

TELJESÍTMÉNYSZINTEK AZ ANGOL NYELVI ÉRETTSÉGI KIPRÓBÁLÁSÁN

Charles J. Alderson
University of Lancaster

Az elmúlt két év során három alkalommal került sor angol nyelvű feladatok kipróbálására a vizsgareform keretében: 1998 decemberében, valamint 1999 áprilisában és júniusában. A vizsgafeladatok kidolgozásának és kipróbálásának teljes folyamatát részletesen feldolgozza *Alderson, Nagy és Öveges* (2000) könyve. Ez a tanulmány a szintek meghatározásáról nyújt részletes, elemző beszámolót.

Az angol érettségi feladatok kipróbálása során nehéz volt megmondani a próbamérés feladatairól, hogy milyen szintűek. Részben mivel a Minisztérium nem határozta meg, hogy mit is jelent valójában az „alap”, „közép” és „felső” szint; másrészt, mert a magyar angolul tanuló populáció rendkívül heterogén (a problémák részletes tárgyalását l. *Fekete, Major és Nikolov*, 1999). A populáció heterogén természete azt jelenti, hogy valószínűleg lehetetlen csak két szintet meghatározni az érettségiben (ami nem más, mint amit az OKI-tól elvárnak) anélkül, hogy egyidejűleg a megfelelt minősítést jelentő jegyről és százalékról is beszéljünk.

Továbbá az sem egyértelmű, mit is értünk a „szint” fogalmán. Amennyiben a 12. évfolyamos tanulóknak középszintet kell elérniük, ez azt jelenti-e, hogy csak a hagyományos magyar skálán 5-öst elérő tanulókat tekinthetjük a megfelelő tudásszintűnek, vagy mindenki, aki legalább 2-est ér el, megfelelő szintűnek minősül? Nyilvánvaló, hogy ez egy kritikus kérdés, főleg a *Fekete és mtsai* háttér tanulmányában (1999) leírt tények tükrében: az érettségiben gyakorlatilag mindenki legalább hármast kap, és a populáció több, mint 85%-a kap legalább négyest. Az a tény, hogy senki sem bukik meg az érettségiben, azt jelenti, hogy a vizsga – gyakorlati szempontból – értéktelen.

Az „alap”, „közép” és „felső” kifejezések egyik jelentése az lehet, ahogyan az állami nyelvvizsgán meghatározott módon utalunk a szintekre. Amennyiben azonban ez a szándék, számos következménnyel kell számolnunk: a fent hivatkozott háttér tanulmány ugyanis kimutatta a vizsgán próbálkozó középiskolások magas bukási arányát az állami nyelvvizsgákon (és akkor természetesen még nem számoltunk azzal a rengeteg tanulóval, akik meg sem kísérelték letenni az állami nyelvvizsgát). Sőt, akár azt is felvethetjük, hogy tartalma és módszerei, továbbá a vizsga megbízhatóságát és érvényességét alámasztó bizonyítékok hiánya ellenére azért van az állami nyelvvizsgának értéke, mert a vizsgázók ott megbukhatnak és meg is buknak.

A háttér tanulmány (*Fekete, Major és Nikolov*, 1999) világosan kimutatta, hogy az angoltanárok ambivalensen viszonyulnak az érettségi megreformálására irányuló tervek-

hez. Egyrészt belátják, hogy reménytelenül elavult tartalma és módszerei miatt a jelenlegi érettségi sürgős reformra szorul; másrészt – és ez fontos szempont – mivel mindenki sikeresen leteszi, a vizsga gyakorlatilag értéktelen. A tanárok azt szeretnék, hogy az érettségi megfelelő értékkel rendelkezzen. Másrésztől, a tanárok kedvelik a jelenlegi érettségit, mert diákjaik sikeresen leteszik azt, továbbá úgy éreznék, hogy a gyenge eredmények az ő tanításukról állítanának ki rossz bizonyítványt! Tudatában vannak annak a dilemmának, hogy ha a vizsgát valós értékkel ruházzuk fel, a diákok egy része szükségszerűen meg fog bukni.

Bár az oktatási kormányzat meg kívánja reformálni a vizsgarendszert, még nem foglalkozott a színvonal kérdésével, és azzal sem, hogy várható-e a sikeres érettségi vizsga arányának a jelenlegihez képest jelentős mértékű csökkenése. Az alapelvek ilyen vákuumában az egyik lehetőség várni a felsőbb döntésekre. Mivel úgy véljük, ez felelőtlen magatartás, a próbamérés néhány feladatán elvégeztük a teljesítményszintek felmérését, a célból hogy lássuk, milyen következtetéseket lehet levonni a szintek meghatározása és esetleg a „megfelelt” szint tekintetében.

A módszer

A próbamérés feladatainak kipróbálása során két különböző eljárást alkalmaztunk az eredményszintek feltárására. Tudnunk kellett, hogy melyik feladat volt vélhetően alap-, közép- és felsőszintű. Maguknak a feladatoknak a nehézségét is meg kellett becsülnünk, mivel egy nehéz feladaton elért alacsony pontszám teljesen más jelentéssel bír, mint egy könnyű feladaton elért alacsony pontszám.

Először az úgynevezett „szemrevételezési” módszert használtuk, amely abból állt, hogy kerestünk kilenc olyan szakembert, akik ismerik mind az Európa Tanács (ET) (*Modern Languages: Learning, Teaching, Assessment. A Common European Framework of Reference*, 1998) szintjeit, mind a hazai angolul tanulók tudását. Felkértük őket, hogy becsüljék meg a feladatok nehézségi szintjét. Az ilyen szemrevételezésből egyrészt potenciális haszna származik a feladatíróknak, másrésztől lehetővé teszi számunkra az eredményszintek feltárását.

Továbbá össze tudtuk hasonlítani a próbamérés feladatainak empirikus nehézségét a CITO referencia-elemeink szintjével (*Noijons és Nagy, 1996*), mivel ezek szintje ismert az Európa Tanács Egységes dokumentumával való összevetésben (a hatfokú európai skála második fokán, A2 szinten állnak). Ez a munka magában foglalta a próbamérési feladatoknak a CITO elemekkel való kalibrálását és a szintek kiszámítását. Ezek után módunkban állt megbecsülni a feladatok nehézségét, valamint kiszámítani a tanulók tudásszintjét jelölő pontokat egyazon skálán. Ezt követően, hogy feltárjuk a vizsgált populáció képességeit, megfigyelhettük a pontszámok megoszlását. Ennek ellenére, mint ahogyan azt a következőkben látni fogjuk, egyedül ezzel a módszerrel nem voltunk képesek megállapítani a szintek közötti ponthatárokat. Ezért van szükség a szemrevételezési adatokra is.

A feladatok szintjének megítélése szemrevételezéssel

A szemrevételezési feladatban való részvételre kilenc szakértőt kértünk fel. Közülük négyen az angolos csoport tagjaiként szorosan közreműködtek az érettségi reformjában; rajtuk kívül tagja volt a munkacsoportnak az alpműveltségi vizsga idegen nyelvi vizsgatárgyi szakértője; a közös egyetemi felvételi vizsga felelőse; a jelenlegi angol érettségi felelőse; az állami nyelvvizsgabizottság angol nyelvi fővizsgáztatója és a Nyelvvizsgákat Akkreditáló Testület korábbi tagja; valamint a tanárok továbbképzésével foglalkozó brit tanácsadó. Mindegyikük kiterjedt, különböző szintű tanítási gyakorlattal rendelkezik, és megbízatásaiknál fogva elvárható tőlük, hogy mértékadó koncepcióval rendelkezzenek az alap-, közép- és felsőszintről. Biztosak vagyunk abban, hogy lehetetlen volna olyan személyt találni, aki náluk jobban tudja, mit is jelentenek ezek a kifejezések. Mellettük szól az is, hogy az Európa Tanács szintekre vonatkozó ajánlásait jól ismerik.

Bár 5 hallás utáni szövegértés, 12 olvasott szöveg értése és 8 angol nyelvhasználati feladat empirikus próbamérési adatait gyűjtöttük össze, úgy döntöttünk, hogy a hallás utáni szövegértés feladatok egy logisztikai nehézség miatt nem képezik a sztenderdizálási folyamat részét: lehetetlen volt ugyanis az összes felvételt az összes bírálónak meghallgatnia a rendelkezésre álló idő alatt. Mivel ez volt az első szabványosításra tett kísérlet, nem kívántuk túlterhelni a bírálókat, míg biztosak nem voltunk abban, hogy az eljárások célravezetők, és rendelkezünk olyan becsült adattal, amely megmutatja, mennyi időbe telik a szemrevételezés.

Ezért úgy döntöttünk, hogy csak az olvasott szöveg értését mérő feladatokat használjuk (beleértve a referencia-feladatot, melyet a CITO projektumból kölcsönöztünk) és a két nyelvhasználati feladatot. Ez utóbbiakat két ok miatt választottuk. Az egyik az volt, hogy az egyik nyelvhasználati feladatot külön tanulmányoztuk egy, a sorrend-megállapítással mint elem-típussal foglalkozó tanulmányban (*Alderson, Pércsich és Szabó, 2000*), és hasznosnak tűnt összegyűjteni külön adatokat erre a feladatra. A tanulmány és a két referencia-feladat tanulmányozása azt sugallta, hogy a feladatok az olvasást éppúgy mérik, mint a nyelvhasználatot, következésképp hasznosnak tűnt kipróbálni, milyen nehézségűnek ítélik meg a bírálók ezeknek a feladatoknak a szintjét (ne felejtjük el, hogy a nyelvhasználatot mérő feladatok kizárólag felső szinten léteznek, alap- és középszinten nem). Amennyiben a sztenderdizálási eljárás hasznosnak bizonyul, a jövőben valószínűleg kiterjesztjük azt a tanulók írásbeli szövegalkotást és a hallás utáni szövegértést mérő feladatokra is.

A sztenderdizálási eljárás anyagai az alábbiakat tartalmazták: részleteket az Európa Tanács Egységes Keretdokumentuma releváns részeiből; a bírálati eljárásokat bemutató utasításokat; az értékelésre váró 15 olvasási és nyelvhasználati feladatot; továbbá a DIALANG elnevezésű Európai Unió projektum által kifejlesztett önértékelési lapot, amely szorosan a Keretdokumentumon alapul (l. az A és B mellékletekben).

- Az 1. lépés során arra kértük a bírálókat, döntsék el, hogy a 15 feladat hol helyezkedik el a három magyar szint – alap, közép és felső – viszonylatában.

- A 2. lépés során az volt a bírálók feladata, hogy ismerkedjenek meg újra a Keretdokumentummal, és döntsék el, a feladatok az Európa Tanács hat szintjének melyikén helyezkednek el.
- A 3. lépés során a bírálóknak az egyes DIALANG önértékelési megállapításokat kellett osztályoznia a hat európai szint valamelyike szerint.
- A 4. lépés két részből állt: a bírálóknak mindkét rész során szem előtt kellett tartaniuk egy középszinten „éppen hogy csak megfelelő” jelöltet – olyan valakit, aki hajszal híján felel meg a szintnek – és minden egyes tesztfeladat minden egyes elemére (összesen 124-re) vonatkozóan el kellett döntenüik, hogy az adott jelöltnek milyen százalékos esélye lenne az elem helyes megoldására. Ezt követően, a második rész során a bírálókat arra kértük, hogy a tesztfeladatot egészében megvizsgálva döntsék el, milyen pontszámokat érne el egy „éppen hogy csak megfelelő középszintű” jelölt az adott feladaton. Ezeket a bírálatokat aztán megismételték egy „éppen hogy csak megfelelő felsőszintű” jelöltre.

Két héttel a fenti lépések befejezése és a lapok visszajuttatása után a szabványosítási feladatok megbeszélését szolgáló értekezlet előkészítéséül egy ötödik feladatot (C melléklet) kaptak ugyanezek a bírálók.

- Az 5. lépés a DIALANG projektum keretei között *Kaftandjjeva, Verhelst és Takala* által kifejlesztett módosított Angoff eljárást használta (*Kaftandjjeva, 1999*). A bírálókat arra kértük, hogy ugyanazon olvasási és nyelvhasználati feladatok minden egyes eleméről döntsék el, melyik elemet találná el egy „éppen hogy csak megfelelő középszintű” jelölt, továbbá ugyanazokra az elemekre vonatkozóan állapítsák meg, melyek azok, amelyeket egy „éppen hogy csak megfelelő felsőszintű” jelölt elronthatna úgy, hogy ettől még mindig felsőszintű tanulónak tekintenénk. Az ötödik lépésnek célja kettős volt: egyrészt, hogy kísérletezzünk egy másfajta, a 4. lépésnél egyszerűbbnek tűnő szabványosító eljárással, másrészt, hogy ellenőrizzük a két eljárás összehasonlíthatóságát és megbízhatóságát, amelyeknek legalább hasonlóan kell lenniük. Az eredményeket a következőkben mutatjuk be.

Ezt követően egy értekezleten (amelyen a kilenc bíráló közül nyolc vett részt) megbeszéltük a szabványosítási eljárás tapasztalatait. Ezen a megbeszélésen a bírálók párokban, majd azt követően plenáris ülésen, megtárgyalták az 5. lépésre adott válaszaikat, továbbá összehasonlították az 5. lépés végrehajtásának folyamatát a 4. lépésével. Ezután a bírálók kézhez kapták a 3. lépés eredményeit (D melléklet), valamint arra kértük őket, hogy hasonlítsák össze saját válaszaikat az összesített eredményekkel. Egy rövid megbeszélést követően arról, hogyan közelítették meg ezt a lépést és mi volt a véleményük az eredményekről, a bírálók megkapták az 1–4. lépés eredményeit (E melléklet), továbbá visszakapták saját válaszaikat, hogy ezeket összevegyék a csoport eredményeivel. Ezt újra az eredményekről szerzett benyomások megbeszélése követte, továbbá a bírálók kicserélték és feljegyezték az egész munka hasznosságáról és nehézségéről alkotott véleményüket. Az 5. lépéshez kapcsolódó adatokat csak a megbeszélést követően számíthattuk ki.

Eredmények

A bírálók megjegyzései

A tapasztalatok megbeszélése során a bírálók egyhangúan úgy érezték, hogy a bírálat megkezdése előtt tréningre illetve jártasság kialakítására nincs sem szükségük, sem igényük. Helyette a többi bíráló, illetve a jártasság kialakítása során kialakuló befolyás nélkül szeretnék volna meghozni saját bírálatukat. Továbbá állításaik szerint mindannyian kellő ismeretekkel rendelkeznek az Európa Tanács dokumentumairól illetve skáláiról.

A ET dokumentumot ért kritika, hogy a feladat nehézsége helyett a szöveg nehézségére koncentrálnak, míg a bírálati folyamat során – nagyonis érthető módon – a feladat nehézségét vették figyelembe. Az 1-es, 4-es és 5-ös lépés során legalább egy bíráló érezte úgy, hogy a feladatok szintjének elbírálásakor négy dolgot kellett szem előtt tartani:

- i) mire kell képesnek lennie egy ideális középszintű diáknak;
- ii) melyek a magyar tanulók előtt álló lehetőségek;
- iii) a próbamérés adatai alapján hol állnak jelenleg a magyar tanulók;
- iv) mire képesek a bíráló saját diákjai.

Érezhető volt, hogy amennyiben más bírálók is eltértek ily módon, ez magyarázata lehet bizonyos, a bírálók közötti véleménybeli különbségeknek, sőt, az egyes bírálók következetlenségének is. Egyik bíráló meglátása szerint azt az egyes feladatoknál tapasztalható egyet nem értést inkább az okozhatta, hogy a bírálókat különböző mértékben befolyásolta a feladatok minősége, mint például a feladat rossz/helytelen elrendezése (10. feladat), vagy a szöveg nehézsége, nem pedig a tanulók tényleges válaszai.

Az összes bíráló egyöntetűen a 3. lépésnél nem utalt vissza az Európa Tanács skálájára, hanem azt kellőképpen elsajátítva képesek voltak magabiztosan osztályozni. Az 5. lépés végrehajtásához már úgy érezték, szinte túlzottan megismerkedtek az ET dokumentummal és a feladatokkal. Néhány bíráló azt mondta, elképzelhető, hogy a korábbi bírálatok emlékei (a két lépés között eltelt két – három hét ellenére!), a bírálati folyamat idegesítő volta, a feladatelem nehézségről való ismeretük, valamint saját türelmetlenségük befolyásolta őket. Meglepő módon néhány bíráló előnyben részesítette a 4. lépést – a százalékos becslés használatát – az 5. lépéssel szemben, míg mások úgy érezték, hogy az 5. lépés könnyebb volt, egyszerűen azért, mert ekkorra már alaposan megismerkedtek a feladatokkal és elemekkel. Várakozásainkkal ellentétben az 5. lépés egyszerűségét illetően nem volt határozott tendencia: a vélemények megoszlottak. Néhány bíráló szívesebben hozott globális bírálatot a feladatokról – 4b lépés –, minthogy elemelelemre mondjon véleményt (4a lépés), nehéznek érezték megítélni, hogy pontosan melyik elemet fogja majd a tanuló eltalálni és melyiket nem, még akkor is, ha elég jó elképzelésük volt azt illetően, milyen pontszámot fog a tanuló elérni az egyes feladaton. Az alapszintű feladatokon elbírálása nehéznek bizonyult, mivel ezek túl könnyűek voltak. Legalább egy bíráló könnyen eldönthetőnek tartotta azt, hogy melyik elemeket nem találná el egy felsőszintű tanuló (5b lépés), mivel a legtöbb feladat egyébként is túl könnyű volt a felsőszintű diákoknak.

A bírálók egyike tudatosan próbált analitikus lenni bírálata során, de úgy érezte, hogy ez kevésbé tette „pontossá”, mint intuitív megérzése. Bár úgy gondolta, hogy analitikusabbnak kellene lennie, ezt a megközelítést mind nehéznek, mind pedig lassúbbnak találta. Egy bíráló számolt be arról, hogy miközben analitikus értékelésre törekedett, akár 10 percet is eltöltött az egyes elemek becslésével! A bírálót megkértük, hogy vezessenek feljegyzést arról, mennyi időt vett igénybe az első négy bírálati lépés: az eredmények hattól több mint 25 óráig terjedtek!

Az összes bíráló arról számolt be, hogy kihívást jelentettek az amúgy érdekes és informatív feladatok. Kíváncsiak voltak arra, hogyan viszonyulnak saját eredményeik a többiekéhez és az átlaghoz, és még több részletet szerettek volna megtudni arról, hogyan tért el véleményük egymásától, illetve hogyan alakult az eredmények szórása. A bírálók a sztandardizálást értékes, bár időnként fáradtságos folyamatnak tartották. Továbbá úgy érezték, hogy egy hasonló munka fölöttébb hasznos lenne a feladatírói tréning során is.

Eltérések a bírálatok terén

Ahogy az várható volt, a szakértő bírálók véleménye a különböző feladatok nehézségét illetően eltérő volt. Az F melléklet az 1, 2 és 3 bírálati lépésekre vonatkozó véleménykülönbségek részleteit mutatja be. A 4. és 5. lépés különbségei szükségszerűen még nagyobbak voltak.

Milyen következtetést lehet ezekből a különbségekből levonni? Mondhatjuk egyszerűen azt, hogy ezek a bírálatok az eltéréseknek köszönhetően nem megbízhatóak. Másrészt érvelhetünk azzal, hogy a különbségek legitimek, mivel szükségszerűen szubjektív kérdésekkel van dolgunk, és a különböző szakértők különböző tapasztalatokkal rendelkeznek. Amennyiben a különbségek legitimek, egyszerűen az átlagot vesszük, olyan alapon, hogy a több jobb, mint a kevesebb, és nincs jobb módja a különbségek feloldásának. Valójában, az összes, az alábbiakban bemutatott eredményben az egyes feladatok bírálatainak mediánját (nem a középértékét) vesszük, ezzel képviselve kilenc szakértő bírálónk átlagát.

Ennek ellenére ezen különbségek – amelyekre később még visszatérünk – egyik lehetséges következménye, hogy amennyiben nem sikerül olyan meggyőző bizonyítékot találni, amelyre a feladatok várt és tényleges nehézsége közötti erős kapcsolatot építhetnénk, meg kell kérdőjeleznünk az ilyen szakértői becslések értékét, talán még azokban az esetekben is, ahol a vélemények megegyeznek egymással. Ezáltal egyedül az empirikus nehézségi adatok maradnak, amelyekre a szintekre vonatkozó döntéseinket alapozhatjuk.

Az 5. lépés eredményeinek összehasonlítása a 4. lépés eredményeivel

Mint ahogyan az már korábban említettük, az 5. lépésre azért volt szükség, hogy ki-
próbáljunk egy egyszerűbb eljárást, valamint lehetővé tegyük a szabványosítási folyamat megbízhatóságának bizonyos fokú felmérését, annak ellenére, hogy belátjuk, az eljárás-

sok közötti különbség a tesztelés–újratesztelés megbízhatóságának felmérését problematikusá teszi. Sőt, az 5b lépés eredményei problematikusak voltak annyiban, hogy szinte minden egyes feladat túl könnyűnek találtatott a felsőszintű tanulók számára, és ezáltal sok bíráló jegyezte meg, hogy a felsőszintű tanulók maximális pontszámokat kapnának – azaz nem tévesztenék el egyik elemet sem.

A G melléklet a 4. és 5. lépések eredményeinek részletes összevetését mutatja be feladatról feladatra. Az eredményeket az alábbi 1. táblázat foglalja össze.

1. táblázat. A 4. és 5. lépések korrelációja

Változók	Középszint	Felsőszint
4b: 5a lépés	0,93	0,94
4a: 5a lépés	0,73	n.s.
A 4a: 5a lépés elsődleges pontszámai	0,95	0,95

Összehasonlítva a 4b lépés eredményeként az egyes feladatokra kapott pontszámokat, ahol az „éppen hogy csak megfelelő” középszintű tanuló pontszámait kellett megbecsülni, az 5a lépésnek az elem-szint bírálatának aggregátójából kapott eredményeivel, létrehozuk a feladat-szintű becsült pontszámokat: az eredmény 0,93 (Spearman rangkorreláció). Felsőszintű tanulókra (5b lépés) vonatkozóan a korreláció 0,94.

Továbbá összehasonlítva a kérdés-szint valószínűségi becsléséből adódó pontszámot (4a lépés) az 5a lépésből adódó pontszámmal (középszint), a korreláció (Spearman) 0,73, de a felsőszintű tanulókra vonatkozóan nincs szignifikáns korreláció, valószínűleg a fent már említett magas pontszám-középtértékek miatt.

Összevetve az 5. lépés eredményeit a 4. lépés eredményeivel, az egyes feladatokra vonatkozó elsődleges pontszám kiszámításával, amelyet a százalékos pontszámok jelölnek, amelyeket egyes feladatelemre kapott sikerességi esély százalékának kiszámításával kapunk meg (ily módon a 8. feladat 24%-os pontszáma 15 feladatelem esetében 3,6, felkerekítve egy becsült 4-es pontértékre, míg egy 10 elemből álló feladat esetén egy 73%-os becsült pontérték 7-es becsült elsődleges pontértéket eredményez). A középszintű tanulókra vonatkozó rangkorreláció 0,95, a felsőszintű tanulóké szintén 0,95.

Úgy tűnik, hogy a különböző bírálati folyamatok között a szabványosítás eredményei konzisztensek és megfelelő mértékben megbízhatóak.

Feladatírók előrejelzései

A 2. táblázat azt mutatja be, hogyan viszonyulnak a feladatírók előrejelzései a szabványosítást végző bírálók véleményéhez:

2. táblázat. Feladatírói előrejelzések és bírálói vélemények összehasonlítása (1. lépés)

Szintek		Bírálok		
		Alap	Közép	Felső
Feladatírók	alap	4	2	
	közép		3	1
	felső	1	1	2

A 14 eset közül öt mutat eltérést, de a maradék kilenc esetben az előrejelzés megegyezik a bírálók véleményével. A legradikálisabb nézetkülönbség az egyik nyelvhasználati referencia feladatnál volt, amelyet a feladatíró felső szintre szánt (3), s amelyet a bírálók alapszintűnek ítélték (1).

A feladatírók előrejelzése és a bírálók véleménye mellett feladatok empirikus nehézsége is fontossággal bír. A 3. táblázatban összehasonlítjuk a feladatírók és szakértő bírálók eredményeit a próbamérés empirikus adataival a megfelelő szintek vetületében.

3. táblázat. A bírálatok összevetése a feladatok nehézségével

	Nehézségi szint-skála	Nehézségi szint középérték	Skála logitok	Középérték logitok
<i>Feladatíró</i>				
Alap	29% – 83%	69%	-2,62-től +0,39-ig	-1,75
Közép	20% – 65%	43%	-0,83-től +0,71-ig	+0,03
Felső	24% – 62%	39%	-0,64-től +1,89-ig	+0,34
Rangkorreláció	0,62		0,67	
<i>Bírálok</i>				
Alap	62% – 81%	72%	-2,62-től -0,64-ig	-1,85
Közép	26% – 83%	50%	-2,30-től +0,60-ig	-0,39
Felső	20% – 44%	32%	-0,51-től +1,89-ig	+0,70
Rangkorreláció	0,75		0,73	

Összevetve a feladatírók előrejelzéseit a nehézségi értékekkel, azt találjuk, hogy alapszinten az értékek 29%-tól 83%-ig terjednek, középszinten 20%-tól 65%-ig, míg felsőszinten 24%-tól 62%-ig. Közép- és felsőszinten a nehézségi középérték különbsége ennek ellenére minimális (43% és 39%). A feladatírói előrejelzések rangkorrelációja a nehézségi értékekkel 0,62.

A bírálók esetében a becslések viszonya a nehézségi értékekkel valamivel jobb: alapszinten a nehézségi értékek 62%-tól 81%-ig terjednek, középszinten 26%-tól 83%-ig,

míg felsőszinten 20%-tól 44%-ig. A nehézségi értékek középértéke határozottan csökken, ahogy az megjósolható volt, egyre növekvő bírált nehézség mellett (72%–50%–29%). Úgy tűnik, a bírálók jobban előre tudják jelezni a nehézségi értékeket, mint a feladatírók, különösképpen alap- és felsőszintű feladatok esetén. A bírálók szint-bírálatának rangkorrelációja a nehézségi értékekkel 0,75, szintén magasabb.

Amennyiben összehasonlítjuk a feladatírói előrejelzéseket az empirikus kipróbálás során nyert kalibrált logit értékekkel, azt látjuk, hogy alapszintre vonatkozóan az adatok $-2,62$ -től $+0,39$ -ig terjednek, a középszintre vonatkozóan $-0,83$ -tól $+0,71$ -ig, míg felsőszinten $-0,64$ -tól $+1,89$ -ig. Logit nehézségi középértékeik tekintetében a középszint és felsőszint között kis különbség áll fenn. A feladatírói előrejelzések logit-értékekkel való rangsor-korrelációja 0,67.

A kalibrált nehézség terén a bírálók esetében különböző bírált szinteken átfedés tapasztalható: alapszintre vonatkozóan a logit-értékek $-2,62$ -től $-0,64$ -ig terjednek, középszinten $-2,30$ -tól $+0,60$ -ig, míg felsőszinten $-0,51$ -től $+1,89$ -ig. Ennek ellenére az empirikus logit nehézségi középérték határozottan emelkedik a bírált nehézségi szinttel, és a bírálók szint-bebecslésének rangsor-korrelációja a logit-értékekkel 0,73.

Bár úgy tűnik, a bírálók jobban előre tudják jelezni az empirikus nehézséget, mint a feladatírók, még mindig jelentős variáció létezik a pontosság terén. Nyilvánvaló, hogy a feladatok empirikus nehézsége nem mindig egyezik meg a szintekről alkotott szakértői véleményekkel. Az empirikus nehézségekre alapvető szükség van, mielőtt még megállapítást tehetnénk bármely feladat valós szintjéről.

A különböző szabványosítási eljárások összehasonlítása

A 3. lépés célkitűzése megvizsgálni, hogy mennyire sikerült a bírálóknak az Európa Tanács szintjeinek beazonosítása. Az adatok a D mellékletben találhatóak, és a bírálók korrelációja az eredeti, kalibrált Európa Tanács szintekkel meggyőző 0,97-es adatot eredményezett. Ez még a DIALANG kalibrációnak az Európa Tanácshoz viszonyított eredményénél is magasabb, ahol a rangsor korreláció 0,85 volt (*Kaftandjieva*, 1999). Érdekes módon, a mi bírálók korrelációja szintén 0,85 a finn önértékelési adatokkal. Úgy tűnik, megbízhatunk bírálók képességeiben azt illetően, hogy megítélik az Európa Tanács ajánlott szintjeit, legalábbis az önértékelési megállapításoknál.

A 4. táblázat azt ábrázolja, hogy a 15 feladat három magyar szinten való osztályozása hogyan felel meg az Európa Tanács hatszintű osztályozásának. Azt láthatjuk, hogy az alapszint vagy A1 vagy A2, a középszint leggyakrabban B1, bár két középfokú feladatot B2 szintre értékelték, és a haladó feladatok vagy C1 vagy B2 szintűek. Egy feladatot sem helyeztek a C2 szintre. Nem meglepő, hogy a magyar szintek egy szélesebb tartományt fednek le, mint az Európa Tanács szintjei, vagy – máshogyan fogalmazva – hogy ezek a feladatok az Európa Tanács szintjeitől eltérnek, tehát nem homogének.

4. táblázat. A bírálók magyar szintjei összevetve az Európa Tanács szintjeivel

Szint	A1	A2	B1	B2	C1
Alap	1	4			
Közép			5	2	
Felső				1	2

Hogyan viszonyulnak az Európa Tanács szintjeinek bírálati a logit értékekhez és a nehézségi értékekhez? Az 5. és 6. táblázatok ezeket az adatokat mutatják be.

5. táblázat. Az Európa Tanács szintjei és a feladatok kalibrált értékei közötti kapcsolat

Szint	Logitok tartománya	Logit középérték
A1	-1,64	-1,64
A2	-2,62-től -0,64-ig	-1,90
B1	-2,30-től +0,39-ig	-0,495
B2	-0,83-től +0,60-ig	-0,24
C1	+0,71-től +1,89-ig	+1,30
Rangkorreláció	+0,73	

Nyilvánvaló, hogy a nehézség területén átfedés van a feladatok között a feltételezhetően különböző szinteken, bár a viszony a várt előjelű. A nehézségi értékekre vonatkozó adatok hasonlóak, amint azt a következőkben bemutatjuk.

6. táblázat. Az Európa Tanács szintjei és a nehézségi értékek közötti kapcsolat

Szint	Nehézségi értékek tartománya	Nehézségi értékek középértéke
A1	76%	76%
A2	62%-tól 81%-ig	71%
B1	29%-tól 83%-ig	54%
B2	26%-tól 65%-ig	45%
C1	20%-tól 24%-ig	22%
Rangkorreláció	+0,76	

Az empirikus nehézség átfedése a szintek között meglehetősen nagy, ami nehezzé teszi az egyes szintek ponthatárainak megállapítását, legalábbis ezen feladatok alapján.

A 4-es lépés úgy került megtervezésre, hogy segítse a ponthatárok megállapítását azáltal, hogy a bírálók megbecsülik az „éppen hogy csak megfelelő” jelölt teljesítményét mind közép, mind felső szinten, mind kérdésről kérdésre, mind pedig a teljes feladat

pontszámára vonatkozóan. A két módszer, azaz az elemről elemre történő és a teljes pontszámra vonatkozó módszer közötti rangkorreláció 0,77 a középszintre, de nem szignifikáns a felsőre.

Az egyes feladatokra vonatkozó lehetséges ponthatárokat úgy számíthatjuk ki, hogy alapul vesszük annak az elbírált valószínűségét, hogy az éppen hogy csak elégséges jelölt egy kérdést helyesen eltalál, majd a feladat minden egyes elemére vonatkozóan összegezzük ezen valószínűségeket, és kiszámítjuk minden egyes bírálóra az „éppen hogy csak elégséges” jelölt sikerének átlagos valószínűségét az adott feladatra vonatkozóan. Ezt követően átlagolhatjuk a bírálók ezen eredményeit, és így megkapjuk az egyes feladatok határpontjának középvértékét. Ezek után összevethetjük ezeket a határpontokat a feladat bírált magyar (1. lépés) és Európa Tanács (2. lépés) szintjeivel. Az eredményeket a 7. táblázatban mutatjuk be.

7. táblázat. Valószínű határpontok az éppen hogy csak elégséges jelöltre, bíralt feladatszint alapján

Szint	„Éppen hogy csak elégséges” közép		„Éppen hogy csak” felső	
	Szórás	Közép- érték	Szórás	Közép- érték
Alap	69% – 79%	74%	90% – 98%	94%
Közép	33% – 69%	55%	66% – 97%	83%
Felső	24% – 29%	27%	61% – 64%	63%
1-es és 4-es lépés közötti sorrend rangkorreláció	0,92		0,81	
<i>Európa Tanács</i>				
A1	74%	74%	94%	94%
A2	69-től 79%-ig	74%	90-től 98%-ig	94%
B1	52-től 69%-ig	60%	79-től 97%-ig	87%
B2	29-től 51%-ig	38%	61-től 81%-ig	69%
C1	24-től 29%-ig	27%	63-től 64%-ig	64%
2-es és 4-es lépés közötti sorrend rangkorreláció	0,94		0,82	

Bár ezek a korrelációk igazán meggyőzőnek tűnnek, nem zárhatjuk ki annak az eshetőségét, hogy az egyes elemek valószínűségének megbecslésekor a bírálókat befolyásolta saját az Európa Tanács szintjein alapuló bírálatuk. Ennek ellenére úgy tűnik, hogy hasznos információra tehetünk szert. Különböző szinteken tudunk feladatokra vonatkozóan határpontokat megállapítani közép- és felsőszintű jelöltekre. A következő lépés az lenne, hogy ezeket a határpontokat alkalmazzuk a próbamérés során szerzett adatokra, hogy ezáltal lássuk a próbamérés populációjára gyakorolt hatást. A 8. táblázat összekapcsolja a 3., 5., 6. és 7. táblázatokat, ezáltal összeveti a valószínű határpontokat a feladatok tényleges nehézségével.

8. táblázat. A tényleges nehézségi értékek összehasonlítása a bírált határpontok középértékével közép és felső szinten

Szint	Bírált közép szintű „éppen hogy csak elégséges” pontszám középértéke	Bírált felső szintű „éppen hogy csak elégséges” pontszám középértéke	Tényleges nehézségi érték
Alap	74%	94%	72%
Közép	55%	83%	50%
Felső	27%	63%	29%
A1	74%	94%	76%
A2	74%	94%	71%
B1	60%	87%	54%
B2	38%	69%	45%
C1	27%	64%	22%

Amit itt látunk, az a középszintű tanulókra vonatkozó megfelelés egy foka a becsült határpont és a feladatokra vonatkozó tényleges nehézségi érték között a három különböző szinten. Természetesen a felsőszintű tanulókra vonatkozó határpontok jóval magasabban, mint a nehézségi érték középértékei: a felsőszintű diákoknak nagyon jól kell teljesíteniük a tesztfeladatokon ahhoz, hogy felsőszintűnek minősüljenek. Ennek ellenére a diákoknak még a középszintű besorolás eléréséhez is nagyjából a populáció középértékének megfelelő pontszámot kell elérniük, hogy a vizsgán átmenjenek. Más szóval, nagyjából a populáció fele nem felel meg a középszinten. Ez komoly problémákat vet fel a döntéshozók számára, mivel amennyiben az érettséginek csak két szintje lesz, középfok és felsőfok, a jelenlegi érettségitől eltérően, sok diák meg fog bukni. Ehhez a fontos ponthoz a későbbiekben még visszatérünk, de addig még szükség van az adatok további elemzésére a feltételezés szilárdságának bizonyítására.

Először vegyük fontolóra a kérdést feladatról feladatra, mivel a 8. táblázat az aggregált adatokat mutatja be. Amennyiben összehasonlítjuk azt a pontszámot, amit egy éppen hogy csak megfelelő jelölt érne el bármely feladaton az adott feladat nehézségének középértékével – a nehézségi értékek középértékének felhasználásával – bizonyos képet kaphatunk a populáció képességéről a feladat nehézségének relációjában. Ez lehetővé teszi számunkra, hogy felbecsüljük az egyes feladatokra vonatkozó, az adott magasságban történő vizsgaszint beállításának következményeit.

Ehhez vizsgáljuk meg az E mellékletet és tekintsük meg az 1-es feladatot. A BA (Barely Adequate: Éppen hogy csak megfelelő) középszintű ponteredmény vagy 33% vagy 29%, a használt módszertől függően, míg a nehézség középértéke 20% volt. Ez azt jelenti, hogy ezen a feladaton megbukna a populáció egy jelentős része. Ezzel szemben mind a feladattírók, mind a bírálók becslései szerint középszintűnek ítélt 4-es feladat határértéke 43% vagy 52% volt az éppen hogy csak megfelelő középszintű jelöltekre vonatkozóan, bár a nehézségi érték 54% volt. Más szóval, a populáció azon aránya, akik sikeresen átmennek, feladatról feladatra fog változni, mivel az empirikus nehézségek nem felelnek meg a bírált nehézségeknek. Ez azt jelenti, hogy a határpontokat, vagy a meg-

felelt értékeket nem lehet előre megállapítani, hanem az adott feladatsorra kell majd kiszámítani az adott vizsga esetében. Mint ahogy korábban említettük, ez a többszintű vizsgarendszer, nem pedig leegyszerűsített kétpólusú megkülönböztetés mellett szól. Továbbá úgy hisszük, ez felveti még a megfelelt jegy általános megkérdőjelezését is. (Az összegzés során még visszatérünk erre a kérdésre.)

Ennek ellenére megjegyezzük, hogy a fenti példában a 4-es feladat az 1-es füzetben volt, míg az 1-es feladat a 3-as füzetben, és tudjuk továbbá, hogy a 3-as füzet populációja gyengébb volt, mint az 1-es füzeté (a részletes elemzést l. *Alderson, Szabó és Pércsich, 2000*). Ez azt jelenti, hogy olyan feladat-nehézségi becslésekre van szükségünk, amelyek függetlenek a feladatot megoldó populáció mintájának képességeitől, ami nem mondható el a fenti adatokról. Ezen adatok értelmezésének nehézsége hangsúlyozza a populációtól független elem-paraméterek, valamint feladattól független személyi képességek mérésének szükségességét. Ezek rendelkezésünkre állnak a jelöltek kalibrált logit ponteredményeinek, valamint a kalibrált logit elemek nehézségi jellemzőinek formájában. Ezt a következő részben fogjuk megvizsgálni.

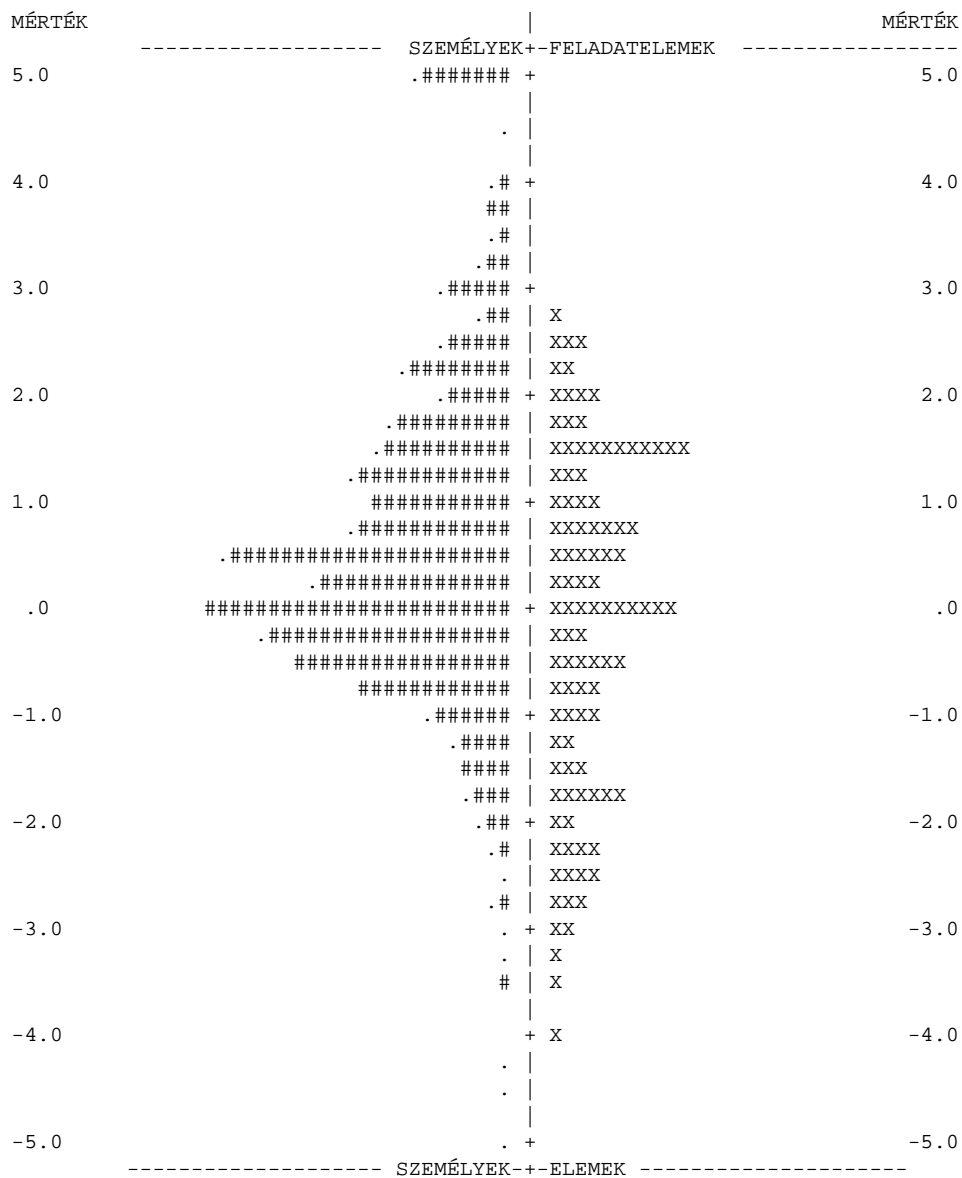
Képességszintek számítása

Mivel a valószínűségi tesztelmélet (IRT, a BigSteps program) használatával referenciaelemeket használtunk az elemek nehézségi fokának kalibrálására, módunkban állt az egyes személyek képességének becsült értékét kiszámítani logitokban, egyazon skálán. Ez lehetővé teszi számunkra, hogy összehasonlítsuk a populáció egyéneinek képességeit még abban az esetben is, ha részben különböző tesztfüzeteket oldottak meg.

Az olvasás, nyelvhasználat és hallás utáni szövegértési feladatok referencia-feladatként való használata segítségével minden egyes személyre kiszámítottuk az átfogó, becsült képesség értékét. Továbbá a hallás utáni szövegértés, olvasás és nyelvhasználati feladatokra vonatkozóan is kiszámítottuk a képesség külön becsült értékeit. A jelen fejezetben kizárólag az olvasás becsült képesség-értékeit adjuk közre, kiemelve a nyelvhasználati feladatokat. Emlékeztetőül: a szabványosító gyakorlat során bírált feladatok közül kettő nyelvhasználati feladat volt. Mivel ezek voltak a referencia-feladatok, és az olvasás és nyelvhasználati képesség becsült értékei az összes nyelvhasználati feladaton alapulva kerültek kiszámításra – amelyeket nem tettünk a jelen szabványosító gyakorlat részévé –, a jelen rész által tartalmazott elemzésből elhagyjuk a két nyelvhasználati feladatot, és az elemzést a 13 olvasási feladatra (beleértve a CITO referenciát), valamint az említett feladatok során nyújtott teljesítményen alapuló kalibrált olvasási képességre korlátozzuk.

Az elemek nehézségének becsült értéke és a jelöltek képességének becsült értéke összetevése révén szokás az IRT-vel egyazon logit skálán megbecsülni azt a mértéket, amennyire az elemek mérték a populáció képességeit. Az 1. ábra az olvasás adatait adja meg.

Charles Alderson



A SZEMÉLY OSZLOPBAN TALÁLHATÓ MINDEN EGYES '#' 4 SZEMÉLYT JELÖL; MINDEN EGYES '.' SZEMÉLYEK 1-TŐL 3-IG TERJEDŐ SZÁMÁT JELÖLI

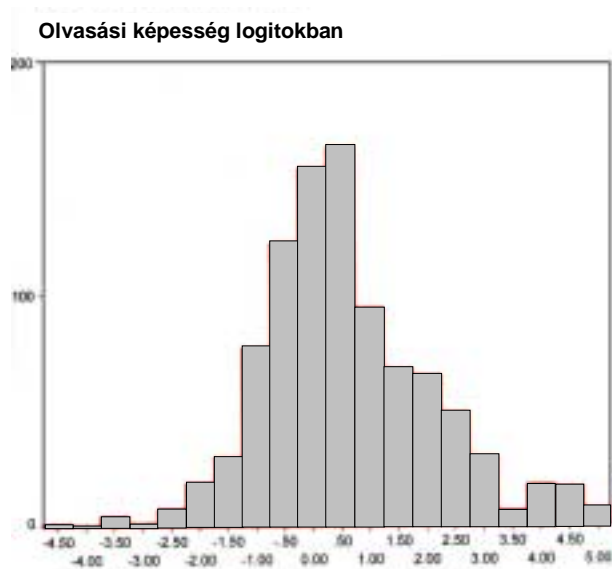
1. ábra
Személyi olvasáskészségek és olvasási elemek térképe

Az a tendencia figyelhető meg, hogy az elemek nehezebbek, mint a feladatokat megoldók képességei. Ennek ellenére a fenti térkép biztosít bennünket afelől, hogy az elemek többsége megfelelő nehézségi szinten volt, és hogy gyakorlatilag az összes diákot a megfelelő módon mértük.

Tudjuk, hogy a fenti populáció olvasási készsége hatalmas eltéréseket mutat – ezt jelzi az 1. és 2. ábra, és a 9. táblázat.

9. táblázat. Kalibrált olvasási készség

Tesztelt populáció <i>n</i>	Eredmény logit középértéke	Szórás	Minimum	Maximum
939	0,5923	1,5107	-4,45	4,78



Az olvasási-képesség logit skálaértékei, 0 várható értékkel

2. ábra
Olvasási képesség logitokban

Nyilvánvaló, hogy az olvasási képesség területén a jelen populáció rendkívül heterogén. A kérdés a következő: a populáció tagjai milyen szinten vagy szinteken állnak? Mely tanulók tudása minősül alapszintűnek, melyeké középszintűnek, és melyeké haladónak? Az látható, hogy a tanulók képességeiket tekintve nagy szórást mutatnak, de hol húzzuk meg a határokat a három szint között, és hogyan jellemezhetjük a skála különböző pontjain elhelyezkedő tanulókat?

Ne felejtjük el, hogy rendelkezésünkre állnak CITO referencia-feladatok, amelyek nehézsége ismert. Mind a hallás utáni szövegértés, mind az olvasás referencia-feladatokról úgy gondoljuk, hogy az Európa Tanács szerinti A2 szinten vannak. A hallás utáni szövegértés és az olvasás empirikus kalibrált értékeit külön-külön fogjuk megtekinteni. A 10 hallás utáni szövegértés referencia-elem logit-értékei $-2,05$ -től $-0,62$ -ig terjednek, középértékük pedig $-1,02$. A 10 olvasás referencia-elem logit-értékei $-2,29$ -től $1,03$ -ig terjednek, középértékük pedig $-0,826$. Amennyiben az egyes referencia-elemekre vonatkoztatjuk ezeket az értékeket, feltételezhetően A2-n, a jelöltek képesség-pontszámainak megoszlásában a következőt látjuk (10. táblázat).

10. táblázat. Az olvasási készség összevetése a referencia-feladatok nehézségével (A2)

CITO	Szórás	Középérték	A minimum alatti populáció	A maximum fölötti populáció	Az A2-n belüli populáció	A középérték fölötti populáció
Hallás utáni szövegértés	$-2,05$ -től $-0,62$ -ig	$-1,02$	12,5%	39%	49%	45%
Olvasás	$-2,29$ -től $1,03$ -ig	$-0,826$	2%	32%	66%	87%

A hallás utáni szövegértés esetében azt találjuk, hogy a populáció 12,5%-a kevesebb, mint $-2,05$ pontértéket, míg 39%-a magasabb, mint $-0,62$ pontértéket ér el. Más szóval, a populáció 48,5%-a az Európa Tanács A2-es szintjén belül esik. A populáció 44,6%-ának becsült képesség-értéke magasabb, mint a $-1,02$ -es középérték.

Az olvasás esetében a referencia-feladatok szélesebb nehézségi szórást mutatnak, és általánosságban valamivel nehezebbek, mint a hallás utáni szövegértés elemei. Ennek ellenére azt látjuk, hogy képességeik tekintetében a populáció egy nagyobb része (66%) tekinthető A2-es szinten lévőnek, bár 87%-uk az A2-es képesség középérték fölött helyezkedik el.

Ezeket az eredményeket nehéz értelmezni. Egyrészt azt jelzik, hogy a populációnak csak 32%-a rendelkezik az A2-es szintnél magasabb olvasási készségekkel – amelyet a bírálóink alapszintűnek tekintenek! Másrészt a populációnak csak 13%-a rendelkezik az A2 szinten szükséges átlagnál alacsonyabb olvasási készségszinttel. Ennek ellenére, a bírálók a CITO olvasási referencia-feladatait B2-es szintűnek becsülték, és középszintűnek a magyar fogalomrendszer szerint. Ez radikális változásokat okozna a jelen populáció olvasási képességeinek megítélésében.

A hallás utáni szövegértés esetében is fennáll ez a kérdés, bár talán nem olyan kiélezetten: a populáció 49%-a A2-es szinten helyezkedik el, és csak 45%-a van az A2 szint középértéke fölött. Amennyiben az A2 valójában alapszint, a populációnak csak valahol 39% és 45% közötti része tekinthető középszinten vagy afölött állónak. A populáció nagyobb része megbukna egy középszintű hallás utáni szövegértést mérő vizsgán.

Nyilvánvaló, hogy a probléma az A2-es szintre szánt (vagy B2-es szintűnek becsült) feladatelemek nehézségi szórásának terjedelmében kereshető. Újra az empirikus nehéz-

ségeknek a szándékunk szerinti nehézséggel való egyeztetési problémájával találjuk magunkat szemben.

Mindezek ellenére az a tény, hogy ezek a referencia-feladatok A2-esek vagy B2-esek, nem határozza meg, hogy hová helyezzük a közép- és a felsőszint közötti határvoalat, azt pedig biztosan nem mutatja meg nekünk, hogy hogyan osztályozzuk a populáció azon tagjait, akik a referencia-feladatok határértékein kívül esnek. Különös figyelmet érdemel az a kérdés, hogy már középszintűnek minősülnek-e azok a tanulók, akik az A2 feladatok középszintűjét vagy annál magasabbat érnek el, vagy csak azok számítanak-e középszintűnek, akik a maximális referencia-feladat értékek fölötti eredményt érnek el. Melyik az a pont, ahol a középszintű tanulók, bármi is legyen ennek a meghatározása, felsőszintűvé válnak? Ezek a dilemmák nyilvánvaló módon mind a hallás utáni szövegértésre, mind az olvasott szöveg értésének eredményeire vonatkoznak.

E kérdések megválaszolásához szükséges a becslésen alapuló adatokra visszautalni, és mivel a hallás utáni szövegértés feladatokról nem gyűjtöttünk becslési adatokat, a probléma következőkben leírt tárgyalását kizárólag az olvasási feladatokra korlátozzuk.

Olvasási készség és az olvasási feladat nehézsége

Szükséges annak a logit-képesség eredményének kijelölése (vagy értelmezése), amelynek elérését egy megfelelő jelöltől elvárhatjuk bármely feladat esetében. Tulajdonképpen arra van szükségünk, hogy kombináljuk az empirikus adatokat a becslési értékekkel. Mivel a tanulók empirikus eredményei és a tényleges elem-nehézségi értékek ugyanazon logit skálán kerülnek kalibrálásra, össze tudjuk vetni a tanulók teljesítményét a feladatok nehézségi értékeivel.

Először csak azokat az elemeket választjuk ki, amelyekről a bírálók úgy érezték (5. lépés), hogy a tanulónak el kellett találni ahhoz, hogy közép- illetve felsőszintűnek nyilváníthassuk őket. Ezt követően megnéztük az egyes kiválasztott feladatelem tényleges logit-nehézségét, majd minden egyes feladatra vonatkozóan kiszámítottuk a logit-érték középszintűjét. Így kapunk egy logit-értéket, amelyet a tanulónak el kellene érnie, hogy elérjen egy „éppen hogy csak megfelelő” szintet az adott feladat során. Ezt követően aggregálhatjuk az eredményeket a feladatokon keresztül általában – így, a feladatvariációtól függetlenül, egy megbízhatóbb illetve reprezentatívabb határértékhez jutva, mivel azt tudjuk, hogy nehézségüket tekintve a feladatok eltérnek. Ezt megtehetjük a becslési, illetve a szándék szerinti feladat-szintre való tekintet nélkül.

A 11. táblázat az összes olvasási elem nehézségének középszintűjét mutatja be, függetlenül a bírálóktól, valamint azon elemek középszintűjét, amelyekről a bírálók úgy gondolták, hogy egy középszintű jelöltnek el kell találnia.

A 11. táblázat azt mutatja, hogy azok a feladatelemek, amelyeket a középszintű tanulónak el kell találni, az összes olvasási feladatelem egy könnyebb részalmazat képezik: más szóval, nem meglepő, de mindenképp megerősítő módon, léteznek olyan, empirikusan nehéz elemek, amelyekről a bírálók azt gondolják, hogy a tanulónak nem kell eltalálnia ahhoz, hogy középszinten állóknak minősüljenek.

11. táblázat. Az olvasott szöveg értését mérő feladatelemek nehézsége logitokban kifejezve

<i>Adat</i>	<i>Összes olvasási elem</i>	<i>Középszintű elemek</i>
Elemek száma	110	62
Középvérték	-0,9420	-1,6082
A középérték standard hibája	0,1435	0,1678
Standard szórás	1,4705	1,3211
Variancia	2,1624	1,7453
Szórás	6,03	5,13
Minimum	-4,18	-4,18
Maximum	1,85	0,95

Ezt követően, amennyiben ezekből az adatokból indulunk ki, és a tanulók képesség-eredményeit ezek függvényében ábrázoljuk (12. táblázat), azt találjuk, hogy a jelen populáció 95%-a a középszintű feladatelemek bírált határértéke fölötti eredményt érte el. A populáció 35%-a a középszintű tanulóknak megállapított maximális határérték fölött lenne, továbbá a populáció 65%-a a lehetséges határértékek szórásának intervallumába esik, attól függően, hogy mely feladatelemeket tekintjük minimum elvárásnak egy középszintű tanuló számára.

12. táblázat. Olvasási készségi szintek. Középvérték-bírálat – 5a lépés (középszintű minimum) összevetve a becsült olvasási készséggel

	<i>Középszintű feladatelemek nehézségi szórása</i>	<i>Feladat-elem nehézség középértéke</i>	<i>A minimum alatti populáció</i>	<i>A maximum fölötti populáció</i>	<i>A középszinten belüli populáció</i>	<i>A középérték feletti populáció</i>
Középszintű olvasás	-4,18-től 0,95-ig	-1,6082	0,3%	35%	65%	95%

Azoknak az olvasott szövegértést mérő feladatoknak a logit középértéke, amelyeket egy középszintű tanulónak el kellene találnia, -1,6082. A populációnak csak 5%-a ér el ennél alacsonyabb eredményt (és ezáltal 95% ér el magasabb eredményt). Amit ezek az adatok mutatni látszanak – amennyiben elhiszük a bírálat és a logit-eredmények közötti összefüggést – az az, hogy a jelen populáció legkevesebb körülbelül 35%-a, és elképzelhető, hogy akár 99,7%-a minősül a középszintnek tekintendő minimum előírt szinten állónak!

Ezek az adatok szemben állnak a CITO olvasási referencia-feladatelemekkel, valamint a jelen populáció olvasási készségének interpretációja hatalmas eltéréseket mutat attól függően, hogy a feladatelemek közül ki mit tekint döntő fontosságúnak. Ne felejtjük el, hogy a bírálók a CITO olvasási referencia-feladatait B2-es, és nem A2-es szintűnek tekintették, bár ennek magyar szintjét középszintre helyezték (E melléklet). Röviden tehát, jelentős nehézséggel találjuk szemben magunkat, mikor ezeket az adatokat bármely, de legfőképp három különböző szint meghatározására használjuk.

Következtetések

Ebben a cikkben olyan eljárások kidolgozását és alkalmazását mutattuk be, amelyek segítségünkre lehetnek abban, hogy megállapítsuk a próbamérés populációjának eredmény szintjét (szintjeit), vagy a magyar rendszer nem definiált elnevezései – alap, közép és felső – vagy a keretdokumentumban lefektetett Európai Tanács szintjei szerint. Szakavatott bírálókat gyűjtöttünk össze olyan kollégáktól, akik várakozásaink szerint elajátították, még ha csak intuitív módon is, a nehézségi szintek fogalmát. Emellett áttekintettük a feladatok nehézségének és a tanulók képességének megállapítására szolgáló empirikus módszert. Az utóbbi a valószínűségi tesztelméleten (Item Response Theory) alapul és referencia-feladatokat használt a próbamérési füzetek különböző feladatainak kalibrálásához. Nyilvánvaló, hogy az empirikus módon megállapított nehézségek a kipróbált feladataink nehézségi jellemzői. A kérdések a következők: hogyan lehet az ilyen nehézségi szinteket a legjobban jellemezni, és hogyan lehet őket előre jelezni? Lehetséges-e a próbamérés populációjának nyilvánvalóan széles szórású készségeinek, illetve a próbamérésen használt, széles nehézségi skálájú feladatoknak megfelelni egy két- vagy akár háromszintű rendszeren belül, mint ahogyan azt jelenleg az Oktatási Minisztérium előírja? Hogyan illeszkedhetnek az ilyen szintek azokhoz a nemzetközileg elfogadott szintekhez, amelyeket a Minisztérium állítólag meg kíván célozni?

Azt tapasztaltuk, hogy a szakértő bírálók által adott empirikus nehézség előrejelzés jobb, mint a feladatíróké. Az előbbiek bírálata a különböző szabványosító eljárások során megbízhatónak és következetesnek bizonyult, továbbá az általuk megállapított nehézségi sorrend nagy vonalakban megegyezett az empirikus nehézség növekedésével. Mindezek ellenére jelentős eltérés volt megfigyelhető a feladat-elemek és a feladatok nehézsége terén. Ez az eltérés oly mértékű volt, hogy annak a biztonsággal történő megállapítása, miszerint egy adott feladat valójában egy adott szinten van-e, nehéz, illetve majdnem lehetetlen volt. Az is lehetetlennek bizonyult, hogy a teljesítményre vonatkozó megfelelő ponthatárt meghatározzunk, akár az egyéni feladatokra, akár egy teljes feladatsorra vonatkozóan, amely meggyőző bizonyítékkal szolgálna arra, hogy a jelöltek elérték az adott szintet. Nyilvánvalóvá vált, hogy akárhogy is kerülnek a szintek végül meghatározásra, a „megfelelt” és „nem felelt meg” döntések hatását nagyon alaposan fontolóra kell majd venni. Jelenleg az érettségi vizsgának nincs értéke, mivel gyakorlatilag minden tanuló átmegy. Amennyiben azt akarjuk, hogy a vizsga értékkel bírjon, vagy néhány tanulónak meg kell buknia, vagy az eredmények jelentésének egy eltérő

rendszerére van szükség. Amennyiben a bukás és a megfelelés fogalma bármilyen formában megmarad, a szintek megállapítását nem lehet a ponthatár figyelembe vételétől függetlenül meghatározni.

Nyilvánvalóvá vált, hogy a feladatírók szándékai vagy a szakértő bírálók véleménye ellenére a feladatok nehézségi szintje eltérő lesz. Ezáltal egy feladatelem vagy feladat, vagy akár egy teljes feladatsor valódi szintje csak post-hoc határozható meg: azaz az eredmények elemzését követően.

Ebből következően, vagy a feladatelemeket és a feladatokat kell majd még az éles vizsga lebonyolítását megelőzően előkalibrálni (ezáltal megismerve empirikus nehézségüket), és a feladatelemeket és feladatokat feladatelem- / feladat-bankokból kell venni; vagy a vizsgaeredményeket csak akkor lehet majd közreadni, mikor a populáció teljesítménye már elemzésre került – amely gyakorlati és logisztikai problémát vet fel.

Az alternatíva: teljes mértékig szakítani kell a két szint, valamint a sikeres vizsga és bukás fogalmával, és fel kell hagyni az 1–5-ös skálával; továbbá szükségessé válik a vizsgaeredményeknek egy olyan, például 0-tól 100-ig terjedő nehézségi skálán történő elhelyezése, amely lehetőleg magába foglalja a populáció készségeinek heterogén természetét. Ezt követően a vizsgaeredmények felhasználói hoznák meg saját döntéseiket azt illetően, hogy a jelöltek megütdtték-e azt a szintet, amely megfelel a jelöltek kiválasztásának céljaira, például valamilyen munkahelyen történő alkalmazásra, egyetemi felvételre vagy bármi másra. Nyilvánvaló, hogy ahhoz, hogy az érettségi megfeleljen a gyengébben teljesítő tanulók szempontjából, akik képesek bizonyos, bár alacsony szintű teljesítményt elérni angol nyelvből, fontos az Európa Tanács A1-es vagy A2-es szintjét elérő tanulók eredményeinek elismerése. A javasolt 0-tól 100-ig terjedő skála pontosat ezt tenné lehetővé.

Ettől függetlenül meg kell jegyezni, hogy a próbamérésünk populációja által elért szinteket a tanulók és tanáraik a jelen próbavizsgára való mindennemű felkészítés nélkül érték el. A feladattípusok és valójában maga a vizsga jellege is teljesen ismeretlen volt számukra. Megfelelő felkészítés mellett tehát nagy biztonsággal várhatjuk, hogy a tanulók teljesítményszintjei emelkedni fognak, mivel más területen bőséges bizonyíték van arra, hogy miután a tanulók megismerkednek azzal, hogy mit is várnak el tőlük a vizsgán, teljesítményük javul. Természetesen ez még fontosabbá teszi, hogy a szintek, szabványok és a megfelelési arányok empirikusan kerüljenek meghatározásra, és ne legyenek előre meghatározottak.

Fordította: Grezsu Katalin

Irodalom

Alderson, C., Nagy, E. és Öveges, E. (2000, szerk.): *English language education in Hungary: Examining Hungarian students' achievements in English*. The British Council, Budapest.

Alderson, C., Pércsich, R. és Szabó, G. (2000): Sequencing as an item type. *Language Testing*. 17. 4. sz. 423–44.

Fekete, H., Major, É. és Nikolov M. (1999, szerk.): *English language education in Hungary: A baseline study*. The British Council, Budapest.

- Kaftandjieva, F. (1999): DIALANG: Some results from data analysis (paper-based piloting, Finnish) Előadás: *2nd Conference of European Language Council*, July 1–3, Jyväskylä, Finnország.
- Modern Languages: Learning, Teaching, Assessment. A Common European Framework of Reference*. (1998). Council of Europe, Strasbourg.
- Noijons, J. és Nagy, E. (1996): Towards a standardized examination system. *Joint Hungarian-Dutch Project (Final Report)*. CITO, OKI, Budapest.

ABSTRACT

CHARLES ALDERSON: LEVELS OF PERFORMANCE AND THE HUNGARIAN MATURA EXAM IN ENGLISH

In the context of the reform of the Hungarian Matura exam and the Council of Europe Common European Framework of reference, this study discusses the problem of identifying levels of performance in reading in English as a foreign language. A pool of pilot tasks was examined by a board of 9 expert judges whose responses on five procedures were then analysed to establish the difficulty of each task. Empirical methods of scaling task difficulty and candidate ability were also explored, the latter based on Item Response Theory. Anchor items were used to calibrate the various tasks in the pilot booklet. The questions posed included (1) How can difficulty levels best be characterised? (2) How can they be predicted? (3) Is it possible to accommodate the evidently wide range of ability of the pilot population, and the wide range of difficulty of tasks piloted, within a scheme of two or three levels, as currently required by educational policy? (4) How might such levels correspond to internationally recognised levels? Expert judges were found to be better at predicting empirical difficulties than were item writers. Their judgements appeared to be reliable and consistent across different standard-setting procedures, and the order of difficulty they predicted corresponded broadly to the increase in empirical difficulty. However, there was considerable variation in difficulty across items and test tasks, such that it was difficult, if not impossible, to establish with confidence that a given task was at a given level. It was found that tasks will vary in difficulty, despite the intentions of item writers or the opinions of expert judges. Therefore the true level of an item, a task or the whole test can only be determined post-hoc, by analysing the results.

Magyar Pedagógia, **100**. Number 4. 423–458. (2000)

Levelezési cím / Address for correspondence:

Charles Alderson

Mellékletek

A melléklet

Szabványosítás az új angol érettségi területén

Kedves Kolléga!

Köszönjük, hogy úgy döntött, munkájával részt kíván venni a jelen tanulmányban. Amint az látni fogja, az Érettségi vizsga két „szinten” közép (Intermediate) és haladó (Advanced) kerül kidolgozásra. Továbbá, a 10. Évfolyam vizsgája az alap (Basic) szintet célozza meg. A jelen tanulmány fő célkitűzése az, hogy útmutatót biztosítson arról, mit is jelentenek ezek a „szintek”, és hogyan is lehet őket meghatározni.

Részletesen meghatározott célkitűzések:

- 1) A tanulók teljesítményére vonatkozó empirikus adatok kiegészítése a határértékekre vonatkozó szemrevételezési adatokkal.
- 2) Az angol vizsgareform project keretein belül kidolgozott tesztfeladatok viszonyítása az Európa Tanács Egységes Keretdokumentumához.
- 3) Jövőbeli szabványosítási gyakorlatok során használható eljárások kidolgozása.

A módszer

Az Ön feladata projektum során kidolgozott és próbamérésen használt adott feladatok nehézségi szintjére vonatkozó szemrevételezés végrehajtása lesz, a magyar szintek és az Európa Tanács Egységes Keretdokumentuma viszonylatában.

Négy, a következőkben bemutatott lépés létezik: az 1. lépés során Önnek az egyes, próbamérésen használt olvasott szöveg értése feladatokat kell alap, közép vagy haladó szintűként osztályoznia. Úgy gondolom, ez nem fog egy óránál több időt igénybe venni. A 2. lépés során arra van szükség, hogy osztályozzon minden egyes olvasási feladatot az Európa Tanács hat szintjének egyikén. Ez akár egy napba telhet, attól függően, hogy Ön mennyire van tisztában az Európa Tanács szintekkel. A 3. lépés során egyszerűen arra van szükség, hogy az Európa Tanács szintjei szerint osztályozza olvasott szöveg értése skálák egy sorát. Úgy gondolom, ez nem fog egy óránál sokkal több időt igénybe venni. A 4. lépés a legkomplexebb és legismeretlenebb, mivel ez azt kéri, hogy képzeljünk el két tanulót, két különböző szinten, majd becsljük meg annak valószínűségét, hogy ezek a tanulók milyen eséllyel fogják sikeresen megoldani az egyes olvasott szöveg értése feladatokat és az egyes feladatelemeket. Ez akár egy vagy két napba is telhet. Mindenesetre, nagyon hálás lennék, ha fel tudnák jegyezni, mennyi időbe telt az egyes lépések elvégzése.

A következőkben megtalálja az egyes eljárásokra vonatkozó részletes instrukciókat.

1. lépés:

Kérjük, olvasson el minden egyes olvasott szöveg értése tesztfeladatot figyelmesen, majd minden egyes tesztfeladatra vonatkozóan döntse el, hogy az alap (Basic), közép (Intermediate) vagy haladó (Advanced) szintű-e. Kérjük, használja tapasztalatát / ítélőképességét vagy "megérzését" arra vonatkozóan, mit is jelentenek, vagy mit kellene, hogy jelentsenek ezek a kifejezések. Minden egyes feladat jobb felső sarkában jelölje azt a szintet, amelyik mellett dönt.

2. lépés: Az Európa Tanács Egységes Keretdokumentuma

A jelen Keretdokumentumban meghatározott hat szint eredetileg a következő néven vált ismertté:

Breakthrough
Waystage
Threshold
Vantage
Effective Operational Proficiency
Mastery

Teljesítményszintek az angol nyelvi érettségi kipróbálásán

De ezeket a verbális címkéket azóta már felváltották (a Keretdokumentum 1998-as második tervezet-változata) a betűből/számból álló szintek:

A1
A2
B1
B2
C1
C2

Kérjük, ismerkedjen meg a Keretdokumentummal – lehetőleg az 1998-as Második verzióval (Version Two) (ennek jobb felső sarkában: CC-LANG (95) 5 rev. V), de ne aggódjon, amennyiben egy korábbi verzióval rendelkezik, mivel ezek nem térnek el nagyban egymástól, és a kulcsfontosságú részeket az alábbiakban megtalálják.

Ebből a szempontból különösképpen releváns az 5-ös és 8-as fejezet, valamint a Függelékek. Mindezek ellenére, az elsajátítás szempontjából ez még mindig sok információ. A 4., 5., 6., 7. és 8. táblázat (8. fejezet, 128–134. o.) a legjobb összefoglalása ezeknek a szinteknek. Ezekből példányokat mellékelek az Önök tájékoztatására. A Függelékek, 167. ff oldal a különböző nyelvi készségek során használt skálákra vonatkozóan nyújtanak példát, és az, amely különösen releváns az Önök eljárására, az olvasott szöveg értése feladatra koncentrááló skálák a 175–77. oldalakon található, továbbá ezekből példányokat mellékelek.

A fenti dokumentumok átolvasását és a szintek interpretálásának elsajátítását követően a következő folyamat annak eldöntése – minden egyes olvasott szöveg értése feladatra vonatkozóan –, miszerint ezek A1, A2, B1 stb. szinten állnak-e. Kérjük, a kiválasztott szintet írja az egyes feladatok bal alsó sarkába.

3. lépés: Önértékelési skálák

A DIALANG-ként ismert, az Európai Bizottság által finanszírozott projekt törekvése az volt, hogy leszűri az Európa Tanács Egységes Keretdokumentuma keretein belüli információ többségét, valamint ezen skálák alapján kifejlesztett a különböző nyelvi készségekre vonatkozó önértékelési megállapításokat. A következőkben mellékelem az olvasási feladat önértékelési skáláinak példányait.

Az Ön feladata az, hogy eldöntse, az egyes megállapítások a hat szint (A1 – C2) melyikét képviselik. Maguknak az olvasott szöveg értése feladatoknak az osztályozása nem tartozik az Ön feladatai közé.

KÉRJÜK, DÖNTÉSE MEGHOZATALÁNÁL NE HASZNÁLJA A 2. LÉPÉS EREDETI SKÁLÁIT.

Kérjük, az olvasásra vonatkozóan a DIALANG minden egyes megállapítására vonatkozóan jelölje értékelését az utolsó oszlopban.

4. lépés: Határértékek és éppen hogy csak elégséges teljesítmények

Gondoljon egy olyan, a középiskola felső évfolyamainak negyedik osztályába járó magyar tanulóra, aki valószínűleg **éppen hogy csak** megfelel a középszintű teszten. Nevezzük ezt a tanulót: Éppen hogy csak megfelelő középszintűnek [Barely Adequate Intermediate (BAI)].

Most gondoljunk egy olyan tanulóra, aki valószínűleg **éppen hogy csak** megfelel a haladó szintű teszten. Nevezzük ezt a tanulót: Éppen hogy csak megfelelő haladó szintűnek [Barely Adequate Advanced (BAA)].

Most tehát minden egyes olvasott szöveg értése feladatra vonatkozóan:

i) először minden egyes feladatelem mellett jelölje, hogy hány százalék az esélye van annak, hogy egy Éppen hogy csak megfelelő középszintű [Barely Adequate Intermediate (BAI)] tanuló helyesen eltalálja a feladatelemet (5%? 40%? 85%? – használjon bármilyen, ön által megfelelőnek tartott értéket), majd ezt követően, másodszer jelölje azt, hány százalék esélye van annak, hogy egy Éppen hogy csak megfelelő haladó szintű [Barely Adequate Advanced (BAA)] tanuló helyesen eltalálja a feladatelemet.

ii) amennyiben ezt már az első feladat minden egyes feladatelemére vonatkozóan végrehajtotta, becsülje meg, milyen összpontszámot érne el egy Éppen hogy csak megfelelő középszintű (Barely Adequate Intermediate) tanuló ebben a feladatban, majd hogy milyen összpontszámot érne el egy Éppen hogy csak megfelelő haladó szintű (Barely Adequate Advanced) tanuló ugyanebben a feladatban.

iii) Csak ezt követően lépjen tovább a következő olvasott szöveg értése feladatra, és ismétlje meg a fenti i) és ii) részfeladatot. Ezeket a lépéseket kell az olvasott szöveg értése feladatcsomag összes feladatán végrehajtani.

Charles Alderson

LEHET, HOGY KÉSZTETÉST ÉREZ ARRA, HOGY ELLENŐRIZZE SAJÁT BÍRÁLATÁT, ÉS FOLYAMATOSAN VISSZAUTALJON A KORÁBBI FELADATELEMEKRE. ERRE AZÉRT NINCS SZÜKSÉG, MIVEL LEHETETLEN EZEKNÉL A BECSÜLT ÉRTÉKEKNÉL 100%-OS HELYES ÉRTÉKET VAGY 100%-OS MEGBÍZHATÓSÁGOT ELÉRNI.

KÉRJÜK, A JELEN ELJÁRÁS EGYIK SZINTJÉN SE KONZULTÁLJON EGYIK KOLLÉGÁJÁVAL SEM. AMENNYIBEN AZ ELJÁRÁSOKAT ILLETŐEN BÁRMILYEN KÉRDÉSE VAN, KIZÁRÓLAG CHARLES ALDERSONNAL LÉPJEN KAPCSOLATBA. A MUNKA VÉGEZTÉVEL EGY ÉRTEKEZLETET TARTUNK AZ EREDMÉNYEK MEGBESZÉLÉSÉRE.

Kérjük, a kitöltött eljárási nyomtatványokat juttassa el Charles Aldersonnak **legkésőbb január 31-ig**.

Köszönjük közreműködését. Munkája felbecsülhetetlen segítséget jelent majd annak eldöntésében, hogy a különböző feladataink milyen szinten helyezkednek el, és milyen szinten kellene elhelyezkedniük.

B melléklet

Szabványosítás az Angol vizsgareform project keretein belül

Köszönjük, hogy úgy döntött, részt vesz a következő szabványosítási gyakorlatban. Kérjük, az instrukciókat tekintse meg a különálló eljárási lapon.

Az ön neve:

Dátum:

Kérjük, jelölje, az egyes megállapítások a hat szint (A1-től C2-ig) közül melyiket jellemzik a legjobban

Szám	Önértékelési megállapítások	CoE szint
1	Képes vagyok specifikus információt megtalálni egyszerű, hétköznapi anyagokban, mint például hirdetésekben, szórólapok, étlapok és menetrendek szövegében.	
2	Képes vagyok felismerni egy szöveg általános vezérgondolatát, de nem szükségképp minden részletéig.	
3	Képes vagyok hírek tartalmát és relevanciáját, továbbá szakmai témákban íródott cikkek és riportok széles körét gyorsan beazonosítani abból a célból, hogy eldöntsem, érdemesek-e ezek további tanulmányozásra.	
4	Képes vagyok felismerni a leghétköznapibb szituációkban szereplő egyszerű feliratokon olvasható ismerős neveket, szavakat és nagyon egyszerű kifejezéseket.	
5	Képes vagyok egy hosszabb vagy több rövidebb szöveg átvizsgálásával olyan specifikus információt fellelni, amelyre egy feladat megoldásához van szükségem.	
6	Képes vagyok megérteni és interpretálni az írott nyelv gyakorlatilag minden formáját, beleértve az elvont, szerkezetileg komplex, vagy rendkívül kollokvialis irodalmi és nem-irodalmi írásokat.	
7	Képes vagyok megérteni hétköznapi nyelven íródott rövid, egyszerű szövegeket.	
8	Képes vagyok megérteni a területemen kívül eső szakmai cikkeket, amennyiben használhatok szótárt a szakmai terminológia leellenőrzésére.	
9	Képes vagyok megérteni egyszerű, információközlő szövegek, valamint rövid, egyszerű leírások általános mondanivalóját, különösképpen, ha azok a szöveget magyarázó képeket tartalmaznak.	
10	Képes vagyok azonosítani egyszerű, írásos anyagokban, mint például levelekben, szórólapokon és eseményeket bemutató rövid újságcikkekben található specifikus információt.	
11	Bár széleskörű passzív szókincsrel rendelkezem, néha kevésbé hétköznapi szavak és kifejezések nehézséget okoznak számomra.	
12	Képes vagyok megérteni az ismerős témakörben írt standard rutin-leveleket és fax-üzeneteket.	
13	Képes vagyok megérteni a mindennapi élet során használt berendezésekre – mint például egy nyilvános telefon – vonatkozó egyszerű instrukciókat.	
14	Képes vagyok megérteni a nyilvános helyeken, például utcán, étteremben, vasútállomáson és munkahelyen látható hétköznapi táblákat és feliratokat.	
15	Képes vagyok megérteni az érdeklődési területemhez kapcsolódó témákban íródott, egyértelmű szövegeket.	
16	Képes vagyok a hétköznapi szövegekben, mint például levelekben, szórólapokon és rövid hivatalos dokumentumokban fellelhető, számomra szükséges általános információt megtalálni és megérteni.	

Charles Alderson

<i>Szám</i>	<i>Önértékelési megállapítások</i>	<i>CoE szint</i>
17	Képes vagyok megérteni rövid, egyszerű, például képeslapokon olvasható üzeneteket.	
18	Képes vagyok olyan, napjaink problémáival foglalkozó cikkeket és riportokat megérteni, amelyekben a cikk írója bizonyos véleményt vagy álláspontot foglal el.	
19	Képes vagyok azonosítani világosan megírt véleménykifejtő szövegek fő következtetéseit.	
20	Képes vagyok megérteni nagyon rövid, egyszerű szövegeket, kikövetkeztetni ismert neveket, szavakat illetve alapkifejezéseket, például a szöveg részeinek újraolvasása által.	
21	Képes vagyok elég jól megérteni magánlevelekben fellelhető események, érzések és kívánságok leírását ahhoz, hogy egy baráttal vagy ismerőssel levelezhessek.	
22	Képes vagyok megérteni a leghétköznapiabb szavakat, beleértve néhány ismert nemzetközi szót tartalmazó rövid, egyszerű szövegeket.	
23	Képes vagyok elolvasni az érdeklődési területeimhez kapcsolódó levelezést, és könnyen megértem annak alapvető értelmét.	
24	Képes vagyok megérteni a munkámhoz kapcsolódó rövid, egyszerű szövegeket.	
25	Sokféle szöveget vagyok képes könnyen, különböző sebességgel és különböző módokon elolvasni, az olvasás céljától és a szöveg típusától függően.	
26	Képes vagyok megérteni rövid, egyszerű magánleveleket.	
27	Képes vagyok követni rövid, egyszerű, írott instrukciókat, különösképpen, ha képeket tartalmaznak.	
28	Képes vagyok felismerni ismerős témákban íródott egyértelmű újságcikkek lényeges pontjait.	
29	Szótár alkalmankénti használatával képes vagyok bármilyen levelezést megérteni.	
30	Képes vagyok részleteiben megérteni egy új berendezésre vagy eljárásra vonatkozó hosszú, komplex instrukciókat, még a saját szakterületemen kívül eső témában is, amennyiben újraolvashatom a nehéz részeket.	
31	Képes vagyok megérteni berendezések világosan megírt, egyértelmű instrukcióit.	

C melléklet

5. Szabványosítási lépés

Ígérem, ez a szabványosítási gyakorlat utolsó lépése!

Kérjük, még a február 29-i megbeszélésen való részvételt megelőzően hajtsa végre a fenti eljárást. Amennyiben valamilyen okból kifolyólag nem tud megjelenni a megbeszélésen, kérjük, postázza a kitöltött eljárási nyomtatványokat az alábbi címre:

Charles Alderson,
The British Council
1068 Budapest
Benczúr utca 26

Két eljárás van. *Az első:* megtekinteni minden egyes olvasott szöveg értése feladatot (mellékelve), és minden egyes feladatelemre vonatkozóan eldönteni, hogy egy középszintű tanulónak – azaz egy olyan tanulónak, aki középszinten éppen hogy csak megfelelő – meg kell-e tudni azt válaszolnia.

Karikázza be az Igen-t vagy a Nem-t.

Az 'Igen' azt jelenti, hogy egy középszintű tanulónak el kellene tudni helyesen találni a kérdést.

A 'Nem' azt jelenti, hogy egy középszintű tanuló elvétheti a kérdést, és még mindig középszintűnek minősül (azaz a kérdés túl nehéz).

A *második* rész-eljárás valamelyest különbözik, és ez alkalommal a haladó szintű tanulóira irányul – azaz olyan tanulóira, akik éppen hogy csak elérik a haladó szintet. Újra vizsgáljon meg minden egyes feladatot, majd minden egyes feladatelemre vonatkozóan döntse el, hogy melyek azok a kérdések, amelyeket egy Haladó szintű tanuló megválaszolhat HELYTELENÜL, és még mindig Haladó szintűnek minősül.

Karikázza be az Igen-t vagy a Nem-t.

Az 'Igen' azt jelenti, hogy egy haladó szintű tanuló elvétheti ezt a kérdést, és még mindig Haladóknak minősül.

A 'Nem' azt jelenti, hogy egy haladó szintű tanulónak el kell tudni helyesen találni a kérdést ahhoz, hogy haladó szintűnek minősüljön.

5a. Szabványosítási lépés

NÉV:

Középszintű olvasás

Egyetért-e azzal, hogy egy, az Olvasás területén középszintű nyelvi készséggel rendelkező személynek helyes választ kellene tudni adni a következő feladatelemekre?

1. feladat „A smashing case”

1-es feladatelem	Igen	Nem
2-es feladatelem	Igen	Nem
3-as feladatelem	Igen	Nem
4-es feladatelem	Igen	Nem
5-ös feladatelem	Igen	Nem
6-os feladatelem	Igen	Nem

Charles Alderson

2. feladat „Girls-only success 'based on selection'”

1-es feladatelem	Igen	Nem
2-es feladatelem	Igen	Nem
3-as feladatelem	Igen	Nem
4-es feladatelem	Igen	Nem
5-ös feladatelem	Igen	Nem
6-os feladatelem	Igen	Nem
7-es feladatelem	Igen	Nem
8-as feladatelem	Igen	Nem
9-es feladatelem	Igen	Nem
10-es feladatelem	Igen	Nem

5b. Szabványosítási lépés

Haladó Olvasott szöveg értése

Kérjük, jelölje meg azokat a feladatelemeket, amelyek HELYTELEN megválaszolása ellenére is Haladó szintűnek tekintendő a válaszadó olvasott szöveg értése készségét.

1. feladat: „A smashing case”

1-es feladatelem	Igen	Nem
2-es feladatelem	Igen	Nem
3-as feladatelem	Igen	Nem
4-es feladatelem	Igen	Nem
5-ös feladatelem	Igen	Nem
6-os feladatelem	Igen	Nem

2. feladat: „Girls-only success 'based on selection'”

1-es feladatelem	Igen	Nem
2-es feladatelem	Igen	Nem
3-as feladatelem	Igen	Nem
4-es feladatelem	Igen	Nem
5-ös feladatelem	Igen	Nem
6-os feladatelem	Igen	Nem
7-es feladatelem	Igen	Nem
8-as feladatelem	Igen	Nem
9-es feladatelem	Igen	Nem
10-es feladatelem	Igen	Nem

(stb.)

D melléklet

Szabványosítási eljárás az angol vizsgareform során

A 3. lépés eredményei

No	Önértékelési megállapítások	Council of Europe szint	DIALANG szint	Szint-közép-érték a 3. lépés során
1	Képes vagyok specifikus információt megtalálni egyszerű, hétköznapi anyagokban, mint például hirdetésekben, szórólapok, étlapok és menürendek szövegében.	A2	A1	A2
2	Képes vagyok felismerni egy szöveg általános vezérgondolatát, de nem szükségképp minden részletéig.	B1	A2	B1
3	Képes vagyok hírek tartalmát és relevanciáját, továbbá szakmai témákban íródott cikkek és riportok széles körét gyorsan beazonosítani abból a célból, hogy eldöntsem, érdemesek-e ezek további tanulmányozásra.	B2	B2	C1
4	Képes vagyok felismerni a leghétköznapibb szituációkban szereplő egyszerű feliratokon olvasható ismerős neveket, szavakat és nagyon egyszerű kifejezéseket.	A1	A1	A1
5	Képes vagyok egy hosszabb vagy több rövidebb szöveg átvizsgálásával olyan specifikus információt fellelni, amelyre egy feladat megoldásához van szükségem.	B1	B1	B1
6	Képes vagyok megérteni és interpretálni az írott nyelv gyakorlatilag minden formáját, beleértve az elvont, szerkezetileg komplex, vagy rendkívül kollokvialis irodalmi és nem-irodalmi írásokat.	C2	C1	C2
7	Képes vagyok megérteni hétköznapi nyelven íródott rövid, egyszerű szövegeket.	A2	A2	A2
8	Képes vagyok megérteni a területemen kívül eső szakmai cikkeket, amennyiben használhatók szótárt a szakmai terminológia leellenőrzésére.	B2	B2	C1
9	Képes vagyok megérteni egyszerű, információközlő szövegek, valamint rövid, egyszerű leírások általános mondanivalóját, különösképpen, ha azok a szöveget magyarázó képeket tartalmaznak.	A1	A2	A1
10	Képes vagyok azonosítani egyszerű, írásos anyagokban, mint például levelekben, szórólapokon és eseményeket bemutató rövid újságcikkekben található specifikus információt.	A2	A2	A2
11	Bár széleskörű passzív szókincsel rendelkezem, néha kevésbé hétköznapi szavak és kifejezések nehézséget okoznak számomra.	B2	B2	B2
12	Képes vagyok megérteni az ismerős témakörben írt standard rutinleveleket és faxüzeneteket.	A2	A2	B1
13	Képes vagyok megérteni a mindennapi élet során használt berendezésekre – mint például egy nyilvános telefon – vonatkozó egyszerű instrukciókat.	A2	A2	A2
14	Képes vagyok megérteni a nyilvános helyeken, például utcán, étteremben, vasútállomáson és munkahelyen látható hétköznapi táblákat és feliratokat.	A2	A1	A2

No	Önértékelési megállapítások	Council of Europe szint	DIALANG szint	Szint-közép-érték a 3 Lépés során
15	Képes vagyok megérteni az érdeklődési területemhez kapcsolódó témákban íródott, egyértelmű szövegeket.	B1	A2	B1
16	Képes vagyok a hétköznapi szövegekben, mint például levelekben, szórólapokon és rövid hivatalos dokumentumokban fellelhető, számomra szükséges általános információt megtalálni és megérteni.	B1	A2	B1
17	Képes vagyok megérteni rövid, egyszerű, például képeslapokon olvasható üzeneteket.	A1	A1	A2
18	Képes vagyok olyan, napjaink problémáival foglalkozó cikkeket és riportokat megérteni, amelyekben a cikk írója bizonyos véleményt vagy álláspontot foglal el.	B2	B1	B2
19	Képes vagyok azonosítani világosan megírt véleménykifejtő szövegek fő következtetéseit.	B1	B1	B1
20	Képes vagyok megérteni nagyon rövid, egyszerű szövegeket, kikövetkeztetni ismert neveket, szavakat illetve alapkifejezéseket, például a szöveg részeinek újraolvasása által.	A1	A1	A1
21	Képes vagyok elég jól megérteni magánlevelekben fellelhető események, érzések és kívánságok leírását ahhoz, hogy egy barátal vagy ismerőssel levelezhessek.	B1	B1	B1
22	Képes vagyok megérteni a leghétköznapiabb szavakat, beleértve néhány ismert nemzetközi szót tartalmazó rövid, egyszerű szövegeket.	A2	A2	A2
23	Képes vagyok elolvasni az érdeklődési területeimhez kapcsolódó levelezést, és könnyen megértem annak alapvető értelmét.	B2	A2	B2
24	Képes vagyok megérteni a munkámhoz kapcsolódó rövid, egyszerű szövegeket.	A2	A1	A2
25	Sokféle szöveget vagyok képes könnyen, különböző sebességgel és különböző módokon elolvasni, az olvasás céljától és a szöveg típusától függően.	B2	B1	C1
26	Képes vagyok megérteni rövid, egyszerű magánleveleket.	A2	A1	A2
27	Képes vagyok követni rövid, egyszerű, írott instrukciókat, különösképpen, ha képeket tartalmaznak.	A1	A1	A1
28	Képes vagyok felismerni ismerős témákban íródott egyértelmű újságcikkek lényeges pontjait.	B1	A2	B1
29	Szótár alkalmankénti használatával képes vagyok bármilyen levelezést megérteni.	C1	B1	C1
30	Képes vagyok részleteiben megérteni egy új berendezésre vagy eljárásra vonatkozó hosszú, komplex instrukciókat, még a saját szakterületemen kívül eső témában is, amennyiben újraolvashatom a nehéz részeket.	C1	C1	C1
31	Képes vagyok megérteni berendezések világosan megírt, egyértelmű instrukcióit.	B1	A2	B1

E melléklet

A szabványosítási gyakorlat eredményei, 1–4. lépések

<i>Feladat</i>	<i>Feladat- író</i>	<i>1. lépés</i>	<i>2. lépés</i>	<i>Logit</i>	<i>Jósági mutató %</i>	<i>4. lépés BAI%</i>	<i>4. lépés BAA%</i>	<i>4. lépés BAI Pontszám</i>	<i>4. lépés BAA Pontszám</i>
1. feladat (3:5) k=6	2	3	C1	+0,71	20%	29	63	2 33%	4 67%
2. feladat (4:5) k=10	3	2	B2	+0,601	26%	33	66	3 30%	8 80%
3. feladat (3:4) k=10	1	1	A2	-1,91	68%	73	90	6 60%	9 90%
4. feladat (1:3) k=7	2	2	B1	-0,187	54%	52	79	3 43%	6 86%
5. feladat (1:2) k=10	1	2	B1	-2,298	83%	69	88	6 60%	10 100%
6. feladat referencia k=10	CITO A2	2	B2	-0,826	65%	51	81	4 40%	9 90%
7. feladat (4:3) k=8	3	3	B2	-0,505	44%	29	61	2 25%	4 50%
8. feladat 1.referencia k=15	3	3	C1	+1,891	24%	24	64	2 13%	10 67%
9. feladat 2.referencia k=4	3	1	A2	-0,635	62%	69	98	3 75%	4 100%
10. feladat (4:2) k=5	1	2	B1	+0,394	29%	60	91	3 60%	5 100%
11. feladat (4:4) k=7	2	2	B1	-0,733	49%	62	97	3 43%	7 100%
12. feladat (3:3) k=10	1	1	A2	-2,428	74%	79	94	7 70%	9 90%
13. feladat (3:2) k=8	1	1	A2	-2,624	81%	76	93	5 63%	8 100%
14. feladat (2:2) k=5	1	1	A1	-1,64	76%	74	94	4 80%	5 100%
15. feladat (2:3) k=9	2	2	B1	+0,349	47%	57	81	6 67%	8 89%

Charles Alderson

F melléklet

A szintek bírálatának eltérései

1. lépés: magyar szintek, Alap (Basic), Közép (Intermediate) és Haladó (Advanced)

	<i>Középérték</i>	<i>Szórás</i>
1. Tesztfeladat	Haladó	Közép 2, Haladó 7
2. Tesztfeladat	Közép	Közép 5, Haladó 4
3. Tesztfeladat	Alap	Alap 7, Közép 2
4. Tesztfeladat	Közép	Közép 7, Haladó 2
5. Tesztfeladat	Közép	Alap 4, Közép 5
6. Tesztfeladat	Közép	Közép 7, Haladó 2
7. Tesztfeladat	Haladó	Közép 3, Haladó 6
8. Tesztfeladat	Haladó	Haladó 9
9. Tesztfeladat	Alap	Alap 5, Közép 4
10. Tesztfeladat	Közép	Közép 9
11. Tesztfeladat	Közép	Alap 3, Közép 6
12. Tesztfeladat	Alap	Alap 7, Közép 2
13. Tesztfeladat	Alap	Alap 7, Közép 2
14. Tesztfeladat	Alap	Alap 7, Közép 2
15. Tesztfeladat	Közép	Alap 1, Közép 8

2. lépés: Az Európa Tanács szintjei

	<i>Medián</i>	<i>Szórás</i>
1. Tesztfeladat	C1	B2 4, C1 5
2. Tesztfeladat	B2	B1 2, B2 4, C1 3
3. Tesztfeladat	A2	A2 8, B1 1
4. Tesztfeladat	B1	B1 5, B2 4
5. Tesztfeladat	B1	A2 3, B1 6
6. Tesztfeladat	B2	B1 4, B2 3, C1 2
7. Tesztfeladat	B2	B1 1, B2 4, C1 3, C2 1
8. Tesztfeladat	C1	B2 4, C1 5
9. Tesztfeladat	A2	A1 1, A2 4, B1 4
10. Tesztfeladat	B1	B1 8, B2 1
11. Tesztfeladat	B1	A2 3, B1 6
12. Tesztfeladat	A2	A1 2, A2 5, B1 2
13. Tesztfeladat	A2	A1 3, A2 5, B2 1
14. Tesztfeladat	A1	A1 5, A2 3, B1 1
15. Tesztfeladat	B1	A2 1, B1 6, B2 2

Teljesítményszintek az angol nyelvi érettségi kipróbálásán

A 3. lépés eredményei

<i>Szám</i>	<i>Önértékelési megállapítások</i>	<i>Az eljárás medián szintje</i>	<i>A szintek szórása</i>
1	Képes vagyok specifikus információt megtalálni egyszerű, hétköznapi anyagokban, mint például hirdetésekben, szórólapok, étlapok és menetredek szövegében.	A2	A1 2, A2 7
2	Képes vagyok felismerni egy szöveg általános vezérgondolatát, de nem szükségképp minden részletéig.	B1	A2 2, B1 6, B2 1
3	Képes vagyok hírek tartalmát és relevanciáját, továbbá szakmai témákban íródott cikkek és riportok széles körét gyorsan beazonosítani abból a célból, hogy eldöntsem, érdemesek-e ezek további tanulmányozásra.	C1	B2 3, C1 5, C2 1
4	Képes vagyok felismerni a leghétköznapiabb szituációkban szereplő egyszerű feliratokon olvasható ismerős neveket, szavakat és nagyon egyszerű kifejezéseket.	A1	A1 9
5	Képes vagyok egy hosszabb vagy több rövidebb szöveg átvizsgálásával olyan specifikus információt fellelni, amelyre egy feladat megoldásához van szükségem.	B1	A2 1, B1 7, B2 1
6	Képes vagyok megérteni és interpretálni az írott nyelv gyakorlatilag minden formáját, beleértve az elvont, szerkezetileg komplex, vagy rendkívül kollokvialis irodalmi és nem-irodalmi írásokat.	C2	C2 9
7	Képes vagyok megérteni hétköznapi nyelven íródott rövid, egyszerű szövegeket.	A2	A2 7, B1 2
8	Képes vagyok megérteni a területemen kívül eső szakmai cikkeket, amennyiben használhatok szótárt a szakmai terminológia leellenőrzésére.	C1	B2 3, C1 5, C2 1
9	Képes vagyok megérteni egyszerű, információközlő szövegek, valamint rövid, egyszerű leírások általános mondanivalóját, különösképpen, ha azok a szöveget magyarázó képeket tartalmaznak.	A1	A1 5, A2 4
10	Képes vagyok azonosítani egyszerű, írásos anyagokban, mint például levelekben, szórólapokon és eseményeket bemutató rövid újságcikkekben található specifikus információt.	A2	A2 7, B1 2
11	Bár széleskörű passzív szókincssel rendelkezem, néha kevésbé hétköznapi szavak és kifejezések nehézséget okoznak számomra.	B2	B1 1, B2 6, C1 2
12	Képes vagyok megérteni az ismerős témakörben írt standard rutin-leveleket és faxüzeneteket.	B1	A2 2, B1 7
13	Képes vagyok megérteni a mindennapi élet során használt be rendezésekre – mint például egy nyilvános telefon – vonatkozó egyszerű instrukciókat.	A2	A1 1, A2 7, B1 1
14	Képes vagyok megérteni a nyilvános helyeken, például utcán, étteremben, vasútállomáson és munkahelyen látható hétköznapi táblákat és feliratokat.	A2	A1 4, A2 5
15	Képes vagyok megérteni az érdeklődési területemhez kapcsolódó témákban íródott, egyértelmű szövegeket.	B1	B1 7, B2 2
16	Képes vagyok a hétköznapi szövegekben, mint például levelekben, szórólapokon és rövid hivatalos dokumentumokban fellelhető, számomra szükséges általános információt megtalálni és megérteni.	B1	A2 1, B1 8

Charles Alderson

<i>Szám</i>	<i>Önértékelési megállapítások</i>	<i>Az eljárás medián szintje</i>	<i>A szintek szórása</i>
17	Képes vagyok megérteni rövid, egyszerű, például képeslapokon olvasható üzeneteket.	A2	A1 3, A2 5
18	Képes vagyok olyan, napjaink problémáival foglalkozó cikkeket és riportokat megérteni, amelyekben a cikk írója bizonyos véleményt vagy álláspontot foglal el.	B2	B2 8, C1 1
19	Képes vagyok azonosítani világosan megírt véleménykifejtő szövegek fő következtetéseit.	B1	B1 5, B2 4
20	Képes vagyok megérteni nagyon rövid, egyszerű szövegeket, kikövetkeztetni ismert neveket, szavakat illetve alapkifejezéseket, például a szöveg részeinek újraolvasása által.	A1	A1 9
21	Képes vagyok elég jól megérteni magánlevelekben fellelhető események, érzések és kívánságok leírását ahhoz, hogy egy barátal vagy ismerőssel levelezhessek.	B1	A2 2, B1 7
22	Képes vagyok megérteni a leghétköznapibb szavakat, beleértve néhány ismert nemzetközi szót tartalmazó rövid, egyszerű szövegeket.	A2	A1 3, A2 6
23	Képes vagyok elolvasni az érdeklődési területeimhez kapcsolódó levelezést, és könnyen megértem annak alapvető értelmét.	B2	A2 1, B1 2, B2 6
24	Képes vagyok megérteni a munkámhoz kapcsolódó rövid, egyszerű szövegeket.	A2	A2 5, B1 4
25	Sokféle szöveget vagyok képes könnyen, különböző sebességgel és különböző módokon elolvasni, az olvasás céljától és a szöveg típusától függően.	C1	B2 2, C1 5, C2 2
26	Képes vagyok megérteni rövid, egyszerű magánleveleket.	A2	A1 1, A2 8
27	Képes vagyok követni rövid, egyszerű, írott instrukciókat, különösképpen, ha képeket tartalmaznak.	A1	A1 6, A2 3
28	Képes vagyok felismerni ismerős témákban íródott egyértelmű újságcikkek lényeges pontjait.	B1	B1 9
29	Szótár alkalmankénti használatával képes vagyok bármilyen levelezést megérteni.	C1	B2 1, C1 8
30	Képes vagyok részleteiben megérteni egy új berendezésre vagy eljárásra vonatkozó hosszú, komplex instrukciókat, még a saját szakterületemen kívül eső témában is, amennyiben újraolvashatom a nehéz részeket.	C1	B2 1, C1 6, C2 2
31	Képes vagyok megérteni berendezések világosan megírt, egyértelmű instrukcióit.	B1	A2 2, B1 6, B2 1

G melléklet

A 4-es és 5-ös lépések összevetése

k = az elemek száma, az összes érték medián

A zárójelben található értékek (3:5) az eredeti tesztfüzetet és részt mutatják (3. Füzet 5. Rész)

<i>Feladat</i>	<i>4a lépés BAI%</i>	<i>4a lépés BAA%</i>	<i>4b lépés BAI Pontszám</i>	<i>4b lépés BAA Pontszám</i>	<i>5a lépés (Közép)</i>	<i>5b lépés (Haladó)</i>
1. feladat (3:5) k=6	29	63	2 33%	4 67%	2	4
2. feladat (4:5) k=10	33	66	3 30%	8 80%	4	10
3. feladat (3:4) k=10	73	90	6 60%	9 90%	8	10
4. feladat (1:3) k=7	52	79	3 43%	6 86%	4	7
5. feladat (1:2) k=10	69	88	6 60%	10 100%	9	10
6. feladat Referencia k=10	51	81	4 40%	9 90%	6	8
7. feladat (4:3) k=8	29	61	2 25%	4 50%	2	5
8. feladat Referencia1 k=15	24	64	2 13%	10 67%	4	12
9. feladat Referencia 2 k=4	69	98	3 75%	4 100%	4	4
10. feladat (4:2) k=5	60	91	3 60%	5 100%	4	5
11. feladat (4:4) k=7	62	97	3 43%	7 100%	5	7
12. feladat (3:3) k=10	79	94	7 70%	9 90%	8	10
13. feladat (3:2) k=8	76	93	5 63%	8 100%	8	8
14. feladat (2:2) k=5	74	94	4 80%	5 100%	5	5
15. feladat (2:3) k=9	57	81	6 67%	8 89%	6	8

Charles Alderson

A téma magyar nyelvű szakirodalmának egységes terminológiáját segítő egy szójegyzék összeállítását tartottam szükségesnek:

Glossary

Pass mark	–	Megfelelt jegy
Pass rate	–	Megfelelt szint
Level of performance	–	Teljesítményszint
Level of achievement	–	Eredményszint
Judgmental method	–	„Szemrevételezési” módszer
Test item	–	Teszt-elem, item
Anchor item	–	Referencia-elem
Facility value	–	Nehézségi érték
Raw score	–	Nyers pontszám
Rank order correlation (Spearman rho)	–	Rangkorreláció (Spearman-féle)
Item response theory	–	Valószínűségi tesztelmélet
Cut-off point	–	Határpont