,99	0,10	0,17	0,03	-4,01	0,98
	[12.]	0,90	0,30	0,29	0,08	-1,53	0,98
	[13.]	0,46	0,50	-0,06	0,01	1,24	1,51
	[17.]	0,88	0,33	0,35	0,10	-1,27	0,98
	[18.]	0,87	0,34	0,49	0,20	-1,16	0,90
3.	[20.]	0,76	0,43	0,29	0,04	-0,35	1,09
	[21.]	0,87	0,33	0,19	0,03	-1,25	1,07
	[24.]	0,28	0,45	0,08	0,01	2,21	1,24
	[26.]	0,90	0,30	0,30	0,06	-1,55	0,99
	[28.]	0,91	0,29	0,35	0,10	-1,62	0,93
	[30.]	0,87	0,34	0,30	0,07	-1,20	1,05
	[33.]	0,87	0,33	0,32	0,09	-1,22	0,98

Az emelt szintű nyelvhelyesség feladatsor problémás itemei

Feladat	Itemszám	Átlag	Szórás	1 pontszám pontbiszeriális korrelációja	R ²	Nehézség	Infit MNSQ
1.	[1.]	0,56	0,50	0,28	0,04	-0,27	1,19
	[2.]	0,85	0,36	0,27	0,06	-1,92	1,07
	[3.]	0,76	0,43	0,31	0,08	-1,29	1,07
	[4.]	0,91	0,29	0,24	0,06	-2,56	1,05
	[5.]	0,83	0,38	0,33	0,12	-1,80	1,03
	[6.]	0,59	0,49	0,42	0,12	-0,34	1,06
	[7.]	0,70	0,46	-0,10	0,02	-0,98	1,48
	[8.]	0,71	0,45	0,24	0,05	-1,05	1,14
	[9.]	0,69	0,46	0,33	0,08	-0,77	1,10

Az emelt szintű hallott szöveg értése feladatsor problémás itemei



A projekt az Európai Unió
támogatásával valósul meg.

