

Szegedi Tudományegyetem, Bölcsészettudományi Kar
Neveléstudományi Doktori Iskola

VÍGH TIBOR

**AZ IDEGEN NYELVI ÉRETTSÉGI MŰKÖDÉSE ÉS HATÁSA
A TANULÓI TELJESÍTMÉNYEK ÉS A TANÁRI NÉZETEK
TÜKRÉBEN**

PhD-értekezés



Témavezető: Dr. Vidákovich Tibor

Pedagógiai értékelés
doktori program

Szeged, 2010

TARTALOMJEGYZÉK

BEVEZETÉS	4
1. A NYELVI MÉRÉS ÉS ÉRTÉKELÉS NEMZETKÖZI TENDENCIÁI	8
1.1. A kommunikatív kompetencia mérésének és értékelésének lehetőségei és problémái	8
1.1.1. A nyelvi mérés és értékelés fogalma	8
1.1.2. A kommunikatív kompetencia szerkezetének modelljei	9
1.1.3. A kommunikatív tesztek minőségi mutatói	15
1.1.4. A kommunikatív tesztek funkciói és típusai	21
1.2. A kommunikatív nyelvi kompetenciák szintjeinek mérése és értékelése	26
1.2.1. A kommunikatív nyelvi kompetenciák szintjei	26
1.2.2. A nyelvi szintek validálási eljárásai, az európai projektek eredményei	28
1.2.3. A kommunikatív nyelvi készségek és az IRT-modellek	31
1.3. A kommunikatív tesztelés hatása a nyelvoktatásra	33
1.3.1. A tesztelés fogalma	33
1.3.2. A tesztelés folyamatát leíró elméleti modellek és tényezők	36
1.3.3. A tesztelés folyamatát leíró empirikus kutatások	43
2. A NYELVI MÉRÉS ÉS ÉRTÉKELÉS HAZAI JELLEMZŐI	47
2.1. A tudáskonceptió változása és tantervi megjelenése	47
2.2. A nyelvoktatás eredményességének vizsgálata	48
2.3. A kommunikatív nyelvi kompetenciák mérése és értékelése az idegen nyelvi érettségin ...	53
2.3.1. A nyelvi érettségi illeszkedése a nyelvtudás-modellhez	54
2.3.2. A nyelvi érettségi illeszkedése a kommunikatív tesztek minőségi mutatóihoz	56
2.3.3. A nyelvi érettségi illeszkedése a kommunikatív tesztek funkcióihoz és típusaihoz	59
2.3.4. Teljesítmények és a diákok nyelvtudás-szintje az idegen nyelvi érettségi vizsgákon	61
3. A NYELVI MÉRÉS ÉS ÉRTÉKELÉS MECHANIZMUSAINAK EMPIRIKUS VIZSGÁLATA	66
3.1. A vizsgálat elméleti háttere és célterületei	66
3.1.1. A vizsgálat elméleti modellje	66
3.1.2. A vizsgálat elméleti modellje az idegen nyelvi érettségi példáján	69
3.2. Az empirikus vizsgálat céljai, kérdései és hipotézisei	72
3.3. A vizsgálat módszerei	76
3.3.1. A vizsgaeredmények elemzésének módszerei	76
3.3.2. A tanári kérdőíves vizsgálat módszerei	80

4. AZ IDEGEN NYELVI ÉRETTSÉGI ELEMZÉSE A VIZSGAEREDMÉNYEK ALAPJÁN.....	96
4.1. Az idegen nyelvi érettségin elért teljesítmények	96
4.1.1. A vizsgázók összetétele a 2007., 2008. és 2009. évi idegen nyelvi érettségin	96
4.1.2. Teljesítmények és a teljesítményt meghatározó változók hatása a középszintű érettségin.....	100
4.1.3. Teljesítmények és a teljesítményt meghatározó változók hatása az emelt szintű érettségin.....	104
4.1.4. Az idegen nyelvi érettségi vizsga jelenlegi működésének jellemzői, vizsgálandó területei	106
4.2. Az idegen nyelvi érettségi belső és külső validitása.....	107
4.2.1. A német nyelvi érettségi feladatok adatbázisának elemzése	107
4.2.2. Az adatbázisok összekapcsolásának, a feladatbank kiépítésének problémái	114
4.3. Az íráskészség mérési-értékelési eljárásainak minőségi mutatói	117
4.3.1. Az emelt szintű érettségi íráskészség vizsgarész mérési-értékelési eljárásai.....	118
4.3.2. A középszintű érettségi íráskészség vizsgarész mérési-értékelési eljárásai.....	127
5. AZ IDEGEN NYELVI ÉRETTSÉGI ELEMZÉSE A TANÁROK VISSZAJELZÉSE ALAPJÁN.....	132
5.1. A tanárok vizsgával kapcsolatos nézetei és attitűdjei.....	132
5.2. Az idegen nyelvi érettségi hatása a nyelvoktatási folyamatra a tanárok nézetei alapján....	137
5.2.1. A tanárok nyelvoktatási folyamatát befolyásoló tényezők	138
5.2.2. A tanárok nyelvoktatási folyamatának jellemzői az érettségi alapján	139
5.3. Az idegen nyelvi érettségi hatása az osztálytermi mérési-értékelési eljárásokra a tanárok nézetei alapján	145
5.4. A vizsgával kapcsolatos nézetek és attitűdök hatása a nyelvoktatás folyamatára, mérési-értékelési eljárásaira	148
6. ÖSSZEGZÉS, KÖVETKEZTETÉSEK ÉS JAVASLATOK.....	152
KÖSZÖNETNYILVÁNÍTÁS	157
IRODALOM.....	158
ÁBRÁK JEGYZÉKE	169
TÁBLÁZATOK JEGYZÉKE.....	170
MELLÉKLETEK JEGYZÉKE.....	172
MELLÉKLETEK.....	173

BEVEZETÉS

A pedagógiai értékelés oktatási rendszerben betöltött szerepét és visszacsatoló mechanizmusait a kibernetika segítségével megalkotott rendszerszemléletű tanítás-tanulás modellek pontosan meghatározzák (Nagy, 1979; Báthory, 1985). A pedagógiai törekvéseknek megfelelő innovációk és reformok, melyek az oktatási rendszert működőképesebbé és hatékonyabbá kívánják tenni, minden esetben visszavezethetők a visszacsatolás jelentőségének felismerésére. Az oktatás eredményességének növelése új visszajelzések beépítésével érhető el. Ezekben az innovációs folyamatokban különösen fontos szerepe van az oktatás bemeneti szabályozása mellett a közoktatásban megfelelően működtetett vizsgáknak, melyek az iskolarendszer kontrollját is jelenthetik és a tanítási-tanulási folyamatra erős hatásuk lehet. Ezt felismerve régi reformstratégia az oktatásban lezajló változások elindításához a mérés és értékelés megváltoztatása (Noble és Smith, 1994). E nagyléptékű és látványos változást előidéző reformok mellett lényeges szerepe van a vizsgán elért eredmények vizsgálatából és a tanítási-tanulási folyamatban résztvevők tapasztalatainak feltárásából kiinduló visszajelzéseknek, melyek hozzájárulnak a vizsga továbbfejlesztéséhez.

A vizsgaeredményekből kiinduló visszacsatolási kör az oktatási rendszer három szintjéről nyújt információt. Egyrészt az eredményekkel számszerűsíthető a tanulók tudása, amely alapot szolgáltat a különböző döntések meghozatalához, és ezek befolyásolhatják a tanulók előmenetelét, továbbhaladását, iskolafokozat-váltását. Másrészt optimalizálhatja az oktatási folyamatot, meghatározhatja, hogy a tanárok a vizsga működéséből milyen következtetéseket vonnak le a saját tanítási gyakorlatukra vonatkozóan, alapja lehet a folyamatban alkalmazott módszerek, eszközök fejlesztéséhez. Harmadrészt pedig a vizsga számszerű eredményei és a meghozott döntések alapot szolgáltatnak a kitűzött célok megvalósulásának, relevanciájának értékeléséhez. A vizsgán elért eredmények újabb igényeket generálhatnak, amelyek befolyásolhatják a vizsgafejlesztési folyamatot. Mindezek mellett különböző felmérésekkel lehetőség van arra, hogy a pedagógiai munkában érdekelt személyek tapasztalatainak és véleményeinek felméréséből és értelmezéséből adjunk visszajelzést a vizsgafejlesztés számára. Ezekkel az információkkal lehetőség nyílik arra, hogy a vizsgafejlesztők megismerjék, milyen a vizsga hatása a tanítási-tanulási folyamatra, és ezek összhangban vannak-e a szándékaikkal.

A vizsgák befolyásolhatják az osztálytermi folyamatokat, tartalmi és módszertani modernizációt indíthatnak el. Ennek oka, hogy a vizsgára való felkészülés központi jelentőségűvé válik a tanárok gondolkodásában, napi munkájában, meghatározza a diákok felkészülési folyamatát, befolyásolja eredményességüket. A vizsgák szerepét a központilag kevésbé szabályozott, kimenet-orientált, angolszász iskolarendszerek példája is mutatja, melyekben az egységes követelményekkel rendelkező magas presztízsű vizsgák jelentős hatást gyakorolnak a tanítási-tanulási folyamatra. Ez a több évtizede ismert jelenség és az ezzel kapcsolatos kérdések gyakran merülnek fel újra új mérési-értékelési rendszerek, érettségi, felvételi vizsgák bevezetésével vagy módosításával. Ezért részletesen kell ismerni a vizsga alapján elinduló hatások folyamatát, működését, mechanizmusait, a folyamatot befolyásoló tényezőket.

A fenti kérdések vizsgálata hazánkban a 2005-ben bevezetett új érettségi vizsga esetében is relevánsak. A vizsgareform azért volt szükséges, mert a korábbi közoktatást lezáró vizsga több mint másfél évszázadig csaknem változatlan volt, a nyolcvanas években jelent meg egyértelműen az igény a vizsga továbbfejlesztésére, majd a kilencvenes évek közepén elinduló reform eredményeként vezették be a kétszintű érettségit (Csapó, 2008). A vizsga egységes a gimnáziumban és a szakközépiskolákban érettségizők számára, közép- és emelt szintből áll, és differenciált: a diákok a kötelező tárgyak

(magyar nyelv és irodalom, matematika, történelem és egy idegen nyelv) mellett körülbelül 130 tárgyból választhatnak; a vizsgán megjelennek a standardizációs törekvések (Horváth és Lukács, 2006). E vizsgáról feltételezik, hogy erős hatást gyakorol a közoktatásra és a tanítási-tanulási folyamatban tartalmi és módszertani modernizációt indít el (Kósa és Simon, 2006). Ugyanakkor kevésbé tisztázott a két vizsgaszint egymáshoz való viszonya, a szintek közötti tartalmi és mennyiségbeli különbségek mértéke. Szükséges a tartalmi standardok továbbfejlesztése, a vizsgaszintek összekapcsolása nemzetközi mérési rendszerekkel. Mindezek mellett az emelt szintű vizsgát kevesen választják, elértéktelenedik (Csapó, 2008). Nem rendelkezünk pontos információkkal arról, hogy az érettséginek milyen a hatása a közoktatásra, illetve mennyiben szolgálja az oktatási folyamat változásait. Az eredmények monitorozása, a vizsgafunkciók működésének ellenőrzése, a vizsga hatásmechanizmusainak feltárása és az ezek alapján történő vizsgafejlesztés a jelenlegi gyakorlatra kevésbé jellemző. A vizsga minőségének javítása érdekében szükségesek a fejlesztő kutatások. Fontos vizsgálandó kérdés, hogyan kell e vizsgát úgy továbbfejlesztetni, hogy kimutatható hatása a fejlesztők szándékával összhangban legyen. Ehhez két feltételt kell biztosítani: egyrészt a vizsgaeredmények alapján ismerni kell a fejlesztés lehetőségeit, másrészt a folyamatban résztvevők véleménye alapján fel kell tárni, hogy a vizsga az oktatási folyamat mely területeire hat.

Az értekezésünk célja, hogy az érettségi kötelező vizsgatárgyai közül az idegen nyelvi érettségi példáján keresztül mutassuk be a mérés-értékelés vizsgafejlesztésben betöltött szerepét, a vizsga közoktatásra gyakorolt hatását. A tantárgycsoport-jellegű nyelvi érettségi vizsgák működésének vizsgálatát négy körülmény teszi indokolttá. Egyrészt az idegen nyelvek tudása iránti igény hazánkban az 1990-es években jelentősen megnövekedett, majd az Európai Unió tagóságunkkal a használható nyelvtudás presztíze tovább nőtt. Másrészt a nyelvi érettségik fejlesztési munkálatai során (lásd Alderson, Nagy és Öveges, 2000; Einhorn, 2004) megjelent a törekvés arra, hogy a vizsga igazodjon a nyelvoktatás és a nyelvi tudásszint-mérés egységesen elfogadott közös európai keretrendszeréhez. A vizsgakövetelményekben a *Közös Európai Referenciakeret (KER, 2002)* szintrendszere alapján definiálják a mérendő nyelvtudást és a két szintet. Harmadrészt a vizsga elhelyezhető és elvileg összemérhető a közoktatási rendszeren kívüli idegennyelv-tudást mérő vizsgákkal, mivel az emelt szintű vizsgán elért eredmény alapján nyelvvizsga-bizonyítvány is szerezhető. Negyedrészt az idegen nyelvi mérés és értékelés elmélete és gyakorlata mind a nemzetközi, mind a hazai viszonylatban jelentős változásokat mutat. Ide sorolható a kompetencia-értékelési rendszerek előtérbe kerülése és a kommunikatív tesztelés alapelveinek követése, amelyek a magyar közoktatási rendszerben a 2000-es évektől elinduló hazai, reprezentatív mintán végzett nyelvtudást mérő kutatásokban (például Csapó, 2001; Nikolov, 2003; Csapó és Nikolov, 2009; Nikolov és Józsa, 2003, 2006) és a 2005-ben bevezetett idegen nyelvi érettségiben is megvalósulnak.

Az idegen nyelvi érettségi számos értékelés- és tesztelési kérdést, vizsgálandó kutatási problémát vet fel. A vizsga működtetésének sok eleme, mint például a nyelvi készségeket feladattipológiailag kötetlenül mérő, autentikus szövegekre és szituációkra épülő, egynyelvű feladatok elkészítése, a vizsga lebonyolítása, kiértékelése kevés minőségbiztosítási elemet tartalmaz. Ezért központi jelentőségű kutatási feladat a különböző évekből származó vizsgatesztek reliabilitásának, validitásának és ekvivalenciájának vizsgálata. Mindezek mellett jelenleg még nem áll rendelkezésre olyan kutatási eredmény, amellyel az idegen nyelvi érettségi két szintjének egymáshoz és a Referenciakeret szintrendszeréhez történő illeszkedését empirikusan igazolhatnánk. Ezen kívül nem rendelkezünk adatokkal arról, hogy a vizsga a nyelvoktatás mely területein, milyen jellegű hatást fejt ki.

A magyar közoktatási rendszerben végzett idegen nyelvi felmérések és az érettségi vizsga a nyelvi tudásszint-mérés nemzetközi tendenciáihoz három szempontból illeszkedik. (1) A kommunikatív tesztelés alapelveinek alkalmazásával kapcsolódik a nyelvtudás mérésének és értékelésének problémáihoz. Ide tartozik például a kommunikatív nyelvi kompetenciák értelmezése és modellezése, valamint a kommunikatív tesztek jószágmutatóinak, a tesztípusok és -funkciók megfelelő alkalmazásának kérdésköre. (2) Az európai nyelvtudásmérési projektekből gyakran vizsgált kutatási terület a nyelvtudást mérő vizsgák szintjeinek bemérése és validálása a Referenciakeret szintrendszeréhez és ezek alapján nyelvi feladatbank létrehozása. (3) A nemzetközi szakirodalomban intenzíven vizsgált terület a kommunikatív tesztelés hatása a nyelvtanításra, a tesztelés folyamatának modellezése és empirikus vizsgálata. A nemzetközi szakirodalom feltárása számos, a hazai kontextusban empirikusan vizsgálható kérdéseket határoz meg. Ugyanakkor e kérdéseket az elmúlt évtizedben a hazai szakirodalomban nem, vagy nem eléggé hangsúlyozzák és vizsgálják.

Az értekezés 1. fejezetében tárgyaljuk az idegen nyelvi mérés és értékelés nemzetközi tendenciáit. Ebben arra törekszünk, hogy az 1970-es években megjelenő, a '80-as években kialakuló és a 2000-es évektől dinamikus fejlődésnek induló kommunikatív tesztelést minél gazdagabb kontextusban mutassuk be. Foglalkozunk a kommunikatív nyelvi kompetenciák összetevőinek és szintjeinek mérési és értékelési lehetőségeivel, problémáival. Elhelyezzük a kommunikatív tesztelést a pedagógiai értékelés új irányzatában, a kompetenciaértékelés rendszerében. A kommunikatív nyelvi kompetenciák mérhetővé tételében bemutatjuk a valószínűségi tesztelméleti modellek alkalmazhatóságát, amely a nyelvtudást mérő vizsgák szintjeinek a Referenciakerethez való illesztésében vált domináns eljárássá. Ebben a fejezetben a nyelvtudást mérő vizsgák hatásának elméleti és gyakorlati kérdéseivel is foglalkozunk.

A 2. fejezet a nyelvi mérés és értékelés hazai jellemzőit mutatja be három csomópont köré csoportosítva. Egyrészt azzal foglalkozunk, hogy e terület hogyan értékelődött föl az elmúlt évtizedben. Másrészt célunk a hazai, reprezentatív mintán végzett idegen nyelv-tudás vizsgálatok tanulságainak szintetizáló áttekintése, mivel vizsgálataink a hazai kontextusba integrálódhatnak. Harmadrészt az idegen nyelvi érettségi jellemzőit mutatjuk be az első fejezetben megalapozott elméleti keretnek megfelelően, és elhelyezzük a vizsgát a kommunikatív tesztelés rendszerében.

A 3. fejezetben két célt tűzünk ki: egyfelől szintetizáljuk az első két fejezetet egy elméleti modellben, másfelől kijelöljük abban és indokoljuk a vizsgálandó célterületeket, bemutatjuk az empirikus vizsgálataink módszereit. A modell magában foglalja a vizsga tanítási-tanulási folyamatra gyakorolt hatásának mechanizmusát, amely szerint a vizsga először az oktatási rendszerben résztvevők nézeteit és attitűdjeit befolyásolja, amelyek meghatározzák az oktatási folyamatot, a tanulás tartalmát és minőségét. A modell továbbá kétféle visszajelentési kört tartalmaz: az egyik a vizsgaeredményekből indul ki, és a vizsgafejlesztési folyamatra, a célok és az igények változásaira irányul. A másik körben különböző felmérések segítségével a tanítási-tanulási folyamatban résztvevők vizsgával kapcsolatos nézeteiről és attitűdjeiről szerezhetünk információkat. Ez utóbbi visszajelentési kör befolyásolja a vizsgafejlesztési folyamatot és elősegíti az eredmények megfelelő értelmezését. Az empirikus kutatásban a modellnek megfelelően két irányból vizsgáljuk az idegen nyelvi érettségi vizsgákat: a vizsgaeredmények és a nyelvtanítók visszajelzései alapján.

A 4. fejezetben az érettségig elért eredményeket a vizsgáról hozzáférhető adatbázisok másodelemzése alapján vizsgáljuk, és a kutatásunk eredményeit három területre vonatkoztatva mutatjuk be. Először elemezzük a vizsga jelenlegi működését, a készségek szerinti teljesítmények változásait és az azokban megjelenő különbségeket. Ehhez

az angol és a német nyelvi érettségi vizsgát választottuk, mivel a középiskolákban ez a két leggyakrabban tanult idegen nyelv, a május-júniusi vizsgaidőszakban körülbelül nyolcvanezren tesznek közép- és emelt szintű vizsgát e két nyelvből. Ezt követően vizsgáljuk a receptív készségeket mérő vizsgarészek belső és külső validitásának kérdéseit. Harmadik területként feltárjuk az íráskészség vizsgarészben alkalmazott értékelési eljárások működését, jellemzőit és problémáit. Ez utóbbi két vizsgálattal rámutatunk arra, hogy a valószínűségi tesztelméleti modellek alkalmazásával hogyan adható visszacsatolás a vizsgafejlesztés számára, empirikusan igazoljuk a vizsgafejlesztés szükségességét és célunk, hogy a vizsgálataink megfelelő alapot szolgáltatassanak a további kutatásokhoz, a vizsgafejlesztés lehetséges irányainak kijelöléséhez.

Az 5. fejezetben a nyelvi érettségit elsősorban az angolt és a németet tanító nyelvtanárok visszajelzései alapján vizsgáljuk. Kérdőíves vizsgálatunkkal feltárjuk, hogy a nyelvtanárok milyen nézetekkel és attitűdökkel rendelkeznek a vizsgával kapcsolatban, s ezek hogyan befolyásolják a nyelvoktatási folyamatra vonatkozó nézeteiket. Arra keressük választ, hogy a nyelvi érettségi a tanárok nézetei szerint milyen jellegű hatást gyakorol a közoktatásra, a vizsga jelentősen megváltozott értékelési alapelvei meghatározzák-e a nyelvtanárok mérési-értékelési eljárásokkal kapcsolatos szemléletét. E visszacsatolási körből származó információkkal a vizsgafejlesztés számára feltárható, hogy a vizsgával közvetlenül kapcsolatban álló nyelvtanárok hogyan látják belülről a vizsga működését és a tanítási-tanulási folyamatban az érettségi hatását, valamint ez mennyiben van összhangban a vizsgafejlesztők szándékaival.

Összefoglalva, az empirikus vizsgálataink az idegen nyelvi érettségi működésének komplex elemzésére épülnek. Különböző adatbázisok másodelemezése alapján vizsgáljuk a korábbi vizsgaévek eredményeit, így elsősorban a tendenciák feltárására és olyan módszertani általánosításokra törekszünk, melyek az elért eredmények alapján a vizsgafejlesztésre és az igények és célok változásaira adnak visszacsatolást. A nyelvi érettségi hatásának vizsgálata már közelebb van a nyelvoktatás és a nyelvi mérés-értékelés módszertani kérdéseire, így vizsgálataink eredményei tartalomfüggőek, terület-specifikusak, a vizsga hatását pedig kizárólag a tanárok gondolkodása alapján vizsgáljuk. Ez a kutatás hozzájárul ahhoz, hogy a vizsgát több szempontból és átfogóan elemezzük. Ugyanakkor a vizsga viselkedését oktatáspolitikai szempontból nem elemezzük, nem törekszünk arra, hogy a vizsga társadalomra, oktatási rendszerre gyakorolt hatását vizsgáljuk, oktatáspolitikai kérdéseket fogalmazzunk meg és konzekvenciákat vonjunk le.

Az idegen nyelvi tudás felértékelődése miatt alapvető fontosságúnak tartjuk az évenkénti, rendszeres és a közoktatási rendszeren belül megszerezhető nyelvtudás szintjét monitorozó felméréseket. Ugyanakkor jelenleg ez megoldatlan, az érettségi ezt a szerepet nem tudja betölteni. Ehhez fontos megismerni a vizsga jelenlegi működését, a vizsgafejlesztés lehetőségeit, a közoktatásra gyakorolt hatását és az ezzel kapcsolatos kérdésekre választ kell adni, mivel ezek alapján tudunk következtetéseket levonni a vizsga továbbfejlesztésére és a felkészítő gyakorlatra vonatkozóan.

1. A NYELVI MÉRÉS ÉS ÉRTÉKELÉS NEMZETKÖZI TENDENCIÁI

1.1. A kommunikatív kompetencia mérésének és értékelésének lehetőségei és problémái

A kommunikatív tesztelés célja a nyelvhasználó kommunikatív kompetenciájának mérése és értékelése. E tesztelés fejlődési szakaszában, a kilencvenes években vált az idegen nyelvi mérés és értékelés elmélete multidiszciplináris tudományággá, amely szoros kapcsolatban áll az alkalmazott nyelvészettel, a statisztikai módszerekkel támogatott mérésmetodikával és a pedagógiai értékeléssel (Alderson, Clapham és Wall, 1995; Bachman és Palmer, 1996; Bárdos, 2001, 2004). A nyelvpedagógiába tartozó kommunikatív tesztelés alkalmazza a tesztelmélet eljárásait.

Ebben a fejezetben arra keresünk választ, hogy a kommunikatív tesztelés miként integrálódik a tesztelmélet egészébe, milyen fogalmakat, tesztelési módszereket alkalmaz. Célunk bemutatni, hogy a kommunikatív kompetencia hogyan modellezhető, méréséből és értékeléséből milyen problémák adódnak és hogyan jelennek meg a modellek teszteké váló fordításában a kompetenciaértékelés minőségi mutatói.

A fejezetben először áttekintjük a kommunikatív tesztelés szempontjából releváns nyelvtudás-modellek fejlődését, taxonómiáit, mivel ezek alapvetően meghatározzák a kommunikatív tesztelés módszereit és eszközeit. Ezt követően azzal foglalkozunk, hogy a kommunikatív teszteknek egyrészt milyen minőségi kritériumoknak kell megfelelniük, másrészt, hogy a tesztek mely típusai alkalmasak a különböző mérési célok elérésére.

1.1.1. A nyelvi mérés és értékelés fogalma

Bachman (1990) szerint az idegen nyelvi mérés és értékelés fogalmát a gyakorlatban gyakran szinonimaként használják, azonban ezek különböző folyamatokat jelölnek (Figueras, 2005). A nyelvi mérés (*assessment/ measurement/ testing*) az adatgyűjtést követő világos szabályok és eljárások szerint történő számszerűsítés, kvantifikálás (Bachman, 1990). E folyamat eredménye adatokat szolgáltat az értékeléshez (*evaluation*), amelynek fő funkciója a döntéshozatal, mely függ a döntést megalapozó információ minőségétől (Bachman, 1990; Lynch, 2001).

Az idegen nyelvi mérés és értékelés eszközei közé tartoznak a különböző típusú és funkciójú nyelvi tesztek. A teszt fogalma tágabb értelemben véve olyan eljárást jelöl, amely előre meghatározott időpontban és követelmények szerint az egyéneket kontrollált feltételek mellett bizonyos cselekvési és viselkedési módokra ösztönzi, és amelyekből következtetéseket lehet levonni az egyén többnyire közvetlenül nem megfigyelhető tulajdonságaira. A nyelvi tesztek által mérendő és értékelendő látens tulajdonságokat, például sajátos tudásokat, képességeket és készségeket konstruktumnak (*construct*) nevezük. A nyelvi tesztek alkalmazásának célja, hogy a konstruktumot mozgósítsa a tesztfeladatok és az egyes tesztelemelek, az itemek segítségével (Davies, Brown, Elder, Hill, Lumley és McNamara, 1999; Grotjahn, 2000a, 2000b). Ezzel egyrészt lehetővé válik, hogy a vizsgázók nyelvtudását mérjük, másrészt pedig, hogy a teszten elért eredmények alapján különböző döntéseket hozzunk, amelyek egyes személyekre, csoportokra vagy egy teljes képzési programra irányulhatnak (Lynch, 2001). A döntések akkor tekinthetők érvényesnek és megbízhatónak, ha meg tudjuk határozni, milyen mértékben egyezik meg a teszten nyújtott teljesítmény más valós nyelvi szituációban, élethelyzet-

ben nyújtott teljesítménnyel (Bachman, 1990; Takala, 1998; Grotjahn, 2002). Ehhez szükséges annak meghatározása, hogy mit értünk nyelvtudáson (Bárdos, 2003, 2004). Olyan modellt kell tehát alkotni, amellyel lehetőség nyílik a nyelvtudás feltárására, azért hogy pontosan meg tudjuk határozni, hogy mit kívánunk mérni és értékelni.

1.1.2. A kommunikatív kompetencia szerkezetének modelljei

A nyelvtudást mérő tesztfeladatokból minden esetben kikövetkeztethető, hogy mi a tesztkészítők felfogása a nyelvtudás fogalmáról és szerkezetéről. A nyelvtudás értelmezése a különböző nyelvtanítási módszerekben eltérő (Bárdos, 2005, 2006; Einhorn, 1998). A nyelvi mérés és értékelés módszertana jelentős fejlődésen ment keresztül (Neuner és Hunfeld, 1993; Bolton, 1996; Bárdos, 2001, 2002, 2004) a szubjektív tesztelési technikáktól kezdve a strukturalista-pszichometrikus szemléleten (Lado, 1961) és az egységes kompetencia hipotézisén (Oller, 1979) keresztül a kommunikatív tesztelésig, amely autentikus szövegekből és feladatokból kiindulva, objektív és szubjektív tesztelési technikákkal méri a nyelvtudást. Ezt a fejlődésmentet részletesen mutatja be Bárdos (2001, 2002, 2004) és Krekeler (2005), az egyes fázisokról Dávid (1992), Katona (1995), Katona és Dörnyei (1991) ad átfogó képet.

A kommunikatív tesztelés nyelvtudásról alkotott fogalmát a szakirodalomban kommunikatív kompetenciaként (*communicative competence*), kommunikatív nyelvi kompetenciákként (*communicative language competencies*), kommunikatív nyelvi képességekként (*communicative language ability*) és kommunikatív nyelvi készségekként (*communicative language skills*) nevezik meg. A továbbiakban különböző nyelvtudásmodelleken keresztül mutatjuk be az egyes fogalmak közötti kapcsolatokat. Ezen túl azzal is foglalkozunk, hogy milyen szerepe van a kommunikatív kompetencia szerkezetéről alkotott elméleti modelleknek a kommunikatív tesztek készítésekor és a teszteredmények értelmezésekor.

A kommunikatív kompetencia fogalmát Hymes (1972) vezette be, hogy szembeállítsa Chomsky (1965) nyelvi kompetencia fogalmával, amely a nyelvi rendszer szabályainak ismeretét, a nyelvileg pontos mondatok létrehozását jelenti. Hymes szerint azonban vannak olyan nyelvhasználati szabályok, amelyek ismerete nélkül a kommunikáció nem valósulhat meg. Értelmezése szerint a kommunikatív kompetenciába beletartozik a nyelvi kompetencia, de meghatározó eleme a szociolingvisztikai kompetencia is, amely a nyelv kontextusban való megfelelő használatának szabályait tartalmazza.

A kommunikatív tesztelés kialakulásához Morrow (1981) jelentősen hozzájárult azzal, hogy meghatározta az autentikus nyelvhasználat tulajdonságait, amelyekből levezette a kommunikatív nyelvoktatás és tesztelés céljait (1. táblázat). Morrow nem fejtette ki, hogy a célokat milyen alapelvek mentén kell figyelembe venni a tesztek készítésekor és az eredmények értelmezésekor. Ehhez szükséges a kommunikatív kompetencia összetevőinek és rendszerének modellezése.

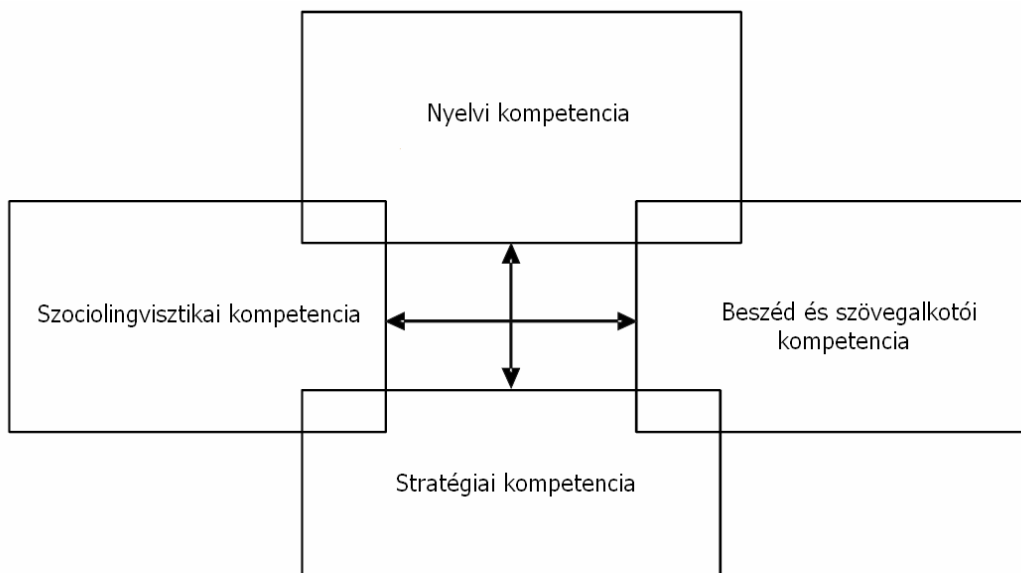
1. táblázat. Az autentikus nyelvhasználat jellemzői és az abból levezethető célok (Morrow, 1981 alapján)

Az autentikus nyelvhasználat	A kommunikatív tesztelés célja
1. interakción alapul,	Tud-e a nyelvhasználó
2. bizonyos kontextusban valósul meg,	1. interakciót létrehozni,
3. valamilyen cél és	2. a kommunikációs szituációnak és
4. hatás elérése érdekében	3. a saját közlési szándékának megfelelően
5. valós nyelvi környezetben történik.	4. sikeresen nyelvileg cselekedni,
	5. a célnyelven előforduló szövegekkel bánni.

A hazai szakirodalomban a kompetenciákat *Nagy József* (2002) „meghatározott funkciót szolgáló motívum- és képességrendszerek”-ként definiálja (32. o.) és a személyiség funkcionális modelljében négy kompetenciát különít el: a kognitív, a szociális, a személyes kompetenciát és a speciális kompetenciákat. Ebben a modellben a nyelvi kommunikáció (a verbális ismeretközlés és -vétel) a kognitív kompetencián belül a kognitív kommunikációs képesség része.

A kommunikatív tesztek alapjául szolgáló nyelvtudás meghatározására több modell is született, amelyek közös jellemzője, hogy a nyelvi kommunikációt meghatározó kommunikatív kompetencia többkomponensű jellegéből indulnak ki, és amelyekben az egyes komponensek szoros összefüggésben, egymással kölcsönhatásban állnak. A bemutatandó modellek és taxonómiák célja a tesztben elért nyelvi teljesítmények és a nyelvtudás közötti összefüggések meghatározása, azaz, hogy a teszten elért eredmények hogyan interpretálhatók.

Canale és Swain (1980) modellje a kommunikatív kompetencia szerkezetét többkomponensűnek, és dinamikusnak írja le. *Canale* (1983) kibővített modellje négy összetevőt tartalmaz (1. ábra).



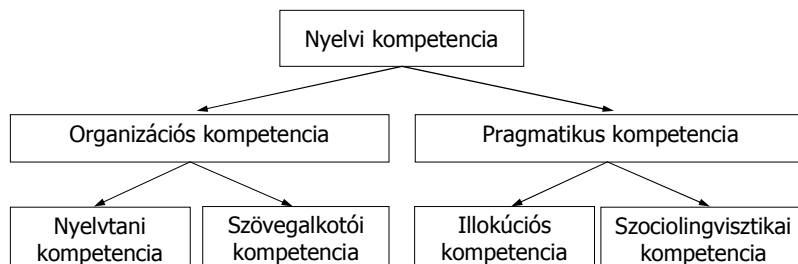
1. ábra

A kommunikatív kompetencia modellje
(*Canale, 1983 alapján: Bárdos, 2004. 150. o.*)

A kommunikatív kompetencia modell egyesíti *Chomsky* (1965) és *Hymes* (1972) elméletét. A nyelvi kompetencia megfelelő szintjével az rendelkezik, aki viszonylag folyamatosan hoz létre nyelvi pontos és szabályos mondatokat. A modellben lévő beszéd- és szövegalkotói kompetencia az előző kiegészítésének tekinthető, mely a szövegkohézióknak és a koherenciának, míg a szociolingvisztikai kompetencia az elfogadhatóságnak és a társadalmi elvárásnak megfelelő kifejezésmódot, nyelvhasználatot jelenti. A stratégiai kompetencia olyan verbális és nem verbális stratégiák ismeretére vonatkozik, amelyek elsősorban a kommunikáció fenntartására és a nyelvi hiányosságok kompenzálására szolgálnak. A modell nem határozza meg pontosan a négy összetevő közötti kapcsolatot és nem írja le a kommunikatív kompetencia és performancia közötti átmenetet. Ennek ellenére a nyelvi mérés és értékelés szakirodalmában a leggyakrabban hivatkozott modell (lásd például *Bachman, 1990; Sigott, 1996; Miyata-Boddy és Langham, 2000; Young, 2000; Grotjahn, 2000a; Tschirner, 2001; Krekeler, 2005; Bárdos, 2002, 2004, 2006*).

Bachman (1990) modellje a kommunikatív nyelvi képességről az előző modell (1. ábra) közvetlen továbbfejlesztésének tekinthető, és sajátossága a nyelvi tesztekkel való kapcsolata. A kommunikatív nyelvi képesség ebben a modellben két összetevőből áll: tudásból vagy kompetenciából és abból a képességből, amely lehetővé teszi, hogy ezt a kontextusba helyezett nyelvhasználat során alkalmazni tudják. A modell három fő komponenset tartalmaz: (I) a nyelvi és (II) a stratégiai kompetenciát, valamint (III) a pszichofiziológiai mechanizmusokat, amelyek olyan neurológiai és fiziológiai folyamatokat jelentenek, amelyek a receptív és produktív nyelvhasználatot segítik. A továbbiakban a modell komponensei közül elsősorban a nyelvi kompetenciát mutatjuk be, a másik kettőt *Bachman* és *Palmer* (1996) modelljénél elemezzük.

Bachman (1990) *Canale* és *Swain* (1980), valamint *Canale* (1983) modelljeivel ellentétben a nyelvi kompetencia szerkezetét két részre bontva hierarchikus rendszerben írja le (2. ábra). A szervezési/ organizációs kompetencia (*organizational competence*) funkciója a nyelvileg megfelelő közlések megértése, alkotása, szóbeli és írott szöveggé szervezése. Ez magába foglalja a nyelvtani és a szövegalkotói kompetenciát. A nyelvtani kompetencia lehetővé teszi a közlések megértését, létrehozását és a nyelv fonológiai, morfológiai, szintaktikai és lexikai ismeretét jelenti; a szövegalkotói kompetencia lehetővé teszi, hogy a közlésekből szöveg jöjjön létre és a kohéziót, retorikát tartalmazza. A nyelvi kompetencia másik összetevője, a pragmatikus kompetencia lehetővé teszi, hogy a közléseket a nyelvhasználó a saját kommunikációs céljaihoz és a nyelvhasználat kontextusához illessze. A pragmatikus kompetencia egyik eleme, a beszédfunkciók kifejezésére alkalmas illokúciós kompetencia, a másik összetevője a szociolingvisztikai kompetencia, amelynek segítségével a nyelvhasználó a közléseket és szövegeket a nyelvhasználat kontextusainak jellemzőihez, a dialektusokhoz, regiszterekhez, idiómákhoz és kulturális utalásokhoz kapcsolja.



2. ábra

A nyelvi kompetencia szerkezete Bachman (1990. 87. o.) modelljében

Bachman (1990) rendszerezésével differenciáltabb képet kapunk a nyelvi kompetencia összetevőiről és a köztük levő kapcsolatról. Ugyanakkor a nyelvi kompetencia hierarchikus rendszerezéséhez kapcsolódik a modell kritikája is. *Bárdos* (2002. 98. o.) szerint a modell egyrészt „zavaró kiemeléseket tartalmaz”, másrészt pedig az „ábrázolásból nem derül ki, hogy a nyelvi kompetencia alcsoportjai nemcsak egymásmelletteiségben, hanem egymás részeiként is léteznek.” Az illokúciós kompetencia például a szociolingvisztikai kompetencia része, míg a szociolingvisztikai kompetencia sem működhet nyelvtani kompetencia nélkül. A modell empirikus validálása nem történt meg (*Sigott*, 1996). Ugyanakkor *Bachman* (1990) modelljében a tesztelés gyakorlati szempontjai is dominálnak, és ezért jelennek meg a receptív és produktív készségek a pszichofiziológiai mechanizmusok között (*Dévény*, 2008).

Bachman és *Palmer* (1996) az előző modellre építve, a nyelvi képességek komponenseit a nyelvi tesztelés gyakorlatához igazodva határozzák meg. Ebben a modellben a kompetencia kifejezés helyett a tudás (*knowledge*) fogalmát használják az elemek megnevezésére. Modelljükben egyik összetevő a nyelvi tudás, a másik a stratégiai kompe-

tencia (2. táblázat). A nyelvi tudás megegyezik *Bachman* (1990) modelljének nyelvi kompetencia fogalmával, az előzőekhez hasonlóan a szervezési és a pragmatikai tudást tartalmazza. A két modell közötti különbség, hogy *Bachman* és *Palmer* (1996) a funkcionális tudás összetevőit részletezik, amely az illokúciós kompetenciával egyezik meg (2. ábra). A beszédfunkciók kifejezéséhez tartozik az információk közlése (kifejezési funkció), a nyelvi tartalmak kitalálása (képzelőerő funkció), az információk megszerzése (heurisztikus funkció) és az emberek befolyásolása (manipulációs funkció). A pragmatikai tudás másik összetevője, a szociolingvisztikai tudás része a nyelvi tudásnak, amely a nyelvoktatási gyakorlatban viszonylag ritkán jelenik meg fejlesztendő területként (*Celce-Murcia*, 2008).

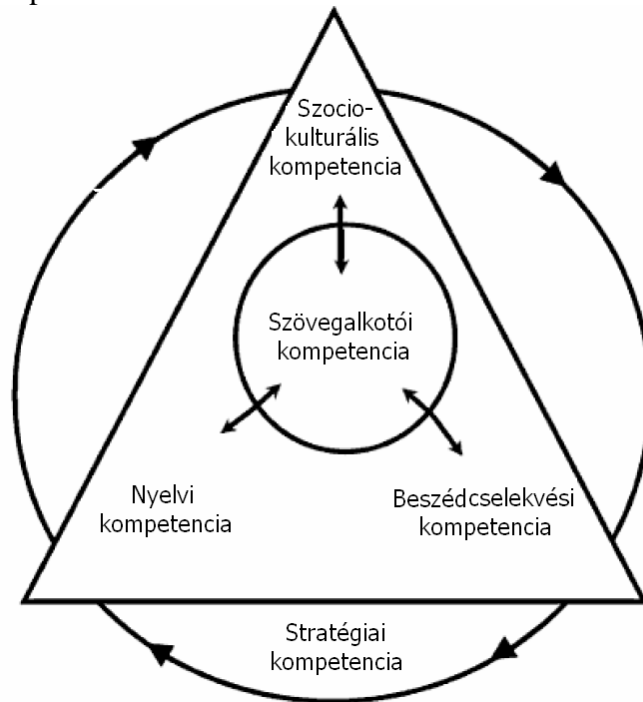
2. táblázat. *Bachman és Palmer (1996) kommunikatív nyelvi képességek modellje*

Nyelvi tudás				Stratégiai kompetencia
Nyelvszervezési tudás		Pragmatikai tudás		
Nyelvtani tudás	Szövegalkotói tudás	Funkcionális tudás	Szociolingvisztikai tudás	
fonológia	kohézió	kifejezési funkció	dialektusok és nyelvi variációk	<i>célok meghatározása</i>
morfológia	retorika	manipulációs funkció a) instrumentális b) reguláris c) interperszonális	regiszter	<i>célok tervezése</i>
szintaxis		heurisztikus funkció	idiómák	<i>célok értékelése</i>
lexika		képzelőerő funkció	kulturális utalások és nyelvi alakzatok	

A modell másik fő komponense, a stratégiai kompetencia értelmezése megegyezik *Bachman* (1990) modelljében. Ennek az összetevőnek nagyobb szerepe van, mint *Canale* és *Swain* (1980) modelljében: míg abban inkább kompenzáló funkciója volt, addig *Bachman* (1990), valamint *Bachman* és *Palmer* (1996) modelljében végrehajtó funkciót tölt be, vagyis olyan metakognitív folyamatokat jelöl, amelyek a beszédfunkciók tervezését, megvalósítását és kiértékelését is tartalmazzák. Ez a komponens nem nyelvi, hanem kommunikatív jellegű: a kommunikációra vonatkozó metatudást és tudatosságot (a célok meghatározását, a folyamat tervezését és kiértékelését) jelöli, mely a nyelvi teljesítmény egyik meghatározó eleme. A modell a két fő komponens közti kölcsönhatást úgy emeli ki, hogy a nyelvi teljesítmény és a nyelvtudás közti kapcsolatot dinamikusnak tekinti. *Bachman* és *Palmer* (1996) modelljükben nem említik *Bachman* (1990) modelljének harmadik komponensét, a pszichofiziológiai mechanizmusokat, helyettük nyelvi készségekről (*language skills*) beszélnek (75–76. o.), amelyek arra szolgálnak, hogy a nyelvet egy adott kontextusban használni tudjuk specifikus nyelvhasználati feladatok végrehajtására. A nyelvi készségeket tekintik kiindulópontnak a nyelvi tesztek tervezésénél és fejlesztésénél.

Bachman és *Palmer* (1996) modelljét különbözőképpen ítéli meg a szakirodalom. *Skehan* (1998) elsősorban azt hiányolja, hogy a modell keveset mond a nyelvi produkció folyamatairól és nem írja le a nyelvi feldolgozó mechanizmusokat, hanem olyan konceptuális sémákat ad, amelyek a tesztek szerkesztésére és a teszteredmények interpretálására szolgálnak. A modell viszont a tesztfejlesztés és -elemzés gyakorlatában jól alkalmazható (*Sigott*, 1996; *Grotjahn*, 2000a; *Krekeler*, 2005), amit az is tükröz, hogy a nyelvvizsgák leírásában ez az egyik leggyakrabban hivatkozott modell (*Dávid*, 2002).

Az idegennyelv-tanítás és vizsgáztatás folyamatának megtervezésére és lebonyolításának megkönnyítésére alkotta meg *Celce-Murcia, Dörnyei és Thurrell* (1995, idézi *Celce-Murcia*, 2008) a nyelvtudás funkcionális-dinamikus modelljét (3. ábra). A modell fő célja, hogy illusztrálja a nyelvi kompetencia elmeinek dinamikus kapcsolatát és egymást támogató szerepét.



3. ábra

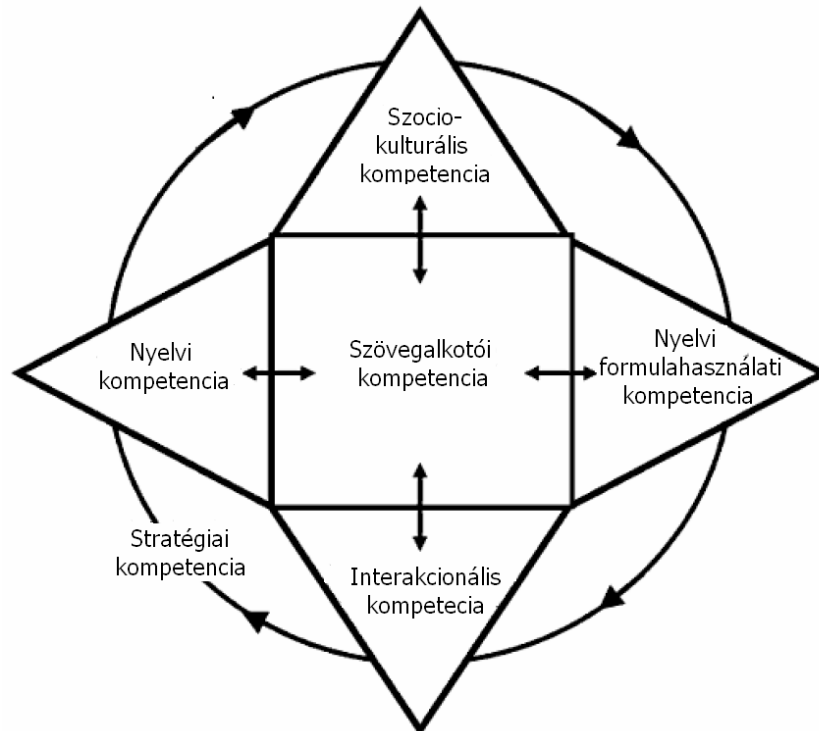
A nyelvtudás funkcionális-dinamikus modellje

(*Celce-Murcia, Dörnyei és Thurrell*, 1995. 10. o., idézi *Celce-Murcia*, 2008. 44. o.)

A modell (1) a szövegalkotói kompetencia központi helyét hangsúlyozza, mely segíti a nyelvhasználó választásait abban, hogy eleget tegyen a szövegkohézió és koherencia követelményeinek. Ez a kompetencia szoros, közvetlen és állandó kapcsolatban van (2) a beszédcselekvési („akcionális”) kompetenciával, amely a kommunikációs szándék megértését és kifejezését jelenti, a cselekvési és a nyelvi szándéknak az összeegyeztetése alapján teszi lehetővé, hogy a nyelvhasználó nyelvi repertoárjából merítve, különböző beszédcselekvéseket hajtson végre. Ehhez járul hozzá (3) a nyelvi kompetencia, amely a nyelvi formák és kódok ismeretét jelenti. (4) A szociokulturális kompetencia a kontextus, megfelelő stílus, non-verbális tényezők, kulturális háttér és a társadalmi normák ismeretére és a kommunikációban való figyelembe vételére vonatkozik. Ezeket a kompetenciákat veszi körül (5) a stratégiai kompetencia, amely a kommunikációs stratégiák kognitív és metakognitív használatát jelenti, és a többi kompetencia segítségére is szolgál. A modellben a szerzők felsorolják az összes kompetencia összetevőit és nagy hangsúlyt fektetnek a szociokulturális és a beszédcselekvési kompetencia részletes bemutatására is.

Celce-Murcia, Dörnyei és Thurrell (1995) modelljének továbbfejlesztett változatában (4. ábra) *Celce-Murcia* (2008) az összetevőket megtartva és a köztük lévő dinamikus kapcsolatot hangsúlyozva, egy újabb elemmel (6) a nyelvi formulahasználati kompetenciával (*formulaic competence*) egészíti ki a korábbi modellt (3. ábra). Ez az összetevő a nyelvi kompetencia ellentétéként határozható meg: míg a nyelvi kompetencia nyílt és szabadon felhasználható rendszer, addig ez a kompetencia zárt rendszer és az állandósult szószerkezeteket, kollokációkat, idiómákat, lexikai szerkezeteket tartalmazza. E kompetencia fontosságát mutatja, hogy a nyelvet folyamatosan használó a nyelvi

formulákat gyakran alkalmazza és ennek megértése, valamint a megfelelő formulák használata a kommunikáció folyamatát, sikerességét alapvetően meghatározza.



4. ábra

A kommunikatív kompetencia funkcionális-dinamikus modellje
(Celce-Murcia, 2008. 45. o.)

Celce-Murcia, Dörnyei és Thurrell (1995), valamint Celce-Murcia (2008) modelljüket elsősorban a nyelvtanítás és nyelvelsajátítás számára készítették, ugyanakkor e modell segítségével a nyelvvizsgáztatási folyamatokat is jól meg lehet tervezni [lásd például Kormos (1999) vizsgálatát]. A modell olyan elhanyagolt részterületekre is hangsúlyt fektet, mint például a beszélgetési struktúrák, a nyelv lexikai és grammatikai elemei, a beszédaktusok és értelmezésük valamint a kommunikációs stratégiák használata (Katona, 1995).

Az idegen nyelvi mérés és értékelés elméletére és gyakorlatára a bemutatott kompetenciamodelleknek nagy hatása volt az 1990-es években. A továbbiakban az ezredfordulót meghatározó európai ajánlásrendszerben használt nyelvtudás-modellt mutatjuk be, amely jelentős mértékben alakította a kommunikatív tesztelés fejlődését. A *Közös Európai Referenciakeret: nyelvtanítás, nyelvtanulás, értékelés*¹ (KER, 2002) dokumentumban definiált és a kutatók körében többnyire egységesen elfogadott kommunikatív nyelvi kompetenciák (*communicative language competences*) fogalma a korábbi modellekhez hasonlóan többkomponensű, közvetlenül nem mérhető és értékelhető konstruktként határozható meg. A Referenciakeret a korábban bemutatott modellekre építve és azokat ötvözve, definiálja a kommunikatív nyelvi kompetenciákat, amelyeknek elemei:

- (1) a nyelvi kompetenciák (amelyek a lexikális, grammatikai, szemantikai, fonológiai, helyesírási és helyes kiejtési kompetenciából állnak),
- (2) a szociolingvisztikai kompetenciák (beleértve a társadalmi viszonyok nyelvi jelölését, az udvariassági szokások, népi bölcsesség és stílusrétegbeli különbségek ismeretét, valamint a dialektusokat és az akcentust), és

¹ A továbbiakban a Referenciakeret fogalmát használjuk, ha a dokumentumról általában van szó. A rövidített hivatkozás: KER, 2002.

- (3) a pragmatikai kompetenciák (ideértve a beszédalkotási, funkcionális kompetenciákat; lásd részletesebben: *KER*, 2002. 5.2. fejezet).

A megfelelő nyelvhasználathoz szükséges a stratégiai kompetencia, amely „szerepet játszik a kommunikatív folyamatok menet közben történő megfigyelésében, [és] irányításában is” (*KER*, 2002. 110. o.). A Referenciakeret (1) a produktív, (2) a receptív, (3) az interaktív és (4) a közvetítő stratégiákat (*KER*, 2002. 4.4. fejezet) határozza meg. A korábbi modellekkel ellentétben a közvetítésnek nagyobb szerepet szán a dokumentum. Ennek oka, hogy míg az előző modellek a kommunikatív kompetencia összetevőit főleg egynyelvű közegben definiálták, addig a *KER* (2002. 102–103. o.) külön foglalkozik a közvetítői tevékenységgel, stratégiákkal.

A Referenciakeret a nyelvtudás hat szintjét határozza meg. Az egymásra épülő, fokozatosan emelkedő nyelvi szintek nagyon széles sávot fognak át. Az A1 szinten lévő nyelvtanuló nagyon korlátozottan, többnyire a kommunikációs partner intenzív segítségével képes a kommunikációra. A C2 szintű pedig egy nagyon magas, majdnem anyanyelvű nyelvhasználóval egyenértékű nyelvtudási szintet jelent. A Referenciakeret szintjeinek elemzésével és empirikus vizsgálatával az 1.2. fejezetben részletesen foglalkozunk.

A nyelvhasználó kommunikatív nyelvi kompetenciái a nyelvi tevékenység során aktivizálódnak. Ez a kommunikatív tesztelés szempontjából azt jelenti, hogy a tesztekkel nem tudjuk a kommunikatív nyelvi kompetenciák összetevőit és szintjeit közvetlenül mérni, mivel csak a nyelvhasználat, azaz a beszédértés, az olvasás, a beszéd, az írás és a közvetítés során tudunk következtetni a konstruktum összetevőire. Ebből következik, hogy a kommunikatív nyelvi kompetenciák mérése és értékelése a receptív, a produktív valamint a közvetítési készség szerint történik jól definiált alapelvek mentén (*Alderson, Clapham és Wall*, 1995). A továbbiakban áttekintjük, hogy a kommunikatív kompetenciák szerkezetét feltáró modellek alapján milyen alapelvek mentén tudjuk az összetevőket mérhetővé tenni.

1.1.3. A kommunikatív tesztek minőségi mutatói

A kommunikatív teszteknek különböző minőségi kritériumoknak kell megfelelniük, amelyek meghatározásában mind a klasszikus tesztelmélet, mind az utóbbi években a nemzetközi és hazai szakirodalomban is megjelenő kompetenciaértékelés minőségi mutatói fontos szerepet töltenek be. A klasszikus tesztelmélet a tesztben mért értéket a valódi érték és valamilyen hiba összegeként határozza meg, és többek között a valódi érték és a hiba paramétereinek meghatározásával, illetve a hiba becslésével foglalkozik. Ez alapján határozza meg egy teszt minőségét, azaz a teszt reliabilitását és validitását. Ugyanakkor a klasszikus tesztelmélet kritériumait a kompetenciaértékelés rendszereinek jellemzésére (például PISA-vizsgálat, országos kompetenciamérések, kompetenciafejlesztő programcsomagok) nem tartják elég átfogónak, ezért ezeket tovább differenciálják és más kritériumokat is megneveznek. Hasonló tendencia figyelhető meg a nyelvi tesztelés szakirodalmában is (például *Bachman és Palmer*, 1996; *Lienert és Raatz*, 1994). Ennek oka, hogy a (kommunikatív) kompetenciák mérése céljából olyan tesztfeladatokat kell alkalmazni, amelyek valamilyen valós helyzetben történő cselekvést, valós (nyelvi) problémák megoldását igénylik. Ebből a szempontból a kommunikatív tesztelés elhelyezhető a kompetenciaértékelés rendszereiben. A kompetenciaértékelés módszerei és eszközei, az alkalmazott feladatok, a kommunikatív tesztekhez hasonlóan különböznek a hagyományos értékeléstől. Ebből következik, hogy a kompetenciaértékelések más szempontokra helyezik a hangsúlyt (*Vidákovich*, 2008; *Vidákovich és Vigh*, 2008).

A továbbiakban először a klasszikus tesztelmélet jóságmutatóit mutatjuk be a kommunikatív tesztelésben. Ezt követően meghatározzuk a kompetenciaértékelés minőségi mutatóit, majd a kommunikatív tesztelés során alkalmazott kritériumrendszereket, végül ezeket viszonyítjuk egymáshoz, hasonlítjuk össze, meghatározzuk a közös pontokat és szintetizáljuk a kritériumokat.

A klasszikus tesztelmélet jóságmutatói a kommunikatív tesztek esetében

A klasszikus jóságmutatók közül a reliabilitás megbízhatóságot jelent, mely annál jobb, minél inkább reprodukálhatóak a teszteredmények, azaz ha az eredmény csak a valódi értéktől függ. A reliabilitásnak három aspektusát különböztetik meg: (1) a teszt belső konzisztenciája (*internal consistency*) az itemek közötti különbségekre vonatkozik. Számításakor különböző felezési módszerekkel (*split-half-method*) a teszt két felét egymással korreláltatják. A reliabilitás számításához reliabilitásmutatókat használnak, amelyek a reliabilitást alulról becslik. A nyelvi tesztek esetében gyakran használt reliabilitásmutató a Cronbach-alfa és a dichotóm itemekre vonatkozó Kuder-Richardson formula. (2) A teszt ismétlése (*test-retest reliability*) a teszteredmények összehasonlíthatóságára vonatkozik. A reliabilitás ebben az esetben a két teszten elért pontszámok korrelációjával határozható meg, ezen kívül a középértékek is összehasonlíthatók. (3) A párhuzamos tesztelés egy teszt különböző verzióira vonatkozik és a tesztek különböző variációinak korrelációjával határozható meg.

A tesztek azon tulajdonságát, hogy azt mérik-e, aminek mérésére felhasználjuk, validitásnak (érvényességnek) nevezzük. A tesztek validitástípusai közül (Nagy, 1975) a prediktív validitás azt jelenti, hogy a teszteredmény alapján hozott döntés megalapozott-e a tanuló jövőbeni teljesítményeit illetően. A tartalmi validitás összetett fogalom, mely akkor teljesül, ha a tesztfeladatok összhangban vannak a tudomány eredményeivel (szakmai validitás), és ha az összes lehetséges feladathalmazból megfelelően kerültek kiválasztásra a tesztfeladatok (mintavételi validitás). A tartalmi validitás további megjelenési formája a strukturális validitás, mely azt mutatja, hogy a tesztben lévő tudáselemek alkalmasak-e a mérendő pszichikus struktúra működtetésére, a feladatmegoldó tevékenység megfelel-e a mérendő pszichikus struktúra működésének (funkcionális validitás), a teszt számszerű eredménye mennyiben tükrözi a mérendő pszichikus struktúra fejlettségi fokát (fejlődési validitás) és a pontozás figyelembe veszi-e a nehézséget, bonyolultságot, fontosságot (skálázási validitás). A validitás harmadik típusa a konstruktvaliditás vagy szerkezeti validitás (*construct validity*), mely a teszt eredményei és a mérendő pszichikus struktúra közötti összefüggéseket mutatja.

A kommunikatív kompetenciák mérésének és értékelésének biztosításában a szerkezeti validitásnak van meghatározó szerepe, amely a kommunikatív tesztek esetében azt mutatja, hogy az eredmények igazolják-e a vizsgált nyelvi jelenségek konstruktumának meglétét. A teszten elért eredmény egyfelől arról ad információkat, hogy a vizsgált személy hogyan teljesített, másrészt arról, hogy a jelölt egy valós nyelvi szituációban hogyan viselkedne. A szerkezeti validitás elsősorban az eredmények interpretálásának alapját képezi, és nem a teszt tulajdonságaként értelmezhető (Messick, 1996; Young, 2000; Bárdos, 2002).

Messick (1996) koncepciójában a szerkezeti validitás hat aspektusát határozza meg:

- tartalmi aspektus (*content aspect*; a konstruktum tartalmának reprezentativitása a tesztben),
- lényegi aspektus (*substantive aspect*; a teszt tartalmának egyezése az elméleti konstruktummal),
- szerkezeti aspektus (*structural aspect*; a teszt szerkezetének egyezése a konstruktum szerkezetével),

- általánosíthatóság aspektusa (*generalizability aspect*; a mérési eljárás reprezentativitása a konstruktumra vetítve),
- külső aspektus (*external aspect*; a mérési eredmények egyezése az elméleti konstruktummal),
- hatás aspektusa (*consequential aspect*; az eredmények interpretációjának hatása).

A szerkezeti validitás egyrészt akkor teljesül, ha a teszt a konstruktum kevés olyan aspektusát tartalmazza, amelyet a teszt nem mér (*construct underrepresentation*), másrészt, ha a teszt a vizsgázó tulajdonságainak kevés olyan aspektusát méri és értékeli, amely a konstruktum számára irreleváns (*construct irrelevance*) (Alderson, 2002, idézi Krekeler, 2005; Brualdi, 1999; Grotjahn, 2000a).

A hagyományos jóságmutatók differenciáltabbá tételéhez Messick (1994, 1995, 1996) jelentősen hozzájárult azzal, hogy koncepciójában a szerkezeti validitást kiterjesztve értelmezte és megjelenítette abban a következményes validitást (*consequential validity*), mely a teszt validitása és hatása közti összefüggéseket írja le. A kompetenciaértékelési rendszerek minőségi mutatóinak kialakulásában és értelmezésében is ezt tekintik kiindulási alapnak (Baartman, Bastiaens, Kirschner és Van der Vleuten, 2006).

A kompetenciaértékelés minőségi mutatói a kommunikatív tesztek esetében

Ebben az alfejezetben először bemutatjuk a kompetenciaértékelési rendszerek minőségi mutatóit, majd a kommunikatív tesztekkel szemben támasztott követelményrendszereket. Feltételezzük, hogy ezek megfeleltethetők a kompetenciaértékelés minőségi mutatóinak, melyekkel lehetőség nyílik a kommunikatív tesztelés kritériumainak értelmezésére.

A kompetenciaértékelési rendszerekben a feladatokhoz és az értékeléshez kapcsolódó kritériumok az autentikusság / életszerűség (*authenticity*), a kognitív komplexitás (*cognitive complexity*) és az igazságosság (*fairness*) szempontja. Az értékelés szereplőjéhez kapcsolódó követelmények az értékelés jelentősége (*meaningfulness*), átláthatósága (*transparency*) és értelmezhetősége (*directness*). A kompetenciaértékelésben jelentős szerepe van még a szakmai és társadalmi környezet szempontjainak: az összehasonlíthatóságnak (*comparability*), a reprodukálhatóságnak (*reproducibility*), a következményeknek (*educational consequences*) és a költséghatékonyságnak (*costs and efficiency*) (Baartman és mtsai, 2006).

A kommunikatív tesztek jellemzésére kétféle kritériumrendszert hangsúlyoznak. Bachman és Palmer (1996) a *usefulness*-t (használhatóság, hasznosság) tartják a legfőbb kritériumnak a nyelvi tesztek tervezésében és fejlesztésében. Kunnan (2004) igazságosság kritériumrendszere (*test fairness framework*) elsősorban a teszt lebonyolítására, az adatfelvétel körülményeire és a teszteredmények értelmezésére vonatkozik.

Bachman és Palmer (1996) a nyelvi tesztek tervezésére vonatkozó követelményét (*usefulness*) mintegy gyűjtőfogalomként kívánták bevezetni a tesztelés jóságmutatói fölé. Ez a törekvésük azonban nem sikerült, a szakma nem vette át a *usefulness* terminust, így a magyar nyelvű szakirodalom nem ad fordítást erre, a német nyelvűben a *Nützlichkeit* kifejezés jelenik meg (Grotjahn, 2000a; Tschirner, 2001; Krekeler, 2005). A továbbiakban a használhatóság kifejezést alkalmazzuk e gyűjtőfogalom jelölésére, amelyet hat egymást kiegészítő tulajdonság: a reliabilitás, a szerkezeti validitás, az autentikusság (*authenticity*), az interaktivitás, (*interactiveness*) a teszthatás (*impact*) és a hatékonyság/ gazdaságosság² (*practicality*) ötvözeteként határoznak meg.

Kunnan (2004) igazságosság kritériumrendszere öt elemet tartalmaz, melyek közül ebben a koncepcióban is megjelenik a validitás és a következmények (*social*

² A *practicality* fordításában a német nyelvű szakirodalom a *Praktikabilität* (Tschirner, 2001) és a *(Test)ökonomie* (Krekeler, 2005) kifejezést használja.

consequences) vizsgálatának követelménye. Ezek mellett az előítélet-mentesség (*absence of bias*) arra utal, hogy a teszt lebonyolítása, valamint az eredmények értékelése és felhasználása ne hozza hátrányos helyzetbe vagy előnybe a tesztelésben részt vevő alcsoportokat. A hozzáférés (*access*) kritériuma azt jelenti, hogy a teszten való részvétel költsége a tesztmegoldó személyek számára legyen megfizethető, helyileg elérhető, a tesztkitöltők ismerjék a tesztelés körülményeit és folyamatát. A teszt lebonyolítására vonatkozó kritérium (*administration*) a megfelelő adatfelvételi körülmények biztosítására, a lebonyolítás egységességére és állandóságára vonatkozik, amely növeli a validitást és a reliabilitást. *Bachman* (2005) szerint az előítélet-mentesség, a hozzáférés és a lebonyolítás kritériumai fontosak, ugyanakkor felveti azt a kérdést, hogyan integrálhatók ezek a minőségi mutatók, milyen összefüggés van a teszt érvényessége és a tesztelés következményei között.

Az egyes koncepciók értelmezéséhez, rendszerezéséhez a kompetenciaértékelés minőségi mutatói egységes alapot nyújthatnak, mivel az egyes kritériumok jelentése alapján e koncepciókban számos hasonlóságot találunk. A 3. táblázatban az egyes kritériumokat úgy helyeztük el, hogy az egy sorban lévő kritériumok jelentése azonosnak tekinthető. Kivételt képez ezalól a *Kunnan* (2004) kritériumai között szereplő lebonyolítás szempontja, mely egyrészt a teszteredmények összehasonlíthatóságát, másrészt a döntések reprodukálhatóságát is szolgálja.

3. táblázat. A kommunikatív tesztelés kritériumai a kompetenciaértékelés rendszerében

Csoportok	A kompetenciaértékelés minőségi mutatói (<i>Baartman és mtsai, 2006</i>)	A használhatóság kritériumok (<i>Bachman és Palmer, 1996</i>)	Az igazságosság-kritériumok (<i>Kunnan, 2004</i>)
I. Feladatokhoz és értékeléshez kapcsolódó kritériumok	autentikusság (<i>authenticity</i>)	autentikusság (<i>authenticity</i>)	
	kognitív komplexitás (<i>cognitive complexity</i>)	interaktivitás (<i>interactiveness</i>)	
	igazságosság (<i>fairness</i>)		előítélet-mentesség (<i>absence of bias</i>)
II. Értékelés szereplőihez kapcsolódó kritériumok	értékelés jelentősége (<i>meaningfulness</i>)		
	értékelés átláthatósága (<i>transparency</i>)		
	értékelés értelmezhetősége (<i>directness</i>)		
III. Szakmai és társadalmi környezethez kapcsolódó kritériumok	összehasonlíthatóság (<i>comparability</i>)		lebonyolítás (<i>administration</i>)
	a döntések reprodukálhatósága (<i>reproducibility</i>)		
	következmények (<i>educational consequences</i>)	teszthatás (<i>impact</i>)	következmények (<i>social consequences</i>)
	költséghatékonyság (<i>costs and efficiency</i>)	hatékonyság, gazdaságosság (<i>practicality</i>)	

A kompetenciaértékelés minőségi mutatói közül *Bachman* és *Palmer* (1996), valamint *Kunnan* (2004) kritériumai a feladatokhoz és értékeléshez, valamint a szakmai és társadalmi környezethez tartozó kritériumokhoz illeszthetők. Az értékelés szereplőihez kapcsolódó kritériumok nem jelennek meg a kommunikatív tesztelés kritériumai között. Ez a kompetenciaértékelés sajátosságaira vezethető vissza, amely az eredmény mellett a

folyamatot is értékeli, célja elsősorban a tanulók fejlődésének regisztrálása, így az értékelés integrálódik a tanítási-tanulási folyamatba, melynek a tanárok és a diákok is aktív részesei. Mindezek mellett a kompetenciaértékelés célja, hogy az értékelés szempontjai a folyamatban részt vevők számára is legyenek világosak, érthetőek. Ennek előfeltétele, hogy a kompetenciaértékelési rendszer legyen átlátható, az értékelés eredményei közvetlenül értelmezhetők a folyamatban részt vevők számára. Ez elsősorban a performancia vagy közvetlen mérések alkalmazásával valósul meg. Ugyanakkor ez a kritérium a kommunikatív tesztelésben is jelentős szerepet tölthet be, például a nyelvi performancia tesztek elsődleges célja is az, hogy életszerű nyelvhasználati szituációban mérje a tanulók nyelvtudását, ebben az esetben az eredmények értelmezése közvetlenül történik.

Az értékelés szereplőjéhez kapcsolódó kritériumok a nyelvi mérés és értékelés területén elsősorban a nyelvi tesztek ellenpólusaként megjelenő fejlesztő értékelési módszerek (például portfólióértékelés, önértékelés) alkalmazásával kerültek előtérbe. A fejlesztő értékelés kritériumait a nyelvi tesztekkel szembeállítva határozzák meg. *Lynch (2001)* például amellet érvel, hogy a fejlesztő értékelés egy új paradigma, így elsősorban a validitást vizsgáló új kritériumokat kell kidolgozni, nem lehet a tesztelési kultúra alapelveit alkalmazni. A kompetenciaértékelés minőségi mutatóinak alkalmazása lehetőséget biztosítana a fejlesztő értékelés integrálására a nyelvi értékelés rendszerében.

A továbbiakban a kommunikatív tesztelés kritériumait részletezzük és javaslatot teszünk ezek szintetizálására. Ehhez meghatározzuk a kommunikatív teszteket jellemző kritériumok összetevőit és jelentését, amelyről áttekintést adunk a 4. táblázatban.

4. táblázat. A kommunikatív tesztelés kritériumai és jellemzői

<i>Kritérium</i>	<i>Jelentés</i>
autentikusság	Milyen mértékben egyezik meg a tesztfeladat megoldása során alkalmazott nyelvtudás azzal a móddal, ahogyan azt valós nyelvi szituációban alkalmaznánk?
interaktivitás	Milyen mértékben vonják be a vizsgázók a tesztfeladat megoldásába a nyelvi készségeiket és egy adott témáról megszerzett előzetes háttértudásukat és ismereteiket?
igazságosság	Mennyiben teszik lehetővé a teszt mérési-értékelési eljárásai, hogy a teszten elért eredmények értékelésében, az eredményekből levont következtetésekben ne kerüljenek egyes tanulócsoportok hátrányos helyzetbe?
összehasonlíthatóság	Mennyiben tekinthető a teszt lebonyolítása konzisztensnek, egységesnek és kontrolláltnak?
döntések reprodukálhatósága	Mennyiben tekinthetők a diákok nyelvtudásáról hozott döntések pontosnak és mennyiben kontrollálják az értékelés körülményeit?
teszthatás	Milyen mértékű és irányú a teszt hatása a tanítási-tanulási folyamatban résztvevőkre, az oktatási módszerekre és az iskolarendszerre?
hatékonyság	Mennyire egyeztethetők össze a tesztfejlesztés és a lebonyolítás forrásai a rendelkezésre álló erőforrásokkal?

Az autentikusság a tesztfeladatok fontos kritériuma. Annak a módnak, ahogy az idegen nyelvet a tesztfeladat megoldásakor a vizsgált személy használja, meg kell egyeznie azzal, ahogy azt valós nyelvi szituációban alkalmazná. Ez az alapelv nemcsak az autentikus (a célnyelven valóban előforduló, és nem pedagógiai céllal a nyelvtanulóknak íródott) szövegek alkalmazásával teljesül, hanem akkor is, ha a tesztfeladatokban egyértelművé tesszük, hogy milyen szituációban, milyen céllal kell elvégezni a nyelvi feladatot, és az mikor tekinthető sikeresnek. Ezt nevezzük interakciós autentikusságnak. A szituációs autentikusság akkor valósul meg, ha a szöveg nyelvi kihívásai, valamint a

feladat jellemző vonásai megegyeznek azokkal a tipikus tulajdonságokkal, amelyek a célnyelvi élethelyzetekben elvárhatók (Sigott, 1996; Bárdos, 2003, 2004). Egy teszt autentikusságát nem lehet statisztikai úton meghatározni. A tesztfeladatban megadott nyelvi szituáció tartalmi elemzésével azonban meg lehet adni az autentikusság bizonyos fokát (Messick, 1996; Takala, 1998; McNamara, 2000). A magas autentikussággal rendelkező kommunikatív teszt nagymértékben járul hozzá a teszt használhatóságához. Azok az autentikus nyelvi szituációk, amelyek a konstruktummal összefüggésben állnak, megkönnyítik a teszteredmények megfelelő értelmezését is, és az autentikus feladatokat a vizsgázók is pozitívabban ítélik meg (Bachman és Palmer, 1996).

Az interaktivitás arra vonatkozik, hogy a vizsgázók milyen mértékben vonják be a tesztfeladat megoldásába a nyelvi készségeiket, és egy adott témáról megszerzett előzetes háttértudásukat és ismereteiket (Bachman és Palmer, 1996). A tesztnek annál nagyobb a megbízhatósága, minél kevésbé befolyásolják a tesztben elért eredményt a vizsgázók emocionális reakciói, illetve azok az ismeretek és képességek, amelyeket a tesztben nem akarunk mérni. A kommunikatív teszt validitása annál nagyobb, minél inkább bevonja a jelölt a nyelvtudását a feladat megoldásába. Egy teszt interaktivitását tartalmi elemzéssel, a vizsgázók viselkedésének megfigyelésével és a teszteredmények elemzésével határozhatjuk meg. Az autentikussághoz hasonlóan csak leíró eredményhez lehet jutni, azaz az interaktivitás bizonyos szintjét tudjuk meghatározni.

Az igazságosság kritériuma azt jelenti, hogy nincs elfogultság az egyes tanulócsoportok esetében a teszt mérési-értékelési eljárásainak alkalmazásában, amely nem befolyásolja hátrányosan a csoportok teljesítményét, az elért eredmények értékelését, az eredményekből levont következtetéseket (Kunnan, 2004).

Az összehasonlíthatóság szempontjának a kommunikatív teszt abban az esetben felel meg, ha azt következetesen és felelősségteljesen bonyolítják le. Az értékelés során azonos értékelési kritériumokat kell alkalmazni minden tanuló esetében.

A döntések reprodukálhatósága arra vonatkozik, hogy a teszt eredménye alapján hozott döntések pontosak-e és az idő múlásával állandónak tekinthetők-e. E szempont arra hívja fel a figyelmet, hogy a diákok nyelvtudásáról hozott döntések nem függhetnek az értékelés körülményeitől. Ennek biztosítására a performancia tesztek eredményeinek értékelésénél különösen nagy figyelmet kell fordítani.

A következmények mindhárom kritériumrendszerben megjelennek, vagyis ez összekötő szerepet tölt be. Míg Baartman és mtsai (2006) elsősorban az értékelés hatását a tanítási-tanulási folyamatra vonatkoztatják, addig Bachman és Palmer (1996), valamint Kunnan (2004) a teszt hatását két szinten különböztetik meg. A mikroszintű teszthatás a tanítási-tanulási folyamatban résztvevőkre és a tanulási módszerekre, a makroszintű hatás pedig a társadalomra és az iskolarendszerre vonatkozik. Mindhárom rendszerben felhívják a figyelmet a szándékolt és nem szándékolt, a pozitív és negatív irányú teszthatások vizsgálatára, melyekkel azt elemzik, hogyan látják a diákok és a tanárok az oktatási folyamat céljait és hogyan igazítják azt a teszthez. E kritérium jellemzőivel, a hatásmechanizmusok modellezésével, az empirikus vizsgálat lehetőségeivel az 1.3. fejezetben részletesen foglalkozunk.

A hatékonyságot – Carroll (1963, idézi Csapó, 1978. 62. o.) iskolai tanulás és tanulás-szervezési modelljével analóg módon – a sikeres tesztfejlesztéshez, lebonyolításhoz rendelkezésre álló források, erőfeszítések és a szükséges erőforrások hányadosával fejezhetjük ki. Ha a tesztfejlesztésre, a kivitelezésre és a lebonyolításra szánt erőfeszítések elérik a szükséges szintet, akkor a teszt hatékonynak, sikeresnek és gazdaságosnak mondható. Ekkor ennek a hányadosnak az értéke egy. Ha ez az érték egynél kisebb, akkor a teszt kevésbé hatékony (Bachman és Palmer, 1996, 2005). Ugyanakkor a gyakorlatban ezzel az eljárással nehezen lehet számszerűsíteni egy teszt hatékonyságát,

mivel a szükséges forrás mennyiségéről sokszor nincs pontos információnk. E kritérium megadásának inkább az a célja, hogy egy tesztet olyan hatékonyra tegyünk, amennyire csak lehet. A kritérium tágabb értelemben véve arra hívja fel a figyelmet, hogy egy vizsgarendszer működtetésekor az optimálisnak megfelelő, legszükségesebb erőforrásokat használjuk fel (Krekeler, 2005), mivel a feladatok előkészítése, a vizsga lebonyolítása jelentősen befolyásolja a vizsga minőségét, megbízhatóságát (Einhorn, 2006b).

A kommunikatív tesztelés kritériumainak biztosítása, kimutatása és elemzése céljából e kritériumrendszerek szintetizálása alapján két fontos alapelvet határozhatunk meg: (1) az összes szempontot kell figyelembe venni a teszt tervezési, fejlesztési, kipróbálási és lebonyolítási fázisában; (2) a kommunikatív tesztelés kritériumaihoz tartozó komponenseket nem lehet egymástól függetlenül megítélni, hanem kombinált hatásukat kell vizsgálni.

1.1.4. A kommunikatív tesztek funkciói és típusai

A kommunikatív tesztek többféle szempont szerint lehet osztályozni. Bachman (1990) (1) a teszt célja, funkciója (*intended use*); (2) tartalma (*content*); (3) viszonyítási rendszere (*frame of reference*); (4) értékelésének típusa (*scoring procedure*); és (5) a bevezetett teszttechnikák és teszt módszerek (*testing method*) szerint tipologizálta a nyelvi tesztek.

A továbbiakban a kommunikatív tesztek először a különböző funkciók és a tartalom szerint osztályozzuk, és meghatározzuk, hogy a teszt típusa mire vonatkozik, majd áttérünk a standardizált és nem standardizált, a norma- és kritériumorientált, a performancia és a kompetencia tesztek készítésének problematikájára, alkalmazásuk kérdéseire.

A kommunikatív tesztek funkciói

A kommunikatív tesztek szerkesztésének fő szempontja a teszt funkciója, amely befolyásolja azt, hogy melyek azok a tartalmak, illetve készségek, amelyeket a tesztben mérni kívánunk. A kommunikatív tesztek funkciójának meghatározásakor egyrészt meg kell állapítani, hogy mely feltételek alapján alkalmazzuk a tesztek bizonyos célok mérésére, másrészt pedig, hogy a tesztek alkalmazásának milyen következményei, hatásai vannak (Perlmann-Balme, 2001; Krekeler, 2005).

A jövőbeli beválásra vonatkozó nyelvi tesztek közé tartoznak a nyelvtanulási képességek tesztjei (*language aptitude tests/ Eignungstests*). Ezek olyan prognosztikai tesztek, amelyek a nyelvtanulók sikerét jelezik előre (Bárdos, 2002). Az ebbe a csoportba tartozó besoroló tesztekre (*placement tests/ Einstufungstests*) jellemző, hogy a tanulói csoportokat változatos, többnyire objektív tesztelési technikával alakítják ki (Albers és Bolton, 1995; Bárdos, 2002). A szelekciós teszt/ felvételi vizsga (*selection tests*, illetve *gatekeeping tests/ Zulassungsprüfungen*) célja a jelentkezők szelektálása, így olyan teszt feladatokat használnak, amelyek jól differenciálnak a jelöltek között. Egy besoroló teszt is lehet szelekciós teszt, azonban ez a teszt típus sokkal átfogóbb, mint a besoroló teszt. Abban az esetben, ha a teszteredmények között tovább differenciálnak, és ez alapján arról is döntenek, hogy a jelöltet melyik kurzusra veszik fel, és mely készségeket kell fejleszteni, akkor inkább besoroló tesztről van szó (Bachman és Palmer, 1996; Perlmann-Balme, 2001; Krekeler, 2005).

A tantervi követelményekre vonatkozó nyelvi tesztek csoportjába tartozó teljesítménytesztek (*achievement tests*) alapján azt vizsgálják, hogy a tananyagot milyen mértékben sikerült elsajátítani és az eredményeket a tantervre vonatkoztatva értelmezik. Ezek a tesztek a tanítási-tanulási folyamat eredményességére utalnak. A teljesítményteszteknek két csoportját lehet elkülöníteni. (1) A haladási tesztek (*progress tests/*

Lernfortschrittstests) célja a tanuló segítése és a további tennivalók feltérképezése. Az elért eredmény következménye nem túl jelentős. (2) Egy tanítási-tanulási folyamatot lezáró tesztek (*exit tests/ Kursabschlussstests*) nagy részletességgel, átfogó módon azt vizsgálják, hogy a tanulóknak milyen mértékben sikerült elsajátítani egy képzési program teljes anyagát. A teszt eredményének következménye messzemenőbb, mivel azt is meghatározzák, hogy egy adott programot meg kell-e ismételni, illetve sikeressége belépőül szolgálhat egy magasabb szintű kurzusba (*Albers és Bolton, 1995; Davies és mtsai, 1999; KER, 2002; Bárdos, 2002; Perlmann-Balme, 2001; Krekeler, 2005*).

A konstruktumra vonatkozó nyelvi tesztek csoportjába sorolhatók a készség szintet mérő tesztek (*proficiency tests/ Feststellungsprüfungen*), amelyek a tanítási-tanulási folyamattól függetlenül mérik a vizsgázó nyelvi készségeit. A tesztek többnyire standardizáltak, a feladatokat központilag készítik és képzett, külső értékelők minősítnek (*Einhorn, 1999*). Ezeknek a teszteknek a célja az elért tudásszint mérése a teljesség igényével, és így a nyelvtudásról többé-kevésbé átfogó képet adnak (*Perlmann-Balme, 2001; Krekeler, 2005*). A készség szinteket vizsgáló tesztek két csoportra bonthatók: (1) az általános, azaz a nyelvi tartalom közvetítését és a készségeket, valamint (2) egy specifikus terület szaknyelvét vizsgáló tesztek (*Bárdos, 2002*).

A nyelvi tesztek ilyen típusú felosztásának célja, hogy a teszt tulajdonságát egy bizonyos funkcióval együtt mutassa be. Ez még akkor is hasznos, ha egy teszt funkcióját nem lehet mindig egyértelműen meghatározni. Ennek oka egyrészt az egyes kategóriák közötti átfedések, másrészt a teszt készítő, -alkalmazók és a vizsgázók eltérő érdekei (*Davies és mtsai, 1999*).

A fentebb ismertetett különböző tesztfunkciók a validitás eltérő aspektusait jelenítik meg (*Bachman, 1990; Brualdi, 1999; Krekeler, 2005*). A szelektív tesztek fő kérdése, hogy hogyan határozzák meg a bekerülés határát. A küszöbértékek meghatározásánál a megfelelési validitás (*criterion-related validity*) játszik központi szerepet, amely egy előre megadott megfelelési (*threshold*) szint alapján a nyelvtudás bizonyos minőségét jelöli. A megfelelési validitás vagy úgy határozható meg, hogy a teszteredményt a teszt megfelelési szintjével korreláltatjuk, vagy egy olyan teszttel, amelynek megfelelési validitása ismert. Ez a validitástípus szakértői véleményezési eljárással is meghatározható (*Davies és mtsai, 1999; Brualdi, 1999; Grotjahn, 2000a; Bárdos, 2002*). A teljesítménytesztek esetében a tartalmi validitás (*content validity*) teljesülése az elsődleges szempont, amely annak mértékét adja meg, hogy a teszt tartalma mennyire méri és reprezentálja a képzési program teljes tartalmát (*Grotjahn, 2000a*). A tartalmi validitást a szakértők véleménye vagy validnak tekintett teljesítménytesztek korrelációja alapján lehet meghatározni (*Alderson, Clapham és Wall, 1995; Bárdos, 2002; KER, 2002*). A készség szinteket vizsgáló tesztek esetében a megegyező/ egyidejű validitás (*concurrent validity*) és az előrejelző validitás (*predictive validity*) játszik központi szerepet. A megegyező validitás a teszten elért eredmény megegyezése más tesztek eredményével. A prediktív validitás pedig a teszteredmény és annak a kritériumnak az egyezése, amely egy jövőbeli valós nyelvhasználati szituációra vonatkozik (*Alderson, Clapham és Wall, 1995*).

A korábban részletesen ismertetett szerkezeti validitás és a triviális-/ látszatvaliditás (*face validity*) minden teszt típusnál megjelenik. E validitás a laikusok véleménye egy teszt érvényességéről. Ennek a validitástípusnak a megítélése a szakirodalomban nem egységes, mivel ezt egyrészt populáris, de a teszt irreleváns aspektusának tartják (például *Stevenson, 1985 „pop validity”-nek* nevezi, idézi *Krekeler, 2005*), másrészt viszont, ha a vizsgázók egy tesztet érvényesnek tartanak, jobb eredményt akarnak elérni rajta (*Bachman és Palmer, 1996; Grotjahn, 2000a*).

Standardizált, informális és nem standardizált formális tesztek

A standardizált tesztek feladatait hivatásos tesztkészítők írják, a legfontosabb jószágmutatókat elemzik, és mind az adatfelvételt, mind a kiértékelést, mind az eredmények értelmezését standardizálják, szabványosítják, és normát állapítanak meg. A szabványosítás csak abban az esetben alkalmazható, amennyiben a teszt jószágmutatóit empirikus elemzések alapján is biztosították.

Az informális tesztek a tanárok készítik, és saját célokra használják fel. E tesztek főleg a tanítási-tanulási folyamatot kísérik, egy korlátozott tananyagrésze, illetve tanulócsoportra vonatkoznak. Ezeknél a tesztekénél a szabványosítás nem jelenik meg kritériumként, az informális tesztek legfontosabb jószágmutatója a szituáció érvényessége, amelyben a tesztet alkalmazzák, valamint a tartalmi validitás (Csapó, 1996; Lienert és Raatz, 1994; Grotjahn, 2000a; Dávid, 1992).

A standardizált és az informális tesztek között helyezhetők el a nem standardizált formális tesztek, amelyek nem felelnek meg a standardizáció követelményének, de céljuk, hogy objektív döntéseket hozzanak a vizsgázók nyelvtudásáról, így az ilyen tesztek elkészítésénél ügyelnek a kommunikatív tesztekkel szemben támasztott kritériumok betartására. Erre a teszt típusra példa a német *Deutsche Sprachprüfung für den Hochschulzugang (DSH)* vizsgarendszer (Krekeler, 2005).

Norma- és kritériumorientált tesztek

A norma- és kritériumorientált tesztek közti megkülönböztetés elsődleges szempontja a teszteredmények értelmezésének módja. Ezen kívül a két tesztforma a tesztfejlesztés céljában, tartalmában és folyamatában is különbözik (Csapó, 1987; Bond, 1996; Grotjahn, 2000a).

A normaorientált tesztek fő jellemzője, hogy az egyén eredményeit vagy egy viszonyítási csoport eredményeihez vagy más vizsgázók átlagteljesítményéhez képest értelmezik (Grotjahn, 2000a; Glaboniat és Müller, 2006). Ezeknek a teszteknek a célja, hogy a vizsgázók teljesítménye között különbséget tegyenek, a tanulók nyelvtudását minősítsék, a csoporthoz viszonyítva kategorizálják, így erősen differenciáló itemeket tartalmaznak (Gronlund, 1988; Krekeler, 2005). E tesztek hátránya, hogy nem adnak részletes információt a vizsgázók készség szintjeiről.

A kritériumorientált tesztek célja az egyéni készség szintek leírása. A teszt a kritériumot reprezentálja, amelynek eredményével az egyéni készség szintet az elvárt készség szinttel hasonlítják össze. A kritériumorientált tesztelés esetében az elérendő célokat szintekre bontják, minden egyes szintnek részletes leírása van, és a mérést követően a jelölt valamelyikbe besorolható. Ebben az esetben meghatározható, hogy a nyelvhasználó az adott szinten mire képes (Glaboniat és Müller, 2006; Perlmann-Balme, 2006). A kritériumorientált tesztelés a *KER* (2002) alapján például magába foglalja a készség szint folyamatosságának és a megfelelő tartományok körének feltérképezését is annak érdekében, hogy a teszten elért egyéni eredmények a teljes kritériumrendszerben elhelyezhetőek legyenek. A kritériumorientált tesztek gyakran kapcsolódnak egy tantervhez és a tanítási-tanulási folyamathoz. Ezek a tesztek előre meghatározott követelmények szerint készülnek, a vizsgált személyek pontosan tudják, hogy mely tartalmakat és készségeket mérnek, ezért magas e tesztek látszatvaliditása és a tanítási-tanulási folyamatra gyakorolt hatása (Grotjahn, 2000a; Krekeler, 2005).

Kompetencia és performancia tesztek

A kompetencia és performancia tesztek megkülönböztetése *Chomsky* (1965) kompetencia és performancia fogalmára vezethető vissza. Az előbbi a nyelv használatának veleszületett képességét, fejleszthetőségét, a nyelv mélystruktúrájának ismeretét, az utóbbi a nyelvi rendszer felhasználását jelenti konkrét helyzetekben és szituációkban. A továbbiakban először az 5. táblázatban bemutatjuk a kompetencia és performancia tesztek általános jellemzőit, eltérő vonásait, alkalmazásának különböző előnyeit és hátrányait. Majd ezt követően részletezzük ezeket a különbségeket, amelyeknek oka a konstruktum leképezésének eltérő módja.

5. táblázat. A kompetencia és a performancia tesztek összehasonlítása (*McNamara, 2000; Grotjahn, 2000a; Krekeler, 2005* alapján)

	Kompetencia tesztek	Performancia tesztek
<i>Mérési cél</i>	A nyelvi készség mérése	A nyelvi készség mérése
<i>Mérési eljárás</i>	A nyelvi készség közvetett leképezése: 1. A nyelvi készség modellezése 2. A készségre jellemző izolált nyelvi elemek mérése 3. A tesztben mért tulajdonság értékelése és következtetés a nem mért készségre	A nyelvi készség közvetlen leképezése: 1. A nyelvhasználati szituáció meghatározása 2. A nyelvhasználat mérése egy adott kontextusban 3. A nyelvi teljesítmény vagy a kommunikációs szituáció megoldásának értékelése
<i>Eredmény értelmezése</i>	Az eredmény általánosítható, a döntések reprodukálhatósága alacsony	Az eredmény nem általánosítható, a döntések reprodukálhatósága magas
<i>Jóság-mutatók</i>	Magas reliabilitás Objektív értékelés Alacsony validitás Alacsony elfogadottság (látszatvaliditás) Esetleges negatív hatás az oktatási folyamatra	Alacsony reliabilitás Szubjektív értékelés Magas validitás Magas elfogadottság (látszatvaliditás) Többnyire pozitív hatás az oktatási folyamatra

A kompetencia tesztek elsődlegesen azt mérik, hogy a vizsgázó a konkrét nyelvi szituációtól függetlenül hogyan tudja a nyelvet használni. Ezek a tesztek abból a feltételezésből indulnak ki, hogy a nyelv különböző, egymással kapcsolatban levő komponensekből áll, amelyeket egymástól függetlenül lehet mérni és értékelni. A kompetencia tesztek szerkesztésének alapja egy, a nyelvi képességek és készségek szerkezetét és a közöttük lévő kapcsolatot is meghatározó átfogó és részletes modell, amely a teszteredmények értelmezésének alapjául szolgál.

A kompetencia tesztek sokoldalúan alkalmazhatók, az eredményekből jól lehet általánosítani más teszten kívüli nyelvi szituációra. Ehhez azonban szükséges, hogy a teszt a modellt minél jobban reprezentálja, ezáltal nagyobb azon nyelvi szituációk száma, amelyekben a vizsgált személy teljesítményét a teszt előre tudja jelezni. E tesztek hátránya, hogy látszólag nem kötődnek a valós nyelvi tevékenységhez (például a tesztben a szavak aláhúzása, néhány szó beírása), így a látszatvaliditásuk is alacsony. Ennek ellenére a zárt feladatok alkalmazása objektív értékelést és magas reliabilitást biztosít (*Niska, 1997; Robinson, 1996; Skehan, 1998; Grotjahn, 2000a; Krekeler, 2005*).

A performancia tesztek mérési és értékelési eljárása abban különbözik a kompetencia tesztekétől, hogy a tesztben egy életszerű nyelvi szituációt képeznek le és azt a képességet mérik, hogy a vizsgált személy hogyan tudja egy bizonyos szituációban a nyelvet használni (*McNamara, 2000; Robinson, 1996; Young, 2000*). *McNamara* (1996, idézi *Grotjahn, 2000a*) az alábbi módon osztja fel a nyelvi performancia teszteseteket. Azokat a teszteseteket, amelyek kizárólag csak a nyelvi teljesítményre vonatkoznak, szűkebb

értelemben vett performancia teszteknek (*weak language performance tests*) nevezi. Ezzel szemben azokat a tesztek, amelyek a feladat megoldását, és így a kommunikációs szituáció megvalósítását is értékelik, tágabb értelemben vett performancia teszteknek (*strong language performance tests*) nevezi.

A performancia teszteknek magas a validitásuk. A valós nyelvi szituációk alkalmazása biztosítja az eredmények értelmezését a konstruktummal való összevetés alapján, s ezek ideális esetben egymással megegyeznek. Az ilyen teszteknek viszonylag magas az elfogadottsága (látszatvaliditása), illetve pozitív a tanítási-tanulási folyamatra gyakorolt hatása. Ugyanakkor a performancia tesztek eredményeiből nem tudunk pontos következtetéseket levonni arra vonatkozóan, hogy egy másik nyelvi szituációban hogyan teljesítene a vizsgázó. Ezeknek a teszteknek a reliabilitása is problémás, mivel az értékelés szubjektivitása megnő. Néhány szerző a közvetlen performancia tesztnél nem tartja elképzelhetőnek a valós nyelvi szituáció pontos leképezését. A tesztfeladat autentikussága meghatározó, mert az alacsony autentikusságú teszt nem közvetlen, és az eredmények értelmezésének validitása is megkérdőjeleződik. Ezen kívül a tesztfeladatban megjelenő nyelvi szituációk kiválasztásának a konstruktumra nézve reprezentatívnak kell lennie (Davies és mtsai, 1999; Skehan, 1998; Takala, 1998; McNamara, 2000; Grotjahn, 2000a; Krekeler, 2005).

Annak meghatározása, hogy performancia vagy kompetencia tesztet alkalmazzunk-e, elsősorban a mérési céltól függ. Abban az esetben, ha egy tesztnek a vizsgázó teljesítményéről konkrét nyelvi szituációban kell információt adnia, akkor a performancia teszt alkalmazása lehetséges (például a produktív készségeket mérő feladatoknál ilyen tesztek lehet alkalmazni). Ha az eredményeket, mint különböző nyelvi szituációkra való utalásokat szeretnénk értelmezni, akkor a kompetencia tesztek megfelelőbbek (Niska, 1997; Robinson, 1996).

Közvetlen, közvetett és félig közvetett tesztek

A közvetlen és a közvetett tesztek az eredmények eltérő értelmezésére vonatkoznak, és a teszt tartalomnak a konstruktumhoz való viszonyát tükrözik. A közvetlen teszteljárásnál a teszteredményből egyből lehet következtetni a konstruktumra. Minél közvetettebb a tesztelési eljárás, annál nagyobb szükség van az eredmények interpretálásához a konstruktum modellezésére. A közvetett tesztek esetében nem magát a mérni kívánt képességet vizsgáljuk, hanem annak egy fontos elemét, és aztán ezt az elemet a teljes képességgel korreláltatjuk (Robinson, 1996; KER, 2002).

A közvetlen és a közvetett tesztek mellett megkülönböztetik még a félig közvetett tesztek (Tschirner, 2001) is, amelyeknél azt a képességet, amelynek mérését célul tűzzük ki magunknak, integrálva teszteljük, azaz a képességet nem bontjuk elemeire, és nem mérjük külön. Ezeknél a tesztnél a mérés sem teljesen autentikus módon történik, mivel az autentikusságot csak szimulálják. A TestDaF vizsga szóbeli részén alkalmazott szimulált szóbeli interjú (*Simulated Oral Proficiency Interview*; SOPI) példa erre a teszt típusra, amelyben a megadott input feladat standardizálásával és a felvett hanganyag központi értékelésével lehetőség nyílik az eljárás objektivitásának, reliabilitásának és a validitásának a növelésére is (Grotjahn, 2000a; Malone, 2000).

A közvetlen és a közvetett tesztek megkülönböztetése azért problémás, mert a tesztelési gyakorlatban egy bizonyos szinten minden teszt közvetett (Krekeler, 2005). A tesztfeladatban megadott nyelvi szituációk (az autentikusságra való törekvés ellenére is) mesterségesek, a mérés nem kívánt beavatkozás; a vizsgahelyzet irreális és nyelvezete nem minden esetben természetes (Bárdos, 2003, 2004). A közvetlennek tartott nyelvi tesztek ezért nem tudják a hozzájuk fűzött elvárásokat teljesíteni (Bachman, 1990; Messick, 1996).

1.2. A kommunikatív nyelvi kompetenciák szintjeinek mérése és értékelése

Egy kommunikatív teszt elkészítésekor, illetve egy vizsga működtetésekor folyamatos mérlegelésre, fejlesztő munkára van szükség. Ehhez vizsgálni kell a kommunikatív tesztek jóságmutatóit, azok egymásra gyakorolt hatását, a különböző vizsgarészekben alkalmazott tesztfeladatokat, értékelési eljárásokat, valamint a teljes vizsgarendszer validitását. Ez a vizsgához kapcsolódó döntések, az eredmények értelmezése és a vizsga következményei alapján ítélt meg (Dávid, 2005). A vizsga belső validitásának kérdésköréhez tartozik még, hogy a vizsgarendszer mérési modellje igazodjon az alkalmazott nyelvtudásmodellhez. Ennek elengedhetetlen feltétele, hogy a vizsgafejlesztők megvizsgálják érvényességét a vizsgarendszer működése során. Ezen kívül a Referenciakeret szintleírása egy olyan külső kritérium, amely alapján az adott vizsgarendszer validitását is értékelhetjük. Ezt nevezzük a vizsga külső validitásának (Dávid, 2005; Glaboniat és Müller, 2006), ami azt jelenti, hogy az adott vizsga szintje megfelelően illeszkedik-e a Referenciakerethez.

A különböző nyelvvizsga-szintek összemérésének igényéből a nyelvtudásmérésben egy újabb kutatási irány bontakozott ki, amelynek célja a vizsgaszintek Referenciakerethez történő validálása. E folyamat alapjául szolgál a *Nyelvvizsgák szintillesztése a Közös európai referenciakerethez* (North, Avermaet, Figueras, Takala és Verhelst, 2003) kézikönyvben leírt szintillesztő eljárás és a hozzá tartozó módszertani segédlet (Banerjee, Kaftandjieva, Takala és Verhelst, 2004). Európai projektek indultak el azzal a céllal, hogy a Referenciakeret szintleírásai alapján nyelvi feladatbankot hozzanak létre. Ide tartozik például a DIALANG-projekt (*Diagnostic Language Assessment System for Learners*) és a 2008-ban lezárult EBAFLS-projekt (*Building a European Bank of Anchor Items for Foreign Language Skills*).

A nyelvi készségek mérésében, a nyelvi szintek bemérésében, valamint a nyelvi feladatbank kiépítésében a kommunikatív tesztelés támaszkodik a valószínűségi (modern) tesztelmélet (*Item Response Theory, IRT*) eszközeire. A nyelvtudásmérésben a valószínűségi tesztelméleti modelleket a nyolcvanas évek közepe óta használják (Nikolov, Pércsich és Szabó, 2000). Az elmúlt 10–15 évben jelentős szerephez jutott a nyelvvizsgák feladatainak elemzésében és kiértékelésében (Eckes, 2003), valamint a nyelvtudást mérő vizsgáknak a Referenciakerethez történő illesztésében vált domináns eszközzé.

Ebben a fejezetben először bemutatjuk a kommunikatív nyelvi kompetenciák referenciaszintjeit. Kitérünk továbbá a nyelvvizsga-szintek egységes meghatározásának szükségességére és folyamatára, melyben különböző európai nyelvtudásmérési projektek eredményeit mutatjuk be. Ezt követően ismertetjük a valószínűségi tesztelmélet alapjait, modelljeit és a nyelvtudásmérésben való alkalmazhatóságának lehetőségeit.

1.2.1. A kommunikatív nyelvi kompetenciák szintjei

A *Közös Európai Referenciakeret* (2002) alapvető célja, hogy „közös alapot teremtsen] Európa-szerte a nyelvi tantervek, tantervkészítési irányelvek, vizsgák, tankönyvek kidolgozásához” (1. o.). A Referenciakeret szándéka, hogy meghatározza a kommunikatív nyelvi kompetenciák szintjeit és ezáltal lehetővé tegye, hogy a tanulók haladása a nyelvtanulás minden szakaszában és az egész életük során mérhető legyen. A továbbiakban a Referenciakeret szintrendszerét mutatjuk be és jellemezzük. Ehhez példaként a *KER*-ben (2002) megadott általános szintrendszert használjuk (6. táblázat).

6. táblázat. Közös Referenciaszintek: általános szintrendszer (KER, 2002. 33. o.)

Függőleges dimenziók		Vízszintes dimenziók
Mesterfokú	C2	Szinte minden hallott vagy olvasott szöveget könnyedén megért. A különböző beszélt vagy írott forrásokból származó információkat össze tudja foglalni, összefüggően tudja felidézni az érveket, és pontosan tudja kifejezni magát, a jelentések kisebb árnyalatai között is különbséget tud tenni, még a bonyolultabb helyzetekben is.
nyelv-		
használó	C1	Meg tud érteni igényesebb és hosszabb szövegeket a különböző típusú szövegek széles körében, és a rejtett jelentéstartalmakat is érzékeli. Folyamatosan és természetes módon tudja kifejezni magát, anélkül hogy túl sokszor kényszerülne arra, hogy keresse a kifejezéseket. [...] Világos, jól szerkesztett, részletes szöveget tud alkotni összetettebb témákban is, és eközben megbízhatóan alkalmazza a szövegszerkesztési mintákat, kötőszavakat és szövegösszekötő elemeket.
Önálló	B2	Meg tudja érteni az összetettebb konkrét vagy elvont témájú szövegek fő gondolatmenetét [...]. Folyamatos és természetes módon olyan szintű normális interakciót tud folytatni anyanyelvű beszélővel, hogy az egyik félnek sem megterhelő. Világos, részletes szöveget tud alkotni különböző témák széles körében, és ki tudja fejteni a véleményét egy aktuális témáról úgy, hogy részletezni tudja a különböző lehetőségekből adódó előnyöket és hátrányokat.
nyelv-		
használó	B1	Megérti a fontosabb információkat olyan világos, standard szövegekben, amelyek ismert témákról szólnak, és gyakori helyzetekhez kapcsolódnak [...]. Egyszerű, összefüggő szöveget tud alkotni olyan témákban, amelyeket ismer, vagy amelyek az érdeklődési körébe tartoznak. Le tud írni élményeket és eseményeket, álmokat, reményeket és ambíciókat, továbbá röviden meg tudja indokolni és magyarázni a különböző álláspontokat és terveket.
Alapszintű	A2	Megért olyan mondatokat és gyakrabban használt kifejezéseket, amelyek az őt közvetlenül érintő területekhez kapcsolódnak [...]. Tud kommunikálni olyan egyszerű és begyakorolt feladatok során, amikor egyszerű és közvetlen információcserére van szükség ismerős és begyakorolt dolgokra vonatkozóan. Tud egyszerű nyelvi eszközöket használva beszélni saját háttéréről, szűkebb környezetéről és a közvetlen szükségleteivel kapcsolatos dolgokról.
nyelv-		
használó	A1	Megérti és használja a számára ismerős mindennapi kifejezéseket és nagyon alapvető fordulatokat, amelyeknek célja konkrét szükségleteinek kielégítése. Be tud mutatkozni és be tud mutatni másokat, meg tud válaszolni és fel tud tenni olyan kérdéseket, amelyek személyes jellegűek [...]. Képes egyszerű interakcióra, amennyiben a másik személy lassan és világosan beszél és segítőkész.

A Referenciakeret szintleírásait függőleges és vízszintes dimenziókra lehet bontani. A függőleges dimenzió egy olyan fokozatosan emelkedő referenciaszint-rendszert vázol, amelynek segítségével a nyelvtanulók kommunikatív nyelvi kompetenciáinak szintje leírható. Az Alapszintű nyelvhasználó szinthez az A1: Minimumszint (*Breakthrough*) és az A2: Alapszint (*Waystage*) tartozik. Az Önálló nyelvhasználó szinthez a B1: Küszöbszint (*Threshold*) és a B2: Középszint (*Vantage*) rendelhető, míg a Mesterfokú nyelvhasználó szintet a C1: Haladó szint (*Effective Operational Proficiency*) és a C2: Mesterszint (*Mastery*) alkotja (lásd részletesebben: KER, 2002. 3. fejezet). A Referenciakeret a szintrendszerét intervallumskálaként feltételezi: „az Alapszintet (A2) a skálán hozzávetőleg ugyanakkora távolság választja el a Küszöbszinttől (B1), mint amekkora az utóbbit a Középszinttől (B2)” (KER, 2002. 23. o.). Ugyanakkor a tanítási gyakorlat azt mutatja, „hogy sok diáknak a Küszöbszint elérése az Alapszintről kétszer annyi idejébe került, mint amennyit az Alapszint elérése igényelt. Feltételezhető, hogy több mint kétszer annyi ideig fog tartani, amíg a Küszöbszintről [...] elérik a Középszintet [...]. Ennek oka, hogy a nyelvi tevékenységek, a készségek és a nyelvi eszközök köre fölfelé haladva szükségképpen bővül” (KER, 2002. 23. o.). Ebből is következik, hogy a függő-

leges dimenzió szintjeit a Referenciakeret további szintekre osztja (lásd részletesebben a *KER*, 2002. „B” függelékét), és kilenc szintet határoz meg. A Referenciakeret megkülönbözteti a már ismertetett hat kritériumszintet (A2; B1; B2 stb.) és három ún. plusz szintet (A2+; B1+; B2+). A kritériumszint és a plusz szint megkülönböztetése szubjektív, azonban így lehetőség nyílik arra, „hogy a szintek és/ vagy szintleírások közös rendszerét különböző felhasználók helyi szükségletei kielégítése érdekében eltérő gyakorlati szintekre oszthatják, és mégis a közös rendszerhez kapcsolódnak” (*KER*, 2002. 41. o.). A plusz szint olyan „ismeretet határoz meg, amely magasabb a kritériumszint által képviselnél, ám nem éri el a következő szintet” (*KER*, 2002. 47. o.).

A Referenciakeret vízszintes dimenziói (6. táblázat) ismertetik azokat a kategóriákat, amelyek a kommunikatív tevékenység és a kommunikatív nyelvi kompetenciák paramétereit (deskriptorait) tartalmazzák. Ezek meghatározzák, hogy mi várható el a nyelvtanulótól a különböző szinteken, így lehetővé teszi annak eldöntését, hogy az adott területeken a tanulók teljesítménye megfelel-e a nyelvtudás adott szintjéhez rendelhető színvonalnak (*KER*, 2002). A Referenciakeret szintleírásai pozitív irányultságúak, cselekvésorientált megközelítésűek, azaz azt jelölik, hogy a tanuló mit tud a nyelvvel kezdeni. A szintek leírásának alapelve, hogy ne legyen pontatlan, túl általános, hanem a szintek között konkrét különbséget lehessen tenni. A szintrendszer további deklarált célja a világos és logikus felépítés, továbbá az egyes szintekhez tartozó deskriptorok átláthatósága, rövid megfogalmazása, azok kontextus nélküli értelmezhetősége, amely lehetővé teszi hogy kritériumlistává alakítva az értékelés során is használhatóak legyenek (*KER*, 2002; *Glaboniat és Müller*, 2006).

A nyelvi tesztelés szempontjából a Referenciakeretet egyre több kritika is éri. A *Dutch CEF Construct Project* (*Alderson, Figueras, Kuijper, Nold, Takala és Tardieu*, 2004) keretében végzett elemzés és a dokumentum német nyelvi fordítása során *Quetz* (2001, 2003) rámutatott arra, hogy a szintek leírása nyelvi nem mindenhol pontos, a vízszintes dimenzió leírásához alkalmazott fogalmak és kategóriák különböző szintekbe is besorolhatók. Ezen kívül a szintek túldimenzionáltak, a Referenciakeret szintleírásai nehezen használhatóak tesztfeladatok készítéséhez, a produktív nyelvi készségek mérésénél javítási és értékelési eljárások kialakításához (*Altmayer*, 2004).

A Referenciakeret fontos és hasznos dokumentum, alkalmazásához viszont szükséges a nyelveken átívelő szintleírások konkretizálása, így megjelennek olyan kiadványok is, amelyek a Referenciakeret alapján egy konkrét nyelvre kidolgozva tartalmazzák többek között a kommunikációs szándékokat, a nyelvtani szerkezeteket és tematikus szókincset, amelyeket a vizsgázóknak vagy receptív, vagy produktív szinten kell tudniuk alkalmazni. Ezeket az információkat tartalmazza például a *Profile Deutsch* (*Glaboniat, Müller, Rusch, Schmitz és Wertenschlag*, 2002, 2005) a Referenciakeret összes szintjére. A program a szintleírások között úgy differenciál, hogy definiálja az általános (*Globale Kann-Beschreibungen*), és a részletes szintleírásokat (*Detaillierte Kann-Beschreibungen*). Az általános szintleírás meghatározza, hogy a vizsgázó milyen jól tudja használni a nyelvet, a részletes pedig megadja, hogy konkrétan mit tud a vizsgázó (*Altmayer*, 2004; *Seiffert*, 2003; *Glaboniat és Müller*, 2006; *Perlmann-Balme*, 2006; *Boócz-Barna*, 2004).

1.2.2. A nyelvi szintek validálási eljárásai, az európai projektek eredményei

A különböző nyelvtudást mérő vizsgákat összekapcsolja a Referenciakeret, amely megteremti az egyes nyelvvizsgáztató központok szintjeinek átláthatóságát, a nyelvvizsgák kölcsönös elismerését és lehetővé teszi a minősítési rendszerek közötti összehasonlíthatóságot (*Bárdos*, 2006; *Boócz-Barna*, 2004; *Boócz-Barna és Majorosi*, 2002). Ha a külön-

böző vizsgákat közös nyelvismereti skálán helyezzük el (7. táblázat), akkor lehetségesnek kell lennie, hogy a vizsgarendszer egyik vizsgáján elért teljesítmény egy másik vizsgán elért eredményhez viszonyítható legyen (KER, 2002).

7. táblázat. *Különböző vizsgarendszerek szintjei (Akkreditációs Kézikönyv, 2007. 28. o.; Bárdos, 2005. 10. o.; University of Cambridge ESOL Exams; Zertifikate und Prüfungen für DaF, 2006. 45. o. alapján)*

KER szintjei	University of Cambridge	Goethe Institut	Magyarországi államilag elismert nyelvvizsgák
C2	CPE (Certificate of Proficiency in English)	ZOP (Zentrale Oberstufenprüfung) KDS (Kleines Deutsches Sprachdiplom)	
C1	CAE (Certificate in Advanced English)	ZMP (Zentrale Mittelstufenprüfung)	Felsőfok
B2	FCE (First Certificate in English)	ZDfB (Zertifikat Deutsch für den Beruf)	Középfok
B1	PET (Preliminary English Test)	ZD (Zertifikat Deutsch)	Alapfok
A2	KET (Key English Test)	Fit in Deutsch 2 Start Deutsch 2	
A1		Fit in Deutsch 1 Start Deutsch 1	

Megjegyzés: A *Fit in Deutsch* a fiatalok, míg a *Start Deutsch* a felnőttek számára készített vizsga. Az Akkreditációs Kézikönyv első kiadásában a magyarországi államilag elismert nyelvvizsga-szintek nem estek egybe a KER szintjeivel.

A Referenciakeret nem ad módszertani ajánlásokat arra vonatkozóan, hogy egy vizsga szintjét hogyan lehet a referenciaszintekhez validálni. Ezt tükrözi Alderson 2002 júliusában egy Helsinkiben tartott konferencián megfogalmazott és mára már klasszikussá vált kérdése: „How do I know that my Level B1 is your Level B1?”. A validálási folyamat alapjául szolgál a *Nyelvvizsgák szintillesztése a Közös európai referenciakerethez* (North és mtsai, 2003) kézikönyvben leírt szintillesztő eljárás és a hozzá tartozó módszertani segédlet (Banerjee és mtsai, 2004). Az illesztő eljárás négy szakaszból áll. (1) Az ismerkedési szakaszban a vizsgáztató központ alaposan megismeri a Referenciakeretet, és a szintek leírásait megbízhatóan elsajátítja. (2) A specifikációs szakaszban a vizsgaközpont tartalmi elemzést készít, amelyben már az esetlegesen meglévő nyelvvizsgaleírást kell a KER rendszerével összevetni, majd a specifikáció során felülvizsgálni, átdolgozni és igazítani. (3) A külső validálás eljárásának standardizációs szakaszában a feladatírók, vizsgáztatók KER-re vonatkozó „megítélésének” standardizálása a fő feladat, illetve az eddigi vizsgafeladatok és az értékelési rendszer standardizálásának felülvizsgálata a KER kalibrált példái alapján. Ebben a szakaszban a vizsgaközpont elindítja az új folyamatok tervezését, beiktatását a korábbi rendszer elemei közé. A vizsgafejlesztők és feladatírók bevonásával történik a KER által definiált követelményeknek megfelelő feladatírás. (4) Az empirikus validálás szakaszában valósul meg a feladatok kipróbálása próbamérés szervezésével és lebonyolításával. A szakasz célja az értékelési rendszer standardizációjának, megbízhatóságának, és a feladatok nehézségi szintjének meghatározása. Ezután a vizsgaközpont megvizsgálja a feladattípusok működését, amelynek tanulságai alapján korrigálja és változtatja a vizsgaleírást (az egyes fázisokról lásd részletesebben: Dávid, 2005; Zeidler, 2009).

A Referenciakeret szintjeihez történő validálási folyamatban a DIALANG és az EBAFLS-projekt célja, hogy a Referenciakeret szintleírásai alapján feladatbankot hoz-

zának létre. A továbbiakban e két európai kutatási projekt eredményeként létrejövő DIALANG tesztrendszer és az EBAFLS-projekt keretében kiépített referenciatankbank jellemzőit tekintjük át.

A DIALANG egy olyan online adaptív nyelvi értékelési rendszer, amely a Referenciakeret diagnosztikai célokra történő alkalmazásaként jött létre. A rendszer a Referenciakeret deskriptorai alapján kalibrált önértékelési skálákat és nyelvi teszteket tartalmaz, és ezek alapján nyújt diagnosztikus információkat. A rendszer eredményként megadja, hogy a tanulók az olvasás és a hallás utáni szövegértés, az írás, a nyelvtan és szókincs területén milyen szinten állnak, valamint visszajelzést ad készség szintjük erős és gyenge pontjairól, továbbá tanácsot ad a készségek fejlesztéséhez (KER, 2002; Alderson és Huhta, 2005).

Az EBAFLS-projekt alapvető célja a Referenciakeret alapján feladatbank létrehozása. A projektet a CITO szervezet koordinálta, és ebben hét európai ország, köztük Magyarország vett részt. A projekt fő célja, hogy angol, francia és német nyelvű olvasott és hallott szövegértés méréséhez B1 szintre hozzanak létre feladatbankot, valamint vizsgálják a mérőeszközök validitását és reliabilitását, illetve kézikönyvet készítsenek, melyben olyan alapelveket fogalmaznak meg, amelyek a feladatbank kibővítésére és új feladatbank létrehozására szolgálnak. A projekt eredményeként olyan referenciatankbankokat akartak létrehozni, melyek tartalmazzák az itemek pszichometriai jellemzőit, és megadják a Referenciakeret szintjeit, hogy elősegítsék az adott nyelvvizsga szintjeinek a Referenciakerethez való illesztését (CITO, 2008).

A validálási folyamatban a feladatok Referenciakerethez történő illeszkedésének vizsgálatához referenciaitemeket tartalmazó feladatokra van szükség. „Referencia olyan item lehet, amelynek nehézségi értékei vagy a referenciaskálához kapcsolt, „kihorgonyzott” mérésben keletkeztek, vagy attól nem függő mérés esetében a bemérendő és a referenciaskála skálatulajdonságainak kell egyeznie” (Dávid, 2005. 290. o.). A referenciaitemek segítségével meghatározhatjuk azt a nehézségi szintet, amelyhez a többi item nehézségét viszonyítani lehet, és így megállapítható, hogy a vizsgafeladatok itemei mennyivel könnyebbek-nehezebbek a referenciaitemeknél. Ebben az eljárásban is szükséges, hogy legyenek közös itemek. A feladatokat összekötő horgonyitemek (*anchor items*) segítségével lehetőség nyílik a feladatsorok együttes elemzésére, a feladatok egy skálára hozására, tehát közvetlenül összehasonlíthatóvá válik az item nehézsége és a diákok képességszintje (Fischer, 1981).

A horgonyitemek és a Referenciakeret szintjeihez kapcsolt, kalibrált referenciaitemek alkalmazásával számos módszertani kérdés (ezekről lásd részletesen: Szabó, 2008) merül fel. Az egyik problémakör a referenciatankbank megfelelő kiválasztása. Az EBAFLS-projekt kutatási beszámolója (CITO, 2008) egyértelműen utal a feladatbank alkalmazásának korlátaira, a projekt részleges kudarcára. Ennek oka, hogy sok item eltérően működött (*differential item functioning*) az egyes országokban. Ez az itemműködés azt jelenti, hogy a különböző országokból származó tanulók más valószínűséggel oldották meg az itemeket. A helyes válasz valószínűsége nemcsak a mérendő tulajdonságtól, hanem más tényezőktől is függ. Ez az itemműködés jelezheti a tanterv és a nyelvoktatási gyakorlat közötti különbségeket, az eltérő mérési és értékelési hagyományokat (egy- vagy kétnyelvű vizsgák alkalmazása). Befolyásoló tényező még a feladattípusok előzetes ismertsége, továbbá az is, hogy az adott nyelv struktúrája mennyire van közel a mérendő idegen nyelvhez. A beszámoló egyértelműen jelzi, hogy a különböző itemműködés okainak vizsgálatához további kutatásokra van szükség a nyelvi mérés és értékelés területén. A problémák kiküszöbölése érdekében az egyes országok szintjén is megadják az itemek nehézségét és diszkriminációs indexét. A másik kérdéskör, hogy mennyi horgony- és referenciaitemre van szükség, amelyekkel az adatbázis-

kat megfelelően össze lehet kötni (Szabó, 2006), illetve megteremtik a kapcsolatot a Referenciakeret szintrendszerével.

A nyelvvizsgaszintek validálása fontos kutatási feladatokat határoz meg. Egyrészt meg kell állapítani, hogy egy vizsgarendszer adott szintje hogyan illeszkedik a Referenciakeret szintjeihez, másrészt pedig, hogy a különböző vizsgák tesztfeladatai, értékelési eljárásai mennyiben felelnek meg a Referenciakeret által meghatározott nyelvtudás szintjeinek. E probléma megoldásában a kommunikatív tesztelés támaszkodik a valószínűségi tesztelméletre. A kommunikatív tesztelésben az IRT eszközöket a DIALANG skálák szintezésénél (KER, 2002. „C” függelék; Figueras, 2005; Perlmann-Balme, 2006), az ALTE (1998) „Mit tud?”-állításainak a Referenciakeret szintleírásával történő megfeleltetése során (KER, 2002. „D” függelék) és az EBAFLS-projekt keretében alkalmazták.

1.2.3. A kommunikatív nyelvi készségek és az IRT-modellek

A Referenciakeretben definiált kommunikatív nyelvi kompetenciák fogalma egy többkomponensű, közvetlenül nem mérhető és értékelhető konstruktumot jelöl. A kommunikatív tesztelés során a kommunikatív nyelvi kompetenciák szintjeinek és összetevőinek mérése és értékelése indirekt módon, a receptív (beszédértés, olvasás) és a produktív (beszéd, írás) készségek alapján történik pontosan definiált alapelvek mentén (Alderson, Clapham és Wall, 1995). A nyelvi készségek mérésével és értékelésével lehetőség nyílik a konstruktum mozgósítására.

A valószínűségi tesztelméleti eszközök alkalmazásának alapfelvetése, hogy egy tesztitemre adott helyes vagy helytelen választ fizikailag nem megfigyelhető olyan látens tulajdonság magyarázza, amely a feladatokat megoldó személy válaszainak konzisztenciájáért felelős (Wainer és Messick, 1983, idézi Nikolov, Pércsich és Szabó, 2000). Ez a látens tulajdonság azt is meghatározza, hogy egy személy az adott itemet milyen valószínűséggel oldja meg (Eckes, 2003).

Az IRT-modellek egyszerre követik nyomon, hogy a vizsgázó milyen választ adott és milyen válasz lett volna a legvalószínűbb. Az így kapott értékeket a modellek folyamatosan összevetik, meghatározzák az eltéréseket és így eljutnak a szisztematikus tényezők mintafüggetlen becsléséig és a képességparaméterek tesztfüggetlen meghatározásáig. Mindegyik IRT-modellre jellemző, hogy az egyes itemeknél megadják a személy helyes válaszána valószínűségét, és ha ismert az itemek nehézségi indexe és a diákok képességparamétere, akkor megadják, hogy minden egyes diák milyen valószínűséggel oldaná meg jól külön-külön az egyes itemeket (lásd például Molnár, 2006). Ezeket a modelleket csoportosíthatjuk az összefüggések típusa és az itemparaméterek száma szerint. A továbbiakban három modell alkalmazási lehetőségét mutatjuk be a nyelvtudásmérésben.

A Rasch-modell (*Rasch's simple logistic model*) dichotóm adatok elemzésére alkalmas. A modell az item nehézségét és a vizsgált személy képességét kezeli egy paraméterként és az itemek diszkriminációs indexét azonosnak veszi. A nyelvi mérés és értékelésben főként a receptív készségek dichotóm itemeinek vizsgálatára alkalmas. Ezek mellett pedig kimutatható, hogy az itemnehézségi indexek mennyiben felelnek meg a vizsgázók nyelvtudás-szintjének. További jellemzője, hogy az adott csoportban megoldott tesztitemek működéséből arról is tudunk következtetést levonni, hogy azok hogyan működnének egy másik csoportban (Nikolov, Pércsich és Szabó, 2000).

A nyelvtudást mérő nemzetközi és hazai vizsgák produktív készségeket mérő vizsgarészében performancia teszteket és analitikus értékelési skálákat használnak, amelyek egységesebb és objektívebb értékelést biztosítanak. Az analitikus értékeléskor először a

feladat megoldásához szükséges komponenseket határozzák meg, amelyekhez értékelési szempontokat rendelnek, majd ezeket különböző skálákkal külön értékelik (*DeAyala, Dodd és Koch, 1991*). A skálákra jellemző, hogy minden pontszámhoz szintleírás tartozik. Az értékelők feladata, hogy a diákok teljesítményét a megfelelő minőségi kategóriákba sorolja be. A produktív készségek fejlettségi szintjének megítélésekor a mért eredményt számos tényező befolyásolja. Ezek közé tartozik a vizsgázók készségének fejlettsége, az értékelők tulajdonságai, az értékelési szempontok jellemzői, melyek pontszámait az értékelési skálában definiálnak (*Engelhard, 1992*).

Míg a klasszikus elemzési módszerekkel főként a produktív készségek fejlettségi szintjét meghatározó háttérváltozók hatását tudjuk kifejezni, addig a valószínűségi tesztelméleti modellekkel számszerűsíthető a vizsgázók készségszintje, egy jól behatárolható intervallumban kifejezhető és jellemezhető a diákok teljesítménye. Ezen kívül megállapítható, hogy az értékelési skálapontok mennyire jól fedik le a vizsgázók produktív készsége alapján meghatározott intervallumot.

Az értékelők megítéléseinek vizsgálata a nemzetközi nyelvtudást mérő vizsgákon a hazainál jóval intenzívebben elemzett terület. Számos kutató (az íráskészség esetében például *Weigle, 1998; Engelhard és Myford, 2003; Eckes, 2005, 2008; Schoonen, 2005*; a beszédkészség esetében például *Vidakovic és Galaczi, 2009*) modellezte az értékelők tulajdonságait. E vizsgálatok tanulságai szerint az értékelők különbözhetnek az értékelési szempontok értelmezésében és alkalmazásában, a vizsgázók nyelvi teljesítményének szigorú vagy enyhe megítélésében, az értékelési skálák megértésében és használatában, a különböző készségszintű diákok értékelésének következetességében. E különbségek oka feltehetően a kognitív és döntéshozó folyamatokban van. Jellemző, hogy különböző értékelők a szóbeli vagy írásbeli produktum egyes komponenseit kiemelve értékelik. *Linacre (1994)* sokoldalú modellje lehetővé teszi, hogy az értékelők csoportjában meghatározzuk az értékelők szigorúságát. A modell alkalmazásával például a TestDaF nyelvvizsgán minden értékelőhöz hozzárendelnek egy szigorúsági mutatót, amit a vizsgázók nyelvi teljesítményének értékelésekor figyelembe vesznek (*Eckes, 2003, 2004*). *Barrett (2001)* és *Weigle (1998)* szerint a vizsgáztatói képzések kevésbé alkalmasak az értékelők egyéni normáinak egységesítésére.

Az értékelők döntési folyamatának szempontjából az értékelési szempontok közvetítenek abban, ahogyan megítélik a vizsgázó produktumát és értékelik pontszámokkal a teljesítményt (*Wolfe, 1997; Lumley, 2002; Eckes, 2008*). Az értékelési szempontok kialakításában alapelv, hogy irányítsák az értékelők figyelmét a készség főbb komponenseinek értelmezésében és a feladatban definiált kommunikatív célok elérésének megítélésében (*Cumming, Kantor és Powers, 2001*).

Az értékelési skálák akkor működnek megfelelően, ha minden szintleírást (deskriptort) azonos módon használnak a vizsgáztatók, és ha a skálák alkalmasak arra, hogy a vizsgázók nyelvi teljesítményének szintjeit pontosan kifejezzék (*Dávid, 2006*). Az analitikus értékelési skálák kritériumorientált értékelést tesznek lehetővé. Ennek jellemzője, hogy a maximális pontszámot leíró deskriptor a vizsgakövetelményekben meghatározott nyelvi teljesítményt, az elérendő kritériumot reprezentálja. A többi értékelési sáv kialakításakor az elérendő kritériumot szintekre bontják és minden szintnek részletes leírása van, amelybe a vizsgázó írásműve besorolható. Így a vizsgázó egyéni teljesítményét az elvárt szinthez viszonyítják (*Vígh, 2008*). A deskriptorok diagnosztikus információt szolgáltatnak arról, hogy a vizsgázónak mennyit kell fejlődnie, hogy a kritériumot elérje (*Bárdos, 2002*). *Masters (1982)* parciális kredit modelljével az értékelési szempontokhoz tartozó skálák empirikus validálása biztosítható (*Park, 2004*). A modellel azt lehet vizsgálni, hogy a vizsgáztatók hogyan alkalmazzák a skálákat, azaz

azt minősíthetjük, hogy minden deskriptor hasznosul-e, képes-e a skála a nyelvi teljesítmények minőségének árnyalt kifejezésére (Dávid, 2006).

A valószínűségi tesztelméleti modellek az idegen nyelvi mérés és értékelés területén széles körben alkalmazhatók, mivel sok a hasonlóság a nyelvtudás mérésének lehetőségei és a valószínűségi tesztelméleti eszközök alapfelvetése között. A nyelvtudást mérő vizsgák szintjeinek kialakításában, egységesítésében, a nyelvi feladatok nehézségi szintjének mérésében, valamint az értékelési eljárások megbízhatóságának növelésében az IRT-modellek fontos szerepet töltenek be.

1.3. A kommunikatív tesztelés hatása a nyelvoktatásra

A nyelvtudást mérő vizsgákkal szemben az utóbbi évtizedben komoly elvárások fogalmazódtak meg. Ezek közé tartozik a kommunikatív nyelvi kompetenciák fejlettségének mérése, a Referenciakeret szintjeihez való illeszkedésük. További elvárások, hogy a vizsga módszereiben kövesse a kommunikatív nyelvoktatást, pozitívan hasson a tanítási-tanulási folyamatra és segítse elő, hogy a nyelvtanuló alkalmazható tudást szerezzen.

A vizsgák tanítási-tanulási folyamatra gyakorolt hatását a pedagógiai értékelésben gyakran természetes folyamatként értelmezik, így míg működésének, mechanizmusainak, a folyamatot befolyásoló tényezőknek a meghatározása háttérbe szorul, addig az idegen nyelvi mérés és értékelés területén igen gazdag nemzetközi szakirodalma van. A *Language Testing* (Alderson és Wall, 1996), az *Assessment in Education* (Rea-Dickins és Scott, 2007) és a *Cambridge ESOL Research Notes* (Barker és Khalifa, 2009) folyóirat különszámában közölt tanulmányokat a témáról. Ezen kívül számos tanulmánykötet (Bailey, 1999; Cheng, 2005; Wall, 2005), PhD-értekezés (például Chen, 2002; Cheng, 1998; Saif, 1999), valamint Cheng, Watanabe és Curtis (2004) könyve mutatta be ezt a pedagógiai értékelés területéről származó jelenséget az idegen nyelvi mérés és értékelés rendszerében. Ugyanakkor a teszthatás mechanizmusainak bemutatását egyetlen hazai publikáció sem tűzte ki célul. Elsősorban a nyelvi mérés és értékelés területén jelenik meg kritériumként (Bárdos, 2002), továbbá a nyelvi tesztelés etikai kérdéseinek kontextusában mutatta be Dévény Ágnes (2003). A hazai kommunikatív szemléletű nyelvtudást mérő vizsgák tanítási-tanulási folyamatra gyakorolt hatásáról jelenleg semmilyen empirikus vizsgálat eredménye nem áll rendelkezésre annak ellenére, hogy ez a jelenség a nemzetközi szakirodalomból megismerhető és az empirikus vizsgálatok eredményei szintetizálhatók.

E fejezet célja a pedagógiai értékelés és az 1990-es években elinduló nyelvi mérés és értékelés területén publikált elméleti és empirikus kutatások eredményeinek szintetizáló áttekintése. Először értelmezzük a nyelvtudást mérő tesztek/vizsgák hatásának fogalmát, a különböző koncepciókban betöltött szerepét. Ezt követően olyan elméleti modellek fejlődését mutatjuk be, amelyek a teszthatás folyamatát és működését értelmezik. Végül az idegen nyelvi mérés és értékelés területén publikált empirikus kutatások eredményeit tekintjük át. Kitérünk arra, hogy milyen módszerrel és a tanítási-tanulási folyamat mely területén vizsgálták ezt a jelenséget, továbbá, hogy a kutatási eredményekből milyen következtetések vonhatók le.

1.3.1. A teszthatás fogalma

Az angol nyelvű szakirodalomban a nyelvtudást mérő tesztek/vizsgák hatását a *washback*, *backwash* és az *impact* fogalmakkal határozzák meg. Bailey (1999) szakirodalmi áttekintésében megállapítja, hogy a vizsgák hatásának fogalmi szintű megnevezése nem egységes, ugyanazt a jelenséget többféleképpen, más-más jelentéstartalommal

definiálják a kutatók. A fejezetben (1) a *washback* és *backwash*, valamint (2) az *impact* kifejezéseket magyarul a teszthatás/ visszahatás fogalmával fordítjuk, és az angol kifejezést alkalmazzuk, mivel az angol nyelvű szakirodalomban jobban elkülöníthető a kifejezések eltérő jelentéstartalma. A vizsgák hatásának értelmezésében (3) az értékelés által vezérelt oktatás (*measurement driven instruction*) és (4) a tantervi összehangolás (*curriculum alignment*) koncepciója is jelentős szerepet tölt be. A visszahatást a tesztek validitásával is összekapcsolják a kutatók, és a kettő viszonyát (5) a visszahatásvaliditás (*washback validity*), (6) a szisztematikus validitás (*systematic validity*), valamint (7) a következményes validitás (*consequential validity*) fogalmával határozzák meg. A továbbiakban a teszthatás egy-egy aspektusát megjelenítő kifejezéseket, koncepciókat tekintjük át.

A *backwash* az általános pedagógiai kutatásokban a teszt negatív hatását jelenti (Cheng és Curtis, 2004), a brit nyelvpedagógiai kutatásokban pedig semleges jelentést tartalommal, a *washback* szinonimájaként használják (Schifko, 2001). Az eltérés ezen a területen abban jelenik meg, hogy a kutatók más szempontokra helyezik a hangsúlyt. Mindkét fogalom a tesztek tanítási-tanulási folyamatra gyakorolt hatására (Hughes, 1989), a folyamatban résztvevőkre (Alderson és Wall, 1993; Messick, 1996), valamint a tananyagra, a tantervi és a tanítási változásokra (Cheng, 1998, 1999), továbbá a tanulók tanuláshoz való hozzáállására vonatkozik (Biggs, 1995, idézi Cheng, 2005. 27. o.).

Néhány szerző az *impact* kifejezést főfogalomnak tekinti, amelynek egyik eleme a *washback* (Taylor, 2005). Hamp-Lyons (1997) szerint a *washback* a teszteknek az oktatásra gyakorolt hatására és befolyásolására utal, míg az *impact* tágabb értelemben véve az oktatási rendszer és a társadalom egészére vonatkozik. Wall (1997) is különbséget tesz az *impact* és a *washback* fogalma között: „Az *impact* olyan hatásokra utal, amelyet egy teszt gyakorolhat az osztálytermen, az iskolán, az oktatási rendszeren vagy a teljes társadalmon belül az egyénekre, az eljárásokra vagy a gyakorlatra. [...] A *washback* (backwash-ként is ismert) fogalma a tesztelésnek a tanításra és a tanulásra gyakorolt hatására vonatkozik.” (291. o.)

Bachman és Palmer (1996) csak az *impact* fogalmát használják, és két aspektusát hangsúlyozzák: a teszt alkalmazása és az eredmények értelmezése hatással van a mikrokörnyezetre (a tanítási-tanulási folyamatban résztvevőkre) és a makrokörnyezetre (a társadalomra és az oktatási rendszerre). A teszthatás ezen értelmezéséhez hasonlóan Shohamy (1999) szerint „az *impact* (backwash vagy washback) olyan viselkedési változásra vonatkozik, amely a teszt alkalmazásának eredményeként történik az oktatási rendszerben és az egész társadalomban” (711. o.). Shohamy (1999, 2001) az *impact* két aspektusát, az oktatásra (*education impact*) és a társadalomra gyakorolt hatást (*social impact*) különbözteti meg, amely elsősorban a nagy téttel rendelkező formalizált tesztekre jellemző. A visszahatás két aspektusának megkülönböztetése egységesen elfogadottá és alkalmazottá vált. Ezt mutatja például, hogy a *The Association of Language Testers in Europe (ALTE)* 2008-ban szervezett konferenciáját a *The Social and Educational Impact of Language Assessment* címmel rendezte.

A vizsgák hatásának értelmezésében az értékelés által vezérelt oktatás (*measurement driven instruction*) koncepciója jelentős szerepet tölt be (Popham, 1987, idézi Cheng és Curtis, 2004. 4. o.; Noble és Smith, 1994; Smith, 1991). E megközelítés arra vonatkozik, hogy amit és ahogyan mérünk és értékelünk, az az oktatási folyamatban értéké válik (McEwen, 1995). E koncepció elsősorban olyan standardizált tesztekre vonatkozik (például érettségi, felvételi vizsga), amelyeknek jelentős tétje, következménye van a tanulóknak, valamint a tanároknak és az iskoláknak (például a közölt teszteredményekből vonnak le következtetéseket az oktatás színvonalára), tehát a tesztek jelentős mértékben meghatározzák a tanítási-tanulási folyamatot.

A vizsgák hatásának más értelmezésében a tantervi összehangolás (*curriculum alignment*) koncepciója a tanterv, a tanítás és a tesztelés közötti kapcsolatot jelöli (Leitzer és Vogler, 1994). A tantervi összehangolásból az is következik, hogy a tantervet a teszteredmények alapján is módosíthatják. Chen (2002) a 8. táblázatban közölt módon összegzi a tantervi összehangolás két formáját: a megelőző (*alignment by frontloading*) és az utólagos (*alignment by backloading*) összehangolás jellemzőit.

8. táblázat. A megelőző és az utólagos tantervi összehangolás jellemzői (English és Steffy, 2001, idézi Chen, 2002. 31. o. alapján)

	Tantervalapú értékelés	Folyamatalapú értékelés
Megelőző	Tantervírás, majd tesztfejlesztés a tanterv követelményeinek értékelésére.	Tanítás, majd tesztfejlesztés a tanítás tartalmának értékelésére.
Utólagos	Nyilvánosan közzétett tesztitemek megszerzése, és ezekre alapozott tantervfejlesztés.	Nyilvánosan közzétett tesztitemek megszerzése, és ezekre épített osztálytermi folyamatok.

A teszt, a tanterv és a tanítás kapcsolatában a teszt szerepe gyakran központi jelentőségűvé válik. Madaus (1988, idézi Spratt, 2005. 5. o.) szerint a tesztelés és nem a „hivatalosan” meghatározott tanterv az, amely egyre inkább befolyásolja az osztálytermi folyamatokat. E kettősséget felismerve, az újabb értékelési irányzatokban (például a kompetenciaértékelésben, vagy a nyelvi mérés és értékelés területén a fejlesztő értékelésben) a tanítási-tanulási folyamatba integrálják az értékelést és ehhez dolgoznak ki minőségi mutatókat (lásd részletesebben: 1.1.3. fejezet).

A terminológiai viták és a teszthatás különböző értelmezése mellett fontos kérdéssé vált az is, hogy milyen a teszthatás és validitás kapcsolata, vagyis a visszahatás pozitív vagy negatív értéke mennyiben határozza meg egy teszt validitását. Az 1960-as évektől végzett pedagógiai kutatások elsősorban a negatív hatásokból indultak ki (például Wiseman, 1961, idézi Saif, 1999. 40. o.; Smith, Edelsky, Draper, Rottenberg és Cherland, 1990), amelyet az értékelés által vezérelt oktatás koncepciója is befolyásolt. Az 1980-as években megfogalmazódott az igény arra, hogy a teszt hatását rendeljék hozzá a validitáshoz, mivel a tanítási-tanulási folyamatban így elérhető pozitív irányú visszahatás. Ennek oka, hogy a vizsgák hatása a nyelvi tesztelés etikai kérdéseivel is összekapcsolódott (Hamp-Lyons, 1997), mivel az új vagy módosított tesztek a nyelvoktatásban tartalmi és módszertani változásokat indíthatnak el. Így fontos biztosítani, hogy a tesztkészítők pozitív irányú, azaz a tesztfejlesztési folyamatban megjelenő szándékoknak megfelelő hatást érjenek el. E cél elérése érdekében több koncepciót fogalmaztak meg.

Morrow (1986) a validitás elsődleges aspektusaként nevezi meg a visszahatás-validitást (*washback validity*), amely a teszt és az ahhoz társított tanítási-tanulási folyamat közötti viszony minőségét mutatja. A kritérium bevezetésének egyrészt az a célja, hogy a tesztfejlesztőket rákényszerítse annak megismerésére, milyen hatása van a teszteknek a tanítási-tanulási folyamatra, másrészt, hogy pozitív hatást érjenek el. Morrow (1986) ugyanakkor utal arra, hogy a visszahatás-validitást nehezen lehet mérni. Saif (1999) szerint ez a megközelítés nem ad választ arra a kérdésre, hogy egy teszt miért nem tekinthető érvényesnek abban az esetben, ha a teszt bevezetésével nem valósul meg a tesztkészítők szándéka, vagy ha a tanítási-tanulási folyamatban nem szándékolt hatások is megjelennek.

Frederiksen és Collins (1989) koncepciójában a szisztematikus validitás (*systematic validity*) „a teszt integrálását jelenti az oktatási rendszerbe és annak a szükségességét bizonyítja, hogy egy új teszt bevezetése tökéletesíti a tanítást és a tanulást. Egy szisztematikus valid teszt (*systemically valid test*) az oktatási rendszerben tantervi és oktatási

változásokat idéz elő, amely így elősegíti a mérendő készségek fejlesztését is” (Frederiksen és Collins, 1989. 27. o.). E megközelítés célja, hogy a teszt egy dinamikus folyamat részévé váljék, amely elősegíti az oktatási rendszer változásait, valamint a tanítási-tanulási folyamat fejlesztését. Ezt a feltételezést szintén kritika érte. Messick (1996) szerint ennek értelmében egy rossz (*poor*) tesztet – amelynek tartalma nem egyezik meg az oktatás céljaival és tartalmával – szintén validnak kell tekinteni abban az esetben, ha pozitív hatása van, tehát ha valamilyen változást indít el az oktatási rendszerben.

Morrow (1986) valamint Frederiksen és Collins (1989) koncepciójának problémáit figyelembe véve megállapítható, hogy a „visszahatás (*washback*) egy olyan komplex jelenség, amelyet nem lehet közvetlenül hozzárendelni a teszt validitásához” (Alderson és Wall, 1993. 116. o.). A teszthatás ezért a tanulás és a tanítás különböző aspektusaira vonatkozik (Cheng és Curtis, 2004), és pozitív vagy negatív értéke nem befolyásolja a teszt validitását.

Messick (1996) validitáskonceptiójában a következményes validitás (*consequential validity*) a teszt validitása és hatása közti összefüggéseket jelenti. Messick (1996) elképzelése szerint a következményes validitás egy aspektusa a szerkezeti validitásnak (1.1.3. fejezet), így sem a tesztelés következménye általában, sem a visszahatás nem áll egyedül. A következményes validitás nemcsak a tesztalkalmazás és a teszteredmények szándékolt és nem szándékolt hatásait jelenti, hanem az oktatási rendszerre és a társadalomra gyakorolt következményeket is. Bachman és Palmer (1996), valamint Kunnan (2004) kritériumrendszerében a teszthatás az egyik összetevő. E kritériumrendszerek meg- egyeznek abban, hogy a teszthatás a validitás egyik aspektusát jeleníti meg.

A vizsgák hatásának fogalmi szintű értelmezéséről és a különböző koncepciókban betöltött szerepéről eltérő, és egymásnak ellentmondó elképzelések alakultak ki. A jelenség pontos megismeréséhez szükséges a következő kérdések tisztázása. Mire vonatkozik a vizsgák hatása: a tanulókra, a tanárookra, a tanuláshoz való viszonyukra, az oktatás tartalmára, a módszerekre, a tantervre, a tankönyvek megfogalmazására, a teszt validitására? Minden tanulóra és tanárra egyaránt érvényesek, azonos mértékűek a hatások (Krekeler, 2002)? A továbbiakban ezekre a kérdésekre keresünk válaszokat a teszthatás elméleti és empirikus kutatásainak bemutatásával.

1.3.2. A teszthatás folyamatát leíró elméleti modellek és tényezők

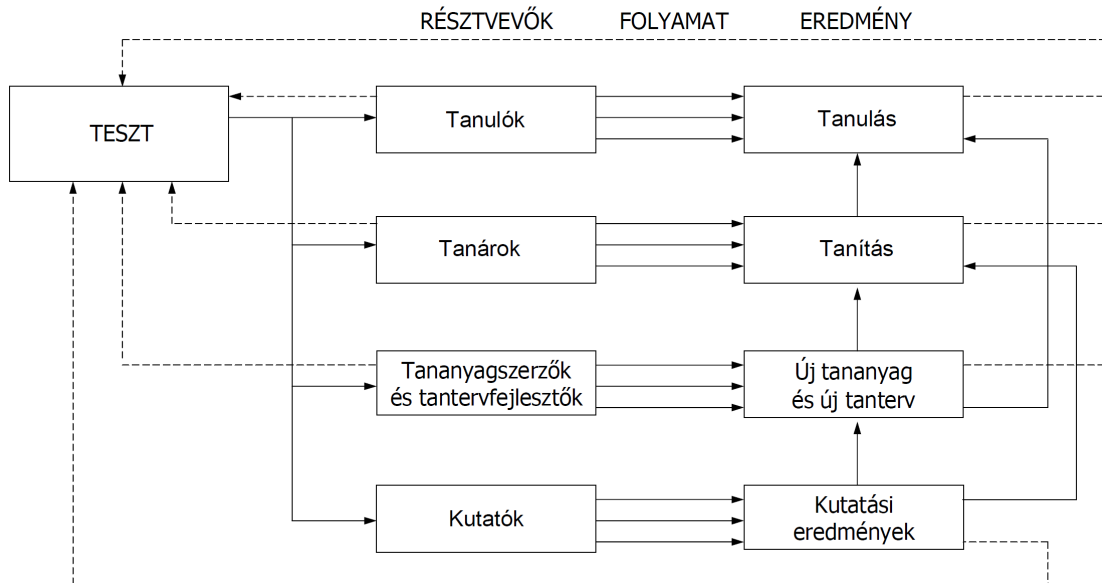
Alderson (1991) a kilencvenes évek elején az idegen nyelvi tesztelés területén végzett kutatások eredményeinek és hiányosságainak összegzése során megállapította, hogy a teszthatás vizsgálata meghatározóvá válik. Ennek egyrészt az az oka, hogy a visszahatás folyamatát, jellemzőit kevésbé kutatták, másrészt az addigi kutatások inkább leíró jellegűek és nem rendszerezett empirikus vizsgálatok voltak.

Az 1990-es évektől elinduló nyelvpedagógiai vizsgálatokban meghatározóak voltak Alderson és Wall (1993) hipotézisei (*washback hypothesis*), amelyek igazolják vagy elvetik a visszahatás létezését, a teszt hatása és validitása közötti kapcsolatot, és az empirikus kutatások alapjait képzik. A szerzőpáros a *washback* fogalmát csak a mikrokontextusra, a tanítási-tanulási folyamatra vonatkoztatva határozta meg. Feltételezésük szerint egy teszt befolyásolhatja a tanítást és a tanulást, valamint meghatározhatja, amit és ahogyan a tanárok tanítanak, ezáltal azt is, amit és ahogyan a tanulók tanulnak. Egy teszt befolyásolhatja a tanítás és a tanulás sebességét és sorrendjét, valamint mértékét és minőségét, hatást gyakorolhat továbbá a tanítás és a tanulás tartalmára, módszereire. A fontos következménnyel rendelkező teszteknek lehet, az alacsony presztízsű teszteknek nem lehet hatásuk.

Az 1990-es évek második felében több kutató modellezte a teszthatás folyamatát. A továbbiakban az elméleti modellek fejlődését, változásait, illetve az egyes modellek közötti összefüggéseket mutatjuk be.

Hughes (1993, idézi *Bailey*, 1996. 262. o.) trichotóm modellje a teszthatás alapmechanizmusát ábrázolja. A modell szerint a teszt először a folyamatban résztvevők, vagyis a tanulók, a tanárok, az adminisztrátorok, a tananyagfejlesztők nézeteire és attitűdjeire van hatással, amelyek befolyásolják és módosítják a tanítási-tanulási folyamatot, ezáltal meghatározzák a folyamat eredményeit: a nyelvtanulás tartalmát, minőségét, eredményességét és a folyamatban elsajátított készségek fejlettségét.

Bailey (1996) visszahatás alapmodellje (5. ábra) kombinálja *Alderson* és *Wall* (1993) hipotéziseit és *Hughes* (1993) trichotóm modelljét.



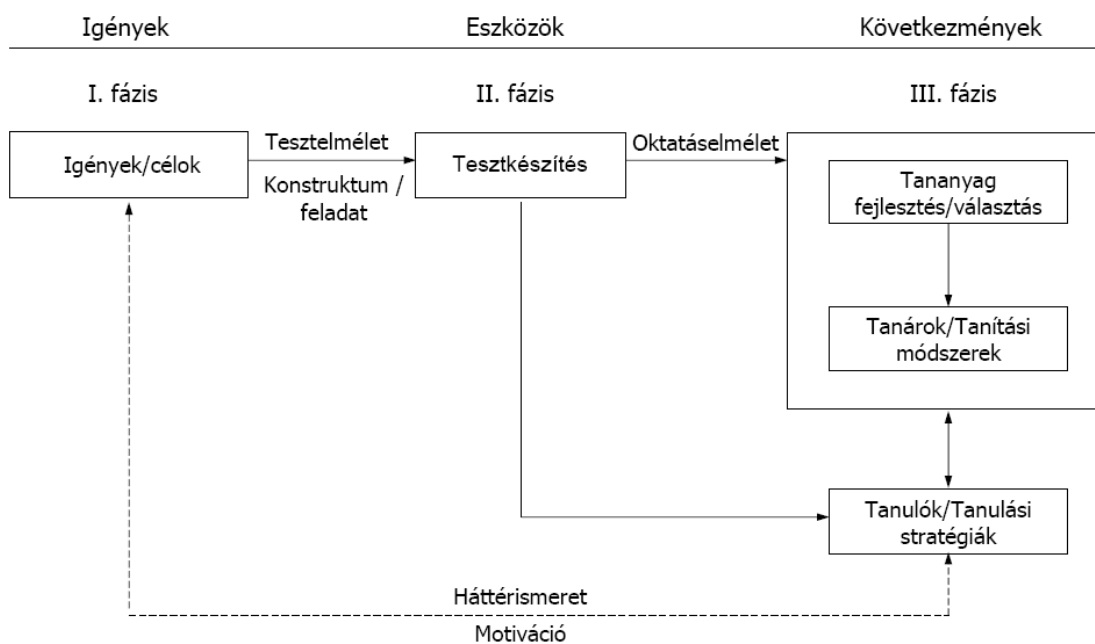
5. ábra
A visszahatás alapmodellje
(*Bailey*, 1996. 264. o.)

A modell – *Hughes* (1993) alapján – egyrészt három tényezőt nevez meg (részrtvevők, folyamat és eredmény), másrészt *Alderson* és *Wall* (1993) hipotéziseit két csoportba sorolja. Ezek alapján *Bailey* (1996) kétféle hatást különböztet meg. A tanulókra irányuló teszthatás (*learner washback*) a teszteredményekből szereshető információkra és a levonható következtetésekre utal. A tanítási-tanulási folyamatra vonatkozó teszthatás (*program washback*) az oktatási rendszerrel közvetlen kapcsolatban állókra irányul, a teszteredmény számukra szolgáltat információkat. A modellben a folytonos vonal a teszt közvetlen hatását jelöli a különböző folyamatokban résztvevőkre, a szaggatott vonal pedig magára a tesztre irányuló lehetséges hatásokat mutatja.

A modell alapvető problémája, hogy nem határozza meg pontosan a folyamat tényező jellemzőinek leírását (*Tsagari*, 2007), valamint annak az eredményekkel való kölcsönhatását (*Wall*, 2005), így *Hughes* (1993) trichotóm modellje dichotómmá válik. A leírásból nem derül ki, hogy a kutatási eredmények miért a tanításra és nem a tanárookra irányulnak. *Saif* (1999, 2006) szerint *Hughes* (1993) és *Bailey* (1996) figyelmen kívül hagyja az igények szerinti tesztkészítést, amely elősegítheti a teszthatás folyamatát. A modell jelentőségét a problémás pontok mellett mégis az adja, hogy a visszahatás folyamatát és működését ábrázolja, és hozzájárul a modellek fejlődéséhez.

Saif (1999) modelljében (6. ábra) kiemeli az igényeknek és a céloknak megfelelő tesztfejlesztést a visszahatás folyamatában. A modell a teszthatás két tulajdonságát mu-

tatja. Egyrészt azt ábrázolja, hogy figyelembe kell venni a társadalmi és az oktatással kapcsolatban álló kliensi igényeket, amelyek közvetlenül és közvetetten határozzák meg a teszt célját, tartalmát és típusát, valamint fejlesztését és kivitelezését. Másrészt pedig azt illusztrálja, hogy a teszt alkalmazása hogyan befolyásolja a tanítási-tanulási folyamatban résztvevők tevékenységét, azaz a tanárok tananyagválasztását, az oktatási módszerek alkalmazását és a tanulók tanulási stratégiáit. A modell ezen kívül még arra is felhívja a figyelmet, hogy a tanulók háttérismereteit és motivációját is figyelembe kell venni a teszt fejlesztésekor.



6. ábra
A visszahatás konceptuális modellje
(Saif, 1999. 69. o.)

Saif (1999) modellje alapján pozitív teszthatás abban az esetben történik, ha a vizsgára való felkészülés megegyezik azzal a valós nyelvi szituációval, amelyben a vizsgázó a nyelvet használni fogja. Ennek biztosításához szükséges a vizsgázók nyelvi igényének (*Bedarfsdiagnose*) felmérése (Tschirner, 2001). Azonban abban az esetben, ha a teszt mérési célja, tartalma és konstruktuma nem egyezik meg a tanítási-tanulási folyamat céljaival, a tanárok csak a vizsgán előforduló feladatokat, feladattípusokat és feladatmegoldó stratégiákat tanítják (Prodromou, 1995; Krumm, 2006; Perlmann-Balme, 2006). Ennek eredményeként a tanulók elsősorban olyan képességeket sajátítanak el, amelyekkel a tesztet meg tudják oldani, de valós, életszerű nyelvi szituációban nem tudják alkalmazni a nyelvtudásukat (Krekeler, 2002, 2005). Saif (1999) modelljének elemzése felveti azt a kérdést, hogy vajon azoknak a teszteknek pozitív-e a hatásuk a tanítási-tanulási folyamatra, amelyeket a valódi igényeknek megfelelően és a kommunikatív tesztelés alapján készítenek, vagy a visszahatás értéke a teszt minőségétől független és más tényezők határozzák meg.

E kérdések megválaszolásához Burrows (2004) jelentősen hozzájárult azzal, hogy a visszahatás folyamatát három paradigma mentén jellemezte, amely alapján értelmezte a teszt és annak hatása közötti összefüggésekről alkotott elképzelések változásait. A tradicionális visszahatás-elmélet (7. a ábra) alapján egy új teszt bevezetése szükségszerűen vezet visszahatáshoz, a tanítási-tanulási folyamat irányítóiból, a tanárokból egyszerű és egyforma reakciót vált ki. A teszthatás lehet pozitív vagy negatív irányú, és ez csak a

teszt minőségétől, nem a tanárok nézeteitől, attitűdjeitől függ. Ennek a tradicionális fel fogásnak megfelel *Morrow* (1986) visszahatás-validitás, valamint *Hughes* (1989) visszahatás-elérés (*working for washback*) koncepciója. *Hughes* (1989. 44–47. o.) – többek között – az alábbi tevékenységet javasolja a pozitív hatás elérése érdekében: (1) a tanítási-tanulási folyamatban fejlesztett készségek mérése; (2) a mérendő készségek széles körű kiválasztása; (3) közvetlen és (4) kritériumorientált tesztelés alkalmazása; (5) a tesztelés alapelveinek megismertetése a tanárokkal.

A tradicionális visszahatás-elméletet *Alderson* és *Wall* (1993), valamint *Messick* (1996) is kritikával illette. Feltételezésük szerint a teszt minőségétől független, hogy a tesztnek pozitív vagy negatív-e a hatása, mivel irányát a tanárok tesztel kapcsolatos nézetei, feltételezései és ismeretei határozzák meg. A tanárok egy új teszt bevezetésére egyéni tanítási reakciókat adnak. *Alderson* és *Wall* (1993) feltételezését kutatási eredmények is alátámasztották. *Wall* és *Alderson* (1993) például megállapították, hogy a tanárok azonos típusú tesztre különböző, míg különböző alapelveket képviselő tesztre nagyon hasonló módon készítették fel diákjaikat. Az egyéni, sokszor eltérő tanítási reakciókat feltehetően az magyarázza, hogy a tesztelés alapelve, filozófiája többnyire rejtve marad, ezért szükséges a tanárok bevonása a tesztfejlesztési és a lebonyolítási folyamatba (*Alderson*, 1999). A tesztfejlesztők által szándékolt hatást befolyásolják a tanároknak a teszt bevezetésére adott egyéni reakciói, amely például függ a tanárok vizsgával kapcsolatos attitűdjeitől (*Turner*, 2001). Ebből következik, hogy a tanár egy olyan médiumként fogható fel, aki közvetít a teszt és a tanulók által elért eredmények között (*Burrows*, 2004). A tesztek hatása így komplex folyamatként értelmezhető (7. b ábra). Ugyanakkor a tanárok egyéni válaszreakcióinak vizsgálata megnehezíti a visszahatás kutatását.

Burrows (2001, 2004) empirikus kutatása alapján csoportosította a tanárok tesztel kapcsolatos reakcióit. Vizsgálatában három csoportot különböztetett meg: az elsőbe tartoztak a tesztet elutasító tanárok (*resisters*), akik a saját tanítási gyakorlatuk és a bevezetett teszt között jelentős és áthidalhatatlan szakadékot érzékeltek. Jellemző volt rájuk, hogy a tanítási képességeiket és a tanítás során szerzett tapasztalataikat jelentősen alábecsülték. A második csoportot az elfogadó tanárok (*adopters*) alkották, akiknek nézetei teljes egészében vagy nagymértékben összhangban voltak a bevezetett teszt alapelveivel. Ők gyakran vettek részt a tesztfejlesztési folyamatban, vállaltak vezető szerepet a teszt lebonyolításában. E két csoport között helyezkedtek el a tesztel alkalmazkodó tanárok (*adaptors*), akik a tesztet ambivalens módon ítélték meg, mivel az oktatási nézeteik, feltételezéseik és ismereteik nem voltak teljes mértékben összhangban a bevezetett teszt mérési-értékelési eljárásaival. *Burrows* (2001, 2004) ez utóbbi csoportot három alcsoportra osztotta: az elsőbe sorolta a tesztel korán alkalmazkodó tanárokat, akik a bevezetett teszt alapján számos ismeretet szereztek például a mérési-értékelés területén vagy a teszt bevezetése megnövekvő irányító szerepet jelentett számukra a tanítási-tanulási folyamatban. A második alcsoportba a tesztel részben alkalmazkodó tanárokat sorolta, akik a teszt módosítását váratlanul tartották és nem profitáltak személyesen a teszt mérési-értékelési eljárásaiból, de érzékelték, hogy más tanár tanításában alkalmazkodott a tesztel. *Burrows* (2001, 2004) a harmadik alcsoportot nevezi az alkalmazkodóknak. Ide azokat a tanárokat sorolta, akiknek a tesztel kapcsolatos attitűdjei a teszt működése során egyre pozitívabbá vált. A második és harmadik alcsoport között jelentős átfedéseket mutatott ki.

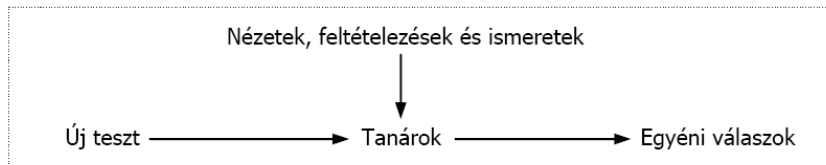
Burrows (2001, 2004) vizsgálatának eredményeként megállapítható, hogy egy új teszt vagy egy új mérési rendszer bevezetésére adott tanári reakciók kvalitatív elemzésével viselkedésmintázatok tárhatók fel. Ezáltal lehetőség nyílik a visszahatás jellemzőinek, az oktatási folyamat változásainak meghatározására és ebből következően a tanítá-

si-tanulási folyamatot leíró innováció modellezésére. Ezek alapján az innovációt meghatározzák a résztvevők nézetei, feltételezései, ismeretei és attitűdjei. A modell figyelembe veszi a tanárok nézetrendszerét és következetes válaszát, amelyek a változásokat idézik elő (7. c ábra). A tanárok viselkedésmintázatai közé tartoznak a teszttel és a változtatáshoz kapcsolatos attitűdök, a teszt minőségével kapcsolatos nézetek, a teszt és a tanítási gyakorlat közötti összefüggések.



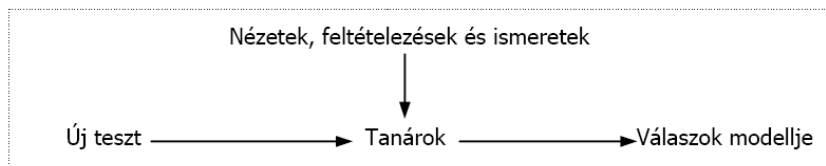
7. a ábra

A tradicionális visszahatás-elmélet: az inger válasz modell



7. b ábra

Az 1990-es évek visszahatás felfogása: a „fekete doboz” modell



7. c ábra

A visszahatás Burrows (2004. 126. o.) által javasolt értelmezése: az innovációs modell

A pedagógus-kutatások alapján megállapítható, hogy a tanárok gondolkodását és gyakorlatát a korábbi személyes és tanítási tapasztalatokból származó nézetek (*beliefs*), feltételezések/feltevések (*assumptions*) és ismeretek (*knowledge*) határozzák meg. Burrows (2001, 2004) e fogalmak értelmezésében Woods (1996) definícióját alkalmazza, mely szerint ezek elsősorban a tanárok gondolkodását meghatározó kognitív dimenzió elemei és jellemzőjük, hogy nem különíthetők el egymástól pontosan. A szakirodalomban egyetértés van arról, hogy a kutatók a nézeteket a tapasztalatokból származó pszichikus konstrukcióknak tekintik. Richardson (1996. 103. o.) szerint a nézetek „a mentális, személyes konstrukciók egy csoportját alkotják, olyan feltételezések, feltevések, propozíciók a világról, amelyeket igaznak vélünk, s amelyek befolyásolják ítéleteinket, mások megítélését, s amelyeket felhasználunk döntéseink során”. Bartels (2006) szakirodalmi áttekintésében a nézetek és ismeretek közötti különbségek és hasonlóságok elemzésében megállapította, hogy ezek egymással szorosan összekapcsolódnak, de a nézetek szűrőfunkcióval is rendelkeznek, melyen keresztül érzékelik a tanárok az új jelenségeket, például egy módosított vagy bevezetett új teszt mérési-értékelési eljárásait. Bár Burrows (2001, 2004) a tanárok attitűdjeit is kiemeli a tesztelés folyamatának jellemzésében, a megalkotott modelljében (7. c ábra) mégsem szán ennek központi szerepet. A nézetek és az attitűdök hasonlóságait és különbségeit Richardson (1996) vizsgálta. Eredményei szerint ezek nehezen különíthetők el egymástól, szorosan összefonódnak. Véleménye szerint a nézetek inkább kognitív, az attitűdök inkább affektív eredetűek és mindkettő befolyásolja a pedagógusok gondolkodását és tanítási gyakorlatát.

Burrows (2004) innovációs modellje Markee (1993) megállapításaira épít, amely azt írja le, hogyan és milyen szempontok figyelembe vételével lehet változásokat elindítani a közoktatási rendszerben és így a tanítási-tanulási folyamatban. A modell az alábbi tényezőket nevezi meg: ki (a résztvevők) adoptál (a folyamat), mit (az innovációt), hol

(a kontextus), mikor (az idő), miért (az indok) és hogyan (az innováció irányításának eltérő módjai). Az innováció nem csak a tanítási-tanulási folyamatra vonatkozik, valamint minőségét jelentősen meghatározza az oktatáspolitikai (Markee, 1993). Andrews (2004) az alábbi javaslatokat teszi a döntéshozóknak a vizsga helyes megváltoztatásának és a pozitív hatás elérésének érdekében:

- kerüljük a kényszerítő körülményeket,
- vegyük figyelembe az innováció komplexitását és előre nem megjósolható hatását, és ne várjunk el előre megjósolt hatásokat,
- segítsük elő a teszt és a tanterv közötti harmóniát,
- segítsük elő a további kutatásokat a jelenség minél jobb megismeréséhez.

A bemutatott és jellemzett modellek leírják és értelmezik a visszahatás folyamatának működését, valamint azt is, hogy hogyan érhető el pozitív irányú teszthatás. Az elméleti megközelítések mindegyikében közös elem, hogy a hatást mint a tanítási-tanulási folyamatban és az oktatási rendszerben bekövetkező változásokat (Cheng, 1998, 2005) ábrázolják. Wall (1996, 2005) ugyanakkor rámutat arra, hogy ezek a modellek nem tartalmazzák azokat a dimenziókat, faktorokat és feltételeket, amelyek elengedhetetlenek a visszahatás elindulásához és minőségét alapvetően befolyásolják.

Watanabe (2004a) a visszahatás folyamatában az alábbi dimenziókat nevezi meg:

- sajátosság: általános (olyan teszthatás, amelyet minden teszt eredményez: például a tanulókat arra motiválja, hogy jobban és többet tanuljanak), specifikus (csak arra az adott tesztre vonatkozó hatások);
- intenzitás: erős (egyedül a teszt határozza meg, hogy mi történik az osztályteremben, a tanárok azonos módon tanítanak), gyenge (a teszt az osztálytermi események csak egy bizonyos részét befolyásolja, a tanárok különbözőképpen tanítanak azonos teszt esetén);
- hosszúság: rövid (a teszthatás csak addig figyelhető meg, amíg a tanulók felkészülnek a tesztre), hosszú (a hatás a teszt megoldása után is tapasztalható);
- szándék: a tesztfejlesztők által szándékos és nem szándékos hatás (ezt a dimenziót a tesztfejlesztés időszakában pontosan meg kell jelölni, hogy legyen hatás);
- érték: pozitív vagy negatív.

A szakirodalomban a visszahatás folyamatának minősége mellett fontos szerepet kapott az intenzitás (*washback intensity*) meghatározása. Gates (1995) szerint az alábbi tényezők határozzák meg a vizsgák hatásának erősségét:

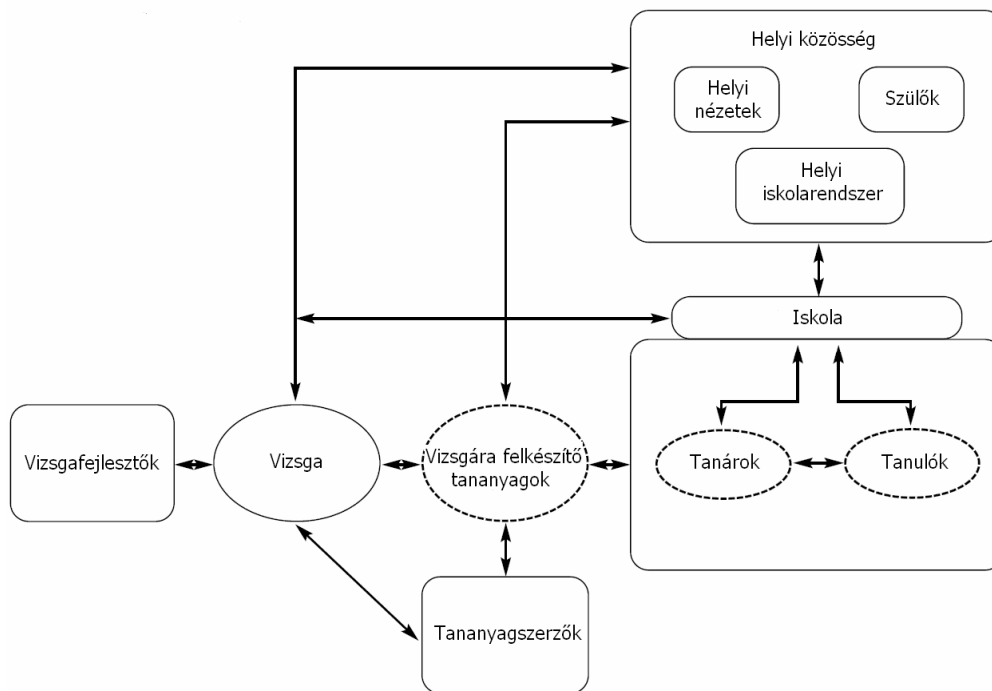
- átláthatóság: minél közelebb áll a tesztkonstrukció a valós igényekhez, annál erősebb a hatása [lásd Saif (1999) modelljét, 6. ábra];
- hasznosság: minél több lehetőséget biztosít a letett vizsga, annál erősebb a hatása;
- monopólium: minél kevesebb versenytársa van az adott vizsgának, annál erősebb a hatása;
- fontosság: minél fontosabb és jelentősebb a tanuló számára, hogy letegye a vizsgát, annál erősebb a hatása.

Watanabe (2004b) szerint az alábbi faktorok befolyásolják a visszahatást:

- tesztfaktorok: a teszt céljai és tartalma, a mért készségek, a tesztelési módszerek, a teszteredmények alapján hozott döntések;
- presztízfaktorok: a teszt fontossága, státusza a közoktatási rendszeren belül;
- személyes faktorok: a tanárok képzettsége, nézeteik az oktatási folyamatról;
- mikrokörnyezeti faktorok: az iskola felszereltsége, amelyben az oktatás, felkészítés zajlik;
- makrokörnyezeti faktorok: az a társadalom, amelyben a tesztet alkalmazzák.

A modellek változásai alapján a teszthatás olyan komplex jelenség, amely közvetlen és közvetett hatást fejt ki szűkebb értelemben véve a tanítási-tanulási folyamatra, a folyamatban résztvevőkre, valamint a tanulásra; tágabb értelemben véve az innovációs folyamatok segítségével az oktatási rendszerre és a társadalomra. A hatás intenzitása, mértéke és értéke előre meghatározhatatlan, mivel számos tényezőtől függ, amelyek közül meghatározó jelentősége van a tanárok tesztel kapcsolatos nézeteinek, feltételezéseinek, ismereteinek és attitűdjeinek. A teszthatás iránya pozitív vagy negatív lehet, amelynek értéke önmagában nem befolyásolja a teszt validitását.

A teszthatás komplexitása miatt nagy az igény egy olyan szintetizáló modellre, amely jelöli az egyes komponensek szerepét és bemutatja az ezek közötti kölcsönhatást (Wall, 2005). Erre a modellre Tsagari (2009) tett javaslatot, melyet empirikus vizsgálata alapján is validált (8. ábra). A kutató a Cambridge Nyelvvizsgarendszer B2 szintű FCE (*First Certificate in English*) angol nyelvű vizsga hatását vizsgálta három tényező alapján: interjúelemzésekkel a tanárok vizsgával kapcsolatos nézeteit és attitűdjeit, tankönyvelemzéssel a vizsgára felkészítő tananyagokat, és naplóelemzéssel a vizsgára felkészülő diákok nézeteit és felkészülési gyakorlatát. A visszahatást komplex, az egyes tényezőkkel folyamatos interakcióban lévő mechanizmusként ábrázolta és számos résztvevőt is bevont ebbe a folyamatba. A 8. ábrán a szaggatott ellipszisek a vizsgált tényezőket, a nyilak nem folyamatot, hanem a tényezők közötti kölcsönhatást jelölik.



8. ábra

A teszthatást szintetizáló koncepcionális modell
(Tsagari, 2009. 8. o.)

A nagy presztízsű és jelentős jogosultságokat biztosító nyelvtudást mérő vizsga elsősorban a tananyagszerzőkre van hatással, akik a vizsgára felkészítő tananyagok elkészítésével a diákok és tanárok igényeinek tesznek eleget. A vizsgára felkészítő tananyagok közvetítenek a tesztelők szándéka és a tanárok nyelvtanítási folyamata között. A tanárok is lényeges elemei a modellnek, hiszen a tananyagok és a diákok között közvetítő szerepet töltenek be. Ugyanakkor Tsagari (2009) nem hangsúlyozza kellőképpen a tanárok vizsgával és a felkészítő tananyagokkal kapcsolatos nézeteinek és attitűdjeinek meghatározó funkcióját, hanem a teszthatást lineáris folyamatként ábrázolja, mert hi-

ányzik a modelltől, hogy a tanárok és a diákok válaszmintázatai hogyan hatnak a vizsgafejlesztési folyamatra. *Tsagari* (2009) lényegesnek tartja modelljébe bevonni az iskola és a helyi közösség hatását. A vizsgálata alapján a teszthatást befolyásolja az iskola légköre, az értékelési kényszer megjelenése és ez hat az osztálytermi felkészítő folyamatokra, de a modelltől és leírásából sem derül ki egyértelműen, hogy ez milyen viszonyban van a vizsgával, és hogyan hat a vizsgafejlesztési folyamatra. A helyi közösség szereplői közül a szülők hatását emeli ki a kutató. A tanárok és diákok vizsgával kapcsolatos nézeteinek vizsgálata alapján a szülői hatás elsősorban a diákok nyelvi motivációjában, a vizsga és az idegen nyelv kiválasztásában jelent meg. Ugyanakkor a modell és leírása sem tartalmazza, hogy a helyi környezet milyen kapcsolatban áll a vizsgára felkészítő tananyagokkal és a vizsgával, továbbá ezek miért nem a tesztfejlesztési folyamatra adnak visszacsatolást.

Tsagari (2009) szintetizáló modelljében a teszthatás komplexitását és az egyes elemek közötti interakciót hangsúlyozza és a korábbi modellekhez képest két tényező: a vizsgára felkészítő tananyagok és a mikrokörnyezet hatását emeli ki. Ugyanakkor ezeket a tényezőket a modell három elemével vizsgálja közvetlenül, de nem hangsúlyozza a tanárok központi szerepét, és nem igazolja kellőképpen a közvetett hatásmechanizmusokat (a környezet és az iskolai légkör visszacsatolását) és ezek általánosíthatóságát. Ez a vizsgálat viszont felhívja a figyelmet arra, hogy a teszthatás komplexitása miatt szükséges olyan empirikus kutatások elvégzése, amelyek segítségével adott környezetben megismerhetjük a visszahatás természetét és mechanizmusait.

1.3.3. A teszthatás folyamatát leíró empirikus kutatások

Alderson és *Wall* (1993) visszahatás-hipotézisei a teszthatás modellezése mellett az empirikus kutatások elindításában is irányadóak voltak. A legelső nyelvpedagógiai hatásvizsgálatban egy Sri Lankán módosított alapszintű angolnyelv-tudást mérő vizsga hatását elemezték (*Wall* és *Alderson*, 1993). A korábbi vizsgálatokkal szemben, amelyek elsősorban önbevalláson alapult adatokat tartalmazó kérdőívekből és interjúkból álltak, osztálytermi megfigyeléseket tartalmazott. A megfigyelésekhez ellenőrző listákat dolgoztak ki, amelyek a tanári tevékenységekre, a tanár-tanuló vagy tanuló-tanuló interakcióra, valamint a felhasznált input szövegtípusokra vonatkoztak. A kétéves, összesen hat megfigyelési szakaszból álló vizsgálat eredményei alapján a vizsga hatást gyakorolt a tanárok által használt tananyagokra, a tanítás tartalmára, a tesztben mért készségek fejlesztésére, és befolyásolta a tanárok mérési és értékelési gyakorlatát. A vizsgának azonban nem, vagy csak nagyon kevés hatása volt a tanárok tanítási módszereire: *Wall* és *Alderson* (1993) nem talált kapcsolatot a módosított vizsga bevezetése előtt és után alkalmazott oktatási módszerekben.

Wall és *Alderson* (1993) az idegen nyelvi mérés és értékelés területén végzett kutatása a visszahatás folyamatáról és működéséről számos további vizsgálatot indított el a világ számos országában (például Amerikai Egyesült Államok, Anglia, Új-Zéland, Ausztrália, Japán, Hong Kong, Kína, Izrael). A vizsgálatok az idegennyelv-oktatás különböző kontextusaira terjedtek ki: nyelvvizsgarendszerek (*Alderson* és *Hamp-Lyons*, 1996; *Hamp-Lyons*, 1998; *Hayes* és *Read*, 2004; *Saville* és *Hawkey*, 2004), középiskolai érettségi (*Qi*, 2004, 2005, 2007), középiskolai vizsgák (*Cheng*, 1997, 1998, 1999, 2004, 2005; *Shohamy*, *Donitsa-Schmit* és *Ferman*, 1996; *Shohamy*, 2001; *Stecher*, *Chun* és *Barron*, 2004; *Andrews*, *Fullilove* és *Wong*, 2002; *Wall* és *Alderson*, 1993), egyetemi felvételik (*Watanabe*, 1996a, 1996b), valamint bevándorlók számára indított nyelvi kurzusok eredményét igazoló vizsgák (*Burrows*, 2001, 2004) hatásvizsgálatára.

A kutatások a vizsgáknak csak mikroszintű (azaz a tanítási-tanulási folyamatra irányuló) hatását tanulmányozták, nem ismert olyan vizsgálat, amely az egyes vizsgák makroszintű hatásait elemezte volna. Ennek lehetséges magyarázata, hogy a hatásmechanizmusok komplex jellegűek, a makroszintű hatások vizsgálata pedig nehezen kivitelezhető. A vizsgálatok ötvözik a kvalitatív és kvantitatív kutatási módszereket, hiszen a kérdőívektől, az interjúkon, a fókuszcsoportos interjúkon, valamint az osztálytermi megfigyelésen keresztül különböző adatgyűjtő eszközökig terjednek. Ilyen például a *Saville és Hawkey (2004)* kutatásában kifejlesztett tananyagokat vizsgáló eszköz, az *Instrument for the Analysis of Textbook*.

Az empirikus kutatások a vizsgarendszerek hatásának vizsgálatában (1) a tantervi változások és az oktatás tartalma, (2) az alkalmazott tananyagok, (3) a tanítás módszertana, (4) az osztálytermi értékelés, (5) a vizsgával kapcsolatos attitűdök és (6) a tanulás területén hoztak eredményeket (*Chen, 2002; Spratt, 2005; Wall, 2005*).

A tantervekre és az oktatás tartalmára gyakorolt hatás. A legtöbb kutatás eredménye (például *Wall és Alderson, 1993; Alderson és Hamp-Lyons, 1996; Cheng, 1998*) szerint a vizsgák jelentősen befolyásolták azt, hogy mely nyelvi készségek fejlesztésére fektettek nagyobb hangsúlyt a tanárok. *Shohamy, Donitsa-Schmit és Ferman (1996)* tanulmánya ettől eltérő képet mutat, mivel az eredményekből kitűnik, hogy az alacsony presztízsű vizsgának a tanítás tartalmára kevés hatása volt. *Watanabe (1996b)* vizsgálata szerint a tanárok nem feltétlenül tanították nagyobb hangsúllyal a vizsgán mért készségeket (például hallás és olvasás utáni szövegértést). *Nikolov (1999)* osztálytermi megfigyelésre és kérdőíves felmérésre épülő hazai vizsgálatában 55 középiskola 118 csoportjának 107 angol szakos nyelvtanárát figyelték meg. A kutatás eredményei szerint a leggyakrabban alkalmazott feladattípusok között „kérdés-felelet, fordítás, hangos olvasás és behelyettesítő nyelvtani gyakorlatok szerepeltek.” (*Nikolov, 1999. 28. o.*) A fordítás és a kitöltéses feladatok tipikus hagyományos szemléletű vizsgafeladatok, amelyek gyakori alkalmazása mutatta a vizsga hatását ezen a területen.

A tananyagokra gyakorolt hatás. A vizsgáknak jelentős hatása volt arra, hogy a tanárok milyen tananyagot használnak mind a tanítás, mind a vizsgára való felkészítés során. A tananyagok gyakran változtak meg az új vizsga bevezetésének hatására (*Cheng, 1998; Andrews, Fullilove és Wong, 2002; Hayes és Read, 2004*), valamint az is előfordult, hogy a tanárok több, saját maguk által a vizsga szemléletének megfelelő tananyagot készítettek (*Watanabe, 1996a, 1996b*). Azonban – mint ahogyan *Wall és Alderson (1993)* és *Nikolov (1999)* rámutatott – a tanárok gyakran vizsgafeladatokat is alkalmaztak a tanulási folyamatban gyakorlófeladatként.

A tanítási módszerekre gyakorolt hatás. Az empirikus kutatások eredményei azt mutatták, hogy a vizsgáknak vagy nem volt hatásuk arra, ahogyan a tanárok tanítottak (*Wall és Alderson, 1993; Cheng, 1998, 2003; Shohamy, Donitsa-Schmit és Ferman, 1996*), vagy a tanárok csak a vizsga szemléletének megfelelő módszerek segítségével készítették fel a tanulókat (*Watanabe, 1996b; Read és Hymes, 2004*). *Ferman (2004)* és *Qi (2004, 2007)* kutatása szerint a tesztkészítők által szándékolt hatás és a tanítási módszerekben megnyilvánuló hatás között nem volt kapcsolat. *Alderson és Hamp-Lyons (1996)* szerint ennek az a magyarázata, hogy tanárfüggő annak a módja és minősége, ahogy a tesztek hatása megjelenik a tanítási módszerekben. Összességében megállapítható, hogy a kutatók egyik vizsgálatban sem tudták egyértelműen kimutatni a vizsgák hatását a tanítási módszerek alkalmazásában. Ennek feltehetően az az oka, hogy a kutatások ezen a területen csak rövid távon vizsgálták a vizsgák hatásait. *Cheng (2003)* arra a következtetésre jutott, hogy a tanítási módszerekre gyakorolt hatás nagy valószínűséggel csak hosszú távon vizsgálható.

Az osztálytermi mérésekre gyakorolt hatás. A teszthatásnak ezt a területét kevés kutatásban vizsgálták (Spratt, 2005). Wall és Alderson (1993) szerint a vizsgának az osztálytermi mérésre és értékelésre gyakorolt hatásának mértéke csekély volt, mivel a tanárok nem ismerték pontosan a vizsgán alkalmazott értékelési eljárásokat.

A tanárok és a tanulók érzelmeire, attitűdjeire gyakorolt hatás. A vizsgák elsősorban a tanulók attitűdjeire, belső motivációjára, érdeklődésére, a vizsgával kapcsolatos érzelmekre lehetnek hatással. A nagy tétellel rendelkező vizsgák elszámoltathatóságuk következtében fejtenek ki hatást a tanárok attitűdjeire. Shohamy, Donitsa-Schmit és Ferman (1996) kutatása szerint a tanárok úgy érezték, hogy a tanulók sikere vagy kudarca a tanításuk minőségét tükrözi. A tanároknak sokszor voltak negatív attitűdjeik az új, vagy módosított, nagy presztízsű vizsgával kapcsolatban (például Alderson és Hamp-Lyons, 1996; Cheng, 1998). Más kutatások azonban pozitív attitűdről számoltak be (például Watanabe, 1996b; Hayes és Read, 2004). Spratt (2005) szakirodalmi tanulmányában rámutat arra, hogy a kutatási eredményekből nem derül ki, milyen hatása volt a tanítási-tanulási folyamatra és eredményére a vizsga által okozott félelem, negatív attitűd.

A tanulásra gyakorolt hatás. Talán ez a legfontosabb, a tanárokat is leginkább érintő kérdés, mégis ez a terület a legkevésbé tanulmányozott. Wall (2000) szerint „olyan teszteredményeket vizsgáló kutatások hiányoznak, amelyek bemutatják, hogy a tanulók többet vagy jobban tanultak-e egy bizonyos tesztre” (502. o.). A kutatási eredmények alapján a tanulásra gyakorolt hatás egyénfüggő, valamint befolyásolja a tanulók motivációja és tanulási stratégiái (Cheng, 1998). Ferman (2004) kutatási eredménye szerint a vizsga elsősorban az alacsonyabb nyelvtudású tanulókra volt nagyobb hatással, a többség esetében csekély hatás tapasztalható.

Az áttekintett kutatások eredményeinek értékelése során fel kell tenni a kérdést, hogy az eredmények mennyire általánosíthatók, illetve milyen következtetéseket lehet levonni. Ezek megválaszolása azért indokolt, mert a kutatások gyakran eltérő, néhány esetben pedig egymásnak ellentmondó eredményeket hoztak.

Az empirikus kutatások eredményei alapján megállapíthatjuk, hogy a vizsgáknak lehet hatása a tantervekre, a tananyagokra, a tanítási módszerekre, a vizsgával kapcsolatos érzelmekre és attitűdökre, valamint a tanulási eredményekre. Ez a hatás azonban nincs mindig jelen, valamint igen különböző formát és intenzitást ölthet (Spratt, 2005). A visszahatás tehát nemcsak egy komplex jelenség, hanem attól is függ, milyen oktatási kontextusban vizsgálják. Alderson (2004) Cheng, Watanabe és Curtis (2004) könyvéhez írt előszavában a kutatások eredményeit, Alderson és Wall (1993) visszahatáshipotéziseire reflektálva az alábbi módon összegzi:

Ma már tudjuk [...], hogy a teszteknek nagyobb hatása van a tanítás tartalmára és az alkalmazott tananyagokra, mint a tanárok módszereire. Tudjuk, hogy a különböző tanárok egy bizonyos tesztre nagyon különböző módon készítik fel [valamint], hogy néhány tanár nagyon különböző tesztre nagyon hasonló módon tanít. Ezen kívül pedig tudjuk, hogy a nagy presztízsű, az egyén számára jelentős következményekkel járó teszteknek nagyobb a hatásuk, mint az alacsonyabb presztízsű teszteknek (Alderson, 2004. ix-x. o.).

A vizsgálatok számos további szempontot is megneveztek. A 9. táblázatban közölt, Spratt (2005) által összeállított tényezők – Burrows (2004) modelljéhez hasonlóan – felhívják a figyelmet a tanárok vizsgával kapcsolatos (1) nézeteinek, (2) attitűdjeinek, valamint (3) képzettségének és (4) személyiségének fontosságára a visszahatás folyamatában.

9. táblázat. A teszthatás mértékét és típusát befolyásoló tényezők (Spratt, 2005 alapján)

<i>Tanárokra vonatkozó faktorok</i>	<i>Források, az iskola, a vizsga</i>
<p><u>A tanárok nézetei</u></p> <ul style="list-style-type: none"> – a vizsga megbízhatóságáról és tökéletességéről – a vizsga presztízséről és hasznosságáról – az eredményes tanítási módszerekről – a vizsga és az aktuális tanítási gyakorlatuk közti ellentmondásról – a vizsga és a tankönyv közti viszonyról – a vizsga és az oktatás tartalma közti viszonyról – a vizsga és az osztálytermi mérés és értékelés közti viszonyról – a tanulók nyelvtanuláshoz való hozzáállásáról <p><u>A tanárok attitűdjei</u></p> <ul style="list-style-type: none"> – a vizsgával kapcsolatban – a vizsgára felkészítő tananyagokkal kapcsolatban – a tanórára való felkészüléssel kapcsolatban <p><u>A tanárok végzettsége és képzettsége</u></p> <ul style="list-style-type: none"> – a tanárok végzettsége és oktatási tapasztalatai – az általános módszertani képzés mennyisége – a vizsgáztató képzésen való részvétel – a vizsga ismerete – a vizsga filozófiájának megértése <p><u>A tanárok egyéb jellemzői</u></p> <ul style="list-style-type: none"> – személyiség – egyéni nyelvtanulási tapasztalataik 	<p><u>Források</u></p> <ul style="list-style-type: none"> – a tananyagok, vizsgaanyagok elérhetősége – a tankönyvek elérhetősége <p><u>Az iskola</u></p> <ul style="list-style-type: none"> – légkör – értékelési kényszer – felkészítési idő és csoportlétszám – kulturális faktorok (például tanulási tradíciók) <p><u>A vizsga</u></p> <ul style="list-style-type: none"> – közelsége – tétje – célja – alkalmazott tesztformátuma – bevezetésének időpontja – elismertsége

Az empirikus kutatások eredményeinek rövid ismertetése, a vizsgált területek és a visszahatás folyamatát, működését befolyásoló tényezők bemutatása lehetővé teszi a további hatásvizsgálatok elindítását. Több publikáció (például *Chen, 2002; Cheng, 1998; Saville és Hawkey, 2004*) közli a kutatás eszközeit, valamint a kutatások pontos lebonyolítását (*Cheng, 1998; Wall és Alderson, 1993*), ezek új kutatási kontextusban alkalmazhatók, így az eredmények összehasonlíthatók. A visszahatás jelenségének komplexitása miatt kiemelten fontos a mérőeszközök validálása. A vizsgák tanítási-tanulási folyamatra gyakorolt hatásának vizsgálatához *Alderson és Banerjee (2001)* tanácsokat fogalmaz meg, ezáltal elősegítik, hogy a jelenség folyamatának, működésének és mechanizmusainak feltárása során megbízható adatokat kapjunk.

Az empirikus vizsgálatok alapján megállapítható, hogy a vizsgák visszahatását befolyásoló tényezők, valamint az ezek hatására bekövetkező tartalmi változások könnyebben azonosíthatók. Ugyanakkor a teszteknek a tanítás módszertani változásaira és a tanulásra kifejtett hatása sokkal összetettebb jelenség, így itt körültekintéssel, szisztematikusan és kritikusan (*Alderson, 2004*) kell végezni az elemzést, valamint a vizsgák hatásmechanizmusainak feltárása mindenképpen igényli a kvalitatív és kvantitatív kutatási módszerek együttes alkalmazását.

2. A NYELVI MÉRÉS ÉS ÉRTÉKELÉS HAZAI JELLEMZŐI

E fejezet célja a hazai nyelvi mérés és értékelés szakirodalmának szintetizáló áttekintése. Ehhez három szempontból mutatjuk be a nyelvtudás mérésére vonatkozó hazai eredményeket: először értelmezzük a nyelvi mérés és értékelés felértékelődésének okait, az ebben a folyamatban megjelenő változásokat. Ezt követően áttekintjük a közoktatási rendszerben történt nagymintás felmérések eredményeit, a tanulók nyelvtudás-szintjét, a nyelvtanulás hatékonyságát és eredményességét befolyásoló tényezőket. A fejezet harmadik részében az új kétszintű érettségi vizsga jellemzőit mutatjuk be az első fejezetben meghatározott elméleti keret alapján. E résszel célunk, hogy a vizsgát elhelyezzük a kommunikatív tesztelés rendszerében, melynek során négy szempontból elemezzük az érettségit. Megvizsgáljuk a vizsga illeszkedését a vizsgakövetelményekben meghatározott nyelvtudás-modellhez (1.1.2. fejezet), a kommunikatív tesztelés minőségi mutatóihoz (4. táblázat), a kommunikatív tesztek funkcióihoz és típusaihoz (1.1.4. fejezet). Ezen kívül pedig bemutatjuk az eddig megismert vizsgaeredményekből levonható következtetéseket.

Ebben a fejezetben a mérés és értékelés fogalmát az 1.1.1. fejezetben megadott definíció szerint értelmezzük, így a hazai felmérések eredményeinek áttekintésekor elősorban arra fókuszálunk, hogy a közoktatásban tanulók nyelvtudás-szintje és az abban megjelenő különbségek hogyan számszerűsíthetők. A nyelvi érettségi jellemzőit tárgyaló alfejezetben a vizsgakövetelmények alapján azt vizsgáljuk, hogy az érettségizők nyelvtudását és szintjeit milyen eljárásokkal mérik. Ebben a részben azt is bemutatjuk, hogy az egyes minőségi mutatók és tesztfunkciók mentén a vizsga számszerű eredményeiből az értékelés során milyen minőségű döntéseket lehet hozni a vizsgázók nyelvtudásáról és ezek az idő múlásával mennyiben tekinthetők állandónak.

2.1. A tudáskonceptió változása és tantervi megjelenése

Magyarországon a nyolcvanas években az idegen nyelvek tudása iránti igény egyre inkább előtérbe került, majd a kilencvenes évek elejétől megnőtt a használható nyelvtudás fontossága és presztízse (Imre, 2007). Ezek a folyamatok minőségi változásokat eredményeztek a közoktatásban megszerezhető nyelvtudással szemben támasztott elvárásokban: a diákok és szüleik az ingyenes, kötelező iskolától azonnal, közvetlenül hasznosítható és a munkaerőpiacon, valamint a felsőoktatásban helyzeti előnyt jelentő tudást várnak el. Az ezredfordulón a nyelvtudás fontossága az Európai Unióhoz való csatlakozás következtében tovább nőtt, melynek hatására az iskolai nyelvtanulás kiemelt jelentőségűvé vált. Hazánknak igazodnia kell az Európai Unió nyelvpolitikai célkitűzéseire és a nemzetközi munkaerő-piaci kihívásokhoz, mely szerint elvárás, hogy a közoktatásból kikerülők az anyanyelvükön kívül további két idegen nyelven tudjanak kommunikálni, képesek legyenek külföldön idegen nyelven tanulni, munkát vállalni (Nikolov, 2007). Az új hazai és nemzetközi elvárásoknak csak olyan személy tud megfelelni, aki értelmezni tudja az idegen nyelveken megszerezhető információt és képes a nyelvtudását tanulmányai során és a munkahelyén használni, továbbfejleszteni, céljainak megfelelő kommunikációt folytatni, azaz alkalmazható tudással rendelkezik. A nyelvtudás nem önmagáért való cél, hanem napjainkra az ismeretszerzés és a kommunikáció elengedhetetlen eszközévé vált.

A társadalmi és a kliensi igényeknek, elvárásoknak megfelelően a *Nemzeti alaptanterv* (1995), és a 2000-ben megjelent *Kerettantervek* deklarált célja „a mindennapi életben használható nyelvtudás” elsajátíttatása a közoktatásban. Ezekben a dokumentumok-

ban azonban még nem határozták meg egyértelműen, hogy mit jelent a használható nyelvtudás fogalma (Bukta és Nikolov, 2002; Petneki, 1998). Ezt a *Nemzeti alaptanterv* 2003-as, majd a 2007-es módosítása pontosan megfogalmazza: a használható nyelvtudás azonos a kommunikatív nyelvi kompetenciák fogalmával, amely „az adott szituációnak megfelelő nyelvhasználati képességet jelenti” (*Nemzeti alaptanterv*, 2003. 39. o.). Ez a nyelvtudás-értelmezés a *Közös Európai Referenciakeret* (KER, 2002) fogalmán alapul. A *Nemzeti alaptanterv* szerint a közoktatásban elérendő szint az első idegen nyelv esetén a 12. évfolyam végére a B1–B2, míg a második idegen nyelv esetén az A2–B1 szint. A KER hatfokú skáláján még meglévő „mesterfokú nyelvhasználói szintek (C1 és C2) elérése a közoktatásban nem tekinthető alapfeladatnak” (*Nemzeti alaptanterv*, 2003. 40. o.).

A tudáskonceptió változásának és tantervi megjelenésének következményeként értelmezhető a hazai nyelvi mérés és értékelés jelentőségének felértékelődése. Ennek megfelelően egyrészt a 2000-es évektől több, reprezentatív felmérést végeztek, melyek közös célja annak megállapítása, hogy a megfogalmazott elvárásoknak a magyar közoktatásban tanulók mennyiben felelnek meg, melyek azok a tényezők, amelyek a nyelvtudás eredményességét alapvetően meghatározzák, és ezek alapján a további fejlesztési feladatok kijelölhetők. Másrészt a *Nemzeti alaptanterv*ben meghatározott célok megvalósulásának mérése miatt a korábbi idegen nyelvi érettségi vizsga reformja is szükségesé vált. A vizsgafejlesztés során létrejött új vizsga célja a kommunikatív nyelvi kompetenciák és szintjeinek mérése a *Nemzeti alaptanterv*vel összhangban. E felmérések közös jellemzője, hogy a kommunikatív tesztelés eljárásainak megfelelően a receptív és a produktív nyelvi készségeket feladatipológiaiilag kötetlen, autentikus szövegekre és szituációkra épülő, egynyelvű feladatokkal mérik. További fontos jellemzőjük, hogy mérési eljárásaik az egységes alapot nyújtó KER (2002) szintleírásaihoz igazodnak.

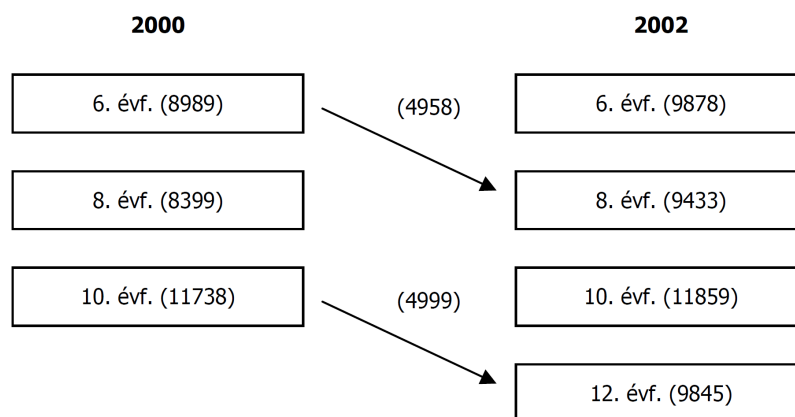
2.2. A nyelvoktatás eredményességének vizsgálata

A közoktatási rendszerben elérhető nyelvtudás szintjét, a nyelvtanulás és -tanítás eredményességét és az azt befolyásoló tényezőket több reprezentatív mintán végzett felmérés keretében vizsgálták. Az alábbiakban bemutatandó felmérések koncepciójában közös, hogy a felhasznált mérőeszközök igazodnak a *Nemzeti alaptanterv*ben meghatározott célokhoz és illeszkednek a kommunikatív tesztelés alapelveihez. A vizsgálok a közoktatásban a két leggyakrabban tanított idegen nyelvre: az angolra és a németre vonatkoznak. Mindegyik felmérésben a feladatsorok szerkezete, felépítése, az alkalmazott szöveg- és feladat típusok és a feladatok itemszáma azonos volt, így a két nyelven elért teszteredmények egymással összehasonlíthatók.

Az első vizsgálatban, 2000. májusában 6., 8., és 10. osztályos tanulók vettek részt (Csapó, 2001, 2003; Csapó és Nikolov, 2001, 2002; Nikolov és Csapó, 2002; Nikolov, 2003). A felmérésben angol és német nyelvet tanulók olvasott és hallott szöveg értését, valamint íráskészségét vizsgálták. E nyelvi készségek mellett a diákok kognitív képességei közül az induktív gondolkodást, az affektív tényezők közül a nyelvtanulással kapcsolatos attitűdöket, motivációt, énképet; továbbá a tanulók családi háttérét és a nyelv választással összefüggő tényezőket vizsgálták.

E vizsgálat közvetlen folytatásának tekinthető a 2002. májusi nyelvi felmérés, melyben 6., 8., 10. és 12. évfolyamos diákok nyelvtudását vizsgálták. A felmérés sajátja, hogy keresztmetszeti és longitudinális vizsgálatra egyaránt lehetőséget nyújt, mivel ugyanazokban az iskolákban a két évvel korábbi 6. és 10. évfolyamos diákok nyelvtudását mérték fel (Csapó, 2003, 2007; Csapó és Nikolov, 2009). A korábbi vizsgálatban

szereplő területek mellé az anyanyelvi szövegértést is bevonták a 6. 8. és 10. évfolyamos diákok körében. A két felmérés kapcsolatát a 9. ábra tartalmazza.



9. ábra

*Az idegennyelv-tudás vizsgálat mintáinak longitudinális kapcsolata
(Csapó, 2007. 341. o.)*

A harmadik nagymintás felmérésre 2003-ban került sor (Nikolov és Józsa, 2003, 2006; Józsa és Nikolov, 2005). A kutatás célja, hogy megvizsgálják az angolt és németet tanuló diákok nyelvi készségeinek fejlettségi szintjét a hatodik és a tizedik osztályban. A vizsgálatba számos iskolai környezetre vonatkozó változó (például a nyelvtanulással töltött évek száma, a heti óraszámok, a csoportlétszám) mellett az iskolán kívüli tényezőket (például a magánórák teljesítményre gyakorolt hatását) is vizsgálták.

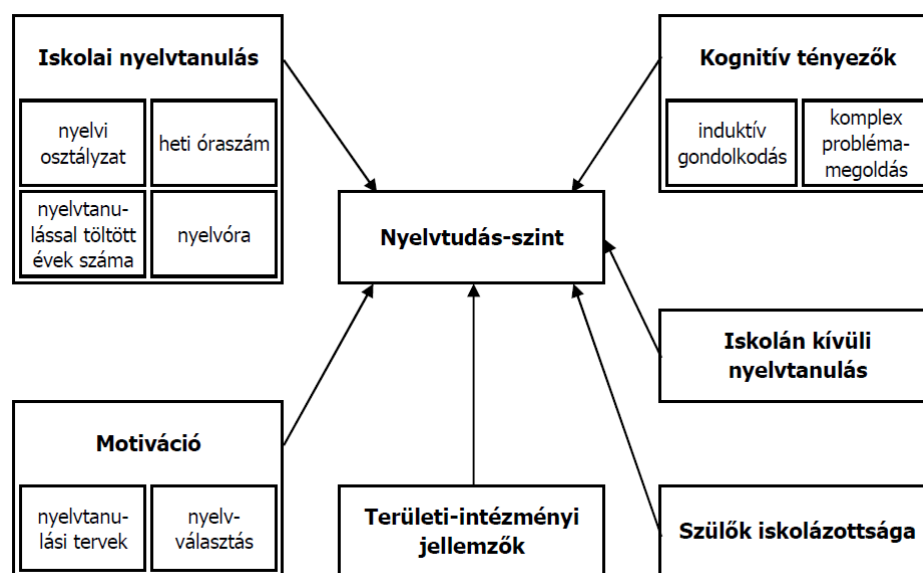
A negyedik felmérésben a 8. és a 12. évfolyamos tanulók angol és németnyelvtudásának szintjét és az azokat meghatározó tényezőket vizsgálták (Csapó és Nikolov, 2008; Nikolov, Tóth és Vígh, 2009; Vígh és Tóth, 2010). A nyelvi felmérés a 2003-ban az 1., 5. és 9. évfolyamon indult *Szegedi Iskolai Longitudinális Programhoz* (HELP)³ kapcsolódik, amely kognitív és affektív területen vizsgálja a magyar oktatás problémáit (Csapó, 2007). A longitudinális program negyedik évében a két idősebb minta érkezett az általános és a középiskola zárásához és ekkor került sor a diákok olvasott szövegértésének vizsgálatára. Ebben a felmérésben az olvasásfeladatok megegyeztek a 2000. és 2002. évben felhasznált feladatokkal, a 8. és 12. évfolyamos diákok tesztjeit egy horgonyfeladat kapcsolta össze. A keresztmetszeti elemzésbe a nyelvi készségek fejlettségét meghatározó tényezők vizsgálatához a longitudinális program keretében felvett tesztek eredményeit alkalmazták.

A hazai, reprezentatív mintán végzett vizsgálatok jelentős különbséget mutatnak a 6., 8., 10. és 12. évfolyamosok nyelvtudásában: az elért teljesítmények széles skálán mozognak. Ezt jól mutatja a legutóbbi, 2007. évi nyelvi felmérés, ahol lehetőség nyílt a 8. és 12. évfolyamos tanulók angol nyelvi olvasott szöveg értése teszten elért eredményeinek közös skálára transzformálására horgonyitemek segítségével. Az olvasást mérő feladatok szintjét két eljárás integrálásával határozták meg: egyrészt a Referencia-keret skáláinak ismeretében szemrevételezéssel, másrészt a feladatokat megoldó diákok képességszintjének megállapításával a valószínűségi tesztelmélet elemzési eljárásának használatával (Vígh és Tóth, 2010). Az eredmények szerint az olvasásfeladatok a két évfolyam olvasási készségének vizsgálatára megfelelő nehézségűek és a KER

³ „Angol megnevezésként a *Hungarian Educational Longitudinal Program* (HELP) megnevezést használjuk, és a [...] HELP rövidítést alkalmazzuk. A „Szegedi Iskolai Longitudinális Program” megnevezés [...] egy hosszabb távú kutatási folyamatra utal, több egyidejűleg vagy egymás után, különböző mintákkal elindított követéses vizsgálatot foglal magában” (Csapó, 2007. 322. o.).

deskriptorai alapján egymástól jól elkülöníthetőek voltak. Így lehetőség nyílt a KER szintjeinek bemérésére, mellyel meghatározhatóvá vált a két évfolyam nyelvtudás-szintje. A 8. és a 12. évfolyamos diákok olvasási készségeinek fejlettségében megjelenő különbségeket mutatja, hogy vannak olyan 8. osztályos nyelvtanulók, akik nem érik el az A2 nyelvi szintet és olyan 12. évfolyamos tanulók, akiknek a nyelvtudása a B2 szint felett helyezkedik el.

A tanulók nyelvi készségeinek fejlettségét meghatározó tényezőket a nagymintás felmérések eredményei alapján a 10. ábrán foglaltuk össze és a továbbiakban az a célunk, hogy e tényezők hatását jellemezzük.



10. ábra

Az idegennyelv-tudás szintjét meghatározó tényezők rendszere

A tanulók nyelvtudásában intézménytípusonként jelentős különbségek vannak. Mindegyik vizsgálat alapján a szakiskolások teljesítménye szignifikánsan alacsonyabb a szakközépiskolások teljesítményénél, mely jelentősen alacsonyabb a gimnáziumban tanulóknál. Ez a különbség az írásfeladatokon volt a legjelentősebb a 2003-as felmérésben a 10. évfolyamon (Nikolov és Józsa, 2003). A vizsgálatok eredményei szerint az angol nyelvi teszten elért eredmények jobbak, mint németből (Csapó, 2001; Nikolov és Józsa, 2003; Csapó és Nikolov, 2008). Minden vizsgálatban a lányok teljesítménye felülmúlja a fiúkét. Ugyanakkor ez a különbség kisebb az olvasás és a hallás utáni szövegértés készségeknél és meghatározó a 6. és a 10. évfolyamon az íráskészséget mérő feladatokon (Józsa és Nikolov, 2005). A 2000-es, 2003-as és 2007-es vizsgálatok tanulságai szerint jelentősek a tanulók nyelvtudásában megjelenő regionális különbségek. A 2007. évi felmérésben például az olvasott szöveg értése teszten elért nyelvi teljesítmények átlagai régiók szerint változóak, és az egyes régiók között jelentős különbségek voltak. A legjobban és a legrosszabban teljesítők közötti különbségek angolból a 8. évfolyamon 19%-pont, a 12. évfolyamon 20%-pont. Német nyelvből pedig a 8. évfolyamon 20%-pont, és a 12. évfolyamosoknál 17%-pont (Nikolov, Tóth és Vígh, 2009).

A nyelvtanulási motivációnak döntő szerepe van a nyelvtanulás eredményességében (Csapó, 2001; Imre, 2007). A diákok motivációját meghatározhatja a nyelvválasztás, mellyel a szülők iskolázottsága áll szoros összefüggésben. A 2000. évi felmérés szerint (Csapó, 2001; Nikolov, 2003) minél magasabb a szülők iskolai végzettsége, annál inkább tanulnak a gyermekeik angolul, míg a szociokulturális szempontból hátrányosabb háttérű diákok gyakrabban tanulnak németül, vagy nem tanulnak idegen nyelveket. A

nyelvtanulási motivációt a diákok céljai, nyelvtanulási tervei is meghatározzák. A 2000. évi felméréshez (Nikolov, 2003) hasonlóan alakultak a 2003-as vizsgálat eredményei (Józsa és Nikolov, 2005), melyek összhangban vannak a 9. évfolyamon 2004-ben végzett idegennyelv-tanulási utak kutatási projekttel (Imre, 2007; Vágó, 2007). E három felmérés azt mutatta, hogy a diákok elsősorban annyira akarják megtanulni az idegen nyelvet, hogy középfokú nyelvvizsgát és érettségit szerezzenek. A motiválatlan diákok aránya magasabb a németet tanulók körében (Nikolov, 2003; Józsa és Nikolov, 2005), mely a 10. évfolyamon a 2003-as felmérésben a 2000-es vizsgálatához képest 11%-ról 18%-ra nőtt. E tényező és a nyelvi teljesítmények között a 2003-as vizsgálat eredménye szoros összefüggést mutatott (Józsa és Nikolov, 2005; Nikolov és Józsa, 2006). Az eredmények szerint a leggyengébb nyelvtudású tanulók fejeznék be leghamarabb a tanulást, míg a legambiciózusabb cél: az egyetemi diploma megszerzését idegen nyelvből a legjobban teljesítő diákok tűzik ki. Mindezek mellett a diákok a nyelvtudásuknak megfelelő, reális célokat tűznek ki.

A kognitív tényezők közül a 2000 és 2002-es vizsgálatban az induktív gondolkodás és az olvasott szöveg értése szoros kapcsolatot mutatott (Csapó és Nikolov, 2002, 2009; Csapó, 2003): a 6., 8. és 10. évfolyamon angol és német nyelvből a korreláció értéke 0,42–0,51; az induktív gondolkodás fejlettsége 10–15%-ban határozta meg az olvasási teszten elért teljesítményeket. Mindezek mellett a 2000. és a 2002. évi felmérés eredményei alapján kimutatható volt, hogy az induktív gondolkodás jól jelzi előre a két évvel későbbi nyelvtudást (Csapó és Nikolov, 2009). Az induktív gondolkodás meghatározó szerepe a 2007. évi nyelvi felmérésben a 8. és a 12. évfolyamon is kimutatható volt (Nikolov, Tóth és Víg, 2009). Az elemzésbe bevont másik kognitív képesség, a komplex problémamegoldás hatása az olvasott szöveg értése teszt eredményére is figyelemre méltó különösen a német esetében. A regresszióanalízis alapján e kognitív képesség fejlettségi szintje is befolyásolja a diákok nyelvi teljesítményeit. Feltehető azonban, hogy a két teszt között lévő korreláció (0,25–0,41) a tesztek jellegével, a feladatok jellemzőivel is magyarázható, mivel az olvasási tesztben főleg olyan feladatok vannak, melyek az információk gyűjtését és egymáshoz rendelését várják el a diákoktól. A 2002. évi felmérésben az anyanyelvi szövegértés és az idegen nyelvi olvasott szöveg értése teszt eredménye közötti összefüggéseket is vizsgálták (Csapó és Nikolov, 2009). Az eredmények szerint a korreláció értéke angoltól és németből a 6. évfolyamon erősebb (angol: 0,49; német: 0,39) és ezek a 8. és 10. évfolyamokon alacsonyabbak. Az egyes nyelvi készségek fejlettsége között szorosabb kapcsolatban van, mint az olvasási teszten elért eredmény és az anyanyelvi szövegértés között. Az idegen nyelvi és az anyanyelvi szövegértés egymástól függetlenül fejlődik annak ellenére, hogy az osztálytermi gyakorlatban sokszor alkalmazzák a fordítási feladatot és a nyelvi jelenségek összehasonlítását (Csapó és Nikolov, 2001; Nikolov, 2003). Az anyanyelvi és idegen nyelvi képességek egymásra gyakorolt hatásának feltárásához, és ezek fejlettségének közös méréséhez további kutatások szükségesek.

A diákok affektív és kognitív jellemzői mellett lényeges az iskolai nyelvtanulás hatékonysága és a nyelvi teljesítmények közötti kapcsolat vizsgálata. A nagymintás mérésekben elsősorban a nyelvi osztályzatok és a teljesítmény kapcsolatát vizsgálták. A 2003-as felmérésben a nyelvtanulással töltött évek száma, a heti óraszámok, a csoportlétszám, a nyelvórai feladatok teljesítményre gyakorolt hatását is elemezték (Nikolov és Józsa, 2003, 2006; Józsa és Nikolov, 2005).

A nyelvi osztályzat és a nyelvi teljesítmény mindegyik felmérésben szoros összefüggést mutatott (Csapó és Nikolov, 2002; Józsa és Nikolov, 2005; Nikolov, Tóth és Víg, 2009), mely azt jelzi, hogy az iskolai értékelés reálisan mutatja a nyelvtudás-szintet. A hatodik és tizedik évfolyamon végzett 2003-as felmérésben a nyelvtanulással

töltött évek, a heti óraszámok és az olvasás, hallás és írás készségek fejlettsége összefüggött, de a korreláció értéke évfolyamonként, nyelvenként és készségenként eltérő volt. A nyelvi készségek fejlettsége és a heti óraszám közötti korreláció közepes (0,39–0,43), míg a nyelvtanulással töltött évek száma változónál alacsony (0,25–0,29). *Nikolov, Pércsich és Szabó (2000)* vizsgálatában hasonlóan összetett kép bontakozott ki, de megállapítható volt, hogy a heti óraszámok alakulása és a teljesítmények javulása közötti kapcsolat az íráskészséget mérő feladatokon mutatható ki legerősebben. Mindezek alapján nem a heti óraszám vagy a nyelvtanulással töltött évek száma határozza meg a leginkább a teljesítményt (lásd 10. táblázat), hanem a nyelvtanulási lehetőségek és a nyelvtanítás minősége (*Nikolov, 1999; Alderson, 2001*), amely azonban nehezebben számszerűsíthető.

A nyelvoktatás jellemzői közül gyakran vizsgált terület, hogy a diákok mennyire ismerik a feladattípusokat és ennek milyen a hatása a nyelvi teljesítményükre. A 2003-as vizsgálatnál összhangban az angol nyelvi érettségi fejlesztési munkálataiban végzett kutatásban főleg az íráskészségnél mutatkozott a legnagyobb különbség a feladattípus ismertsége szerint (*Józsa és Nikolov, 2005*), és az írásfeladat ismerősége nagyobb mértékben befolyásolta a teljesítményeket, mint az olvasás feladatok esetében (*Nikolov, Pércsich és Szabó, 2000*).

A nyelvi készségek fejlettségét meghatározó iskolán belüli és kívüli háttérváltozók hatását a 10. táblázat tartalmazza. A regresszióanalízis alapján az iskolán kívüli nyelvtanulást jellemző különórák szerepe kevésbé meghatározó, általános iskolában nem mutatható ki a teljesítményre gyakorolt hatása. Ennek oka az lehet, hogy a diákok elsősorban korrepetálás céljából járnak különórára (*Józsa és Nikolov, 2005*). *Nikolov (2007)* szerint az iskolán kívüli nyelvtanulás formái közül a nyelvi teljesítményt jelentősebben határozhatja meg a célnyelvi környezet, az önálló nyelvfejlesztés, a nem szinkronizált célnyelven hallgatott filmek és tv-műsorok, az interneten elérhető információk és tevékenységek. Bár ezek hatása empirikusan nem vizsgálható, egy horvát-magyar nyolcadikos tanulók angolnyelv-tudását mérő vizsgálatban a horvát diákok szignifikánsan jobban teljesítettek, melyet – sok más, a korábbiakban említett tényezők mellett – azzal is lehetett magyarázni, hogy Horvátországban a filmeket és tv-műsorokat nem szinkronizálják, és ez kedvezően befolyásolhatja a horvát diákok nyelvi teljesítményét (*Djigunović, Nikolov és Ottó, 2006*).

10. táblázat. A nyelvi készségek fejlettségét meghatározó néhány háttérváltozó hatása, regresszióanalízis, $r\beta$ (%)

Háttérváltozó	6. évfolyam		10. évfolyam	
	Angol	Német	Angol	Német
Szülők iskolázottsága	25	18	24	17
Heti óraszám	13	10	14	13
Nyelvtanulással töltött évek	3	3	4	4
Iskolán kívüli nyelvtanulás	n.s.	n.s.	2	2
Megmagyarázott variancia (%)	41	31	44	36

Forrás: *Józsa és Nikolov, 2005. 334. o.*

A diákok nyelvtudását az iskolai, iskolán kívüli, valamint a kognitív és affektív tényezők mellett a szülők iskolázottsága határozza meg jelentősen (10. táblázat). A szülők iskolai végzettsége és a nyelvi teszten elért teljesítmény között szoros kapcsolat van: a szülők iskolázottságának növekedésével a nyelvi teszteken elért eredmény is növekszik (lásd például *Józsa és Nikolov, 2005*). Az angol nyelvet tanulók szocioökonómiai státusza szignifikánsan kedvezőbb és erősebb összefüggést mutat a nyelvi teljesítménnyel,

mint a németül tanuló diákok esetében. Ez az összefüggés a 2003-as vizsgálaton kívül a 2000. és 2007. évi nyelvi felmérésben is hasonlóan alakult. E felmérésekben végzett regresszióanalízis szerint a szocioökonómiai státusznak van hatása a nyelvi teljesítményre, de nem játszik döntő szerepet a tanulók nyelvi készségeinek alakításában. Ez a szerep minden életkori mintában jelentősebb angolból, mint németből (Csapó és Nikolov, 2002; Nikolov, Tóth és Vígh, 2009).

A közoktatásban megszerezhető nyelvtudás szintjét meghatározó kognitív, affektív, iskolán belüli és iskolán kívüli háttérváltozók hatásának áttekintése alapján szükségesnek tűnik olyan vizsgálatok elvégzése, melyek e tényezőket komplexen és átfogóan vizsgálják. Mindezek mellett szükséges a tanítás minőségét összetett módszerekkel elemezni. Ezek a vizsgálatok lehetővé tennék a nyelvi teljesítményt meghatározó változórendszer feltérképezését, mely alapján lehetne diagnosztizálni a nyelvtanításban megjelenő problémákat és ki lehetne jelölni a jövőbeli feladatokat.

A hazai nyelvi mérés és értékelés területén jelentős előrelépést jelentenének az évenkénti és rendszeres nyelvtudás-szintet monitorozó felmérések. Ezt a célt szolgálhatná a 2005-ben bevezetett kétszintű nyelvi érettségi. A továbbiakban az empirikus vizsgálatok célkitűzéseinek pontos definiálásához e vizsga jellemzőit mutatjuk be.

2.3. A kommunikatív nyelvi kompetenciák mérése és értékelése az idegen nyelvi érettségiben

Ebben az alfejezetben az idegen nyelvi érettségi jellemzőit elemezzük az 1.1. fejezetben megadott elméleti keretnek megfelelően. A minden élő idegen nyelvre érvényes vizsgakövetelmények és vizsgaleírás (A 2010. január 1-től hatályos, Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet) alapján vizsgáljuk az érettségi nyelvtudás-modellhez való illeszkedését, a kommunikatív tesztelés kritériumainak érvényesülését (4. táblázat), a tesztfunkciók megvalósulását. Ezt követően az eddig megismert vizsgaeredmények alapján mutatjuk be a vizsgázók nyelvtudás-szintjét és meghatározó tényezőit.

A magyar közoktatási rendszerben az idegen nyelvi mérés és értékelés területén jelentős fejlődésként értelmezhető a 2005-ben bevezetett kétszintű érettségi vizsga, amelyet 1996 óta az akkori Országos Közoktatási Intézet Követelmény- és Vizsgafejlesztő Központjában alakítottak ki. A vizsgareform munkálatairól több tanulmánykötetben számoltak be a vizsga fejlesztői (az angol érettségiről: Alderson, Nagy és Öveges, 2000; a német érettségiről: Einhorn, 1998, 2004). A vizsgareformot a 2.1. fejezetben bemutatott tudáskonceptió változása és tantervi megjelenése tette szükségessé. Az 1995-ben bevezetett Nemzeti alaptantervben olyan új célokat határoztak meg, amelyek a nyelvtanítás didaktikai-módszertani fejlődésének, és a '90-es évek közepén hazánkban is megjelenő kommunikatív szemlélet következménye. A kommunikatív nyelvtanítás céljainak mérésére az akkori érettségi vizsga követelményei nem feleltek meg (Petneki, 1998, 2009). A vizsgareform eredményeként létrejött kétszintű érettségi érvényes vizsgának tekinthető a Nemzeti alaptantervben megfogalmazott célok és kijelölt fejlesztési feladatok megvalósulásának mérésére.

A kétszintű érettségi jellemzőinek összegzéséhez viszonyítási pontként az 1. mellékletben megadjuk a vizsgafejlesztés kiinduló állapotát: a korábbi érettségi jellemzőit és problematikáját. A két vizsga közötti szembevetendő különbség, hogy míg a korábbi vizsga elsősorban hagyományosabb módon, a nyelvi kompetencia összetevőit izolált nyelvi elemek segítségével mérte, addig a kétszintű érettségi célja a kommunikatív nyelvhasználat mérése. Ezért a korábbi vizsga egyik eleme sem maradt meg a kétszintű

érettségiben. E fejezet célkitűzései miatt nem elemezzük részletesen a korábbi vizsga problémáit, hanem a továbbiakban a kétszintű nyelvi érettségit vizsgáljuk a kommunikatív tesztelés szempontjai alapján.

2.3.1. A nyelvi érettségi illeszkedése a nyelvtudás-modellhez

A kommunikatív nyelvi kompetenciák értelmezése a vizsgadokumentumban

Az élő idegen nyelvekre vonatkozó vizsgakövetelmények dokumentum a nyelvi érettségi mérési célját az alábbi módon definiálja:

Az idegen nyelvi érettségi vizsga célja a kommunikatív nyelvtudás mérése, azaz annak megállapítása, hogy a vizsgázó képes-e kommunikációs céljait megvalósítani. A vizsga [...] a négy nyelvi alapkészséget méri: olvasott szöveg értése, hallott szöveg értése, beszédkésztség és íráskésztség. A kommunikatív készségek alkalmazásához a nyelvhasználónak rendelkeznie kell megfelelő szókinccsel, és ismernie kell a nyelv struktúráját is. Ezért a nyelvtani és lexikai kompetenciát mindkét szinten külön vizsgarész keretében is mérjük. [...] A vizsga középszintje az A2–B1, az emelt szint pedig a B2 szintnek felel meg [A 2010. január 1-től hatályos, Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet].

Az idegen nyelvi érettségi mérési céljának definiálásában a *Közös Európai Referenciakeret* (2002) nyelvtudás-értelmezése jelenik meg (1.1.2. fejezet), vagyis célja a kommunikatív nyelvi kompetenciák mérése, melyet kommunikatív nyelvtudásként határoz meg a dokumentum. E komplex rendszer összetevőiből egyet emel ki a rendelet: a nyelvi kompetenciákat, de nem említi a szociolingvisztikai és a pragmatikai kompetenciákat, valamint a kommunikatív folyamatok megfigyelésében és irányításában szerepet játszó stratégiai komponenst. A kommunikatív nyelvi kompetenciák összetevői helyett a mérendő készségeket határozza meg a dokumentum. Ennek feltehető oka, hogy a nyelvi mérés és értékelés során a kommunikatív nyelvi kompetenciákat és összetevőit nem lehet közvetlenül mérni, hanem a feladatokkal tudjuk a készségeket mozgósítani, működtetni és következtetni a komponensrendszer elemeire, fejlettségére.

Az élő idegen nyelvekre vonatkozó vizsgakövetelmények és vizsgaleírás a nyelvtani és lexikai kompetenciát a szókinccsre és a nyelv struktúrájának ismeretére vonatkoztatja és ezt mint a kommunikatív készségek alkalmazásának egyik meghatározó összetevőjeként definiálja. Ugyanakkor a dokumentum a továbbiakban ezt az összetevőt készségnek nevezi. A követelmények meghatározása alapján e vizsgarész a nyelvhasználati készséget méri funkcionálisan autentikus – a célnyelven valóban előforduló – szövegekbe ágyazva.

Az idegen nyelvi érettségiben mért összetevők minden élő idegen nyelvre azonosak, mindkét szint öt vizsgarészből áll. Míg emelt szinten azonos (20%) az öt vizsgarész aránya, addig középszinten a receptív és a produktív készségeket mérő vizsgarészek aránya 22%, a nyelvhelyesség vizsgarésze 12%. Ennek oka, hogy e vizsgarész középszinten azt méri, hogy a vizsgázók képesek-e alapvető, illetve a kommunikáció során gyakran használt nyelvtani szerkezetek és lexikai egységek felismerésére, kiegészítésére és létrehozására [A 2010. január 1-től hatályos, Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet].

A Referenciakeretben definiált stratégiai komponens a készségekre lebontott követelményeknél jelenik meg: a receptív készségeknél az olvasási és hallási célnak megfelelő stratégiákat méri a vizsga oly módon, hogy a feladatokban meghatározzák, hogy a szöveg globális, szelektív vagy részletes megértését várják-e el. A produktív készségeknél pontosan megadják a kommunikatív keretet: a vizsgázó és a vizsgáztató szerepét, a szituációt és az elérendő célt. Ez pedig a megfelelő produktív és interaktív stratégiahasználatot határozza meg.

Az élő idegen nyelvre vonatkozó vizsgakövetelmények további lényeges eleme, hogy a meghatározott mérési célokhoz készségekre lebontva definiálja (1) a követelményeket, (2) a feladatokhoz felhasználható szövegek jellemzőit és (3) a szövegfajtaikat a Referenciakeret alapján. Ezt követik (4) a témakörök, (5) a kommunikációs helyzetek és szándékok, valamint (6) a szókincsre és nyelvtanra vonatkozó információk. E felsorolás is mutatja, hogy e kötelező vizsgatárgyhoz minden élő idegen nyelvre érvényes és részletes vizsgakövetelmény tartozik, amelyet minden tanított idegen nyelvre vonatkozó dokumentumok egészítenek ki. Az egyes idegen nyelvekre érvényes vizsgakövetelmények megegyeznek az (1)–(4) pontban, míg a kommunikációs helyzetek és szándékok mindegyikére hoznak az adott nyelven példákat, az elvárt szintek alapján meghatározzák a szókincsre és nyelvtanra vonatkozó jellemzőket. E részletes dokumentum a tanítási-tanulási folyamat tervezését is segítheti.

A kommunikatív nyelvi kompetenciák szintjeinek mérése az idegen nyelvi érettségin

Az idegen nyelvi érettségi két szintje a Referenciakeret három szintjét méri: míg a középszintű vizsga az A2–B1, addig az emelt szintű vizsga a B2 szinthez igazodik (a vizsgakövetelményekben hivatkozott általános szintrendszert a 6. táblázat tartalmazza). Ez a szintmeghatározás összhangban van a *Nemzeti alaptanterv* (2007) fejlesztési feladataival, mely szerint a 12. évfolyam végére az első idegen nyelvnél a B1–B2, míg a második idegen nyelvnél az A2–B1 szintet kell a tanulóknak elérni. A továbbiakban a receptív és produktív készségeket mérő vizsgarészek példái alapján mutatjuk be, hogy e szintek mérése hogyan különül el egymástól.

A Referenciakeret három szintjét mérő kétszintű érettségin komoly kihívást jelent az egyes vizsgarészekben a szintek mérése, valamint, hogy ezek az évek múlásával ne változzanak, ne eredményezzen nehézségben eltérő feladatsorokat. A középszintű vizsgán az írásbeli vizsgarészekben belül a lépcsőzetesség elve érvényesül: a feladatok és itemek egyre nehezednek. A feladatkészítők szándéka, hogy az első feladat legyen a legkönnyebb, az utolsó a legnehezebb. Ugyanakkor jelenleg nem rendelkezünk adatokkal arról, hogy e célt mennyiben lehet viszonylag kevés itemmel úgy elérni⁴, hogy a feladatokról jelenleg nem tudjuk meghatározni pontosan, hogy melyik szintet mérik. Az emelt szintű vizsgán a receptív készséget mérő vizsgarészekben a B2 szint mérése első sorban a szöveg és az ahhoz tartozó feladat jellemzői alapján valósul meg.

A produktív készségeket mérő vizsgarészekben a *KER* (2002) három szintje közötti különbségek a megadott témákban és a témakifejtés mértékében jelennek meg: míg középszinten a vizsgázónak személyes nézőpontból, hozzá közel álló, mindennapi témáról kell írnia vagy beszélnie, addig emelt szinten a témát általános szempontból kell megközelíteni és a problémákat tárgyalni. Ebből következik, hogy a feladatok a vizsgázó által létrehozandó szövegfajtaiban, szöveghosszban és az értékelési rendszerben is elkülönülnek. E szempontok alapján foglalja össze a 11. táblázat az íráskészség példáján keresztül a feladatsorban lévő szintbeli különbségeket. A produktív készségek értékelésében az értékelési szempontok és skálák töltik be a legfontosabb szerepet, amelyek célja, hogy az értékelők figyelmét irányítsák a vizsgarészben mért komponensekre. Az íráskészség feladatokkal például a kommunikatív, a szövegalkotói és a nyelvi komponenseket mérik és értékelik.

⁴ Az olvasott szöveg értése vizsgarészben hasonló a vizsgaleírásban előírt itemek száma a két szinten (25–30), míg a nyelvhelyességnél középszinten 18–30, emelt szinten 30–45 item; a hallott szöveg értésénél 18–25, míg emelt szinten 20–30 item szerepel.

11. táblázat. A szintek közötti különbségek az íráskészség feladatokban a létrehozandó szövegfajta, szöveghossz és az értékelési kritériumok szerint

Középszint 1. feladat (A2)	Középszint 2. feladat (B1)	Emelt szint 1. feladat (B2)	Emelt szint 2. feladat (B2)
személyes közlés: e-mail vagy üdvözlőlap	magánjellegű levél	magánjellegű vagy intézményeknek szóló levél	véleménynyilvánítás olvasói levél vagy cikk (diák)újság számára
50–80 szó	100–120 szó	120–150 szó	200–250 szó
Tartalom: a feladat teljesítése, szöveghossz	Tartalom, formai jegyek és hangnem Szövegalkotás	Tartalom, formai jegyek és hangnem Szövegalkotás	Tartalom Szövegalkotás
Nyelvhelyesség, érthetőség	Nyelvtan, helyesírás Szókincs, kifejezőmód	Nyelvtan, helyesírás Szókincs, kifejezőmód	Nyelvtan, helyesírás Szókincs, kifejezőmód

Megjegyzés: A szövegfajta és a szöveghossz minden élő idegen nyelv esetében egységes, a feladatok értékeléséhez tartozó szempontok viszont nyelvenként lényegesen eltérhetnek. Angol nyelvből például az Íráskép is szerepel értékelési kritériumként. A táblázat így az egyes nyelvekre vonatkozó közös elemeket tartalmazza. Az értékelési szempontokhoz rendelt pontszámok is jelentősen eltérnek nyelvenként.

Az értékelés másik lényegi eleme, hogy a vizsgán mindkét produktív készség esetében értékelési skálákat kell használni. Ezek a skálák az egyes idegen nyelvi érettségi vizsgáknál nem egységesek, de közös jellemzőjük, hogy általában négy deskriptorból állnak és előfordul, hogy a skálaszerkezetük nem azonos: egy deskriptorhoz két pont is tartozhat. Erre az eltérő skálaszerkezetre mutatunk példát a 2. mellékletben az emelt szintű német érettségi íráskészség vizsgarészhez tartozó szövegalkotás szempontnál. Az íráskészségben mért nyelvi szintek az értékelési kritériumokon belül is elkülönülnek egymástól. A nyelvi komponenseket mérő szempontoknál elsődlegesen a szöveg nyelvi minősége alapján kell megítélni az írásművet. Az A2 szinten belül elvárás, hogy a vizsgázó általában helyesen használja az alapvető nyelvtani struktúrákat, szókincset; B1 szinten azt kell mérlegelni, hogy a vizsgázó helyesen használja-e az egyszerű nyelvtani struktúrákat; B2 szinten pedig az az elvárás, hogy a vizsgázó változatos struktúrákat és szókincset használjon. Ebből következik, hogy a javítás során a szövegben előforduló hibákat a mért szintnek megfelelően kell differenciálni és az értékelésnél a szintnek megfelelően a szöveg nyelvi megformáltságát, minőségét kell figyelembe venni (erről lásd részletesebben: *Einhorn, 2006a; Vigh, 2007*).

2.3.2. A nyelvi érettségi illeszkedése a kommunikatív tesztek minőségi mutatóihoz

Ebben az alfejezetben a kommunikatív tesztelés kritériumait figyelembe véve a vizsgakövetelmények és a vizsgaleírás alapján azt elemezzük, hogy mennyire teljesíti a vizsga az autentikusság, az interaktivitás, az igazságosság, az összehasonlíthatóság, a döntések reprodukálhatósága, a tesztelés és a hatékonyság követelményeit. A kommunikatív tesztelés kritériumainak jelentését a 4. táblázatban összegeztük.

Az autentikusság kritériumának való megfelelés a vizsgán egyrészt a feladatok jellemzőiben, másrészt az értékelés jellegében testesül meg. A receptív és a produktív készségekhez, valamint a nyelvhelyesség vizsgarészhez felhasználható feladatok autentikus, vagyis a célnyelven valóban előforduló szövegeket tartalmazhatnak. Ezek azonban lehetnek kis mértékben szerkesztettek és adaptáltak, a hallott szöveg értéke vizsgarészben autentikus hangzású stúdiófelvételek. Az interakciós autentikusságot a receptív és a produktív készséget mérő feladatokban úgy próbálják meg biztosítani, hogy a feladatokban pontosan meghatározzák, milyen szituációban és céllal kell elvégezni a feladatot, amely a receptív készségeknél meghatározza az olvasási vagy hallási stratégiát; a produktív készségeknél pedig a kommunikatív keret a tranzakció és interakció jellegét

és hangnemét. Ugyanakkor a produktív készségek mérésénél megadott kommunikatív keret jelentős mértékben behatárolja az íráskészség vizsgarészben a tanulók által létrehozandó szövegfajtákat (Einhorn, 2006a, 2007b). Ezen kívül a receptív készségeket mérő feladatokban gyakori, hogy tartalmi irányító szempontokat adnak meg, amelyek egyfelől irányítják a tanulók figyelmét, másfelől biztosítják a produktumok tartalmi összehasonlíthatóságát. Ez az eljárás azonban zárttá teszi a feladatokat, hiszen az érettségizőknek a megadott szempontokat kell megfelelő minőségben kifejteni. Ugyanakkor a feladatok jellemzői nem minden esetben egyeznek meg a célnyelvi környezetben várható tulajdonságokkal, vagyis a szituációs autentikusságot nehezen lehet biztosítani. E probléma megjelenési formája például az, hogy új, a vizsgázó által esetleg még relevánsnak tartott irányító szempont figyelembevételére és értékelésére az útmutató nem ad lehetőséget.

Az interaktivitás biztosítására a vizsgakövetelmények és a vizsgaleírás minden készség mérésénél felhívja a figyelmet, vagyis csak olyan szövegek fordulhatnak elő a vizsgán, melyek megértéséhez nincs szükség az érettségi általános műveltségi szintjét meghaladó ismeretekre. A szövegek és témák megfelelnek a korosztály életpaszatalatainak. A Referenciakeret szintleírásaira épülő vizsgakövetelményekben egyértelműen definiálják azokat a kontextusokat és kommunikációs helyzeteket, amelyben az érettségizőknek nyelvi cselekedniük kell. Ennek azért van jelentősége, mert ez az eljárás az interaktivitás biztosítását erősíti, mivel a konstrukció definiálásának ez a másik módja a nyelvtudás elméleti modellezése mellett. Ennek azonban előfeltétele, hogy a vizsga feladatai minél több, elsősorban a nyelvtudás bevonásával megoldható kommunikációs helyzetet reprezentáljanak (Tschirner, 2001). A produktív készséget mérő vizsgarészekben a vizsgázók néhány esetben két feladat közül választhatnak, melyekben különböző kommunikációs célokat kell elérni. A választhatóság alapelve azért fontos, mert a vizsgázóknak a produktív készségük fejlettsége és nem az adott témáról megszerzett háttértudásuk alapján kell teljesíteniük (Einhorn, 2007b). Mindezek mellett biztosítani kell, hogy a megadott témákat azonos feltételek mellett dolgozhassák ki a diákok, a két választható téma megoldottságában ne legyen szignifikáns különbség.

Az igazságosság kritériuma akkor teljesül, ha a vizsgán elért eredményekben megjelenő teljesítménykülönbségek például nemek, iskolatípusok, a diákok szocio-ökonómiai státusza szerint nem a vizsga hatására jönnek létre, vagyis az érettségi nem hoz hátrányos helyzetbe egyes vizsgázói csoportokat. A 2.2. fejezben áttekintett nyelvi felmérések alapján jelentős különbségek vannak a tanulók teljesítményében, így ehhez a vizsgaszinteknek és a feladatsoroknak igazodniuk kell. Ebből következik, hogy az idegen nyelvi érettségi egy rendkívül heterogén közoktatási rendszer záróvizsgálója (Einhorn, 2007b). Az igazságosság szempontja a középszintű vizsgán úgy teljesül, hogy a feladatokat úgy állítják össze, hogy az A2 nyelvi szinten lévő tanulók is le tudják tenni a vizsgát.

Az összehasonlíthatóság szempontja úgy érvényesül, hogy a vizsgakövetelmények és vizsgaleírás minden élő idegen nyelvre érvényes és ugyanaz: minden nyelvnél azonosak a vizsga részei, a használható segédeszközök, az időkeret. Ez vizsgaidőszakonként sem változik. Ugyanakkor az egyes nyelvenként a produktív készséget mérő feladatok értékelési kritériumaiban, az analitikus skálákban vannak különbségek. Az összehasonlíthatóság kritériumának a középszintű szóbeli vizsgarész kevésbé felel meg, mivel a feladatokat a vizsgázók tanárai készítik és a teljesítményüket is ők értékelik központi útmutató és értékelési skálák alapján. Az emelt szintű íráskészség vizsgarészben pedig a nyelvvizsgákhoz képest magas az értékelők aránya (Dávid, 2006).

A döntések reprodukálhatósága elsősorban arra vonatkozik, hogy a teszt eredményei alapján hozott döntések az idő múlásával állandónak tekinthetők-e. Az idegen

nyelvi érettségin a nemzetközi standardokhoz való igazodás lehetővé tenné, hogy az egyes feladatsorok nehézsége állandó legyen, a feladatkészítés és értékelés körülményei ne befolyásolják az eredményeket. Ugyanakkor sem a feladatok előzetes kipróbálása, sem a Referenciakerethez való illesztése, sem az értékelési skálák empirikus validálása nem történt meg. Így feltételezhető, hogy az egyes években meghozott döntések nem tekinthetők pontosnak, konzisztensnek, állandónak. Ehhez szükséges lenne a vizsgaeredmények monitorozása és a vizsgafejlesztésbe való visszacsatolása.

Az idegen nyelvi érettségi jelentős hatással van a közoktatásra, mivel a *Nemzeti alaptanterv*vel összhangban iskolai követelménnyé teszi a Referenciakeret adott szintjének elérését (Fazekas, 2006). Ugyanakkor a vizsga hatásának mértékéről, irányáról, a tanítási-tanulási folyamatban résztvevőkre, a folyamatra és az eredményekre gyakorolt hatásról nem rendelkezünk empirikus adatokkal. Így kutatást igényel annak feltárása, hogy az idegen nyelvi érettséginek milyen a hatása a közoktatásra, hogyan szolgálja az idegen nyelvek tanításában bekövetkezett szemléleti és tartalmi modernizációt, és milyen mértékben játszik szerepet abban, hogy a kommunikatív tesztelés ne csak a minősítő értékelésben, hanem a mindennapi tanítási-tanulási folyamatban is elterjedjen.

A hatékonyság az érettségi esetében azért lényeges, mert a vizsgára fordított költségek növelik a vizsga minőségét, megbízhatóságát. Ugyanakkor összhangot kell teremteni a vizsgával kapcsolatos igények és lehetőségek között. A továbbiakban ennek a kritériumnak a fontosságát két példán szemléltetjük *Einhorn* (2006b) tanulmánya alapján.

A produktív készségek mérésekor a korábbi vizsgán a felkészítő tanár értékelt központi útmutató nélkül. Ez alacsony ráfordítást igényelt, viszont az eredmény nem volt összemérhető, megbízható. A jelenlegi középszintű vizsgán a felkészítő tanárok értékelik a dolgozatokat központi útmutató alapján. Ez az előző megoldáshoz képest magasabb ráfordítást igényel (a vizsgáztatói képzés, az értékelési útmutató kidolgozása, az analitikus skálák empirikus validálása miatt), de ez növeli és összemérhetőbbé teszi az eredményeket. Az emelt szintű vizsgán külső értékelő értékeli központi útmutató alapján, amely növeli az értékelés megbízhatóságát, viszont a költségeket is emeli. Az emelt szintű vizsgán ez az eljárás azzal indokolható, hogy e vizsga a középfokú oktatást lezáró funkciója mellett nyelvvizsgaként és felvételiként is működik. A nyelvvizsgákon jellemző, hogy képzett külső értékelők dolgoznak, vagy kettős értékelést végeznek, illetve a valószínűségi tesztelméleti modellekkel vizsgálják az értékelők szigorúságát és ezt veszik figyelembe a teljesítmények értelmezésekor (erről lásd részletesebben: 1.2.3. fejezet). Ezek az eljárások az érettséginél is fontosak lehetnek, de nem biztos, hogy megéri a ráfordítást.

A receptív készségek mérésekor nem feltétlenül tekinthetjük hatékonyak az érettségit. Jellemző, hogy a korábbi érettséginek megfelelő feladatkészítési folyamat érvényesül: adott vizsgaidőszakra egy tantárgyi bizottság készíti a feladatsorokat, előzetes kipróbálás nélkül (*Einhorn*, 2008c), a vizsga minőségét pedig 2008 óta nem monitorozzák. Ez a klasszikus vizsgakészítési modell a feladatok lektorálásához, a többlépcsős feladatfejlesztési folyamathoz, a kipróbáláshoz és a feladatok empirikus nehézségének beméréséhez képest kisebb költségeket igényel. Ugyanakkor ebben a folyamatban figyelembe kell venni a kockázatokat, a következményeket és a felmerülő problémákat (ezekre a nyelvi érettségin elért eredményeket bemutató 2.3.4. fejezetben részletesebben térünk ki).

A kommunikatív tesztelés kritériumainak elemzése az idegen nyelvi érettségi példáján megmutatta, hogy a vizsgaleírás és a vizsgakövetelmények alapján az érettségi törekszik az autentikusság, az interaktivitás, a hatékonyság és az igazságosság kritériumainak megvalósítására. Ugyanakkor nincsenek meggyőző adatok arról, hogy a vizsga mennyiben felel meg az összehasonlíthatóság és a döntések reprodukálhatósága kritéri-

umoknak, mennyiben hatékony, és milyen jellegű hatást fejt ki az oktatási rendszerre, a tanárok munkájára, a diákok nyelvtanulására. E kritériumok biztosítására és elemzésére nagyobb hangsúlyt kell fektetni.

2.3.3. A nyelvi érettségi illeszkedése a kommunikatív tesztek funkcióihoz és típusaihoz

A fejezetben célunk, hogy az érettségi tulajdonságait egy adott tesztfunkcióval együtt mutassuk be és meghatározzuk, hogy a kommunikatív tesztek mely típusába sorolható. E vizsga elemzésében az 1.1.4. fejezet felépítésével és logikájával analóg módon haladunk.

Az emelt szintű idegen nyelvi érettségi vizsga a jövőbeli beválásra vonatkozó nyelvi tesztek közé sorolható, mivel felvételi vizsgaként is funkcionál. A hatályos rendelkezések szerint [100/1997. (VI. 13.) Korm. Rendelet az érettségi vizsga vizsgaszabályzatának kiadásáról] emelt szinten kell vizsgáznia annak, aki nyelvszakon kíván a felsőoktatásban továbbtanulni. A vizsga e funkciónak akkor felel meg, ha olyan feladatokat és itemeket tartalmaz, amelyek jól differenciálnak, magas az elkülönítés-mutatójuk. Ez a bekerülés kritériumának meghatározásában fontos. Ebben a funkcióban az emelt szintű vizsgán a megfelelési és a prediktív validitásnak is teljesülnie kell.

Az idegen nyelvi érettségi mindkét szintje elsősorban a tantervi követelményekre vonatkozó nyelvi tesztek közé tartozik. Az idegen nyelvi érettségi a *Nemzeti alaptanterv*ben megjelölt célok és feladatok megvalósulását méri, vagyis alapvetően a középfokú oktatást lezáró vizsgáról van szó, mely az iskolai oktatás lehetőségeihez igazodik. A vizsga eredménye ebben a funkcióban arról szolgáltat információt, hogy a tanulmányok befejezésekor milyen mértékben sikerült elérni a kitűzött célokat és a vizsga sikeressége belépőül szolgálhat egy magasabb képzési programba. Ilyen értelemben az emelt szintű vizsga felvételi szerepe összeegyeztethető ezzel a funkcióval. Ebben az esetben az érettséginek a tartalmi validitás kritériumát kell teljesítenie, azaz a vizsga tartalmának összhangban kell lennie a tantervi célokkal és fejlesztési feladatokkal. Ugyanakkor e funkció megvalósíthatóságával feltehetően ellentétes az előrehozott vizsga lehetősége, melynek során a tanulók a középfokú oktatás lezárása előtt teszik le a vizsgát. Ez a vizsgafajta azonban azért ellentmondásos, mert a vizsgázók ekkor feltehetően nem a tanítási-tanulási folyamat hatására érik el a kitűzött célokat, hanem más módon. Szükséges megvizsgálni, hogy a diákok milyen arányban választják ezt a vizsgafajta és milyen a hatása a nyelvtanulási folyamatra, mennyiben van összhangban a *Nemzeti alaptanterv*ben megjelenő európai nyelvpolitika célkitűzéseivel.

Az idegen nyelvi érettségi besorolható a konstruktumot mérő nyelvi tesztek csoportjába. Ennek oka, hogy igazodik a *Közös Európai Referenciakeret* (2002) szintrendszeréhez, amely alapján a függőleges dimenzió (azaz a nyelvhasználó milyen szinten van), illetve vízszintes dimenzió (azaz a nyelvhasználó mit tud megvalósítani) mentén jellemezhető a vizsgázó nyelvtudása. Az érettségi mindkét szintje így besorolható a nyelvi készség szinteket mérő tesztek kategóriájába is. Ezekben a tesztekben elsősorban az egyező validitásnak kell megfelelni. Ez azt jelenti, hogy az emelt (B2) szinten vizsgázó eredményének meg kell felelnie egy másik nyelvtudást mérő vizsga B2 szintjén elért eredménnyel (Dávid, 2006). Ez különösen fontos, hiszen az emelt szintű érettségi vizsga állami nyelvvizsgával is egyenértékű: ha a vizsgázó 60%-os teljesítmény fölött teljesít, középfokú C típusú nyelvvizsga-bizonyítványt is szerez. Ha a teljesítmény 40–59% között van, akkor az eredmény alapfokú C típusú államilag elismert nyelvvizsgának felel meg [100/1997. (VI. 13.) Korm. Rendelet az érettségi vizsga vizsgaszabályzatának kiadásáról]. Ezen intézkedés lehetőséget biztosít arra, hogy a középfokú nyelvvizsga a

közoktatás végére mindenki számára ingyen megszerezhető legyen (*Kapitánffy*, 2001). Ezért a vizsga a tanítási-tanulási folyamattól független nyelvvizsgák rendszerében is elhelyezhető. Mindezek mellett a vizsgának igazodnia kell a meghatározott nyelvtudásmodellhez (belső validitás), valamint a Referenciakeret szintjeihez (külső validitás). Nem rendelkezünk azonban adatokkal arról, hogyan viszonyul a két vizsgaszint egymáshoz és a vizsgán mért három nyelvi szint a Referenciakeret deskriptoraihoz. Ez a nemzetközi szakirodalomban alaposan elemzett terület (lásd 1.2.2. fejezet), melyben a valószínűségi tesztelméleti eszközök alkalmazása elengedhetetlen (lásd 1.2.3. fejezet).

Az idegen nyelvi érettségi vizsga a standardizáció követelményének akkor felel meg, ha ugyanarra a nyelvtudásra bármelyik vizsgán ugyanazt az eredményt érik el a diákok (*Csapó*, 1996 alapján). E kritérium teljesülésének előfeltétele a nyelvtudás pontos meghatározása, a feladatokkal való lefedés során az előtesztelés a validitás biztosítása céljából, majd a vizsgafeladatok kalibrálása azzal a céllal, hogy a különböző években vizsgázók azonos nyelvtudással ugyanazt az eredményt érik el. Ezeknek a szigorú feltételeknek a jelenlegi vizsga nem felel meg. Ugyanakkor az érettségin megjelennek a standardizációs törekvések (*Horváth és Lukács*, 2006): (1) részletes és minden élő idegen nyelvre egységes a vizsgakövetelmény és a vizsgaleírás, (2) a közép- és emelt szinten az írásbeli vizsgafeladatok központilag készülnek, (3) a receptív és produktív feladatok értékeléséhez részletes javítási-értékelési útmutató, a produktív készségek értékeléséhez skálák tartoznak. A vizsga bevezetését segítették a vizsgáztatói képzések, melyek célja az értékelők egyéni normáinak egységesítése. A hatékonyság kritériumának tárgyalásakor azonban utaltunk rá, hogy a feladatírás módja nem szabályozott, a középszintű érettségi a középfokú oktatást lezáró vizsgaként működik, míg az emelt szintű vizsga a külső értékeléssel a nyelvi készségszintet mérő vizsgák alapelvét követi (*Einhorn*, 2007b). Mindezek alapján az érettségi a nem standardizált formális tesztek közé sorolható, melyek fő jellemzője, hogy a vizsgakészítők ügyelnek a kommunikatív tesztelés kritériumainak biztosítására.

A kritériumorientált értékelés a vizsgán két szempontból érvényesül: egyrészt a Referenciakeret három szintjének mérésével lehetőség nyílik arra, hogy a vizsgázók egyéni készségszintjét pontosan le tudjuk írni. Ennek azonban előfeltétele, hogy a teszt a követelményszinteket minél jobban és pontosabban reprezentálja, mert így lehet a teszt eredményeiből következtetni a vizsgázók nyelvtudásának szintjére. Másrészt pedig a produktív készségeknél alkalmazott analitikus értékelési skálák kritériumorientált értékelést tesznek lehetővé. A skálákban a maximális deskriptor az elérendő szintet reprezentálja, mely a vizsgakövetelményekre és a KER szintjeire épül. Ezt a szintet részekre bontják, minden szintnek részletes leírása van, melybe az értékelők a teljesítmény értékelésénél a vizsgázót besorolják. Ezzel az értékeléssel diagnosztikus információt is lehet adni a produktum megítéléséhez kapcsolódó értékelési szempontok alapján a vizsgázók produktív készségeinek fejlettségéről, így ez az eljárás a tanítási-tanulási folyamatban is felhasználható. A kritériumorientált értékelés jól alkalmazható az érettségi felvételi funkciójánál a bekerülés határának megállapításához.

Az idegen nyelvi érettségin a receptív és a produktív készségek mérésekor elsősorban performancia tesztekkel használnak. Ezek közös célja a nyelvi készség mérése, melyhez meghatározzák a nyelvhasználati szituációt. A receptív készséget mérő feladatokban megadják, hogy milyen céllal és stratégiával kell a szöveget megérteni. A produktív készségek mérésekor egy nyelvhasználati szituációt képeznek le és adott kontextusban mérnek a teszttel. Ebben az esetben az autentikusságot szimulálják, így a vizsga nyelvezete sokszor nem természetes, a vizsgahelyzetben nem lehet valós nyelvi környezetet teremteni. Ebben az esetben a vizsga célja a nyelvi teljesítmény mérése és értékelése, ezért a vizsgán szűkebb értelemben vett performancia tesztekkel alkalmaznak. A

nyelvhelyesség vizsgarész a kompetencia tesztek logikáját követi, de nem izolált, a szövegkörnyezetből kivett mondatok és célzottan a vizsgára készülő szövegek alapján történik a mérés, hanem autentikus szövegekkel vizsgálják a nyelvhasználati képességet.

A vizsga eltérő funkciókat valósít meg: elsősorban a középfokú oktatást lezáró vizsga és egyben nyelvtudásszintet mérő vizsga is, így a vizsga e két típus keverékének tekinthető. A vizsgán megfigyelhetők a standardizációs törekvések, de a vizsgaminőség növelésére nagyobb figyelmet kell fordítani.

2.3.4. Teljesítmények és a diákok nyelvtudás-szintje az idegen nyelvi érettségi vizsgákon

Az idegen nyelvi érettségin elért teljesítményekről három publikált forrás áll rendelkezésre: (1) A 2005. és 2006. évi érettségi eredményeiről külön tanulmányokban számoltak be, melyekhez az Oktatáskutató és Fejlesztő Intézetben (OFI) feladat- és itemszintű elemzéseket végeztek (Einhorn, 2007a, 2008a, 2008b; Major, 2007, 2008). (2) Az Oktatási Hivatal (OH) minden évben nyilvánosságra hozza a legfontosabb eredményeket, teljesítmény-eloszlásokat⁵. (3) Az Oktatási és Kulturális Minisztérium (OKM) felmérést végzett a diákok nyelvtudásáról azzal a céllal, hogy a 2008. évi érettségi eredményei alapján megállapítsák, hogy a diákok középszinten milyen arányban érik el a B1 szintet és emelt szinten a B2 szintet. Az idegen nyelvi érettségiről a publikált eredményeken kívül az OH adatbázisai nyilvánosak és hozzáférhetők⁶, az OFI adatbázisain másod-elemzések is végezhetők. A továbbiakban az eddig publikált adatok alapján mutatjuk be az érettségin elért eredményeket és viszonyítjuk a 2.3.2. és 2.3.3. fejezetekben tárgyalt problémakörökhöz. Ezek alapján a fejezet végén meghatározzuk, hogy az eredmények milyen további tennivalókra és elemzésekre utalnak.

Az első, 2005. évi érettségi nem tekinthető szabályszerűnek, mivel a diákok ebben az évben nyelvvizsga-bizonyítvánnyal válhatták ki az emelt szintű vizsgát. 2005-ben angolból az emelt szintű vizsgára jelentkeztettek 88%-a, míg németből 91%-a rendelkezett nyelvvizsgával (Einhorn, 2007b). Emelt szintre így feltehetően csak azok jelentkeztettek, akiknek ez a továbbtanuláshoz feltétlenül szükséges volt. Mindezek alapján arra lehetett számítani, hogy növekedni fog az emelt szintre jelentkezők száma, de ez a 2006. évi vizsgán nem következett be: angolból az összes vizsgázó 8–9% választotta az emelt szintet, németből a vizsgázók 6%-a jelentkezett emelt szintre. Feltételezhető, hogy az emelt szintű vizsgát megírók köre ebben a két évben nem fedte le az emelt szint tényleges célcsoportját, tehát a nyelvszakos felsőoktatásba jelentkezőkön kívül a nem nyelvszakos felsőoktatásba jelentkezők és a nyelvigényes munkahelyre készülők vagy a szakmát tanulók körét (Einhorn, 2007b).

A vizsgaeredmények az angolt és a németet illetően 2005 és 2006-ban hasonlóan alakultak (12. táblázat).

12. táblázat. A 2005. és 2006. május-júniusi érettségi adatai angolból és németből

	Angol, 2005		Angol, 2006		Német, 2005		Német, 2006	
	K	E	K	E	K	E	K	E
Vizsgázók száma	32721	3075	38315	3530	23836	1470	26381	1645
Átlag (%)	54,01	64,38	52,38	62,98	56,45	70,52	56,84	69,39
Jegyátlag	3,29	4,28	3,24	4,34	3,38	4,41	3,43	4,47

Jelölések: K: Középszint, E: Emelt szint; Forrás: Einhorn, 2007b. 85. és 87. o. táblázatai.

⁵ Forrás: <http://www.oh.gov.hu/kozoktatas/korabbi-erettsegi/korabbi-erettsegi>

⁶ Az adatbázisok elérhetősége: <https://www.ketszintu.hu/publicstat.php>

Középszinten 2005-ben és 2006-ban mindkét nyelvből 50% körüli volt a vizsgázók átlageredménye, az emelt szint jobb eredményei alapján a vizsga az érettségizők számára könnyűnek bizonyult. E két vizsga elemzése alapján feltételezhető volt, hogy ha az emelt szintre többen jelentkeznek, mindkét szint vizsgázói csoportja átrendeződhet és feltehetően a további vizsgákon közép- és emelt szinten alacsonyabb teljesítmény várható (Einhorn, 2007b). Az átrendeződés nem következett be. Az OH tájékoztatója alapján nőtt a vizsgázók száma mindkét nyelvből és szinten és ezzel együtt a teljesítmények is javultak: középszinten angolból jelentősebben, mint németből (13. táblázat). A teljes vizsgára vonatkozó teljesítmény-eloszlás alapján a 2005 és 2006-os érettségihez képest mindkét nyelvből nőtt azon vizsgázók száma, akik 80–100% között teljesítettek. Míg 2006-ban angolból a vizsgázók 10,28%-a teljesített 80% fölött, addig 2009-re már 36,63%. Németből is hasonló tendencia figyelhető meg: 2006-ban a vizsgázók 17,39%-a teljesített a 80–100%-os teljesítmény-sávban, 2009-ben már a vizsgázók egynegyede került ide. Ennek több oka lehet: (1) feltételezhető, hogy a feladatsorok a korábbi vizsgához képest jóval könnyebbek, (2) a felkészülés hatása mutatható ki, a diákok megismerik a feladattípusokat és ez javítja teljesítményüket, vagy (3) változik a vizsgázók pontszerző stratégiája: középszinten is lehetnek olyan diákok, akik megszerezték a nyelvvizsga-bizonyítványt, és így nyelvtudásuk magasabb B1 szintnél.

Emelt szinten a középszinthez hasonló tendencia figyelhető meg. Minden évben nőtt azon vizsgázók száma, akik 60% fölött teljesítettek és jeles érdemjegyet szereztek: 2006-ban angolból 59,09%-ról 2009-re 73,94%-ra, németből folyamatosan emelkedve 70,46%-ról 85,43%-ra nőtt a 60% fölötti teljesítményt elérő vizsgázók aránya. Az emelt szintű vizsga felvételiként is funkcionál, így az egyes készségek esetében is érdemes megnézni a teljesítmény-eloszlást mindkét nyelvből (Érettségi 2009. május-június. Tapasztalatok, 2009). Ezek közül a beszéd-készség vizsgarész megbízhatósága alacsony: angolból 2009-ben a vizsgázók 30%-a, németből 31%-a teljesített a 90–100%-os teljesítménysávban, ugyanez a tendencia kevésbé szélsőségesen a nyelvhelyesség vizsgarész kivételével a többi vizsgarészre is igaz. Az eredményekből arra következtethetünk, hogy az emelt szintű vizsga nem tölti be megfelelően a receptív és a produktív készségeknél a felvételi funkcióját. Az egyes vizsgaévekben az eredmények változnak, így a kommunikatív tesztelés kritériumai közül a döntések reprodukálhatósága sem valósul meg. Ez feltehetően a feladatsorok kipróbálásának, a feladatok részletes elemzésének és szintillesztésének hiányával magyarázható, valamint a produktív készségeknél az értékelési skálák validálásának szükségességére utalnak.

13. táblázat. A 2007., 2008. és 2009. május-júniusi érettségi vizsga adatai angol és német nyelvből

<i>Angol</i>	2007		2008		2009	
	<i>K</i>	<i>E</i>	<i>K</i>	<i>E</i>	<i>K</i>	<i>E</i>
Vizsgázók száma	44638	3668	45261	3701	50206	4052
Teljesítményátlag (%)	61,94	70,97	61,27	71,25	65,16	69,48
Jegyátlag	3,69	4,65	3,66	4,68	3,83	4,60
<i>Német</i>	2007		2008		2009	
	<i>K</i>	<i>E</i>	<i>K</i>	<i>E</i>	<i>K</i>	<i>E</i>
Vizsgázók száma	28321	1812	26256	1679	26364	1942
Teljesítményátlag (%)	61,41	69,45	57,28	73,44	57,22	74,96
Jegyátlag	3,65	4,49	3,46	4,70	3,46	4,77

*Jelölések: K: középszint, E: emelt szint;

Forrás: Érettségi 2009. május-június. Tapasztalatok, 2009.

A 2007-es érettségit követő évekből az egyes vizsgarészekben elért teljesítményekről nincs információink. Erre az OH adatbázisainak elemzéséből lehetne következtetni, mely minden vizsgázóra vonatkoztatva tartalmazza az egyes készségek vizsgapontszámait, de nem történt meg ezek részletes elemzése. Rendelkezésre állnak viszont a 2005. és 2006. évi érettségi klasszikus tesztelméleti módszerekkel végzett elemzésének eredményei (Einhorn, 2007a, 2008a; Major, 2007, 2008). A 14. táblázat az egyes vizsgarészek megoldottságát mutatja közép- és emelt szinten angol és német nyelvből.

14. táblázat. A közép- és emelt szintű vizsga eredményei 2005-ben és 2006-ban készségenként angoltól és németből (%)

Vizsgarész	Angol, 2005		Angol, 2006		Német, 2005		Német, 2006	
	K	E	K	E	K	E	K	E
Olvasott szöveg értése	66,85	80,21	49,72	61,33	60,27	82,86	61,04	78,56
Nyelvhelyesség	50,15	66,78	39,42	63,73	50,06	83,14	53,74	59,91
Hallott szöveg értése	42,90	62,55	51,39	53,97	69,70	76,30	59,88	75,35
Íráskészség	49,75	77,40	53,52	69,90	42,88	73,31	50,50	64,55
Beszéd-készség	58,89	73,01	63,29	67,40	56,30	76,49	61,04	70,46
Teljes vizsga	54,01	72,39	52,38	62,98	56,45	70,52	56,84	69,39

Megjegyzés: A 2005. évi érettségi eredményei becslésen alapulnak, mert a rendelkezésre álló adatbázisban nem lehetett elkülöníteni a nyelvvizsgával rendelkezők és a valóban vizsgázók adatait.

Jelölések: K: Középszint, E: Emelt szint; Forrás: Einhorn, 2008a; Major, 2008.

Angoltól középszinten mindkét évben a nyelvhelyesség vizsgarész bizonyult a legnehezebbnek. Ez a feladatok újszerűségével, a diákok esetleges felkészületlenségével magyarázható, és mutatja a fejlesztés szükségességét. A feladatszerű elemzések szerint a 2006. évi vizsgán ez a feladatsor nehéz volt, de tesztelméleti szempontból többnyire megfelelően mért. Az íráskészség esetében a tanórai fejlesztés szükségessége jelent meg: a próbaérettségik tanulságai alapján is ez a vizsgarész jelentette a diákok számára a legnagyobb kihívást. A 2005-ös vizsga feladatszerű elemzései azt mutatták, hogy a hallott szöveg értése feladatsor nehéz volt a diákoknak, az olvasott szöveg értése vizsgarész pedig túl könnyű feladatokat tartalmazott. Ebből is adódhat a két évben a feladatok teljesítettségének ingadozása. Míg 2005-ben nem tért el jelentősen a beszéd-készség vizsgarész a többi készség megoldottságától, addig 2006-ban magas teljesítményt értek el a vizsgázók. Ez a helyi értékelés problémáit mutathatja. Emelt szinten 2005-ben 60% fölötti átlageredmények születtek. Ennek okait az előzőekben elemeztük. A 2006-os érettségi tapasztalatai alapján az olvasott szöveg értése és a nyelvhelyesség vizsgarészek kevés olyan ítemet tartalmaztak, amelyek a magasabb képességtartományban tennének különbséget a vizsgázók között. A hallott szöveg értése feladatsor 2006-ban nagyon egyenetlen volt: a feladatszerű elemzések tanulságai szerint egy feladat túlságosan nehéz volt a vizsgázóknak. Az íráskészséghez tartozó vizsgadolgozatok elemzése azt mutatta, hogy bizonytalanságok fordultak elő a javítási folyamatban a külső javítás standardizált eljárásai ellenére.

Németből középszinten mindkét évben az íráskészségnél volt a legalacsonyabb a teljesítményátlag, és a legmagasabb a szövegértést mérő vizsgarészekben és a szóbeli vizsgán. A nyelvhelyesség vizsgarész az íráskészség után bizonyult a legnehezebbnek. A feladatsorok minőségét a feladat- és ítemszintű elemzések alapján megfelelőnek ítélték (Einhorn, 2007a). Emelt szinten a középszinthez hasonló tendencia látható, de egyenetlen a feladatsorok teljesítettsége, elsősorban a nyelvhelyesség és az íráskészség tesztben. A 2006-os érettségi eredményei minden készségnél alacsonyabbak, a feladatsorok minősége nem volt megfelelő, amely a feladat-készítési folyamat hiányosságaira

utal: elsősorban a nyílt végű feladatoknál nem volt jó minőségű az értékelési útmutató. Az íráskészség vizsgarészben ellenőrző javítást is végeztek, amely alapján főleg emelt szinten pontatlanságok, durva értékelési hibák fordultak elő (Einhorn, 2008a).

A 2005. és 2006. évi érettségi eredményei alapján a teljes vizsgára vonatkoztatva az iskolatípusok közötti különbségeket ismerjük. Középszinten angolból a két év átlagában jelentős, 13%-pontnyi különbség mutatható ki a gimnáziumok és szakközépiskolák között. Ez az eredmény az emelt szinthez nagyon hasonló. Németből középszinten a két év átlagában 14%-pont különbség mutatható ki. A két vizsgázói csoport teljesítménygörbéje egymás tükörképe. Emelt szinten roppant ellentmondásos a két év eredménye németből: míg 2005-ben 17%-pont volt a különbség a gimnáziumok és a szakközépiskolák között, addig 2006-ban nem volt túl jelentős: 7%-pont. A további vizsgaeredmények alapján célszerű megvizsgálni, hogyan alakul a két iskolatípus közötti különbség: a vizsga mennyiben konzerválja, növeli, vagy csökkenti a meglévő különbségeket.

Az OKM felkéréséből a Nyelvvizsgát Akkreditáló Testület (NYAT) bevonásával a 2008. évi érettségi vizsgát elemezték szakértői véleményezési eljárással, melynek célja az volt, hogy a KER skálái alapján állapítsák meg, hogy a középszintű olvasott és hallott szöveg értése, valamint az íráskészség feladatokon hány pontot kell minimálisan teljesíteni a vizsgázónak a B1, emelt szinten pedig a B2 szint eléréséhez. A vizsgát a módosított Angoff-módszerrel végezték (Tannenbaum és Wylie 2005, 2008), melynek lényege, hogy a szakértők a feladatsorokat először egymástól függetlenül, majd egymással egyeztetve minősítik a KER szintrendszere alapján úgy, hogy egy 13 fokozatú skálán megállapítják, hogy középszinten a B1, emelt szinten a B2 szinten lévő vizsgázó milyen valószínűséggel oldaná meg helyesen az itemet és az így kialakított értékek átlagait tekintik a készségenkénti határpontnak. Az íráskészség vizsgarészben a feladatokat és az analitikus értékelési skálákat használták a szintminősítéshez. Mindezek alapján határozták meg a határoló pontértéket, mely azt jelöli, amit középszinten a B1, emelt szinten a B2 szint eléréséhez teljesíteni kell. Ezek alapján megállapítható, hogy a 2008. évi érettségien a vizsgázók hány százaléka éri el ezt a szintet (15. és 16. táblázat).

15. táblázat. A 2008. évi május-júniusi vizsgaidőszak szintillesztési eljárásának eredménye angolból

	Olvasás		Hallás		Írás		Összesen	
	K	E	K	E	K	E	K	E
Maximális vizsgapont	33	30	33	30	33	30	99	90
Határoló pontérték	18	17	23	18	29	21	70	56
Vizsgázók átlagpontszáma	18	21	22	23	20	22	60	66
Határoló pontértéket elért vizsgázók (%)	50	80	47	86	31	68	39	79

Jelölések: K: Középszint, E: Emelt szint; Forrás: *Felmérés a 2008-ban Magyarországon érettségizettek idegennyelv-tudásáról*, 2009.

A NYAT-vizsgálat alapján középszinten angolból és németből az olvasási készségnél és angolból a hallott szöveg értése vizsgarészben a vizsgázók közel fele érte el a B1 szintet. Angolból a vizsgázók magasabb aránya érte el ezt a szintet, mint németből. Emelt szinten magas a receptív készségeknél a B2 szinten teljesítők aránya mindkét nyelvből, de összességében itt is angolból érték el többen a B2 szintet. A vizsgálat eredményei mindezek mellett arra is felhívják a figyelmet, hogy a nyelvtudás nem egyenletes: a vizsgázók általában a receptív készségekből magasabb szinten vannak, mint az íráskészségnél. Feltételezhető tehát, hogy az emelt szintű vizsgázók a szövegértés esetében C1 szinten vannak. A középszinten elért teljesítményeket torzíthatja, hogy lehetnek olyanok, akik B2 szintű nyelvvizsga-bizonyítvány megszerzését követően tet-

tek érettségét. A kutatási beszámolóban a nyelvvizsgázók számának kontrollálásával állapították meg, hogy a vizsgázók 36%-a érte el a B1 szintet, 7%-uk a B2 szintet és 56%-uk van A2 szinten. Ugyanakkor a kutatási beszámoló nem ad választ arra a kérdésre, hogy iskolatípusonként hogyan oszlik meg a vizsgázók száma, vagyis a szakközépiskolában ezeknek az adatoknak a figyelembevételével a tanulók hány százaléka éri el a B1 szintet.

16. táblázat. A 2008. évi május-júniusi vizsgaidőszak szintillesztési eljárásának eredménye németből

	Olvasás		Hallás		Írás		Össz.	
	K	E	K	E	K	E	K	E
Maximális vizsgapont	33	30	33	30	33	30	99	90
Határoló pontérték	21	21	21	20	27	24	69	65
Vizsgázók átlagpontszáma	21	24	19	22	17	22	55	67
Határoló pontértéket elért vizsgázók (%)	49	83	39	69	24	51	33	65

Jelölések: K: Középszint, E: Emelt szint; Forrás: *Felmérés a 2008-ban Magyarországon érettségizettek idegennyelv-tudásáról*, 2009.

Az idegen nyelvi érettségi a korábbi vizsgához képest előrelépést jelent, hiszen igazodik a nyelvi mérés és értékelés nemzetközi tendenciáihoz, elhelyezhető a kommunikatív tesztelés rendszerében, az érettségizők pontosan meghatározott követelmények alapján zárják le a középiskolai tanulmányaikat. Ugyanakkor az érettségivel szemben támasztott minőségi elvárásokat és a vizsgafunkciók teljesülését az eredmények folyamatos monitorozásával és a vizsga tanítási-tanulási folyamatára gyakorolt hatásának elemzésével lehet vizsgálni. Lényeges továbbá ezeknek az információknak a vizsgafejlesztésbe való visszacsatolása. Ez a folyamat azonban a bemutatott és eddig megismert eredmények alapján kevésbé jellemző, a vizsga működése alapján számos vizsgálandó területet, kutatási problémát azonosíthatunk. Kevés adattal rendelkezünk az érettségizők nyelvi készségeinek fejlettségéről, az azt meghatározó tényezőkről; a receptív készséget és a nyelvhelyességet mérő vizsgarészekben alkalmazott feladatok és itemek nehézségi szintjéről, valamint a produktív készségek mérési-értékelési eljárásainak minőségéről: a performancia feladatok jellemzőiről, az értékelési szempontok és skálák működéséről. Mindezek mellett nem rendelkezünk adatokkal arról, hogy a vizsga hogyan és milyen mértékben fejti ki hatását a tanítási-tanulási folyamatra. Szükséges elemezni, hogy a vizsga hatására hogyan változik a tanárok munkája, hogyan jelennek meg a vizsgán alkalmazott mérési eljárások és értékelési alapelvek a tanítási-tanulási folyamatban.

3. A NYELVI MÉRÉS ÉS ÉRTÉKELÉS MECHANIZMUSAINAK EMPIRIKUS VIZSGÁLATA

A fejezet elsődleges célja a szakirodalmi áttekintés és az empirikus vizsgálataink közötti kapcsolat megteremtése. E cél elérése érdekében egy modellt készítettünk az 1.3. fejezetben bemutatott teszthatás folyamatát leíró elméleti és empirikus kutatások eredményei alapján. A nyelvi mérés-értékelés mechanizmusainak modellezését azért tartjuk szükségesnek, mert célul tűztük ki az érettségi komplex és többszemponútú elemzését, mellyel a 2.3. fejezetben azonosított problémákat vizsgálhatjuk és ennek segítségével a modell relevanciáját igazolhatjuk, a modellt empirikusan validálhatjuk. A fejezetben először a szakirodalmi vizsgálat alapján általánosan jellemezzük a visszahatás működését szintetizáló modellt, majd a nyelvi érettségi példáján keresztül értelmezzük. Ezt követően kijelöljük a modellben a vizsgálandó területeket, és ez alapján fogalmazzuk meg a vizsgálatunk céljait, a kutatási kérdéseket és hipotéziseket. Végül a kutatási módszereket, a vizsgálat mintáját, eszközeit és módszereit mutatjuk be.

3.1. A vizsgálat elméleti háttere és célterületei

3.1.1. A vizsgálat elméleti modellje

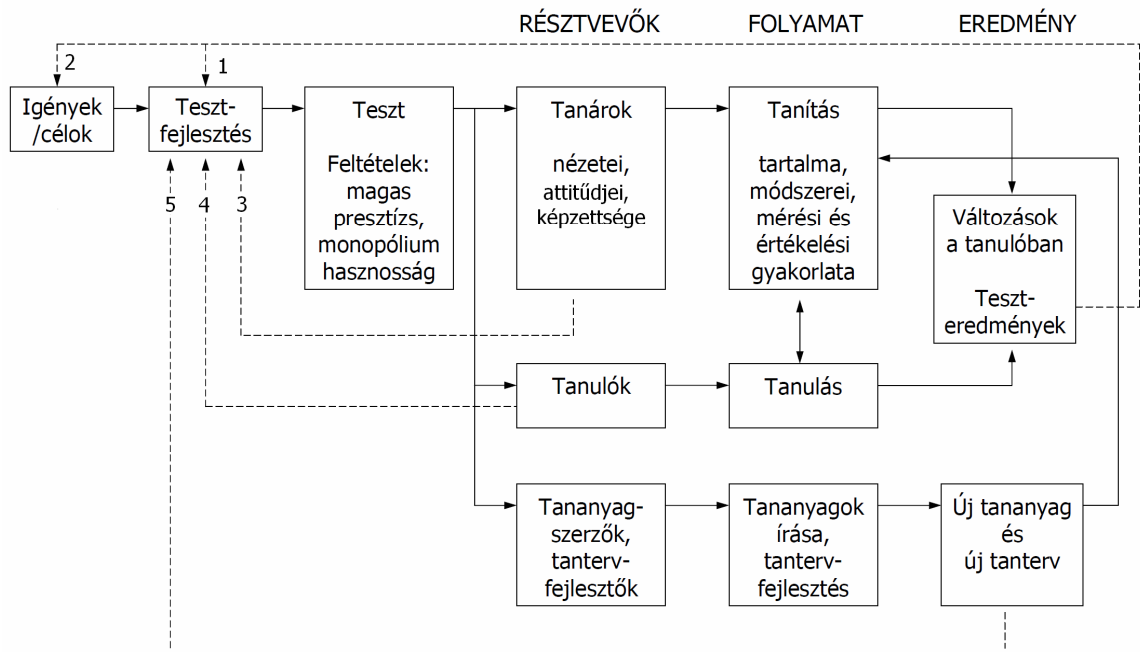
Az angol nyelvű szakirodalomban a teszthatást a *washback*, a *backwash* és az *impact* fogalmakkal határozzák meg (1.3.1. fejezet). A fogalomértelmezések szintetizálása alapján a vizsgálatunkban a teszthatást egy komplex folyamatként értelmezzük, és két-féle típusát különböztetjük meg. (1) A tesztek hatásmechanizmusai szűkebb értelemben véve a tanárookra és a tanulóakra (Alderson és Wall, 1993; Messick, 1996), a tanítási-tanulási folyamatra, azaz a mikrokörnyezetre (Bachman és Palmer, 1996) vonatkozik. Ezt az angol nyelvpedagógiai szakirodalomban a *washback* vagy a *backwash* fogalommal, az *impact* egy dimenziójának megnevezésével (Hamp-Lyons, 1997; Wall, 1997) jelölik a kutatók. (2) Tágabb értelemben véve az oktatási rendszerre, illetve a társadalom egészére (Hamp-Lyons, 1997; Shohamy, 1999; Wall, 1997), azaz a makrokörnyezetre (Bachman és Palmer, 1996) gyakorolt hatásról van szó, amely elsősorban a nagy tétellel rendelkező formalizált tesztekre jellemző. A kutatásban a mikroszintű teszthatások feltárására törekszünk.

A vizsgálatunk szempontjából lényeges kérdés, hogy mely tényezők határozzák meg a teszthatás folyamatát, hiszen ennek intenzitása, mértéke és értéke előre meghatározhatatlan. Ez a kérdés a nyelvpedagógiai szakirodalomban több paradigma mentén értelmezhető (1.3.2. fejezet). Ezek közül a kutatásunkban a teszthatást innovációs folyamatként értelmezzük, mely szerint a tesztek segíthetik a tanítási-tanulási folyamat változásait. A hatás számos tényezőtől függ, melyek közül meghatározó jelentősége van a tanárok teszttel kapcsolatos nézeteinek, feltételezéseinek, ismereteinek és attitűdjeinek (Burrows, 2004). A vizsgálat alapjául szolgáló szintetizáló modell (11. ábra), a teszthatást a tanítási-tanulási folyamatban bekövetkező változásként ábrázolja, és célja a folyamat elindulásához szükséges tényezők megnevezése, a nemzetközi szakirodalomból ismert empirikus kutatások eredményeinek integrálása és a folyamat modellálása.

A szintetizáló modellben Saif (2006) alapján az első fázis az igényeknek és a céloknak megfelelő tesztfejlesztési folyamat, amely három fő szakaszból áll. A tesztfejlesztés egyes szakaszai jól definiált alapelvek és minőségi elvárások szerint történnek (például Milanovic, 1997). A tervezési szakaszban a felmerült igények és célok, valamint a létező gyakorlat elemzése történik. A társadalmi és kliensi igények figyelembe vétele azért

fontos, mert ez közvetlenül és közvetetten határozza meg a létrejövő teszt célját, tartalmát és típusát, valamint fejlesztését és kivitelezését. A teszt lebonyolítására szolgáló tesztmodell kialakításának szakaszában központi elem a célcsoport meghatározása és tudásszintjük rögzítése, amelyhez igazítva történik a követelmények definiálása a nemzetközi standardokhoz való igazodás alapján (1.2.2. fejezet). Ezt követi a követelményszinteknek megfelelő mérési-értékelési eljárások kipróbálása, minőségük nyomon követése és a kipróbálás során szerzett tapasztalatok folyamatos beépítése a fejlesztési folyamatba. A szervezési fázisban a tesztfeladatok elkészítése, a feladatbank kiépítése, az értékelők képzése, majd az adminisztráció kidolgozása történik.

A tesztfejlesztési folyamat eredményeként egy olyan tesztnek kell létrejönnie, amelynek magas a presztízse, státusza és elismertsége vagy monopóliuma van a közoktatási rendszeren belül, továbbá a teszt teljesítése a jelölt számára fontos lehetőségeket, jogosultságokat jelent (Gates, 1995). Ez utóbbit fejezi ki a hasznosság fogalma. Ezek a tényezők a hatás intenzitását és mértékét alapvetően meghatározhatják (Watanabe, 2004a), így elengedhetetlenek a hatásmechanizmusok elindulásához. Ugyanakkor egy teszt minőségétől többnyire független annak hatása, mivel azt elsősorban a tanárok teszttel kapcsolatos nézetei, attitűdjei és a tanárok képzettsége határozzák meg (9. táblázat). A nemzetközi empirikus vizsgálatok tanulságai szerint (Cheng, Watanabe és Curtis, 2004) a tanárok egy új, vagy módosított tesztre egyéni, sokszor eltérő tanítási reakciókat adnak. A tanárnak kulcsfontosságú szerepe van abban, hogy a teszthatás hogyan és milyen mértékben jelenik meg (Spratt, 2005).



11. ábra
A visszahatás működését szintetizáló modell

A modellben a folytonos vonal Hughes (1993) és Bailey (1996) alapján a teszt közvetlen és közvetett hatását mutatja a folyamatban részt vevő személyek tesztrel kapcsolatos nézeteire és attitűdjeire. Ezek meghatározhatják a tanítás tartalmát, módszereit és értékelési gyakorlatát, a tanulás tartalmát, minőségét és eredményét (Hughes, 1993), amiről a teszteredmények alapján kapunk információkat. A tanítási folyamatra vonatkozó empirikus kutatások eredményei alapján (1.3.3. fejezet) a tesztek főként a tanítás tartalmára vannak hatással (Cheng, 1997; Shohamy, Donitsa-Schmit és Ferman, 1996; Watanabe, 1996b). A modellben lévő visszajelentési kör azt mutatja, hogy a tanárok

gyakran alkalmazzák a tesztek hatására létrejövő új tananyagokat, tankönyveket és feladatgyűjteményeket, amelyek befolyásolják a tanítás tartalmát. A teszteknek ritkán van azonnali hatása a tanítási módszerekre, így ezt csak hosszabb távon lehet vizsgálni (Cheng, 2005). A teszt befolyásolhatja a tanárok mérési-értékelési gyakorlatát (Wall és Alderson, 1993), amely ritkán vizsgált terület.

A modell öt visszacsatolási kört tartalmaz, amelyet szaggatott vonallal jelöltünk. A visszacsatolási körök hármasként funkcionálnak: egyrészt feloldják a teszt normatív, egyoldalú irányítását, másrészt elősegítik az innovációs folyamatok elindulását, harmadrészt lehetővé teszik a tesztelés folyamatának empirikus vizsgálatát. A tesztelésből kiinduló szaggatott vonal a tesztfejlesztési folyamatra, az igények és célok változásaira irányuló visszacsatolást illusztrálja, amely főleg a teszt eredményességéről szolgáltat információkat a tesztfejlesztőknek és a közoktatásban résztvevőknek. (1) Az eredményeknek a tesztfejlesztési folyamatba történő visszacsatolása főként a teszt mérési-értékelési eljárásainak vizsgálatát és a követelményszintek ellenőrzését szolgálják; (2) az igényekre és a célokra irányuló visszacsatolás elsősorban a teszt céljának, tartalmának és típusának bevalóságáról szolgáltat információt. A teljesítményekben megjelenő jelentősebb különbségek például új igényeket generálhatnak a teszt differenciáltabbá tételében, amely újabb tesztfejlesztési folyamatot indíthat el. A teszt tanítási-tanulási folyamatra gyakorolt hatásának vizsgálatában különböző felmérésekkel a tanítási-tanulási folyamatban résztvevők tesztrel kapcsolatos nézeteiről és attitűdjeiről is kaphatunk információkat. (3) A tesztelés folyamatának empirikus vizsgálatával tanári viselkedésmintázatok tárhatók fel, amelyek a tesztelés minőségével kapcsolatos nézetekre, a teszt és a tanítási folyamat közti összefüggésekre és a változtatásokhoz kötődő attitűdökre irányulhatnak. (4) A tanulók tesztrel kapcsolatos véleményeinek feltárásával lehetőség nyílik annak ellenőrzésére, hogy a teszt a tanulói célcsoport igényeivel összhangban van-e. Ezek mellett pedig arról is szerezhető információ, hogy a létrejött új teszt megoldására való készülés során a diákok a korábbi tesztnél többet vagy jobban tanulnak-e. E tényező vizsgálata a nemzetközi szakirodalomban kevésbé jelenik meg (Wall, 2005), kevés empirikus vizsgálat áll rendelkezésre (például Cheng, 1998; Ferman, 2004). (5) A teszt hatására létrejövő új tananyagok és tantervek elemzésével megvizsgálható, hogy ezek mennyire vannak összhangban a tesztkészítők szándékával. A visszacsatolási körökből származó információk befolyásolhatják a tesztfejlesztési folyamatot, elősegítik az eredmények megfelelő értelmezését és ezek alapján a teszt esetleges módosításával innováció érhető el.

A kutatások a vizsgáknak csak mikroszintű (azaz a tanítási-tanulási folyamatra irányuló) hatását tanulmányozták, melyekben ötvözik a kvalitatív és kvantitatív módszereket. A visszahatás jelenségének komplexitása miatt kiemelten fontos a mérőeszközök validitásának biztosítása. Ehhez Alderson és Banerjee (2001) tanácsokat fogalmaz meg, amelyeket a vizsgálatunk megtervezése során kiindulási alapnak tekintettünk.

A fejezet további részében vizsgálatunk tárgyán, az idegen nyelvi érettségi vizsga példáján keresztül határozzuk meg a modell elemeinek jellemzőit, az elemek közötti kapcsolatot, majd bemutatjuk a vizsgafejlesztési folyamatot és a vizsga alapján elindult és eddig megismert hatásmechanizmusokat. A modell segítségével azonosítjuk a vizsga működésében megjelenő problémákat, ellentmondásokat, a vizsga közoktatásra gyakorolt hatása szempontjából az eddig feltáratlan területeket. Ezután kiválasztjuk az empirikus vizsgálat szempontjából releváns területeket, indokoljuk az elemzés szükségességét. Célunk ezzel az empirikus vizsgálat elméleti megalapozása.

3.1.2. A vizsgálat elméleti modellje az idegen nyelvi érettségi példáján

Igények és célok változásai

Az 1980-as években egyre inkább előtérbe kerültek a magyar közoktatásban megszerzhető nyelvtudással szemben támasztott minőségi elvárások. A kilencvenes évek elejétől megjelent az a társadalmi igény, hogy a diákok közvetlenül hasznosítható nyelvtudást szerezzenek. Az ezredfordulót követő években az európai nyelvpolitika célkitűzéseihez és a nemzetközi munkaerőpiaci kihívásokhoz való igazodás elengedhetetlenné teszi Magyarországra számára, hogy minden állampolgár az anyanyelvén kívül legalább két idegen nyelven tudjon beszélni (Nikolov, 2007). Az idegen nyelvek tudásának presztízse az elmúlt évtizedekben nagymértékben felértékelődött.

A társadalom és a civil szféra igényeinek megfelelően ezek a prioritások megjelentek a *Nemzeti alaptanterv* 1995-ös változatában, megfogalmazódtak a *NAT* 2003-as módosításában és megerősödtek a 2007-es módosításban, amely szerint az iskolai nyelvtanulás alapvető célja a mindennapi életben használható nyelvtudás elsajátítása. A *Nemzeti alaptanterv* (2003, 2007) a *Közös Európai Referenciakeret* (2002) alapján definiálja a közoktatásban elérendő szintet (e folyamatról lásd részletesebben: 2.1. fejezet).

A vizsgafejlesztés folyamata

E változásoknak köszönhetően az idegen nyelvi mérés és értékelés területén jelentős fejlődésként értelmezhető az új idegen nyelvi érettségi vizsga, amelynek bevezetését a *Nemzeti alaptantervben* (2003, 2007) megjelenő kommunikatív nyelvi kompetenciák és szintjeinek mérése tette szükségessé. A vizsgareform során (lásd Alderson, Nagy és Öveges, 2000; Einhorn, 2004) az idegen nyelvi érettségit úgy alakították ki, hogy mérési és értékelési eljárásaiban kövesse a kommunikatív tesztelés alapelveit, valamint illeszkedjék a Referenciakeret szintleírásaihoz.

A kommunikatív tesztelés alapelveinek követése abban érvényesül, hogy a vizsga a négy nyelvi alapkészséget, valamint a nyelvhelyességet méri, feladatai autentikus szövegekre épülnek, feladattipológiailag kötetlenek és a feladatsorok teljesen egynyelvűek (Einhorn és Major, 2006). A vizsgakövetelményekben egyértelműen definiálják azokat a kontextusokat és kommunikációs helyzeteket, amelyben az érettségizőknek nyelvileg cselekedniük kell. Az új érettségi a kommunikatív nyelvhasználat mérése miatt alapvetően újfajta értékelési kultúrát jelent a korábbi vizsgához képest, amely elvárja a vizsgáztató tanároktól, hogy ismerjék és tudják alkalmazni a kommunikatív tesztelés eljárásait (az érettségi jellemzőit a 2.3. fejezetben összegeztük).

A *Nemzeti alaptantervben* (2003, 2007) deklarált céloknak megfelelően az idegen nyelvi érettségi vizsga két szintjével (közép- és emelt szint) a Referenciakeret három szintjét, az A2–B1 és a B2 szintet méri. A középszintű vizsgának széles képességterületet kell átfogni, amely a nyelvtudás mérését megnehezíti. A vizsgázók nyelvi készségeinek minél teljesebb mérése érdekében minden feladatsor a lépcsőzetesség elve alapján épül fel (Einhorn, 2007b). Komoly kihívást jelent a különböző vizsgaidőszakokban keletkező feladatsorok nehézségi szintjének, állandóságának biztosítása. Ezt nehezíti az a tény, hogy jelenleg nincs mód a feladatok előzetes kipróbálására, azok nehézségi szintjét csak utólag lehet(ne) meghatározni. Az emelt szint a Referenciakeret B2 szintjét méri, így mérésmetodikai szempontból megfelelőbb. A hazai nyelvtudásvizsgálatok eredményei azonban azt mutatják, hogy a diákok a C1 szintet az iskolarendszeren belül a nyelvi tagozatokon, a két tanítási nyelvű iskolákban el tudják érni, így a jelenlegi vizsgaszint nem jelent számukra kihívást (Vámos, 2007).

A vizsga jellemzői

A kommunikatív nyelvhasználatot a Referenciakeret alapján mérő vizsga működtetésekor a nyelvvizsgákhoz hasonlóan felmerül és fontos a külső és belső validitás biztosítása. Az idegen nyelvi érettségi azonban több ponton eltér a nyelvtudást mérő vizsgáktól, hiszen elsődleges szempontja a magyar közoktatási rendszerbe való illeszkedése, így funkcióját tekintve alapvetően a középfokú oktatást lezáró vizsga. Emellett még nyelvtudásszintet mérő vizsga is, hiszen a továbbtanuláshoz a szelekció e vizsga eredménye alapján történik, továbbá az emelt szintű vizsga nyelvvizsgával egyenértékű lehet (2.3.3. fejezet). Az idegen nyelvi érettségi így e két vizsgatípus keveréke és e kettőség megjelenési formája, hogy az írásbeli vizsga mindkét szintje és az emelt szintű szóbeli vizsga központi feladatsorokból áll, az érettségizők teljesítményét központi útmutató alapján javítják és értékelik, a középszintű szóbeli feladatai viszont iskolai szinten készülnek. Az emelt szintű vizsga külső vizsga, a középszintű vizsgát pedig külsőleg kontrollált belső vizsgának tekinthetjük.

A vizsga működtetésének sok eleme (például a feladatok készítése, a vizsga lebonyolítása és értékelése) jelenleg kevés minőségbiztosítási elemet tartalmaz (ennek jellemzőit a 2.3.2. fejezetben és következményeit a 2.3.4. fejezetben összegeztük). A vizsgaminőség növelése szempontjából jelentős előrelépést jelentett, hogy a nemzetközi feladatfejlesztési szabványoknak megfelelően (feladatok írása, lektorálása, kipróbálása, véglegesítése) készítettek körülbelül 300 mérőfeladatot 2005 és 2007 között a Nemzeti Fejlesztési Terv keretében az akkori Országos Közoktatási Intézetben. Ugyanakkor a feladatok csak az elméleti paramétereik mentén kapcsolhatók össze a létrejött feladatállományban, az empirikus paraméterek meghatározására [például a feladatok nehézsége, illeszkedésük a *KER* (2002) szintjeihez] nem volt lehetőség. Átgondolatlan és szabályozatlan továbbá ezeknek a feladatoknak a felhasználása a további vizsgaidőszakokban (*Einhorn*, 2008c). A legfőbb probléma, hogy jelenleg nem áll rendelkezésre olyan feladatbank, amely a tanított idegen nyelvekből előzetesen kipróbált és az egyes szintekhez kalibrált feladatokat tartalmaz. Fontos kutatási feladat tehát a nyelvi érettségi vizsgák két szintjének egymáshoz és a Referenciakerethez való illesztése alapján feladatbank létrehozása és a vizsga minőségének fejlesztése.

A kétszintű (nyelvi) érettségi sok problémája, hiányossága mellett erős hatást gyakorol a közoktatásra, hiszen a vizsgára való felkészülés központi szerepet kap a szakközépiskolák és a gimnáziumok napi munkájában. Az idegen nyelvi érettségi hatása abból ered, hogy a kommunikatív nyelvhasználatot méri, így igazodik a nyelvtanulók valós igényeihez. Emellett magas presztízzsel rendelkezik, mert záróvizsgaként, felvételi vizsgaként és nyelvvizsgaként is működik, így az érettségizőknek fontos, hogy letegyék a vizsgát. Ezek alapján megállapítható, hogy a vizsga hatásmechanizmusainak elindulása szempontjából minden fontos tényező megjelenik.

A vizsga hatása a tanítási-tanulási folyamatban résztvevőkre, a folyamatra és eredményére

A kétszintű idegen nyelvi érettségi vizsga mérési-értékelési módszerei jelentősen megváltoztak a korábbi vizsgához képest (1. melléklet), ugyanakkor az új vizsgamodell és a vizsgakövetelmények elfogadtatása a tanárok körében nem jelentett nagy problémát a vizsgafejlesztés folyamatában. Ennek oka, hogy a korábbi nyelvi érettségi tartalmában és módszereiben elavult és értéktelen volt, továbbá a korábbi vizsga és a nyelvoktatásban megjelenő új igények és elvárások között szakadék húzódott. Ennek a vizsgának negatív hatása volt a tanítási-tanulási folyamatra, mivel a felkészítés során a tanárok gyakran alkalmaztak hagyományosabb nyelvtanozó-fordíttató módszerű nyelvtanítást (*Nikolov*, 1999). Ugyanakkor a kommunikatív nyelvoktatás már a nyolcvanas években

megjelent hazánkban is, sok nyelvtanár már a kétszintű érettségi bevezetése előtt használt kommunikatív nyelvkönyveket és módszereket, a vizsga csak követte ezt a gyakorlatot. Így a vizsga a készségek mérésével megerősítheti a készségfejlesztésre irányuló kommunikatív nyelvoktatást. Míg az új vizsga koncepciója kevésbé váltott ki szakmai ellenállást a tanárok körében, addig a vizsga értékelési rendszere már jelentős vitákat generált a vizsgareform során (Einhorn, 2007b). Ennek oka, hogy míg a korábbi vizsgán egy nagyon kidolgozatlan, rövid és több nyelvre is érvényes értékelési útmutatót kellett használni, addig az új vizsgán részletes, előíró-jellegű értékelési rendszer a jellemző. Ez jelentős változást jelent a tanárok értékelési attitűdjeiben, hiszen a tanár nem térhet el a megadott javítókulcstól, a nyelvi hibákat funkció szerint kell megítélnie az egyes vizsgarészekben és a produktív készségeknél értékelési skálákat kell használnia. A vizsga egységes értékelési kultúrája feltételezhetően a mindennapi osztálytermi mérések gyakorlatává válhat.

A vizsga jellemzői közül a vizsgaszintek kialakítása és a vizsga típusai gyakorolnak hatást a nyelvtanulási folyamatra. A vizsgaszint húzóerőt jelenthet és növelheti a motivációt, ezért a tanulók reális nyelvtudás-szintjéhez kell igazítani a vizsgát. A vizsgatípusok közül az előrehozott vizsga lehetősége nagyon problémás és feltehetően negatívan hat a nyelvtanulási folyamatra. A tanulók a közép- és emelt szint választásával hatnak vissza elsősorban a vizsgafejlesztési folyamatra. Sok tényező (például a továbbtanulási szándék, a tervezett munkakör betöltése, a pontszámítások módja, a szülői és tanári vélemények és hatások) alapján döntenek a vizsgázók arról, hogy melyik vizsgaszintet választják. A vizsga egészére, így az idegen nyelvi érettségi vizsgákra is jellemző, hogy az emelt szintet kevesen választják, a kétciklusú bolognai felsőoktatás bevezetése is az emelt szint funkciójának elértéktelenedését mutatja, amely újabb vizsgafejlesztési igényeket generálhat (Csapó, 2008).

A vizsga közvetlen és erős hatása a tananyagszerzőkre már egyértelműen kimutatható. Tömegesen jelentek meg a tankönyvkínálatban az érettségihez hasonló mérőfeladatokat tartalmazó feladatgyűjtemények. A hazai nyelvoktatási gyakorlatban nagy hagyománya van a mérőfeladatok felhasználásának a vizsgatréning során. Ezek a feladatok azonban nem alkalmasak a nyelvhasználati készségek fejlesztésre. Ha a fejlesztés céljából akarják a nyelvtanárok a feladatgyűjteményeket felhasználni, akkor ezek negatív hatást gyakorolhatnak a tanítási-tanulási folyamatra.

A vizsga működésének és hatásának empirikus vizsgálata

A vizsgálat elméleti modellje öt visszacsatolási kört tartalmaz (11. ábra), amellyel a vizsga működéséről és tanítási-tanulási folyamatra gyakorolt hatásáról szerezhetünk információkat. Az empirikus vizsgálatban a modellnek megfelelően ezek közül három vizsgálatát tűzzük ki célul: a vizsgaeredmények visszacsatolását (1) a vizsgafejlesztési folyamatra és (2) az igények és célok változásaira, (3) a tanárok tesztel kapcsolatos nézeteinek és attitűdjeinek visszacsatolását a fejlesztési folyamatba. Így két irányból vizsgáljuk az idegen nyelvi érettségi vizsgákat: a vizsgaeredmények és a nyelvtanárok visszajelzései alapján. Ugyanakkor a vizsgálatunkban közvetlenül nem elemezzük a 4. visszacsatolási kört, vagyis a tanulók vizsgáról alkotott véleményét és a felkészülési folyamatot, a közép- és emelt szintű vizsga választását meghatározó tényezőket. Kutatásunkban arra sem vállalkozunk, hogy a modellben lévő 5. visszacsatolási kört, azaz a vizsga hatására létrejövő új tankönyveket és tanterveket elemezzük és megvizsgáljuk, hogy ezek mennyiben vannak összhangban a kommunikatív nyelvoktatás alapelveivel és a vizsgafejlesztők szándékaival, valamint, hogy ezeket a nyelvtanárok hogyan alkalmazzák a tanítási-tanulási folyamatban. Ennek oka, hogy a nemzetközi empirikus vizsgálatok alapján (1.3.3. fejezet) feltételezzük, hogy ezek csak hosszabb távon vizsgálha-

tók. A nyelvi érettségi működésének és hatásmechanizmusának minden területre kiterjedő vizsgálatát nem tűztük ki célul, hanem a vizsgafejlesztés és -minőség javítása érdekében, valamint a modell igazolása céljából olyan területeket választottunk ki, amelyek elemzésével átfogó visszajelzést adhatunk a vizsgafejlesztés számára és megfelelő alapot szolgáltatathatunk a további kutatásokhoz.

Az értekezés empirikus fejezeteiben a vizsgaeredmények felőli a vizsgafejlesztési folyamatra és a célok, igények változásaira irányuló visszacsatolási kör vizsgálata az idegen nyelvi érettségi esetében azért releváns, mert a vizsgának jelentős szerepe lehet a közoktatásban megszerezhető nyelvtudás-szint monitorozásában, változásának nyomon követésében, hiszen jelenleg ez megoldatlan (*Nikolov, 2007*). A magyar érettségizők nyelvi készségeinek fejlettségéről, az azt meghatározó tényezőkről, továbbá az idegen nyelvi érettségi feladatsorok nehézségi szintjéről és állandóságáról, az értékelési eljárások minőségéről jelenleg kevés adat áll rendelkezésre. Ezen a területen számos fejlesztő kutatásra van szükség. Egyrészt szükséges az érettségizők nyelvtudás-szintjének vizsgálata annak érdekében, hogy ahhoz tudják igazítani a vizsgát minden tanított idegen nyelvből. Másrészt vizsgálni kell, hogy a vizsgafeladatok szintje hogyan viszonyul a követelményekben rögzített készségszintekhez (*Einhorn, 2007b*). Harmadrészt vizsgálni kell, hogy a meglévő feladatok szintjeinek meghatározásával hogyan építhető ki feladatbank, amely a további vizsgafeladatok fejlesztését is segítené. Negyedrészt pedig szükséges megvizsgálni a produktív készségek mérésénél és értékelésénél alkalmazott újfajta eljárásokat. Ezek közé tartozik a performancia feladatok, az értékelési szempontok és skálák működésének jellemzői.

A tanárok visszajelzése felőli és a vizsgafejlesztési folyamatra irányuló visszacsatolási kör vizsgálatát azért tartjuk szükségesnek, mert az új szemléletű vizsga bevezetésétől a nyelvoktatás tartalmi és módszertani modernizációja várható (*Einhorn, 2007b; Eckes, Ellis, Kalnberzina, Pizorn, Springer, Szollás, és Tsagari, 2005*). Ugyanakkor jelenleg nem áll rendelkezésre olyan kutatási eredmény, amellyel az érettségi vizsga tanítási-tanulási folyamatra gyakorolt hatását értelmezhetnénk. A kutatásunk annak feltárására törekszik, hogy a kétszintű idegen nyelvi érettségi a tanárok véleménye szerint milyen jellegű hatást gyakorol a közoktatásra, hogyan szolgálja a nyelvtanárok tanításában bekövetkezett szemléleti és tartalmi modernizációt és milyen mértékben játszik szerepet abban, hogy a kommunikatív tesztelés ne csak a minősítő értékelésben, hanem a mindennapi tanítási-tanulási folyamatban is elterjedjen. A tanárok visszajelzése alapján a modellben az alábbi területeket vizsgáljuk: (1) a tanárok vizsgával kapcsolatos nézetei, attitűdjei és a tanárok képzettsége; (2) ezek hatása a nyelvtanítási folyamat tartalmára, módszereire, a kommunikatív tesztelés mérési-értékelési eljárásainak alkalmazására; (3) a vizsga alapján létrejövő új tananyagok hatása a nyelvoktatásra.

3.2. Az empirikus vizsgálat céljai, kérdései és hipotézisei

A korábbi érettségi vizsgákon elért eredmények alapján három szempontból vizsgáljuk az érettségit. Egyrészt elemezzük a jelenlegi vizsga működését, a készségenkénti teljesítmények változását és az azokban megmutatkozó különbségeket. Másrészt a receptív készséget mérő vizsgarészek empirikus vizsgálata alapján elsősorban az érettségi belső, másodsorban pedig a külső validitását elemezzük. Harmadrészt az íráskészség vizsgarészben alkalmazott értékelési eljárások jellemzőit vizsgáljuk. Ezek mellett pedig negyedik területként az érettségit a tanárok visszajelzései alapján is elemezzük. A fejezet további részében először indokoljuk e négy terület vizsgálatának szükségességét és

megfogalmazzuk a kutatás céljait, majd ezekhez kapcsolódva a kutatás kérdéseit, végül a szakirodalmi szintézis eredményei alapján fogalmazzuk meg a hipotéziseket.

I. A jelenlegi idegen nyelvi érettségi vizsgákon elért teljesítmények alapján célunk megvizsgálni, hogyan alakul és változik az érettségizők teljesítménye az egyes nyelvi készségeket mérő vizsgarészekben, milyen különbségek mutathatók ki a vizsgázók nyelvi készségének fejlettségében iskola- és képzéstípusok és nemek szerint. Bár a hazai nyelvpedagógiai vizsgálatok (2.2. fejezet) alapján a nyelvtanítás eredményességében nem a képzés- és iskolatípusok közötti különbségek jelentik a fő töréspontot, az érettségi vizsga működése szempontjából mégis lényeges szempont, hogyan szolgálja a vizsga a két iskolatípus közötti különbségek konzerválódását, mérséklését vagy megnövekedését. A nemek közötti különbségek vizsgálata azért indokolt, mert ez az eredményességet meghatározó egyéni különbségek egyik tényezője. A teljesítménykülönbségeket több szinten elemezzük. A két leggyakrabban tanított idegen nyelv összehasonlítása alapján azt is megvizsgáljuk, milyen eltérések és hasonlóságok vannak az angol és a német nyelv, valamint az érettségizők készségének fejlettségében feltárt különbségek jellemzőinél a 2007., 2008. és 2009. években. Ezt a vizsgálatot a vizsgaeredményekből kiinduló és az igényekre és célokra irányuló visszacsatolás elemzésében tartjuk szükségesnek. Megvizsgáljuk, hogy a kétszintű idegen nyelvi érettségi hogyan válik a közoktatás részévé, milyen tendenciák figyelhetők meg ebben a folyamatban és hogyan kezeli a vizsga a tanulók nyelvtudásában meglévő különbségeket. Ezzel célunk a további kutatási problémák felvetése és elemzési területek kijelölése.

II. A receptív készségeket mérő feladatsoroknak a nyelvtudás-modellhez és a tanulók nyelvtudás-szintjéhez való illeszkedése alapján célunk a következő területek vizsgálata: A feladatsorok minősége szempontjából meghatározzuk a feladatok, itemek nehézségi szintjét, és azt vizsgáljuk, hogy a feladatkészítők által megállapított nehézségi sorrend mennyire esik egybe az empirikus nehézség növekedésével. A feladatok nehézségével kapcsolatosan egyrészt azt elemezzük, hol van ugrás az egyes itemek nehézségi indexei között, másrészt azt, hogy a feladatok megfelelő nehézségűek-e a vizsgázók számára, vagyis az itemnehézségi indexek mennyire fedik le a vizsgázók nyelvi készségszintjei által meghatározott intervallumot. A feladatfejlesztés szempontjából arra keresünk választ, hogy melyik itemet milyen okból lehet elhagyni a vizsgált tesztekben. A vizsgálat célja a feladatsorok, a feladatok és az itemek működésének teljesebb megismerése, a nehézségi szintek meghatározása, az abból levonható következtetések megfogalmazása. Ezek mellett pedig annak lehetőségét is megvizsgáljuk, hogyan lehet kiépíteni egy nyelvi feladatbankot és milyen problémák, kérdések merülnek fel a folyamatban. Ezt a vizsgálatot elsősorban a vizsgafejlesztési folyamatra irányuló visszacsatolás miatt a vizsgaminőség javítása érdekében tartjuk indokoltnak.

III. Az íráskészség vizsgarész külön elemzésének alapvető célja, hogy definiáljunk olyan mutatókat, amelyekkel az íráskészség mérési-értékelési eljárásainak minősége jellemezhető. Az elemzés alapjául szolgáló *Engelhard* (1992) modelljének megfelelően a vizsgázók íráskészség-szintjét, az értékelési szempontok jellemezőit és az értékelési skálák szerkezetét vizsgáljuk. Az értékelési szempontoknál meghatározzuk az azok közötti összefüggéseket, nehézségi szintet. Az érettségizők íráskészségével kapcsolatban azt vizsgáljuk, hogy a feladatok és értékelési eljárások milyen mértékben képesek a vizsgázók teljesítményét különböző szintekbe különíteni. Az értékelési skálák fejlesztése szempontjából azt elemezzük, hogy a skálák mennyire megfelelően működnek, azaz, mennyire jól rendezettek az egyes értékelési szempontokhoz tartozó pontszámok és az értékelők minden szintleírást azonos mértékben használnak-e. Az érettségizők íráskészség-szintjének, az értékelési szempontok nehézségének és az egyes analitikus skálák szerkezetének vizsgálatával célunk olyan minőségi mutatókat jellemező paraméterek

definiálása, amelyek mentén a különböző évekből származó eredmények összehasonlíthatók és változásai nyomon követhetők. Így elősegíthetjük azt a folyamatot, amelyben az elért eredmények alapján adható visszajelzés a vizsgafejlesztés számára.

IV. A tanárok vizsgával kapcsolatos visszajelzései során a szakirodalom elemzése és az érettségi jellemzői alapján kutatásunk célja, annak vizsgálata, hogy a tanárok vizsgával kapcsolatos nézetei, attitűdjei milyen összefüggésben vannak a vizsga működéséről szerzett tapasztalataikkal és ezek mennyiben határozzák meg a nyelvoktatás tartalmát, és módszereit. Mivel a tanárok a vizsgadolgozatokat javítóként és értékelőként, valamint feladatíróként (a középszintű szóbeli vizsgán) vesznek részt a vizsga lebonyolításában, így valószínűsíthető a vizsga közvetlen hatása az osztályszintű mérési-értékelési gyakorlatra. Célunk megvizsgálni, hogy a kommunikatív mérési-értékelési eljárások megjelennek-e a tanítási-tanulási folyamatban, a kommunikatív tesztelés kritériumait a tanárok mennyire tartják fontosnak. A nyelvtanulási folyamat eredményességében számos, a vizsgától független tényező is szerepet játszhat (például a tanulók tanuláshoz való hozzáállása, a diákok motivációjáról és nyelvtudásáról kialakított tanári nézetek). Ugyanakkor a vizsga hatással lehet arra, hogy a tanárok tesztmegoldó stratégiákat tanítanak-e korábbi vizsgafeladatok és az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatok segítségével vagy inkább a nyelvi készségek fejlesztésére helyezik a hangsúlyt. A tanárok vizsgával kapcsolatos nézeteiről és attitűdjeiről szereshető információk befolyásolhatják a vizsgafejlesztési folyamatot, elősegíthetik az eredmények megfelelő értelmezését.

Mindezek figyelembe vételével az értekezés empirikus fejezeteiben az alábbi kutatási kérdésekre adunk választ:

I. A vizsgán elért készségenkénti teljesítmények változásaihoz és meghatározó tényezőihez kapcsolódó kutatási kérdések:

1. Hogyan változnak az idegen nyelvi érettségien elért teljesítmények az egyes készségeknél?
2. Milyen különbségek találhatóak az érettségizők nyelvi készségeinek fejlettségében iskolatípusok, képzési típusok és nemek szerint?
3. Milyen eltérések és hasonlóságok vannak az angol és a német nyelv esetében, valamint az érettségizők készségének fejlettségében, és a feltárt különbségek jellemzőinél az egyes érettségi vizsgaévekben?

II. A receptív készséget mérő vizsgarészek elemzéséhez kapcsolódó kutatási kérdések:

4. Mennyire esik egybe a feladatkészítők által megállapított nehézségi sorrend az empirikus nehézség növekedésével?
5. Mennyire fedik le a feladatok itemnehézségi indexei a vizsgázók nyelvi készségszintjei által meghatározott intervallumot?
6. Hogyan építhető ki egy, a Referenciakeretre épülő nyelvi feladatbank az érettségi feladatai alapján?

III. Az íráskészséget mérő vizsgarészek elemzéséhez kapcsolódó kutatási kérdések:

7. Milyen nehézségi szintekbe sorolhatók az egyes értékelési skálapontok?
8. Mennyire fedik le az értékelési skálapontok a vizsgázók íráskészség-szintjei által meghatározott intervallumot?
9. Mennyire jól rendezettek az egyes értékelési szempontokhoz tartozó pontszámok az értékelési skálákban?

IV. A tanárok visszajelzésének elemzéséhez kapcsolódó kutatási kérdések:

10. Milyen nézetekkel és attitűdökkel rendelkeznek a nyelvtanárok a vizsgával kapcsolatban? Milyen vizsgával kapcsolatos tapasztalatokkal van ez összefüggésben?
11. Milyen tényezők határozzák meg a tanárok nyelvoktatási folyamatát és ebben milyen szerepe van a vizsgakövetelményeknek és a vizsgafeladatoknak?
12. Meghatározzák-e a nyelvoktatás tartalmát, módszereit, mérési-értékelési gyakorlatát a tanárok vizsgával kapcsolatos nézetei és attitűdjei? Ha igen, mennyiben?
13. Mennyiben segíti elő az idegen nyelvi érettségi a kommunikatív tesztelés mérési és értékelési eljárásainak elterjedését a mindennapi tanítási-tanulási folyamatban? A kommunikatív tesztelés mely kritériumait tartják a tanárok fontosnak, melyeket kevésbé a mérőfeladatok önálló elkészítésében?
14. Milyen nézetekkel rendelkeznek a tanárok a vizsga hatására keletkező érettségi-re felkészítő tananyagok, korábbi vizsgafeladatok felhasználásának céljairól?

A feltett kutatási kérdésekre az alábbi hipotéziseket fogalmazzuk meg a szakirodalmi háttér alapján és célunk ezeknek az empirikus igazolása vagy megcáfolása.

- I. A 2000-es évektől elinduló hazai, reprezentatív mintán végzett nyelvtudást mérő kutatások (2.2. fejezet) és a nyelvi érettségig eddig megismert vizsgázói eredmények (2.3. fejezet) alapján feltételezzük, hogy
 1. a gimnáziumi tanulók teljesítménye szignifikánsan jobb a szakközépiskolában tanulóknál mindkét nyelv esetében;
 2. a nappali tagozaton tanulók szignifikánsan magasabb teljesítményt mutatnak a felnőttoktatásban tanulóknál mindkét nyelv esetében;
 3. a lányok szignifikánsabban jobban teljesítenek a fiúknál és ez a különbség a produktív készségeknél jelentősebb, mint a receptív készségeknél.
- II. A nyelvi érettségi vizsga működésének jellemzői, mérési-értékelési eljárásai alapján (2.3. fejezet) feltételezzük, hogy
 4. a receptív készséget mérő vizsgafeladatok itemnehézségi indexei nem fedik le teljesen a vizsgázók nyelvi készségszintjei által meghatározott intervallumot;
 5. a magasabb készségtartományból hiányoznak itemek.
- III. Az íráskészség vizsgarészben alkalmazott mérési-értékelési eljárások (2.3. fejezet) alapján feltételezzük, hogy
 6. az értékelési skálapontok nem fedik le teljesen a vizsgázók íráskészség-szintjei által meghatározott intervallumot;
 7. az egyes értékelési skálák működése eltér szempontonként és a skálákban definiált pontszámok alapján.
- IV. A teszthatást vizsgáló nemzetközi empirikus vizsgálatok tanulságai (1.3. fejezet) és a vizsga jellemzői (2.3. fejezet) alapján feltételezzük, hogy
 8. a tanárok vizsgával kapcsolatos nézetei és attitűdjei összefüggésben vannak a vizsgázatói tapasztalatukkal, valamint, hogy mely vizsgaszintre készítik fel a tanulókat és vettek-e részt az érettségire vonatkozó vizsgázatói képzésen;
 9. a tanárok ismerik és elismerik a vizsga jelentőségét, a korábbi vizsgánál megbízhatóbbnak és érvényesebbnek tartják;
 10. a tanárok vizsgával kapcsolatos nézetei és attitűdjei meghatározzák a nyelvoktatás tartalmát, de attól kisebb hatást fejtenek ki a módszerekre;
 11. a vizsga értékelési szemlélete befolyásolja a tanárok mérési-értékelési eljárásokkal kapcsolatos nézeteit.

3.3. A vizsgálat módszerei

A vizsgán elért készségenkénti teljesítmények változásaihoz és a nyelvi készségek szintjét meghatározó tényezőkhöz, a receptív és a produktív készségeket mérő vizsgarészek elemzéséhez kapcsolódó kutatási kérdések vizsgálatában a nyelvi érettségiről rendelkezésre álló adatbázisokat használtuk fel és azokon végeztünk másodelemzéseket. A tanárok visszajelzésének elemzéséhez kapcsolódó kutatási kérdések vizsgálatához önállóan dolgoztunk ki kérdőívet, szerveztük a mintát, bonyolítottuk le a vizsgálatot és elemeztük az adatokat. Ez az oka annak, hogy a vizsgaeredmények és a nyelvtanárok visszajelzései alapján történő empirikus vizsgálat módszereit bemutató fejezet két része részletezettségében eltér egymástól.

A vizsgaeredmények elemzésének módszereinél a felhasználható adatbázisok alkalmazásának előnyeit és hátrányait mutatjuk be, majd indokoljuk az adatbázisok kiválasztását. Ezt követően a feltett kutatási kérdéseknek megfelelően az adatelemzés módszereit ismertetjük. A tanárok visszajelzésének elemzési módszereinél bemutatjuk az adatgyűjtő eszköz választásának indoklását, a minta szervezésének szempontjait, a minta jellemzőit, a kérdőív fejlesztési folyamatát, tartalmi-tematikus egységeit és tesztelési mutatóit, valamint az adatfelvétel lebonyolítását és az adatelemzés módszereit.

3.3.1. A vizsgaeredmények elemzésének módszerei

Az adatbázisok jellemzői

A vizsga minőségéről, a különböző vizsgaidőszakban elért teljesítmények változásáról, a feladatok és értékelési eljárások minőségéről két forrás áll rendelkezésre: (1) az Oktatási Hivatal (OH) minden vizsgázónak az egyes vizsgaidőszakban elért pontszámait tartalmazó publikus adatbázisai; (2) az Oktatáskutató és Fejlesztő Intézet (OFI) által a Nemzeti Fejlesztési Terv keretében felvett, az egyes készségeket mérő feladatsorok részpontszámait tartalmazó adatbázisok.

Az OH nyilvánosan közzétett adatbázisai minden érettségizőről minden vizsgaidőszakban tartalmazzák az eredményeket régió, iskolatípus, településtípus, nemek, képzési típusok és vizsgatípusok szerinti bontásban. A nyelvi érettségiknél rögzítik az egyes vizsgarészekben elért vizsgapontszámot és az összteljesítményt. Ez alapján megállapítható az érettségizők készségének fejlettségi szintje, az azokban megjelenő különbségek mértéke, alkalmas a teljesítménytrendek dokumentálására, továbbá a két leginkább tanított idegen nyelv összehasonlító elemzésére. Ugyanakkor az adatbázisok alkalmatlanok a trendek okainak felderítésére, mert nem teszik lehetővé a feladat- és itemszintű elemzéseket, nem adnak információt a vizsgázók nyelvi készségszintjét meghatározó háttérváltozókról (például a kognitív és affektív tulajdonságokról, az iskolán belüli és iskolán kívüli tényezőkről), nem lehetséges továbbá elemezni az egyes feladatsorok és értékelési eljárások minőségét, a különböző vizsgaidőszakokban keletkező feladatoknak a vizsgakövetelményekhez és a Referenciakerethez való illeszkedését. Ezek alapján nem lehetséges utólagosan, kutatói szinten a vizsga minőségét részletesen elemezni.

Az OFI által felvett kisebb adatbázisok a fent említett hiányosságok kiküszöbölése érdekében lehetővé teszik a feladat- és itemszintű elemzéseket. Az idegen nyelvi érettséginél a közép- és emelt szintű vizsgafeladatokhoz tartozó adatbázisok az olvasott és hallott szöveg értése, a nyelvhelyesség feladatsorok itemszintű adatait és az íráskészség vizsgarészen elért részpontszámokat tartalmazzák. Ezek alapján információ szerezhető a feladatsorok minőségéről, a feladatok nehézségéről, az értékelési eljárások alkalmazásáról, a fejlesztés lehetőségeiről. Az adatbázisok hátránya viszont, hogy kevés, körülbelül 1000 vizsgázó adatait tartalmazzák, amelyekből nem lehet automatikusan következtetni

a vizsgázói populáció teljesítményére. Ezen kívül az adatbázisokat csak külön elemezhetjük, hiszen nincs olyan közös item vagy személy, amely összekötné azokat. További probléma, hogy ilyen adatbázisokat csak a 2005. és 2006. évi érettségikhez hoztak létre. E két vizsga azonban még nem rutinszerűen zajlott, így kevésbé modellálja jól és pontosan a további vizsgákat, valamint e két évben az érettségi eredményeinek részletes elemzése mindkét nyelvből megtörtént (*Einhorn, 2007a, 2008a; Major, 2007, 2008*).

A receptív készségeket és a nyelvhelyességet mérő vizsgarészekben elemezték az egyes feladatok itemeinek működését a különböző teljesítménycsoportokban. A vizsgálat során az itemek nehézségét az összpontszámhoz viszonyítva adták meg, amely alapján meghatározható, hogy a mintában és a teljesítménycsoportokban milyen arányban válaszolták meg helyesen a vizsgázók az adott itemet. Ez az elemzés lehetővé teszi a kiugróan könnyű és nehéz itemek azonosítását, de az nem állapítható meg, hogy az itemek nehézségi indexei között mekkora különbségek vannak, és ezek mennyiben felelnek meg a vizsgázók nyelvtudás-szintjének. Egy item nehézségi szintje a klasszikus tesztelméleti módszerekkel számolva nemcsak az itemet, hanem a mintát is jellemzi és a vizsgázók készségének fejlettségéről kapott kép függ a teszt jellemzőitől. Az adott feladatsor nehézsége főként az átlagos készségű vizsgázók eredményét befolyásolja (*Molnár, 2006*). A nem megfelelőnek ítélt feladat vagy item esetén így nem tudjuk biztosan meghatározni, hogy a populáció sajátosságáról vagy teszthibáról van szó. E problémák kezelésében alkalmazható a valószínűségi tesztelmülethez tartozó Rasch-modell, amelyvel elemző szoftverek logaritmikus transzformációt hajtanak végre az item- és személyadatokon, az ordinális skálán lévő adatokat áttranszformálják intervallumskálára, így a modell alkalmazásával lehetőség nyílik egy objektív itemnehézségi sorrend felállítására, amely egybevethető a vizsgált személyek képességeinek értékeivel (*Rasch, 1960*). A vizsgázók adott csoportjában megoldott itemek megfelelőnek vagy alkalmatlannak ítéltetők egy teljesen más csoportban is. Vagyis a feladatsorok minőségét célszerű a valószínűségi tesztelmélet adta lehetőségek segítségével is megvizsgálni, amelyhez az adatbázisokban szereplő adatmennyiség elegendő.

A produktív készségek közül az OFI által összeállított kisebb adatbázisokban csak az íráskészség részpontszámai találhatóak, a beszéd-készséget csak az OH adatbázisai alapján elemezhetjük. Az OFI adatbázisainak elemzésekor az íráskészségnél az egyes szempontoknál elért átlagos megoldottságot vizsgálták (*Einhorn, 2007a, 2008a; Major, 2007, 2008*) és kvalitatív elemzést végeztek mindkét évből kb. 100 kimásolt dolgozat alapján, amelyben a tipikus tanulói megoldásokat és a javításokat ellenőrizték. Ugyanakkor az adatbázisok alapján a valószínűségi tesztelméleti eszközökhöz tartozó *Masters* (1982) parciális kredit modelljével (*partial credit model*) lehetővé válik a vizsgázók íráskészségének és az egyes, különböző skálaszerkezetű értékelési szempontok nehézségének, valamint az analitikus skálában definiált szintleírások működésének vizsgálata. E modellre is jellemző a paraméterek szeparálhatósága és a specifikus objektivitás (*specifically objectivity*) megvalósítása a személyek és az értékelési szempontok nehézségének összehasonlítása területén. A modell alkalmazása azért tűnik ígéretesnek, mert lehetőséget biztosít a mért eredményt befolyásoló tényezők egyenkénti és egymásra gyakorolt hatásának elemzésére.

Az OH és az OFI által összeállított adatbázisok alkalmazásának előnyeit és hátrányait figyelembe véve mindenféleképp utalni kell arra az ellentmondásra, hogy az eredményeknek a vizsgafejlesztési folyamatba való visszacsatolása a jelenlegi gyakorlatban nem jelenik meg. Ugyanakkor az érettségi vizsga egészére vonatkozóan is megfogalmazódott annak szükségessége, hogy a vizsga továbbfejlesztésre szorul, szükséges a tartalmi standardok továbbfejlesztése, pontosítása és karbantartása. Ezek mellett pedig szükséges az érettségi vizsga szintjeinek összekapcsolása nemzetközi mérési rendsze-

rekkel. Ezzel a folyamattal „meg lehet akadályozni a tesztpontszámok inflálódását, vagyis azt a jelenséget, hogy a pontszámok különböző okokból – például a „tesztre tanítás” – formálisan nőnek, miközben a valódi tudás mértéke nem változik ugyanilyen ütemben” (Csapó, 2008, 75. o.). Ez a nyelvi érettségivel azt jelenti, hogy a vizsga szintjeinek a *Közös Európai Referenciakeret*hez (2002) való illesztése elengedhetetlen (Nikolov, 2007). „A vizsgafejlesztés folyamata nem zárult le a modell kialakításával és a rutinszerű működtetés megkezdésével [...]. Az érettségi vizsga tehát a jelenlegi állapotában még mindig átmenetnek tekinthető” (Einhorn, 2008c, 235. o.). Ezért szükséges az eredmények nyomon követése, ellenőrzése és az ezek alapján történő vizsgafejlesztés. Ugyanakkor a 2006. évi érettségi elemzése után lényegében semmiféle információval nem rendelkezünk a vizsga jelenlegi működéséről és minőségéről. A vizsga továbbfejlesztése szempontjából komoly probléma, hogy kizárólag a tendenciákat rögzítő 2007., 2008. és 2009. évi érettségi eredményei érhetőek el, de ezek részletes elemzése, a tendenciák feltárása sem történt meg. Ettől súlyosabb, hogy a feladat- és itemszintű elemzések ezekben a vizsgaidőszakokban nem lehetségesek, nem áll rendelkezésre a megfelelő adatbázis, pedig a vizsga minőségének javítása érdekében szükségesek a fejlesztő kutatások. E hiány csökkentéséhez járulunk hozzá a vizsgálatainkkal, amelyek célja, hogy empirikusan igazolják a vizsgafejlesztés szükségességét és példaként szolgáljanak a további kutatásokhoz.

Az empirikus vizsgálathoz felhasznált adatbázisok másodelemzésének módszerei

A vizsgán elért készségenkénti teljesítmények változásaihoz és meghatározó tényezőihez kapcsolódó kutatási kérdések (1–3) vizsgálatában az OH adatbázisait használtuk, amelyek a 2007., 2008. és 2009. évi májusi angol és német nyelvi érettségire vonatkoznak. Az adatelemzés első fázisában alapstatisztikai elemzéseket végeztünk, azaz a teljesítményeket vizsgáltuk, majd hipotézisvizsgálatok segítségével a nyelvi készségek fejlettségében megjelenő iskolatípus, képzési programok és nemek szerinti különbségek feltárásában kétmintás t-próbát végeztünk, ezek teljesítményre gyakorolt hatását eta-négyzettel vizsgáltuk. Ez a mutató hatáserősséget jelent, a külső variancia és a teljes variancia hányadosaként határozható meg, értéke 0–1 közötti szám, amely azt mutatja, hogy a csoportok (esetünkben például az iskolatípusok) közötti különbség hány százalékban határozza meg az összteljesítményt (lásd például Csapó, Székely és Tóth, 2009). Az eta-négyzet lényeges tulajdonsága, hogy a különbségek mértékét mutató F-értékkel ellentétben nem feltételez lineáris kapcsolatot a függő (a teljesítmények) és a független változó (például a különböző iskolatípusok) között, így például ezzel a mutatóval a szelekció mértéke is kifejezhető (lásd például Csapó, Molnár és Kinyó, 2009). Ezt követően az eredményeket hasonlítottuk össze az angol és a német nyelvnél.

A receptív készséget mérő feladatlapon elemzéséhez kapcsolódó 4. és 5. kutatási kérdések vizsgálatában a 2006. évi középszintű német érettségi adatbázisában szereplő két receptív készség, valamint a nyelvhelyesség (*Középszintű írásbeli vizsga, Német nyelv*, 2006) Rasch-moddal történő elemzését mutatjuk be. Az adatbázis önálló vizsgálatát indokolta az a tény, hogy a vizsga célja a Referenciakeret két szintjének mérése (A2–B1), így a vizsgafeladatoknak egy nagyon széles képességterületet kell átfogni. Mivel a vizsgafeladatok előzetes kipróbálása, nehézségi szintjük vizsgálata nem történt meg, így célszerű ezt az adatbázist az alapján megvizsgálni, hogy a feladatsorok empirikus nehézségi szintje mennyire fedi le a vizsgázók képességtartományát. Az elemzés kérdéseinek megválaszolására és a célok elérésére az adatfájlban úgy rögzítettük az adatokat, hogy a helyes válasz 1, a helytelen 0 pontot ért. A nyelvhelyesség vizsgarész feleltetválasztós feladataiban a disztraktorokat is kódoltuk, így ezek elemzésére szintén lehetőség nyílt. Az adatbázisban lévő feladatok legfontosabb adatait és megbízhatósági érté-

két a 17. táblázat tartalmazza, amelyeket a nyerspontoszám alapján számoltuk ki, melynek maximális értékét minden feladatsor esetén az itemszám mutatja.

17. táblázat. A feladatlapok átlaga, szórása és Cronbach- α -ja (N=1092)

Feladatlap	Átlag	Szórás	Cronbach- α
Olvasott szöveg értéke (itemszám=25)	17,88	4,31	0,80
Hallott szöveg értéke (itemszám=24)	16,50	4,17	0,81
Nyelvhelyesség (itemszám=30)	19,57	5,17	0,81

A 4. és 5. kutatási kérdés megválaszolására a három feladatlapot először a Rasch-modell alkalmazásával egy közös skálára kalibráltuk a ConQuest (Wu, Adams és Wilson, 1998) szoftverrel, majd a feladatlapok nehézségi indexét vizsgáltuk, amelyet a következő lépésben a mintába került diákok képességszintjével összekapcsolva közös képességskálán három dimenzióban a minta- és itemtérképen ábráztuk. A feladatbank kiépítésének lehetőségére utaló 6. kérdés megválaszolására a 2005. és 2006. évi közép- és emelt szintű feladatlapok minta- és itemtérképét vizsgáltuk egyenként. Ennek célja az volt, hogy empirikus úton igazoljuk az egyes időszakokban keletkező feladatlapok ekvivalencia-vizsgálatának, és az érettségi két szintjének egymáshoz és a KER-hez való illesztésének szükségességét. Ezt követően európai projektek eredményeit tanulmányoztuk, amelyek a nyelvi feladatbank kiépítésére vonatkoznak. Ezek alapján mutatjuk be a feladatbank kiépítésének folyamatára tett javaslatunkat.

Az íráskészséget mérő feladatlapok elemzéséhez kapcsolódó (7–9.) kutatási kérdések megválaszolására a 2006. évi emelt szintű német érettségi adatbázisát használtuk, amely a vizsgázói populáció (N=1640) 44,88%-át, összesen 736 vizsgázó adatát tartalmazza (18. táblázat). A minta és a vizsgázói populáció teljesítmény-eloszlása között nincs szignifikáns eltérés ($\chi^2=0,014$; $p=0,99$), a kettő átlagteljesítménye is megegyezik.

18. táblázat. A minta eloszlása a teljesítmény szerint (N=736)

Teljesítmény (%)	Vizsgázói populáció*		Minta	
	Elemzészám (db)	Arány (%)	Elemzészám (db)	Arány (%)
0–9	48	2,9	21	2,9
10–19	88	5,4	40	5,4
20–29	64	3,9	29	3,9
30–39	71	4,3	32	4,3
40–49	124	7,6	56	7,6
50–59	197	12,0	88	12,0
60–69	202	12,3	91	12,3
70–79	247	15,1	111	15,1
80–89	251	15,3	113	15,3
90–100	348	21,2	155	21,1
Összesen	1640	100,0	736	100,0
Átlagteljesítmény (%-pont)	64,6		64,7	

*Forrás: Einhorn, 2008a, 17. ábra

A 2006. évi emelt szintű érettségi íráskészség feladatlap két feladatot tartalmaz, az elsőben a tanulócsoporthoz kellett egy versenyre jelentkezni, a másodikban két téma közül választhattak az érettségizők: vagy olvasói levelet írhattak a kortárs csoportjukról vagy pedig egy cikket diákújság számára az olvasásról (*Emelt szintű írásbeli vizsga, Német nyelv*, 2006). A két feladat paramétereit a 19. táblázat tartalmazza, amelyben az

adatokat a nyerspontok alapján adjuk meg. Ezek alapján megállapítható, hogy az értékelési skálák használata magas megbízhatóságot eredményez. A két feladatot azonos nyelvi szintre készítették, azonban az első feladatban a vizsgázók szignifikánsan jobb átlagteljesítményt értek el ($t=2,14$; $p=0,03$). Ennek oka az egyes értékelési szempontok nehézségében, az értékelési skálák és a feladatban megadott kommunikációs célok, valamint a létrehozandó szövegfajta különbségében keresendők. A második feladatban lévő két téma paraméterei közül sem az átlagteljesítményben, sem az egyes szempontoknál nincs szignifikáns különbség, amely megfelel a tesztkészítők elvárásainak.

19. táblázat. A feladatok és a feladatlap átlaga, szórása, Cronbach- α -ja a nyerspontok alapján

	Elemsszám	Átlag	Szórás	Cronbach- α
1. feladat (20 pont)	736	13,08	5,04	0,87
2. feladat (20 pont)	736	12,65	5,87	0,94
1. téma	263	12,68	6,16	0,94
2. téma	472	12,64	5,72	0,94
<i>Feladatlap (40 pont)</i>	736	25,84	9,80	0,92

A 7–9. kutatási kérdésekre a valószínűségi tesztelméleti eszközök közül *Masters* parciális kredit modelljét alkalmaztuk, mivel az egyes értékelési szempontokhoz tartozó skáláknak eltérő a szerkezete. Először az adatokat közös skálára kalibráltuk a ConQuest programmal és az értékelési szempontok átlagos nehézségi indexét és modellilleszkedését vizsgáltuk. A következő lépésben az értékelési skálapontok küszöbértékeit a mintába került diákok íráskészség-szintjével kapcsoltuk össze és elemeztük. Ezt követően pedig a skálák karakterisztikus görbéit vizsgáltuk és azt állapítottuk meg, hogy az értékelők hogyan használják az egyes pontszámokat, azaz, hogy minden pontszám és deskriptor azonos mértékben hasznosul-e. A vizsgálatunk célja, hogy e lépésekkel definiáljunk és empirikus adatok alapján értelmezzünk olyan paramétereket, amelyekkel az íráskészség mérési-értékelési eljárásainak minősége megállapítható. Ezek felhasználásával a különböző vizsgaidőszakokból származó eredmények összehasonlíthatók és változásai nyomon követhetők. Így elősegíthetjük azt a folyamatot, amelyben az elért vizsgázói eredmények alapján adható visszacsatolás a vizsgafejlesztés számára.

3.3.2. A tanári kérdőíves vizsgálat módszerei

Az idegen nyelvi érettségi vizsga tanítási-tanulási folyamatra gyakorolt hatását a hazai nyelvtudásmérésben még szisztematikusan nem vizsgálták. Ezért kutatásunkban célul tűztük ki a vizsga hatását meghatározó tényezők átfogó vizsgálatát. A nemzetközi szakirodalom alapján a tanárok vizsgával kapcsolatos nézetei és attitűdjei meghatározó jelentőségűek, így ezeket kell elsődlegesen megvizsgálni. Ennek érdekében egy kérdőíves felméréssel azt vizsgáljuk, hogy a középiskolai nyelvtanárok milyen hatások meglétéről nyilatkoznak. A nemzetközi teszthatásokat feltáró kutatások a tanárok kérdőíves felmérése mellett egyéni, illetve fókuszcsoportos interjúkat, valamint osztálytermi megfigyeléseket is tartalmaznak. Vizsgálatunkban nem vállalkoztunk a különböző adatgyűjtő eszközök kombinált alkalmazására, mert egyrészt ez meghaladta volna az értekezés kereteit, másrészt elsődleges célunk az volt, hogy a vizsga hatásmechanizmusaira vonatkozó legfontosabb területeket vizsgáljuk a modell alapján. Ezen információkkal kívánjuk segíteni a vizsgafejlesztési folyamatot és a további, a vizsga hatását egyes területeken részletesen és más adatgyűjtő eszközök segítségével elinduló kutatásokat.

A minta jellemzői

A vizsgálat mintavétele célzott volt, olyan nyelvtanárokat akartunk közvetlenül elérni és a vizsgálatban való részvételre felkérni, akik az érettségi működését belülről látják, gyakran vesznek részt vizsgáztatásban. A lebonyolítás során 1261 középiskolai nyelvtanár e-mail címére küldtük ki közvetlenül a felkérő levelet (5. melléklet). Az elektronikus adatfelvétel előnye, hogy megfelelően biztosítja az anonimitást, a kérdőív kitöltése egyszerű, az adatok azonnal az adatbázisba kerülnek, hátránya viszont, hogy tapasztalatok szerint csupán 10–15%-os válaszolási arányra lehet számítani. A kutatásban való részvételre felkérték 14,59%-a, 184 tanár töltötte ki a kérdőívet. A tanárok önkéntesen vettek részt a felmérésben, javadalmazást a kérdőív kitöltéséért nem kaptak.

Az elektronikus kérdőívet (4. melléklet) hiánytalanul 160 nyelvtanár töltötte ki, 24 kérdőívben többnyire az utolsó oldalon elhelyezett háttérváltozók megadása hiányzott. A mintát e csökkentett elemszám alapján tudjuk jellemezni, amelynek az iskolatípus és a nyelvek szerinti eloszlását a 20. táblázat tartalmazza. A 2008. évi *Közoktatási Statisztikai Kiadvány* szerint a németet tanító középiskolai tanárok felülreprezentáltak, ezért a némettanárok adataiból véletlenszerűen elhagyva a létrejött 149 fős minta (21. táblázat) reprezentatívnak tekinthető a tanított idegen nyelv és az iskolatípus szerint ($\chi^2=0,80$; $p=0,98$). A 10–14. kutatási kérdések vizsgálatához ezt a mintát alkalmazzuk.

20. táblázat. A mintába került nyelvtanárok iskolatípus szerinti eloszlása (N=160)

Iskolatípus	Elemszám				
	Összesen	angol	német	francia	olasz
Gimnázium	78	37	30	6	5
Gimnázium és szakközépiskola	17	7	10	0	0
Gimnázium, szakközépiskola és szakiskola	7	2	5	0	0
Szakközépiskola	58	33	25	0	0

21. táblázat. A csökkentett elemszámú minta jellemzői (N=149)

Kategória	Populáció*		Populáció 1,5%-a	
	N	%	N	%
Angoltanár gimnáziumban	3081	31,3	46	30,9
Némettanár gimnáziumban	2169	22,0	34	22,8
Franciatanár gimnáziumban	523	5,3	6	4,0
Olasztanár gimnáziumban	280	2,8	5	3,4
Angoltanár szakközépiskolában	2216	22,5	33	22,1
Némettanár szakközépiskolában	1590	16,1	25	16,8
Összesen	9859	100	149	100

*Forrás: *Közoktatási Statisztikai Kiadvány* (2008)

A minta jellemzéséhez a kérdőív több demográfiai kérdést tartalmaz, amely a válaszadó nemére, végzettségére, a tanított évek számára és a vizsgáztatói tapasztalataira vonatkozik. A megkérdezettek 87,9%-a nő. A minta településtípus szerinti eloszlását a 22. táblázat mutatja. A tanárok 80,5%-a egyetemi diplomával rendelkezik a tanított idegen nyelvből, a minta 13,4%-a hároméves nyelvtanári végzettséggel tanítja az idegen nyelvet és 6,0% rendelkezik főiskolai végzettséggel. A minta eloszlása a tanított évek száma szerint heterogén (átlag=17,8; szórás=10,0; relatív szórás=56,2%), több módussal jellemezhető.

22. táblázat. A minta eloszlása településtípus szerint (N=149)

Kategória	Gyakoriság	Arány (%)
Nem város	2	1,3
Kisváros	46	30,9
Megyei jogú város	63	42,3
Főváros	38	25,5
<i>Összesen</i>	149	100,0

A válaszadó pedagógusokat a szakmában eltöltött évek száma alapján négy egyenlő nagyságú csoportra osztottuk (23. táblázat). A legfeljebb 9 éve tanítókat a pálya elején járóknak tekintettük, a másik csoport tagjai 10–19 évet, a harmadik 20–29, végül az utolsó, az előbbiekkal azonos létszámú tanári csoportot azok a nyelvtanárok alkották, akik 30 évnél hosszabb időt töltöttek el a pályán. Ugyanakkor e felosztással az első kategóriába a pályakezdők (általában 0–5 év gyakorlattal rendelkezők) és a tapasztaltabb tanárok egyaránt kerülnek. A megadott csoportosítás mellett szóló érv az, hogy a pályakezdő tanárok általában még nem érettségiztetnek, a vizsgával kapcsolatos nézetek és attitűdök hosszabb távon formálódnak, továbbá a vizsgálatunkban nem cél az, hogy a pályakezdők gondolkodását a többi csoporthoz képest vizsgáljuk.

23. táblázat. A minta eloszlása a tanári pályán eltöltött évek száma szerint (N=149)

A pályán eltöltött évek száma	Gyakoriság	Arány (%)
1–9 éve tanít	40	26,8
10–19 éve tanít	43	28,9
20–29 éve tanít	33	22,1
30–39 éve tanít	33	22,1
<i>Összesen</i>	149	100,0

A mintába került nyelvtanárok 50,3%-a középszintű érettségire készíti fel diákjait, mindössze 3,4% csak emelt szintre, míg a válaszadók 46,3%-a mindkét szintre, ezért a tanárok vizsgáról való megítélése tekintetében a minta két nagyobb csoportra osztható. A nyelvtanárok 47,7%-a normál óraszámú csoportban, 28,2%-a normál- és emelt óraszámú csoportban készíti fel a diákokat az érettségire; a minta 14,1%-a csak emeltszintű, míg 10,1%-a kéttannyelvű csoportban is tanít. A mintába került nyelvtanárok 58,4%-a eddig középszinten, ettől kevesebben (34,2%) a középszint mellett emelt szinten is érettségiztetett. Kisebbségben vannak azok a tanárok, akik még egyik szinten sem (4,0%), vagy csak emelt szinten (2,7%) vizsgáztattak.

A tanárok továbbképzéseken való részvételét a kérdőív három kérdésével vizsgáltuk (64., 66. és 67. kérdés). Eredményeink szerint a tanárok 66,4%-a vett részt érettségire vonatkozó vizsgáztatói képzésen, ettől kevesebben (40,3%) más nyelvvizsgáztatói képzésen is részt vettek. Ezek mellett azt is megkérdeztük, hogy milyen módszertani továbbképzéseken vettek részt a tanárok a kétszintű idegen nyelvi érettségi bevezetése óta (24. táblázat). A kérdést *Nikolov, Ottó és Öveges (2008)* vizsgálatából adaptáltuk, amelyben az idegennyelv-tanulás és tanítás helyzetét és fejlesztési lehetőségeit vizsgálták a szakképző intézményekben. Eredményeink szerint a külföldi továbbképzésen részt vettek aránya az említett kutatáshoz képest a mintában magasabb. Az egyes kategóriák a továbbképzések mennyiségében egymástól közel azonos távolságra lévő skálán helyezhetők el, így ezekhez megfelelő súlyokat rendeltünk (például 1-es súly a 0–30 órás továbbképzés, 2-es súly a 31–60 órás továbbképzés esetén, stb.), kivéve a 120+ órás to-

vábbképzéshez és a külföldi továbbképzéshez, amelynél azonos (5-ös) súlyt adtunk. Ezzel a módszerrel képeztünk egy összevont változót.

24. táblázat. A továbbképzések megoszlása időtartam, gyakoriság és hely szerint (%)

	0–30 óras tovább- képzés	31–60 óras tovább- képzés	61–90 óras tovább- képzés	91–120 óras to- vábbképzés	120+ óras tovább- képzés	Külföldi tovább- képzés
Soha	26	78	94	95	92	65
Egyszer	33	13	5	5	7	26
Kétszer	15	7	1		1	9
Háromszor	11	2				1
Négyszer	5	1				
Ötször	3					
Hatször	1					
Hétszer	1					
Hétnél többször	5					

A minta eloszlása a településtípus, a tanított évek száma, a vizsgára való felkészítés és a vizsgáztatói tapasztalat alapján arányos: minden kategóriába közel azonos mértékben kerültek a vizsgált személyek, a minta megfelelő a különböző statisztikai vizsgálatok elvégzésére. Ugyanakkor a kapott eredmények csak tájékoztató jellegűek, mivel nem állnak rendelkezésre adatok e változók reprezentativitásának ellenőrzésére.

A kérdőív létrehozása

A kérdőív fejlesztő munkájának legfontosabb állomásait a 25. táblázat tartalmazza. A fejlesztésnek két intenzívebb időszaka volt: (1) a kérdőív koncepciójának kialakítása, a kérdőív kismintás kipróbálása és elemzése, valamint (2) a véglegesített kérdőív elektronikus változatának létrehozása és az adatfelvétel lebonyolítása.

A kérdőív fejlesztésében először szakirodalmi elemzés és szintetizálás módszerével hoztuk létre a vizsgálatunk alapjául szolgáló modellt (11. ábra) és tártuk fel azokat a tényezőket, amelyek a hatás mértékét és típusát alapvetően meghatározhatják (9. táblázat). A kérdőív fejlesztési folyamatának célja az volt, hogy a tanárookra vonatkozó faktorokat vizsgáljuk. Ezek adták a kérdőívünk szerkezetét, amelyen a fejlesztés során nem módosítottunk.

25. táblázat. A kérdőív fejlesztési folyamata

Időpont	Fejlesztési feladat
2007. június–augusztus	Szakirodalmi elemzés és szintetizálás
2007. szeptember–október	A nemzetközi szakirodalomból ismert kérdőívek adaptálása
2007. október	A kérdőív első változatának szakmai lektorálása
2007. november 8–19.	A kérdőív kismintás kipróbálása
2007. november–2008. január	A kérdőív adatelemzése, továbbfejlesztése
2008. május–december	A kutatási koncepció véglegesítése, a minta szervezése
2008. december	A módosított kérdőív szakmai lektorálása
2009. január–április	A kérdőív elektronikus változatának létrehozása
2009. április 15–22.	Az elektronikus kérdőív pilot mérése
2009. május 3–22.	A kérdőíves felmérés lebonyolítása

A kérdőívtételek megírásában először több nemzetközi vizsgálatban alkalmazott tanári kérdőívet elemeztünk és adaptáltunk (Chen, 2002. 175–180. o.; Cheng, 1998. 348–351. o.; Caine, 2005. 109–115. o.). Ez alapján jött létre egy 55 itemből álló kérdőív. A megfogalmazott kérdőívtételeket az érettségi fejlesztési folyamatában résztvevő öt kolléga lektorálta. Ennek során a validitást ellenőrizve arra kértük őket, hogy minden állításhoz jelöljék, megfelelőnek tartják-e az itemet a kutatás céljai szempontjából, továbbá érthető-e az item, vagyis megfelelő-e a szövegezés és a nyelvezet.

A kérdőív átdolgozása után került sor a kismintás kipróbálásra, amelyben német nyelv-tanárok vettek részt (N=31). A teljes kérdőív reliabilitása (Cronbach- $\alpha=0,68$) alacsonynak bizonyult. A kérdőívtételek elemzése azt mutatta, hogy több nem működött megfelelően. Problémásnak bizonyultak azok az állítások, amelyekben a nyelvoktatási gyakorlatot és az érettségi hatását egyben fogalmaztuk meg (pl.: *A hallott szöveg értésének fejlesztése az új érettségi hatására nagyobb szerephez jut az órán*). A nem megfelelő itemműködés oka feltehetően az volt, hogy sok tanár nem az érettségi hatására váltott módszereket, hanem már a bevezetése előtt, hiszen a kommunikatív nyelvoktatás alapelvei korábban már megjelentek, az érettségi csak követte ezt a szemléletet. Ezért döntöttünk úgy, hogy a vizsga és a nyelvoktatási gyakorlat közötti összefüggések vizsgálatához két skálát dolgozunk ki, melyek közül az egyik az egyiket jelöli a tanárok, hogy mennyire jellemző az adott tanítási gyakorlat, és a másikon azt, hogy mennyire jellemző abban az érettségi hatása.

2008 második felében került sor a kérdőív javítására, továbbfejlesztésére, a kutatási koncepció részletes kidolgozására és a minta szervezésére. Ebben a fázisban külön problémát jelentett a tanárok e-mail címéhez való hozzáférés, amelyet végül úgy oldottunk meg, hogy az Oktatási és Kulturális Minisztérium honlapján elérhető iskolai adatbázis alapján az iskolák honlapján nyilvánosan közzétett nyelv-tanárok e-mail címét gyűjtöttük össze. Az így létrejött adatbázisban nem rögzítettük külön a nyelv-tanár nevét, iskolájának címét, hanem kizárólag a szakot és az e-mail-címet, amelyre a kérdőív kitöltésére felkérő levelet küldtük.

A módosított kérdőívet újabb szakmai lektorálás követte, majd 2009. januárjában kezdődött a kérdőív elektronikus rögzítése. Az elkészült elektronikus kérdőív pilot mérésének fő célja a rendszer tesztelése volt, továbbá arra kértük a tanárokat, hogy jelöljék, mennyire szólítja meg őket a felkérő e-mail, mennyire tartják reálisnak a megadott időkeretet a kérdőív kitöltésére, hol maradtak a kérdőívben olyan állítások, amelyek nem egyértelműek és javítani kellene. A visszajelzések alapján kisebb javításokat végeztünk.

Az elektronikus kérdőívben (4. melléklet) egy képernyőrészen tartalmilag összefüggő és egy gondolatmenettel kitölthető kérdőívtételek jelentek meg. Célunk az volt, hogy a lehető legegyszerűbb legyen a kitöltés, és ezért csak a választ kellett a tanároknak megjelölniük. Ugyanakkor ez a megoldás monotónná tette a kérdőív kitöltését, de egy oldalon 13 kérdőívtételnél több sehol nem szerepelt, görgetni elsősorban a háttér-változók kitöltésénél kellett. Mindezek figyelembe vétele miatt is szükséges a kérdőív működésének részletes elemzése.

A kérdőív szerkezete

A kérdőívvel az érettségi vizsgával kapcsolatos (1) nézeteket, (2) attitűdöket, és (3) a tanárok képzettségét vizsgáljuk (26. táblázat). Mind a nézeteket, mind az attitűdöket ötfokozatú Likert-skálán mértük. A páratlan számú középső értékkel rendelkező skálát azért választottuk, mert így a tanárok a vizsgához való semleges álláspontjukat is kifejezhetik, így a közömbös megkérdeszettek nem kényszerítettük döntésre. Ennek oka, hogy a nézetek és attitűdök viszonylag hosszabb időn át formálódnak.

26. táblázat. A tanári kérdőív változórendszere

Kérdőívrész	Vizsgált tényező	Item	Item-szám
	– a vizsga megbízhatóságáról és tökéletességéről kialakított nézetek	2, 8, 9, 10	4
I. A tanárok vizsgával kapcsolatos nézetei, attitűdjei	– a vizsga presztízséről és hasznosságáról kialakított nézetek	1, 3, 7	3
	– a vizsga filozófiájáról kialakított nézetek	4, 5, 6	3
	– a vizsga következményeivel kapcsolatos attitűdök	12–16	5
	– a vizsgára felkészítő tananyagokkal kapcsolatos attitűdök	17, 19	2
	– a tanórára való felkészüléssel kapcsolatos attitűdök	11, 18	2
II. A tanárok nézetei az oktatási folyamatot befolyásoló tényezőkről	– a képzettséggel kapcsolatos tényezők hatásáról kialakított nézetek	20–22	3
	– a külső tényezők hatásáról kialakított nézetek	23–26	4
III. A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról	– az eredményes tanítási módszerekről kialakított nézetek	28, 30, 35	3
	– a vizsga és a tanítási gyakorlat közti kapcsolatról kialakított nézetek	29, 34, 36	3
	– a vizsga és az oktatás tartalma közti kapcsolatról kialakított nézetek	27, 31, 32, 33	4
	– a vizsga és az osztálytermi értékelési eljárások közti kapcsolatról kialakított nézetek	37	1
IV. A tanárok nézetei az idegen nyelvi tesztelés kritériumairól	– a klasszikus tesztelmélet jóságmutatóinak fontosságáról kialakított nézetek	38, 41, 46	3
	– a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek	39, 40, 43, 45, 48–50	7
	– a kompetencia tesztek kritériumainak fontosságáról kialakított nézetek	42, 44, 48	3
V. A tanárok nézetei a vizsgafeladatok oktatási folyamatban való alkalmazásának céljairól	– a vizsgafeladatok alkalmazásának előrejelző funkciójáról kialakított nézetek	51, 53, 54	3
	– a vizsgafeladatok motiváció és készségfejlesztés céljából való alkalmazásáról kialakított nézetek	52, 55	2
VI. A tanárok képzettsége	– végzettség, diploma megszerzésének éve, tanított évek	59, 60, 61	3
	– tanárok oktatási tapasztalatai	62	1
	– vizsgáztatói képzésen való részvétel, vizsgáztatói tapasztalatok	64, 65, 66	3
	– módszertani képzésben való részvétel mennyisége	67	1
VI. Független változók	– iskolatípus, az iskola helye	56, 57	2
	– nem	58	1
	– tanított idegen nyelv	0*	1
	– nyelvi csoport	63	1

*Megjegyzés: A tanított idegen nyelv változót a kérdőív legelején kellett a tanároknak megadni. Ennek oka, hogy a nyelvtanárok a válaszaikat csak a kiválasztott idegen nyelvre vonatkoztassák, így ezt a kérdőív csoportazonosító kérdéseként a validitás növelése céljából alkalmaztuk.

A kérdőív első részében az 1–10. kérdőív-tétellel a tanárok vizsgával kapcsolatos nézeteit, a 11–19. állítással a tanárok vizsgára vonatkozó attitűdjeit vizsgáltuk. Az 1–8. kérdőív-tételt *Cheng* (1998), míg a 12–15. állításokat *Chen* (2002) kérdőívéből vettük át és alakítottuk a saját kontextusunkra. *Chen* (2002) kérdőívének néhány részét *Al-Jamel*

és Ghadi (2008) is felhasználták a jordániai vizsgálatukban. A kérdőívhez felhasznált forrásokat a 3. melléklet tartalmazza.

A második részben Cheng (1998) és Caine (2005) kérdőívei alapján a tanárok nyelvtanítási gyakorlatát befolyásoló tényezőket soroltunk fel a 20–26. kérdőív-tételben. Ebből a részből arról is információt kaphatunk, hogy a felsorolt tényezőkhez képest a tanítási gyakorlatot a kétszintű érettségi vizsgakövetelményei és vizsgafeladatai mennyiben határozzák meg.

A harmadik részben két alskálán vizsgáltuk a tanárok nyelvtanítási gyakorlatát az érettségi jellemzői alapján. A kérdőív-tételek létrehozása során Petneki (2007) alapján meghatároztuk az idegen nyelvi érettségi vizsga jellemzőit és a nyelvoktatási gyakorlatra kifejtett valószínűsíthető hatását (27. táblázat). Ezekhez írtuk a kérdőív-tételeket, amelyeknél a tanároknak azt kellett jelölniük, hogy mennyire jellemző az adott nyelvoktatási gyakorlat és mennyire jellemző abban az érettségi hatása. Az egyes kérdőív-tételekhez tartozó kettős skála használatát a kérdőív kipróbálása során szerzett tapasztalataink tették szükségessé. Ehhez a részhez sorolhatóak még azok a kérdések, amelyekkel azt vizsgáltuk, hogy a tanárok váltottak-e tankönyvet az új érettségi hatására (68. kérdés), továbbá melyik tankönyvet (69. kérdés) és feladatgyűjteményt (70. kérdés) használják leggyakrabban az érettségire való felkészülési folyamatban. Ezekhez nyílt végű kérdéseket írtunk és a kérdőív VI. részében helyeztük el.

27. táblázat. „A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról” kérdőív-rész létrehozásának szempontjai (Petneki, 2007. 121–122. o. alapján)

Az érettségi jellemzője	Következménye a nyelvoktatási gyakorlatra	Item
Kétszintű, amely a Közös Európai Referencia-keret három szintjét méri.	A tanulók differenciált felkészítése.	28.
Egynyelvű: a feladatmeghatározásokat és a feladatokat a célnyelven fogalmazták meg.	A nyelvórákon a célnyelv használata.	29.
<i>Kommunikatív szemléletű:</i>		
– kommunikatív szándékokat és helyzeteket kell megvalósítani,	– nyelvi szituációkra épülő nyelvórák változatos munkaformákban,	34. 35.
– autentikus szövegekre épül,	– autentikus szövegek használata az órán,	31.
– változatos, szövegközpontú feladattípusokat tartalmaz,	– változatos, szövegközpontú feladatok alkalmazása az órán,	36.
– „új” készség a hallott szöveg értése,	– hallott szöveg értésének fejlesztése,	27.
– a korábbi vizsgához képest nem tartalmaz fordítási feladatot.	– fordítás helyzete.	32.
<i>Nyelvhelyesség szerepe új:</i> A vizsga nem a formális nyelvtani ismereteket méri, hanem a nyelvhasználat képességét funkcionálisan.	A nyelvtan gyakoroltatása autentikus szövegekkel és szituációkkal.	33.
<i>Standardizált:</i>		
– A vizsgakövetelmények egységesen érvényesek, leírásuk részletes.	– A tanulók folyamatos felkészítése a készségek kommunikatív használatára.	30.
– A javítási és az értékelési útmutató részletes, a produktív készségeknél értékelési szempontok és skálák vannak.	– A produktív készségek értékelésénél szempontok és értékelési skálák alkalmazása.	37.

A kérdőív negyedik részében azt vizsgáltuk, hogy a kétszintű idegen nyelvi érettségiben követett kommunikatív tesztelés kritériumait a tanárok mennyire tartják fontosnak (28. táblázat). A kérdőív-tételek megírásában abból indultunk ki, hogy a vizsgán alkalmazott alapelvek hatással lehetnek az osztálytermi mérés és értékelés során alkalmazott, a nyelvtanárok által készített mérőfeladatokra is. A kérdőív-tételek létrehozásakor elő-

ször meghatároztuk a kommunikatív tesztelés kritériumait és jelentését, amelyek a feladatok elkészítésére vonatkoznak, majd értelmeztük ezeket a tanítási-tanulási folyamatban alkalmazott mérőfeladatokra. Az interaktivitásra vonatkozó 45. állítást negatív formában fogalmazzuk meg, így ezt az adatelemzés során átkódoltuk. Ebben a részben a klasszikus tesztelmélet jóságmutatóira vonatkozó állítások mellett (38., 41., 46.) a kompetencia tesztekkel kapcsolatban is megfogalmazzunk kérdőívteteleket (42., 44., 47.).

28. táblázat. „A tanárok nézetei az idegen nyelvi tesztelés kritériumairól” kérdőív rész létrehozásának szempontjai a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek vizsgálatához

<i>Kritérium</i>	<i>Jelentés</i>	<i>Item</i>
külső validitás	Mennyire illeszkedik a teszt egy külső kritériumhoz?	48., 49.
látszatvaliditás	Mennyire tartják a tesztmegoldók a tesztet érvényesnek?	50.
autentikusság	Milyen mértékben egyezik meg a tesztfeladat megoldása során alkalmazott nyelvtudás azzal a móddal, ahogyan azt valós nyelvi szituációban alkalmaznánk?	39., 43.
interaktivitás	Milyen mértékben vonják be a vizsgázók a tesztfeladat megoldásába a nyelvi készségeiket, és egy adott témáról megszerzett előzetes háttértudásukat és ismereteiket?	45.
teszthatás	Milyen mértékű és irányú a teszt hatása a nyelvtanulókra?	40.

Az ötödik részhez tartozó öt kérdőívtelet (51–55) Cheng (1998) kérdőíve alapján dolgoztuk ki, amellyel azt vizsgáltuk, hogy a nyelvtanárok szerint melyek a vizsgafeladatok alkalmazásának céljai a tanítási-tanulási folyamatban (3. melléklet). Több vizsgálat alapján (Kiss-Gulyás 2001; Nikolov, 1999) megállapítható, hogy Magyarországon nagy hagyománya van a mérőfeladatok felhasználásának a tanítási-tanulási folyamatban, amelyeknek alapvetően más a célja és a funkciója, mint a nyelvoktatási folyamatban alkalmazható gyakorlófeladatoknak (Einhorn, 2008b). A megfogalmazott állításoknál az egyetértés mértékét kellett a tanároknak jelölni és ezzel azt vizsgáltuk, hogy a tanárok mennyire (jól) látják a két feladatforma közötti különbségeket.

A kérdőív segítségével három függő változót vizsgálunk: (1) az érettségi hatása a nyelvoktatási folyamatra, (2) a kommunikatív tesztelés kritériumainak fontossága és (3) a vizsgafeladatok felhasználása a nyelvoktatási folyamatban. E három függő változót kétféle független változócsoporthoz befolyásolhatja: (1) a tanárok végzettsége és képzettsége, a tanítási tapasztalat, a tanított évek száma, a módszertani és vizsgáztatói továbbképzéseken való részvétel és a vizsgáztatói tapasztalatok. Ezek mellett az adatelemzés során célunk a vizsga megítélésére vonatkozó további háttérváltozók létrehozása. (2) A tanítási környezetre vonatkozó független változók közül a függő változókat befolyásolhatja az iskola típusa, az iskola helye, a nyelvi csoport jellege, amelyben az érettségire való felkészítés történik, vagy a nyelvtanulók elvárásai.

A kérdőív tesztelméleti mutatói

Az egyes kérdőívtetelek tesztelméleti mutatóinak vizsgálatában először a teljes kérdőív reliabilitását, valamint az egyes itemek működését itemkihagyásos reliabilitással elemeztük. Ezután azt vizsgáltuk, hogy az előzetes, szakirodalmi vizsgálatra épülő elméleti változórendszer (26. táblázat) az empirikus adataink mennyire adják vissza. Ennek eszköze a faktoranalízis volt, amelyet a kérdőív validitásának ellenőrzésére használtunk. A nemzetközi teszthatást vizsgáló kutatások a jelenség komplexitása miatt felhívják a validálási folyamat fontosságára a figyelmet (például Alderson és Banerjee, 2001). Ez az oka annak, hogy a kérdőívünk validitásának biztosítására nagy hangsúlyt fektettünk. Először kérdőívreszenként végeztünk faktoranalízist, és azt vizsgáltuk, hogy mely

itemeket kell elhagyni a kérdőívbelől. Ezt követően a teljes kérdőívre is elvégeztük a faktoranalízist és az ennek eredménye alapján létrejött változórendszerrel hasonlítottuk össze az előzetes struktúrával. Ezt követően újra megvizsgáltuk a teljes kérdőív és az egyes kérdőívreszek reliabilitását.

A teljes kérdőív reliabilitása (Cronbach- $\alpha=0,88$) megfelelő, az egyes kérdőív-tételek elemzésekor egy esetben találtunk nem megfelelő elkülönítés-mutatóval rendelkező itemet ($r=-0,13$; $p=0,58$). A 10. állítás (*Feleslegesnek tartom a középszintű vizsgán az A2 szintű feladatokat, a B1 szint elérhető az érettségizőknek*) nem működött megfelelően. Ez az item feltehetően azért rossz, mert a vizsgaszintek és a tanulók nyelvtudásának megítélése egyszerre jelenik meg benne, amelyet a KER (2002) szintrendszere alapján határoztunk meg. Célszerű így ezt az itemet elhagyni a kérdőív további elemzéséből és a kérdőív esetleges továbbfejlesztése során módosítani.

A kérdőív validitásának ellenőrzésére a tanárok válaszai mögött meghúzódó látens struktúrák feltárása céljából az öt kérdőív-reszben szereplő változókkal faktorelemzést végeztünk. Ennek legfőbb célja, hogy az egyes kérdőív-reszekben belül a tanárok válaszai alapján tárjunk fel olyan rejtett változókat, amelybe a kérdőív manifeszt változói besorolhatók, majd a besorolt változók jellemzőivel a faktorok interpretálhatók. A faktoranalízis elvégzése előtt minden esetben kiszámoltuk a Kaiser-Meyer-Olkin-mutatót, amely azt adja meg, hogy a változórendszer mennyiben alkalmas faktoranalízisre. Csak akkor tekintettük a változórendszerrel alkalmasnak a látens struktúrákeresésre, ha a KMO index legalább 0,6. Minden kérdőív-resz vizsgálatakor a varimax rotáció utáni faktorsúlyokhoz 0,4-et választottunk faktorsúlyhatárként. Ezt követően definiáltuk a faktorokat és hasonlítottuk össze a szakirodalmi kutatás során létrehozott kérdőívstruktúrával (26. táblázat) és az egyes kérdőív-reszek tartalmával (27. és 28. táblázat).

A tanárok vizsgával kapcsolatos nézeteivel és attitűdjeivel végzett faktoranalízis eredményeit a 29. táblázat tartalmazza. A KMO-mutató értéke 0,82, a parciális korrelációk elfogadhatóan kicsik és a változók páronként nem függetlenek (a Bartlett-teszt szignifikanciája: $p<0,001$). A táblázat a varimax rotáció utáni faktorsúlyokat tartalmazza, amelyek azt mutatják, hogy az adott kérdőív-tétel milyen szoros összefüggésben van a létrejött faktorialal. A táblázatban csak a 0,4-nél nagyobb faktorsúlyokat tüntettük föl. Az áttekinthetőség kedvéért az állításokat nem a kérdőívben szereplő sorrendjükben, hanem a megfelelő faktorok szerint csoportosítva adtuk meg. Az első kérdőív-reszhez tartozó 18 változóból az elemzés 5 dimenziót tárt fel, amelyek a teljes variancia 63%-át magyarázzák.

Az 1. faktor nyolc változót tartalmaz. A kérdőívstruktúra alapján ide kerültek a vizsga megbízhatóságáról és tökéletességéről (2. és 8.), presztízséről és hasznosságáról (1., 3., 7.), illetve filozófiájáról (4., 5., 6.) alkotott tanári nézetek. A 10. kérdőív-tételt töröltük, a 9. állítás a 4. faktorba került. Míg a szakirodalomban e három tényezőt egymástól elkülönítve határozzák meg (Spratt, 2005), addig a tanárok válaszai alapján ezek nem különülnek el egymástól. Az 1–8. állításokat Cheng (1998) kérdőívbelől adaptáltuk. Ez a változórendszer egyrészt azt jelöli, hogy a vizsgaeredmények hogyan szolgálnak döntések alapjául és hatnak közvetlenül és közvetetten a diákokra, másrészt pedig arra utal, hogyan érzékelik a tanárok a kétszintű nyelvi érettségi státuszát, amely befolyásolhatja a nyelvoktatási folyamatukat. A továbbiakban ezt az összevont változórendszerrel a tanárok nézetei a vizsga jelentőségéről faktorként interpretáljuk.

A 2. faktor négy, vizsgával kapcsolatos attitűdöt jelöl, amely nem tartalmazza az előzetes struktúrában lévő 16. kérdőív-tételt. A faktorba sorolt változók arra utalnak, hogy a nyelvtanárok a vizsgaeredményekből milyen következtetéseket vonnak le a saját tanítási gyakorlatukra vonatkozóan. Chen (2002) ezt a változórendszerrel tanári hivatástudatnak nevezi, amely Alderson és Wall (1993) szerint azt jelenti, hogy a tanárok

mennyire aggódnak a vizsga eredményei miatt, mennyire érzik magukat felelősnek és kínosan, ha a diákok rosszul teljesítenek a vizsgán.

29. táblázat. A tanárok vizsgával kapcsolatos nézeteivel és attitűdjeivel végzett faktoranalízis eredménye

Nr.	Kérdőív-tétel	Faktor				
		1.	2	3	4	5
<i>A kétszintű idegen nyelvi érettségi vizsga...</i>						
5.	a Nemzeti alaptantervben deklarált célokot a korábbi vizsgánál jobban méri.	0,85				
6.	elősegíti a nyelvi készségek fejlesztését a nyelvoktatásban.	0,83				
3.	a korábbi vizsgánál jobban szolgálja a felsőoktatás igényeit.	0,80				
2.	a diákok nyelvtudását a korábbi vizsgánál jobban méri.	0,80				
7.	a nyelvoktatásban tartalmi és módszertani modernizációt indított el.	0,73				
8.	a korábbi vizsgánál jobban segíti a javítói és értékelői munkámat.	0,70				
4.	illeszkedik a Közös Európai Referenciakeret szintrendszeréhez.	0,69				
1.	a diákoknak a korábbi vizsgánál nagyobb jogosítványokat nyújt.	0,60				
13.	Az érettségi vizsgaeredmények befolyásolják, hogy a tanulók és a kollégák engem jó tanárnak tartanak-e.		0,83			
15.	Felelősnek érzem magam, ha a tanulók nem sikeresek az érettségi vizsgán.		0,72			
12.	Kínosan érzem magam, ha az érettségin a diákjaim rosszabbul teljesítenek, mint a kollégám diákjai.		0,60			
14.	Úgy érzem, hogy az érettségi megbízható visszajelzést ad a tanításom minőségéről.	0,53	0,56			
18.	Szükségesnek érzem a tankönyvi feladatok kiegészítését az érettségre felkészítő feladatgyűjteményekből származó mérőfeladatokkal.			0,78		
11.	Mióta bevezették a kétszintű érettségit, több időre van szükségem az óráimra való felkészüléshez.		0,42	0,58		
9.	Jónak tartom, hogy az érettségin van külön nyelvhelyeséget mérő vizsgarész.				0,85	
16.	Szükségesnek tartom, hogy a közvetítési készség mérése is megjelenjen az emelt szintű érettségin.				0,54	
19.	Az érettségre felkészítő feladatgyűjteményekben szereplő mérőfeladatokkal megfelelően tudom szimulálni a vizsgaszituációt.					0,84
17.	Az óráimon használt tankönyv alkalmas az érettségin mért készségek fejlesztésére.			0,49		0,53
<i>Megmagyarázott variancia (%)</i>		27,48	12,17	9,14	7,16	6,80

Megjegyzés: A táblázatban a 0,40-nél nagyobb faktorsúlyokat tüntettük föl. A KMO-index: 0,816; a Bartlett-teszt szignifikanciája: $p < 0,001$; a teljes megmagyarázott variancia: 62,75%.

A 14. kérdőívtételt többdimenzionalitás jellemzi, mert 0,53-as faktorsúllyal az első faktorba is besorolható. A kérdőív esetleges továbbfejlesztése során érdemes mérlegelni a kérdőívtétel átfogalmazását (például: *A kétszintű idegen nyelvi érettségi vizsga megbízható visszajelzést ad a tanításom minőségéről*), amely alapján ez az állítás az első fak-

torba kerülne és a nyelvoktatásra vonatkozó nézetek egyik változóját jelölné. E kérdőív-tétel helyét a teljes kérdőívre vonatkozó faktoranalízis eredménye alapján döntjük majd el. A továbbiakban a 2. faktort a *vizsga következményeivel kapcsolatos tanári attitűdök*-nek nevezzük.

A 3. faktorhoz tartozó két item megegyezik a kérdőív struktúrájában elfoglalt hely-lyel. Ez a faktor a *tanárok attitűdjeit tartalmazza a tanórára való felkészüléssel kapcsola-tban*. Ugyanakkor a 11. kérdőív-tétel összefüggésben van a 2. faktorról is.

A 4. faktorba két állítás tartozik, melyek közül a 16. item egy vizsgálással kapcsolatos attitűdöt, a 9. item egy, a vizsga megbízhatóságára vonatkozó nézetet jelöl. Közös a két állításban, hogy egy-egy vizsgarészére vonatkoznak és a vizsga mérési-értékelési eljárásainak megbízhatóságát jelölik. Ez a két állítás a *tanárok nézetei a vizsga megbízható-ságáról* faktorként értelmezhető.

Az 5. faktor a tanárok attitűdjeit tartalmazza a vizsgára felkészítő tananyagokkal kapcsolatban, amely egyrészt a tankönyv (17. item), másrészt az érettségire felkészítő feladatgyűjtemények alkalmasságára (19. kérdőív-tétel) vonatkozik. E két változó a kér-dőívstruktúrának megfelelően a *vizsgára felkészítő tananyagokkal kapcsolatos attitűdö-ket* jelöli, ugyanakkor a 17. itemet többdimenzionalitás jellemzi, mivel hasonló faktor-súllyal a tanórai munka tervezése faktorba is besorolhatjuk. A 4. és 5. faktorba sorolt változók helye bizonytalan, feltételezzük, hogy ezek rontják a kérdőív validitását. Így a teljes kérdőívre vonatkozó faktoranalízis alapján határozunk arról, hogy e két faktor változóiból képezünk-e összevont mutatót.

A második kérdőív-részben lévő nyelvtanítási gyakorlatot befolyásoló tényezőkkel végzett faktoranalízis eredményét a 30. táblázat mutatja. A KMO-index az éppen elfo-gadható 0,6 értéknél helyezhető el, érdemes megvizsgálni a létrejött faktorokat. Ezek alapján két faktor definiálható: az elsőbe a *nyelvtanítási folyamatot befolyásoló külső, tárgyi tényezőket*, míg a másodikba a *nyelvtanítási folyamatot befolyásoló tanári ténye-zőket* sorolhatjuk. Ez a változóstruktúra nagyjából megegyezik a kérdőív szerkezetével (26. táblázat), ugyanakkor a 26. item (nyelvtanulók elvárásai) az előzetes elvárásainkkal ellentétben a 2. faktorba került, melyben a legkisebb faktorsúllyal rendelkezik.

30. táblázat. „A tanárok nézetei az oktatási folyamatot befolyásoló tényezőkről” kérdőív-részben lévő változókkal végzett faktoranalízis eredménye

Nr.	Kérdőív-tétel	Faktor	
		1	2
	<i>A nyelvtanítási gyakorlatot befolyásolják...</i>		
25.	a kétszintű nyelvi érettségi korábbi vizsgafeladatai.	0,87	
24.	a kétszintű nyelvi érettségi vizsgakövetelményei.	0,84	
23.	az aktuálisan alkalmazott nyelvkönyv.	0,50	
20.	a felsőoktatásban, a nyelvtanárképzésben tanultak.		0,72
21.	a módszertani továbbképzésen tanultak.		0,65
22.	a korábbi nyelvoktatási tapasztalataim.		0,60
26.	a nyelvtanulók elvárásai.		0,45
	<i>Megmagyarázott variancia (%)</i>	27,06	23,80

Megjegyzés: A táblázatban a 0,40-nél nagyobb faktorsúlyokat tüntettük föl. A KMO-index: 0,591; a Bartlett-teszt szignifikanciája: $p < 0,001$; a teljes megmagyarázott variancia: 50,85%.

A harmadik kérdőív-részben az állításokat kétféle módon jelöltük. A szám jelöli a kérdőív-tételt, az „a” betű azt az alskálát, amelyben a tanároknak jelölniük kellett, hogy az adott nyelvtanítási gyakorlat mennyire jellemző. A „b” betű az érettségi hatásának mértékét reprezentáló alskálát jelöli (31. táblázat). A változórendszerben a parciális kor-

relációk elfogadhatóan kicsik (KMO=0,81), és a változók páronként nem függetlenek (Bartlett-teszt szignifikanciája: $p < 0,001$). A faktoranalízis 6 dimenziót hozott létre, amelyek a változórendszer varianciájának 64%-át magyarázzák.

31. táblázat. „A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról” kérdőív rész változóival végzett faktoranalízis eredménye

Nr.	Kérdőív tétel	Faktor					
		1	2	3	4	5	6
35b	Változatos munkaformákat alkalmazok.	0,84					
30b	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.	0,83					
34b	A nyelvórákat a nyelvi szituációk gyakorlására építem.	0,79					
36b	Változatos, szövegközpontú feladattípusokat használlok.	0,78					
29b	A nyelvórákon a célnyelvet használom.	0,77					
31b	A nyelvórákon autentikus szövegekkel dolgozom.	0,75					
33b	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.	0,69					
27b	A hallott szöveg értésének fejlesztése nagy szerephez jut az óráimon.	0,63					
28b	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.	0,62					
28a	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.		0,73				0,44
35a	Változatos munkaformákat alkalmazok.		0,64				
36a	Változatos, szövegközpontú feladattípusokat használlok.		0,58				
31a	A nyelvórákon autentikus szövegekkel dolgozom.			0,82			
33a	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.			0,75			
37b	A produktív készségek mérésénél értékelési szempontokat és skálákat használlok.				0,78		
37a	A produktív készségek mérésénél értékelési szempontokat és skálákat használlok.				0,75		
32b	A diákoknak fordítási feladatokat adok.				0,44		
32a	A diákoknak fordítási feladatokat adok.				-0,51		
29a	A nyelvórákon a célnyelvet használom.					0,80	
30a	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.		0,50			0,64	
34a	A nyelvórákat a nyelvi szituációk gyakorlására építem.					0,50	
27a	A hallott szöveg értésének fejlesztése nagy szerephez jut az óráimon.						0,75
Megmagyarázott variancia (%)		24,21	9,60	8,40	8,30	7,77	6,19

Megjegyzés: A táblázatban a 0,40-nél nagyobb faktorsúlyokat tüntettük föl. A KMO-index: 0,805; a Bartlett-teszt szignifikanciája: $p < 0,001$; a teljes megmagyarázott variancia: 64,47%.

Az első, legnagyobb megmagyarázott varianciával rendelkező faktorba kilenc olyan változó került, amelyek az érettségi nyelvoktatási gyakorlatra kifejtett hatását jelölik. Ezek az állítások elkülönülnek az első alskálával vizsgált tanítási gyakorlatra vonatkozó kérdőív tételaktől, melyeket a vizsga jellemzői alapján fogalmaztunk meg (27. táblázat).

A 2. és 5. faktor jellemzőit egyben tárgyaljuk. A második faktorba 4 változó került, ugyanakkor ez a változórendszer átfedésben van az ötödik faktorial. A két faktort összekapcsolja a nyelvi készségek folyamatos fejlesztését tartalmazó 30a állítás, amely a kommunikatív nyelvoktatás alapvető célját jelöli. A 2. faktor a nyelvoktatás módszereihez kapcsolható változókat (a változatos munkaformában történő szövegközpontú feladatokra épülő differenciált készségfejlesztő nyelvoktatás jellemzőit) tartalmazza. Az 5. faktor a nyelvoktatási gyakorlat főbb jellemzőire, a célnyelv órai használatára és a nyelvi szituációk alkalmazására vonatkozik.

A 3. faktor a nyelvoktatás tartalmát jelölő két kérdőívtételt tartalmazza. Ugyanakkor a kérdőívstruktúra alapján (26. táblázat) ide tartozna a 27. állítás is, amely önállóan a 6. faktorba került. Feltételezzük, hogy itt validitás-problémáról van szó, amelynek az lehet az oka, hogy a példában a 27. állításhoz hasonló megfogalmazást alkalmaztunk (*Az olvasott szöveg értésének fejlesztése nagy szerephez jut az óráimon*), így a kérdőív esetleges továbbfejlesztése során mindenféleképp megfontolandó a példa szükségessége vagy cseréje. Külön faktorba került még a 32. állítás is, amely az első alskálával negatív faktorsúllyal szerepel a 4. faktorban. Így célszerű a 27a és 32. item elhagyása a kérdőívből.

A 4. faktorba került a 32. és 37. állítás mindkét alskálával. Amennyiben a faktoranalízist a törölt állítások kihagyásával végezzük el (8. melléklet), akkor a 37. item mindkét alskálával egy faktorba kerül. Ennek feltehető oka, hogy a 37. kérdőívtételnél a két skálán mért állítás faktorsúlya nagyon hasonló értékű (31. táblázat), a kettő között erős korreláció van ($r=0,49$; $p<0,001$). Ebből arra következtetünk, hogy az értékelési szempontok és skálák használata a produktív készségek mérésénél és értékelésénél az érettségi hatására vált többnyire jellemzővé, ezért kerül ez az állítás külön faktorba.

A negyedik kérdőív részben a tanároknak különböző értékelési kritériumoknál kellett eldönteniük, hogy mennyire tartják azokat fontosnak a saját mérőfeladataik elkészítésében. A faktoranalízis során három faktor jött létre, amelyekkel a változórendszer 51%-a magyarázható (32. táblázat).

32. táblázat. „A tanárok nézetei az idegen nyelvi tesztelés kritériumairól” kérdőív részben lévő változókkal végzett faktoranalízis eredménye

Nr.	Kérdőívtétel	Faktor		
		1	2	3
	<i>A saját magam által készített mérőfeladat...</i>			
48.	épüljön a Közös Európai Referenciakeret szintrendszerére.	0,77		
39.	épüljön autentikus szövegekre.	0,66		
49.	igazodjon a kétszintű idegen nyelvi érettségi szemléletéhez.	0,62		
50.	épüljön olyan szövegekre, szituációkra, amelyeket a diákok szívesen oldanak meg.	0,59		
45.	csak a nyelvtudást, az általános műveltséget, háttérismeretet ne mérje.	0,60		
42.	mérje külön az egyes nyelvi készségeket.	0,57		
43.	tükrözzön reális, a valóságban is várható nyelvi szituációt.	0,53		
47.	a tanuló formális nyelvtani ismereteit mérje.		0,76	
46.	mérje a tananyag elsajátításának mértékét.		0,75	
44.	a tanuló szókincsét célzottan mérje.		0,68	
41.	legyen objektíven értékelhető.			0,72
38.	mérjen megbízhatóan, pontosan.			0,68
40.	eredménye segítse elő a nyelvtanuló előrehaladását.			0,64
	<i>Megmagyarázott variancia (%)</i>	21,63	15,57	14,23

Megjegyzés: A táblázatban a 0,40-nél nagyobb faktorsúlyokat tüntettük föl. A KMO-index: 0,741; a Bartlett-teszt szignifikanciája: $p<0,001$; a teljes magyarázott variancia: 51,43%.

Az 1. faktorba sorolt változók többnyire a kommunikatív tesztelésre jellemző kritériumokat (28. táblázat) jelölik, bár ez a faktor nem tartalmazza a 40. itemet és ebbe került a kompetencia tesztekre jellemző 42. állítás is. A 2. faktor a validitás különböző aspektusait jelöli, amely gyakran megjelenik a tanítási-tanulási folyamatban, a 3. faktor pedig a jóságmutatókat tartalmazza.

Az ötödik kérdőív-részben a központilag összeállított vizsgafeladatok tanítási-tanulási folyamatban való alkalmazásának célját leíró állításokat soroltuk fel. Ezek alapján csak egy faktor jött létre (33. táblázat), azaz a tanárok válaszmintázatai alapján – előzetes elvárásainkkal ellentétben – nem osztható két részre a vizsgafeladatok felhasználásának lehetséges valódi funkciói (51, 53, 54) és diszfunkciói (motiváció és készségfejlesztés).

33. táblázat. „A tanárok nézetei a vizsgafeladatok oktatási folyamatban való felhasználásának céljairól” kérdőív-rész változóival végzett faktoranalízis eredménye

Nr.	Kérdőív-tétel	Faktorsúly
<i>A központilag összeállított vizsgafeladatok tanítási-tanulási folyamatban való alkalmazásának célja, hogy...</i>		
52.	a diákokat tanulásra motiválja.	0,82
54.	megállapítsa, az adott készséggel milyen minőségben rendelkezik a diák.	0,74
55.	a diákok nyelvi készségeit fejlessze.	0,72
53.	a diákokat felkészítse az érettségi vizsgán a konkrét teendőkre.	0,71
51.	a diáknak a vizsgán várható teljesítményéről nyújtson információkat.	0,70

Megjegyzés: A táblázatban a 0,40-nél nagyobb faktorsúlyokat tüntettük föl. A KMO-index: 0,732; a Bartlett-teszt szignifikanciája: $p < 0,001$; a teljes megmagyarázott variancia: 54,37%.

Az egyes kérdőív-részek elemzését követően a teljes kérdőív-re vonatkoztatva is elvégeztük a faktoranalízist, amelynek eredményét a 9. melléklet tartalmazza, az egyes faktorokról a 34. táblázatban adunk áttekintést. A faktoranalízisbe nem vontuk be az alacsony megbízhatósággal rendelkező II. kérdőív-rész változóit. Az alacsony reliabilitás (Cronbach- $\alpha=0,56$) oka, hogy a 20–26. kérdőív-tételeknél a nyelvtanárok nem használták ki a teljes ötfokozatú skálát, homogén a minta, összességében ez a rész nem illik a teljes kérdőív-be, így ebből a kérdőív-részből nem hozunk létre összevont változót. Kihagytuk a faktoranalízisből a nem megfelelő elkülönítésmutatóval rendelkező 10. itemet és a validitás-problémát mutató 27a, 32a és 32b itemet. A faktoranalízis komponenseibe nem került be az alacsony faktorsúllyal rendelkező 9. item és kihagytuk a negatív faktorsúlyú (-0,4) 17. itemet. Ezek az állítások a vizsgával kapcsolatos nézetekhez és attitűdökhöz tartoznak (29. táblázat). A 16. és 19. kérdőív-tételek önállóan külön faktorba kerültek, ezért ezeket is kihagytuk a faktoranalízisből. A 66 itemből összesen 14 kérdőív-tételt hagytunk el. Az így lebonolyított faktoranalízis 13 komponensből áll (KMO-index=0,73), amelyekkel a változórendszer 67%-ban magyarázható (9. melléklet).

A teljes kérdőív-re vonatkozó 34. táblázat adatainak a 26. táblázattal való összehasonlítása alapján megállapítható, hogy a tanárok nézetei a vizsga jelentőségéről nem differenciálódnak. Az első kérdőív-részben kérdéses 14. állítás (*Úgy érzem, hogy az érettségi megbízható visszajelzést ad a tanításom minőségéről*) egyértelműen a tanárok nézetei a vizsga jelentőségéről faktorba került, így a továbbiakban oda soroljuk. A harmadik kérdőív-részből a faktoranalízis alapján néhány itemet ki kellett hagyni, de a kérdések összességében többnyire megegyeznek az előzetes változórendszerrel. A 8. és 10. faktort összekötő 30a állítást (*A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel*) mindkét faktorba besoroltuk. A 36a item az előzetes kérdőív-struktúrával (26. táblázat) és a kérdőív faktoranalízisével ellentétben a vizsga és az okta-

tás tartalma közti kapcsolatról kialakított nézetek faktorba került a 31a és 32a változók mellé (9. melléklet). Így a továbbiakban e három kérdőívteletből hozunk létre összevont változót. Jelentősebb különbségeket találtunk a negyedik és ötödik kérdőív részben. A negyedik részben az itemek helyét újra be kellett azonosítanunk, más rendezőelvet kellett használnunk. Feltételezzük, hogy ennek a kérdőív résznek alacsonyabb a validitása. Ebben három probléma jelenhet meg: (1) feltételezhető, hogy a nyelvtanárok az adott kritériumokon nem a szakirodalomban megadott jelentéstartalmakat értik, (2) elképzelhető, hogy az adott kritérium átfordítása a tanítási-tanulási folyamatban alkalmazott mérőeszközökre kevésbé sikerült, (3) valószínűsíthető, hogy a kérdésfeltevés nem volt egyértelmű. Az 5. és 11. faktor között átfedések vannak, így ezekből közös változót hozunk létre. Az ötödik kérdőív részben az előzetesen feltételezett kétféle kérdőív szerkezet empirikusan nem igazolható. A további elemzéshez a 34. táblázat alapján hozunk létre összevont változókat.

34. táblázat. A teljes kérdőívre végzett faktoranalízis eredménye

Kérdőív rész	Vizsgált tényező	Item	Item-szám	Faktor
I. A tanárok vizsgával kapcsolatos nézetei, attitűdjei	– a tanárok nézetei a vizsga jelentőségéről	1–8, 14	9	2
	– a vizsga következményeivel kapcsolatos attitűdök	12, 13, 15	3	6
	– a tanórára való felkészüléssel kapcsolatos attitűdök	11, 18	2	13
III. A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról	– a vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézetek	27b–31b, 33b–36b	9	1
	– az eredményes tanítási módszerekről kialakított nézetek	28a, 30a, 35a	3	8
	– a vizsga és a tanítási gyakorlat közti kapcsolatról kialakított nézetek	29a, 30a, 34a	3	10
	– a vizsga és az oktatás tartalma közti kapcsolatról kialakított nézetek	31a, 33a, 36a	3	4
IV. A tanárok nézetei az idegen nyelvi tesztelés kritériumairól	– a vizsga és az osztálytermi értékelési eljárások közti kapcsolatról kialakított nézetek	37a, 37b	2	9
	– a klasszikus tesztelmélet jóságmutatóinak fontosságáról kialakított nézetek	38, 40, 41	3	12
	– a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek	39, 42, 43, 45, 48–50	7	5, 11
V. A tanárok nézetei a vizsgafeladatok oktatási folyamatban való alkalmazásának céljairól	– a validitás fontosságáról kialakított nézetek	44, 46, 47	3	7
		51–55	5	3

A kérdőív validitásának ellenőrzését követően a teljes kérdőív reliabilitása (Cronbach- $\alpha=0,89$) megfelelő, az egyes kérdőív részek megbízhatósági értékeit a 35. táblázat tartalmazza. A nyelvtanítási gyakorlatra kérdező itemek reliabilitását a két alskála együttes vizsgálata alapján adtuk meg a 35. táblázatban. Az első alskálával vizsgált nyelvtanítási gyakorlatra vonatkozó állítások alacsonyabb megbízhatósággal (Cronbach- $\alpha=0,74$) mérnek, mint az érettségi hatásának megítélése vonatkozó (Cronbach- $\alpha=0,89$). Az első esetben lévő alacsonyabb megbízhatóság feltehető oka, hogy a nyelvtanárok ismerik a kommunikatív nyelvoktatás alapelveit, amelyekhez a válaszaikban igazodnak, így önmagukban ezek az állítások kevésbé adnak megbízható képet a nyelvtanítási gyakorlatról. Több hazai, az osztálytermi megfigyelések tapasztalatait vizsgáló és a diákok véleményét is feltáró kutatások (Nikolov, 1999, 2003; Nikolov, Ottó és Öveges, 2008) általában a tanárok nyilatkozatainál kedvezőtlenebb

eredményeket mutatnak. Ugyanakkor főként azt vizsgáltuk, hogy a tanárok milyen hatások meglétéről nyilatkoznak, célunk az érettségi és a nyelvoktatási gyakorlat közötti kapcsolat feltárása, amelynek vizsgálatára a felsorolt állítások megbízhatónak tűnnek.

35. táblázat. A kérdőív egyes részeihez tartozó itemek száma és megbízhatósága

<i>Kérdőív-rész</i>	<i>Itemszám</i>	<i>Cronbach-a</i>
I. A tanárok vizsgával kapcsolatos nézetei, attitűdjei	14	0,83
III. A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról	19	0,82
IV. A tanárok nézetei az idegen nyelvi tesztelés kritériumairól	13	0,77
V. A tanárok nézetei a vizsgafeladatok oktatási folyamatban való alkalmazásának céljairól	5	0,79

A vizsgálat lebonyolítása

A kérdőív internetes adatfelvételére 2009. májusában került sor. Ebben az időszakban indult el a 2009. évi érettségi vizsgaidőszak. Azért ezt választottuk az adatfelvétel lebonyolítására, mert ekkor az érettségizető nyelvtanárok már nem tanítanak a 12. évfolyamos osztályokban, az írásbeli dolgozatok megírása ugyan már elkezdődött, de a tanároknak még nem kellett javítaniuk a dolgozatokat. Ez az időszak továbbá biztosította a kérdőíves felmérésünk aktualitását.

Először a nyelvtanároknak egy felkérő levelet küldtünk ki (5. melléklet), amelyben felhívtuk a figyelmüket a kérdőívre. Az internetes kérdőívet a Szegedi Tudományegyetem Neveléstudományi Intézet honlapján helyeztük el⁷. A kérdőív kitöltése névtelen és önkéntes volt, átlagosan 25 percet vett igénybe. Úgy alakítottuk az adatfelvételt, hogy a tanárok a kérdőívet az adatfelvétel időszakában elmenthették és bármikor folytathatták a kitöltést. A névtelenséget úgy garantáltuk, hogy a kérdőív hozzáférésehez egy előzetesen generált kódszám megadására kértük a kollégákat. A kódszám azt szolgálta, hogy egy válaszadó csak egyszer töltsse ki a kérdőívet, amelyet az adatfeldolgozás és -elemzés során semmilyen formában nem hoztunk kapcsolatba a kérdőívet kitöltő személlyel, a kapott egyéni adatokat nem dolgoztuk fel külön, és harmadik fél részére nem adtuk ki.

A felkérő levélben megadott adatfelvétel lejárta előtt 107-en töltötték ki a teljes kérdőívet, ezért döntöttünk úgy, hogy az adatfelvételt egy héttel meghosszabbítjuk és erről a kollégákat egy tájékoztató levélben (6. melléklet) értesítettük. Ennek hatására az adatfelvétel zárásának időpontjáig 160-an töltötték ki a kérdőívet hiánytalanul.

Az adatfelvétel lezárásakor a 7. mellékletben található köszönő levelet küldtük ki a kollégáknak. A kutatásban részt vett nyelvtanárok számára röviden összegeztük a vizsgálat legfontosabb eredményeit.

⁷ http://www.edu.u-szeged.hu/nt/_tanarikerdoiv/index.php

4. AZ IDEGEN NYELVI ÉRETTSÉGI ELEMZÉSE A VIZSGAEREDMÉNYEK ALAPJÁN

Ebben a fejezetben az értekezés koncepcionális részében definiált négy területből hármat elemzünk: (1) az érettségien elért készségenkénti teljesítmények változásait és az azt meghatározó tényezőket, melyhez az OH nyilvánosan közzétett adatbázisait használtuk; (2) a receptív készséget mérő vizsgarészekben az érettségi külső és belső validitását, valamint (3) az íráskészséget mérő vizsgarészben alkalmazott mérési-értékelési eljárások minőségét az OFI adatbázisai alapján. Közös jellemzője ezeknek a vizsgálatoknak, hogy az elméleti modellnek megfelelően az eredményekből indulnak ki és az igények és célok változásaira, valamint a vizsgafejlesztési folyamatra irányulnak.

4.1. Az idegen nyelvi érettségien elért teljesítmények

Ebben az alfejezetben elemezzük a jelenlegi vizsga működését, a készségenkénti teljesítmények változását és az azokban megmutatkozó különbségeket a két leggyakrabban tanított idegen nyelv: az angol és a német esetében. E fejezetben bemutatandó elemzések kapcsán szükséges utalni arra, hogy az OH az adatbázisokból csak általános tendenciákra lehet következtetni, de ezek okait nem tudjuk vizsgálni, az adatbázisok alkalmatlanok a vizsga minőségének részletes elemzésére (erről lásd részletesebben a 3.3.1. fejezetet). Vizsgálatunkba csak azoknak az érettségizőknek az adatait vontuk be, akik mind a vizsga írásbeli, mind a szóbeli részén részt vettek. Így az adatbázisokból először töröltük azon vizsgázók adatait, akik valamely vizsgarészen nem jelentek meg, illetve az Országos Középiskolai Tanulmányi Versenyen (OKTV) vagy a Szakmai Előkészítő Érettségi Tárgyak Versenyen (SZÉTV) elért helyezések alapján kaptak mentességet, továbbá azokat is, akik az írásbelit szóban vagy a szóbelit írásban tették le. Hiányzó adatként kezeltük azoknak a vizsgázóknak az adatait, akik valamely vizsgarész alól mentességet kaptak (ez elsősorban a hallott szöveg értése vizsgarészt érintette).

A továbbiakban először áttekintést adunk a 2007., 2008. és 2009. évi május-júniusi angol és német nyelvi érettségi vizsgázói összetételéről, majd a középszintű és végül az emelt szintű vizsga eredményeit elemezzük a három évben. Ebben megvizsgáljuk a készségenkénti teljesítményeket és -változásait és az átlagteljesítményeket iskolatípusok, képzéstípusok és nemek szerinti bontásban mindkét nyelv esetében. Végül összehasonlítjuk a közép- és emelt szintű vizsga elemzésének eredményeit.

4.1.1. A vizsgázók összetétele a 2007., 2008. és 2009. évi idegen nyelvi érettségien

A vizsgázói összetétel változásának elemzése az értekezés célkitűzései szempontjából azért releváns, mert így megvizsgálhatjuk, hogy az idegen nyelvi érettségien részt vevő tanulók mely vizsgaszintet választják, az egyes vizsgafajtáknál megfigyelhető-e átrendeződés, megjelennek-e újabb vizsgázói csoportok. Ezzel a vizsgálattal az elméleti modellben lévő negyedik, a tanulóktól a vizsgafejlesztésre irányuló visszacsatolási körről szerezhetünk információkat (11. ábra). Ugyanakkor ezt nem jelöltük meg vizsgálandó célterületként, mivel az OH adatbázisai alapján kizárólag a vizsgázók választásának eredményéről és a vizsgázói összetételről szerezhetünk információkat különböző háttérváltozók szerint, a kutatásunk így nem terjedt ki a vizsgaszintek és -fajták választását meghatározó tényezők feltárására. Ez egy következő kutatás része lehet, amelyhez a továbbiakban bemutatandó eredmények kiindulási pontot jelenthetnek.

A 36. és 37. táblázat mutatja a három vizsgált évben a vizsgázók eloszlását a különböző vizsgafajták szerint közép- és emelt szinten. A két vizsgaszint közül az emelt szintet angolból a vizsgázók 8%-a választotta mindhárom évben, németből kevesebben, 2007 és 2008-ban 6%, míg 2009-ben 7%. Ez az arány nem változott a 2005. és 2006. évi vizsgákhoz képest (2.3.4. fejezet). A két nyelv közül többségében az angolt választották az érettségizők: középszinten a három év átlagában a vizsgázók 64%-a, emelt szinten 68%-a. Az angolból érettségizők száma kissé növekszik az évek múlásával.

36. táblázat. A középszinten vizsgázók száma és aránya a vizsgafajták szerint angolból és németből

Vizsgafajta	Angol						Német					
	2007		2008		2009		2007		2008		2009	
	N	%	N	%	N	%	N	%	N	%	N	%
Előrehozott	4891	11,0	8007	17,8	9008	18,1	3260	11,6	4570	17,5	4628	17,7
Ismétlő	137	0,3	272	0,6	336	0,7	45	0,2	75	0,3	90	0,3
Javító	87	0,2	105	0,2	236	0,5	44	0,2	72	0,3	77	0,3
Kiegészítő	38	0,1	63	0,1	79	0,2	8	0,0	26	0,1	24	0,1
Pótló	26	0,1	18	0,0	32	0,1	4	0,0	10	0,0	4	0,0
Rendes	39223	88,3	36509	81,2	40178	80,6	24793	88,1	21312	81,8	21352	81,6
Összesen	44402	100,0	44974	100,0	49869	100,0	28154	100,0	26065	100,0	26175	100,0

37. táblázat. Az emelt szinten vizsgázók száma és aránya a vizsgafajták szerint angolból és németből

Vizsgafajta	Angol						Német					
	2007		2008		2009		2007		2008		2009	
	N	%	N	%	N	%	N	%	N	%	N	%
Előrehozott	405	11,1	525	14,2	427	10,5	199	11,0	217	12,9	177	9,1
Ismétlő	183	5,0	121	3,3	114	2,8	39	2,2	29	1,7	29	1,5
Javító	7	0,2	9	0,2	45	1,1	5	0,3	2	0,1	8	0,4
Kiegészítő	140	3,8	122	3,3	133	3,3	40	2,2	35	2,1	24	1,2
Pótló	5	0,1	2	0,01	5	0,1	0	0,0	3	0,2	0	0,0
Rendes	2330	63,6	2065	55,9	2021	49,8	1189	65,9	1000	59,7	1061	54,7
Szintemelő	592	16,2	847	22,9	1310	32,3	333	18,4	390	23,3	640	33,01
Összesen	3662	100,0	3691	100,0	4055	100,0	1805	100,0	1676	100,0	1939	100,0

Középszinten mindkét nyelvből a többség a középiskolai tanulmányok befejezése után, a tanulói jogviszony fennállása alatt vagy megszűnését követően első alkalommal tett (rendes) vizsgát (36. táblázat). Ugyanakkor a középiskolai tanulmányok teljes befejezése előtt, első alkalommal vizsgázók száma mindkét nyelvnél középszinten növekszik: míg 2007-ben a vizsgázók 11%-a választotta az előrehozott vizsgát, addig 2008 és 2009-ben 18%. Emelt szinten átlagosan 11–12% tett előrehozott vizsgát (37. táblázat). Az OH adatbázisaiban a vizsgázó évfolyamát is megadják, amely alapján kiderül, hogy például 2009-ben a középszintű vizsgán mindkét nyelvből a vizsgázók fele, 11. osztályban tett előrehozott vizsgát, tőlük kevesebben 10.-ben, de előfordul, hogy már a 9. évfolyamon leérettségiztek ebből a tárgyból. Feltételezzük, hogy ezek a tanulók az iskolán kívüli nyelvtanulás során érték el a (nem túl magas) A2/B1 szintet. Kérdéses és vitatható ennek a vizsgafajtának a választhatósága az idegen nyelvi érettségi esetében, hiszen egy olyan kötelező tárgyról van szó, amelyet a középiskolai tanulmányok befejezéséig

tanulnak a diákok. Mindezek mellett a hazai vizsgálatok következetesen azt mutatják, hogy a diákok csak annyira akarják megtanulni az idegen nyelvet, hogy leérettségizzenek belőle (lásd 2.2. fejezet). Részletes elemzést igényelne, hogy a tanulók miért választják ezt a vizsgafajtát, továbbá milyen hatása van ennek a lehetőségnek a nyelvtanulási folyamatra. E vizsgálat alapján lehetne módosítási igényeket megfogalmazni a vizsgaszabályozás és -fejlesztés számára.

Emelt szinten a vizsgázók összetétele elsősorban három vizsgafajtára épül (37. táblázat). Az előrehozott vizsga mellett mindkét nyelvre jellemző, hogy az emelt szintű vizsgára a középiskolai tanulmányok befejeztével egyre kevesebben jelentkeztek, de nőtt azoknak a száma, akik korábban már érettségiztek és azután tettek szintemelő vizsgát, vagyis az adatokból átrendeződés figyelhető meg a két vizsgafajta között.

Mindkét szinten a javító és a pótló vizsgafajtát választók száma nagyon alacsony, ismétlő és kiegészítő vizsgát elsősorban emelt szinten tettek az érettségizők, de a teljes vizsgára vonatkoztatva e vizsgafajtákra való jelentkezés kevésbé jellemző.

A közép- és emelt szinten vizsgázók többsége nappali képzésben vett részt, a nyelvi érettségizők száma középszinten a felnőttoktatást befejezők körében alacsony (38. táblázat) és emelt szinten elhanyagolható. A vizsgázók alacsony részvételének az lehet a magyarázata, hogy az idegen nyelvi érettségi a felnőttképzés levelező és esti tagozatán nem kötelező tárgy, ehelyett az érettségizők a vizsgált időszakokban más tárgyból vizsgázhattak [100/1997. (VI. 13.) Korm. Rendelet az érettségi vizsga vizsgaszabályzatának kiadásáról 60/B. § (1) bekezdés]. Középszinten a felnőttoktatási intézményekben angolból átlagosan 6%, míg németből 8% tett érettségi vizsgát. A felnőttoktatásban a nappali és levelező/esti képzésben részt vettek aránya a középszintű vizsgán közel azonos, így ennél a két almintánál meg lehet vizsgálni, hogy a felnőttoktatásban melyik képzéstípus eredményesebb. Ezt a vizsgálatot két körülmény is indokolta teszi: egyrészt a felnőttoktatásban az idegen nyelvek tanításának eredményességéről nincs adatunk, másrészt a felnőttoktatásban érettségizők aránya az új szabályozást követően emelkedni fog. Ennek oka, hogy 2009/2010. tanév május-júniusi vizsgaidőszakától az esti és a levelező képzésben is kötelező vizsgatárgy lesz az idegen nyelv.

38. táblázat. A közép- és emelt szinten vizsgázók százalékos eloszlása a képzéstípusok szerint angolból és németből

Képzéstípus	Angol						Német					
	2007		2008		2009		2007		2008		2009	
	K	E	K	E	K	E	K	E	K	E	K	E
Nappali	93,4	84,3	93,2	87,8	94,0	87,1	91,9	90,0	91,4	91,9	93,1	93,1
Felnőttoktatás nappali	2,5	0,4	2,7	0,4	2,2	0,4	4,7	0,7	5,2	0,7	4,1	0,1
Felnőttoktatás esti/levelező	3,7	1,1	3,4	0,7	3,0	0,3	3,1	0,9	2,9	0,8	2,4	0,5
Nincs jogviszonya	0,4	14,2	0,7	11,2	0,9	12,2	0,3	8,5	0,4	6,7	0,5	6,2

Jelölések: K: középszint, E: emelt szint.

Tanulói jogviszony nélkül középszinten kevesen, míg emelt szinten a három vizsgált vizsgaidőszak átlagában angolból az érettségizők 13%-a, németből 7%-a tette le a vizsgát. Ennek oka, hogy a hatályos rendelkezések szerint [100/1997. (VI. 13.) Korm. Rendelet... 45. § (3) a) és b) bekezdés] az emelt szintű idegen nyelvi érettségien B1 vagy B2 szintű nyelvvizsga-bizonyítvány szerezhető. Feltehetően ezek a vizsgázók az emelt szintű érettségi nyelvvizsgával egyenértékű funkcióját használják ki vagy nyelv-

szakon kívánnak felsőoktatási intézményben tanulni. Érdeemes megvizsgálni, hogyan alakul e vizsgázók koreloszlása és milyen céllal választják az emelt szintű vizsgát. Ugyanakkor 2006 és 2009 között átlagosan 63 ezer 14–19 éves tanuló tett nyelvvizsgát, így ez a szabályozás nem változtatta meg lényegesen a közoktatásban tanuló korosztály részvételét a nyelvvizsgákon⁸.

Angol nyelvből a három év átlagában a gimnáziumot és a szakközépiskolát végzett érettségizők aránya középszinten közel azonos, míg németből gimnáziumban kevesebben, átlagosan a vizsgázók 43%-a érettségizett (39. táblázat). Emelt szinten a gimnáziumot befejezők aránya magasabb mindkét nyelvből: angolból átlagosan 62%, míg németből 59%. Az emelt szinten vizsgázók hiányzó adatai elsősorban azokra vonatkozik, akik nem rendelkeznek tanulói jogvisztonnyal.

39. táblázat. A közép- és emelt szinten vizsgázók százalékos eloszlása az iskolatípusok szerint angolból és németből

Iskolatípus	Angol						Német					
	2007		2008		2009		2007		2008		2009	
	K	E	K	E	K	E	K	E	K	E	K	E
Gimnázium	47,8	61,3	49,7	61,7	52,8	61,7	42,1	57,2	42,9	59,2	44,9	59,3
Szakközépiskola	51,9	24,6	49,6	25,0	46,4	25,0	57,7	33,9	56,7	31,6	54,6	33,2
Nincs adat	0,4	14,2	0,7	13,3	0,8	13,3	0,3	8,9	0,4	9,1	0,5	7,5

Jelölések: K: középszint, E: emelt szint.

Középszinten mindkét nyelvből nem változott lényegesen a vizsgázók számának regionális eloszlása (40. táblázat).

40. táblázat. A közép- és emelt szinten vizsgázók százalékos eloszlása régiók szerint angolból és németből

Régió	Angol						Német					
	2007		2008		2009		2007		2008		2009	
	K	E	K	E	K	E	K	E	K	E	K	E
Nyugat-Dunántúl	7,1	7,3	7,6	8,6	7,6	9,1	16,9	14,0	17,0	16,7	17,2	19,2
Közép-Dunántúl	9,7	9,0	9,5	8,4	9,7	9,4	12,3	7,5	11,9	7,8	12,3	7,4
Dél-Dunántúl	7,3	9,1	7,2	7,9	7,3	7,9	12,2	11,1	12,2	13,4	12,2	11,8
Észak-Magyarország	12,6	9,7	13,0	9,6	13,0	11,0	10,6	9,3	10,4	5,7	10,9	7,7
Észak-Alföld	17,3	10,6	17,0	10,7	17,1	10,5	14,9	12,6	15,6	12,9	15,0	10,3
Dél-Alföld	14,5	11,6	14,4	11,6	14,6	10,9	12,3	11,2	12,3	13,2	12,3	14,6
Közép-Magyarország	31,6	42,8	31,2	43,2	30,7	41,2	20,9	34,2	20,6	30,2	20,0	28,9

Jelölések: K: középszint, E: emelt szint.

A Közép-magyarországi régióban vizsgáztak legtöbbször mindkét vizsgaszinten. Angolból középszinten ebben a régióban négyszer többen érettségiztek, mint a Nyugat és Dél-dunántúli régióban, emelt szinten a vizsgázók többsége ebben a régióban tette le a vizsgát. Németből arányait tekintve közép- és emelt szinten a dunántúli régiókban vizsgáz-

⁸ Forrás: http://www.nyak.hu/doc/statisztika.asp?strId=_21_ (megtekintés: 2010. 02. 14.). A hivatkozott adat minden élő idegen nyelvre vonatkozik.

tak többen. Ez az arány összhangban van azzal, hogy a középiskolát megkezdő tanulók Dél- és Nyugat-Dunántúlon inkább németet tanulnak (Imre, 2007). Emelt szinten mindkét nyelvből főleg a Nyugat-dunántúli régióban növekszik a vizsgázók száma, míg a Dél-dunántúli és Észak-magyarországi régióban ingadozik.

4.1.2. Teljesítmények és a teljesítményt meghatározó változók hatása a középszintű érettségien

Az egyes készségek teljesíthettségét angol nyelvből a 41. táblázat, német nyelvből a 42. táblázat tartalmazza, melyben a vizsgapontszámot százalékos skálára transzformáltuk és megadtuk a teljesítményátlagot és a szórást. Az átlag standard hibája (melyet a továbbiakban std. hibaként rövidítünk) minden esetben 0,2 alatt van. Közvetlenül összehasonlíthatjuk az egyes vizsgaéveken belül az adatokat, mivel a tesztek ugyanazon vizsgázók oldották meg, így képet kaphatunk a vizsgarészekben mért készségek fejlettségéről. Közvetetten vizsgálhatók az egyes évek közötti teljesítmények, mivel az adatbázisokat nem köti össze közös személy vagy item, de minden feladatsor a vizsgakövetelményekre épül, a feladatok szerkezete, a feladatírás alapelve azonos, mindkét nyelvből a vizsgázói populáció nagy és összetétele a három évben nem változott jelentősen.

A vizsgázók átlagteljesítményét két egymással összefüggő tényező határozza meg leginkább: a diákok nyelvtudás-szintje és a feladatok nehézsége. Ez utóbbi főleg az átlagos képességszintű diákok teljesítményét befolyásolja. Az egyes vizsgaévekben az átlagteljesítményben megjelenő ingadozások eltérő nehézségű feladatokra utalnak, a teljesítménynövekedésben a felkészítés hatása mutatható ki, a csökkenő teljesítményátlagok azt jelezhetik, hogy a vizsgafeladatokat összeállító tantárgyi bizottság nehezíti a feladatsorokat. Ugyanakkor ennek pontos megállapítására feladat- és itemszintű elemzésekre van szükség, mely a 4.2. fejezet alapját fogja képezni. A 41. és a 42. táblázat adatainak értelmezésében először a vizsgaéveken belül elemezzük a készségek teljesíthettségét, majd a teljesítményátlagok változásaiból levonható következtetéseket fogalmazzuk meg.

Angol nyelvből a 2007. és 2009. évi középszintű érettségien a legkönnyebb vizsgarész az olvasott szöveg értése volt, ugyanakkor a 2008. évi vizsgán belül a többi készséghez képest itt érték el a vizsgázók a legalacsonyabb átlagteljesítményt (41. táblázat). Feltételezhetjük, hogy a feladatsorok nehézsége eltérő.

41. táblázat. A középszintű vizsga eredményei készségenként és a teljes vizsgán angoltól (%)

Vizsgarész	2007			2008			2009		
	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értése	71,25	16,70	0,08	55,95	23,96	0,11	73,01	23,01	0,10
Nyelvhelyesség	45,03	26,40	0,13	56,03	23,51	0,11	53,03	27,00	0,12
Hallott szöveg értése	62,30	20,37	0,10	65,21	19,55	0,09	62,02	19,22	0,09
Íráskészség	60,16	31,01	0,15	60,94	32,73	0,15	65,04	31,66	0,14
Beszéd-készség	64,89	26,15	0,12	67,58	26,62	0,13	68,82	26,71	0,12
<i>Teljes vizsga</i>	61,97	20,31	0,10	61,32	22,13	0,10	65,19	22,33	0,10

Az érettségizők könnyen teljesítették a belső vizsgaként működő szóbeli érettségét, amelyben elért átlagteljesítmény mindhárom évben szignifikánsan különbözik a teljes vizsgától. A hallott szöveg értése vizsgarésznél a többihez képest homogénebb a vizsgázói populáció, az íráskészségnél viszont heterogén, azaz ennél a vizsgarésznél egyes

háttértényezők szerint jelentős különbségek lehetnek. A többi készséghez képest mindhárom évben a legnehezebb feladatsor a nyelvhelyesség, ahol a legalacsonyabbak a teljesítményátlagok.

A középszintű érettségén német nyelvből mindhárom évben az írás- és beszédképességnél heterogén a populáció, míg a nyelvhelyességnél homogén (42. táblázat). A legkönnyebben teljesíthető vizsgarész az olvasott szöveg értése és a beszédképesség. A hallott szöveg értése feladatsor teljesítettsége németből 2007-ben a második legmagasabb, 2008-ban a harmadik, míg 2009-ben a többi készséghez képest ezt teljesítették a legkevésbé az érettségizők. Feltételezzük, hogy e feladatsor nehézsége a három évben jelentősen eltér egymástól. Ez a teszt kevés itemet tartalmaz, így komoly kihívás a vizsgarész nehézségének állandóságát úgy biztosítani, hogy a feladatokat előzetesen nem próbálják ki.

42. táblázat. A középszintű vizsga eredményei készségenként és a teljes vizsgán németből (%)

Vizsgarész	2007			2008			2009		
	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értése	73,45	20,51	0,12	62,26	20,87	0,13	60,18	23,39	0,14
Nyelvhelyesség	47,60	17,52	0,10	52,55	20,23	0,13	55,92	21,36	0,13
Hallott szöveg értése	70,48	19,89	0,12	54,20	23,25	0,14	50,33	18,20	0,11
Írásképesség	48,05	33,26	0,20	51,90	33,59	0,21	54,25	33,61	0,21
Beszédképesség	62,94	26,67	0,16	65,01	27,21	0,17	66,52	27,18	0,17
Teljes vizsga	61,45	20,68	0,12	57,31	22,22	0,14	57,26	21,70	0,13

A nyelvhelyességnél és a két produktív készségnél figyelhető meg teljesítménynövekedés a 2005–2009. évi érettségi vizsgákon (14. táblázat). Az írásképességnél – különösen a német nyelv esetében – valószínűsíthető az érettségi közvetlen hatása a felkészülési folyamatra, melyben a diákok megismerik az írásművek megírásának folyamatát, az értékelési rendszert. Hazai vizsgálatok (például *Józsa és Nikolov, 2005; Nikolov, Pércsich és Szabó, 2000*) azt mutatják, hogy a feladatok előzetes ismertsége, a nyelvórákon való gyakorlásuk mértéke és nehézségének megítélése erőteljesebben befolyásolja a teljesítményeket, mint az olvasásnál. Kutatást igényel annak feltárása, hogy a tanárok e készség fejlesztésére nagyobb hangsúlyt fektetnek-e vagy inkább feladatmegoldói stratégiákat tanítanak korábbi vizsgafeladatok segítségével. Míg angolból változik az olvasott és hallott szöveg értése feladatsor átlagos megoldottsága, addig németből szignifikánsan csökkennek az átlagteljesítmények. Feltételezzük, hogy a tantárgyi bizottság az eddigi tapasztalatok alapján nehezítette az olvasott és hallott szöveg értése feladatsorokat. Ugyanakkor a teljes vizsgán angolból 2007 és 2008-ban nem változott jelentősen az átlagteljesítmény, míg 2009-ben nőtt; németből pedig 2008 és 2009-ben azonos volt, mely a 2007. évi vizsgához képest alacsonyabb. Ezek az információk azt mutatják, hogy az azonos készséget mérő feladatsorok nehézsége feltehetően különböző, valamint a vizsgára való felkészülés során a diákok megismerik a vizsgán várható feladattípusokat, ami befolyásolhatja teljesítményüket.

A továbbiakban áttérünk a középszintű érettségi beszédképesség vizsgarészének elemzésére, amelyről nagyon kevés adat áll rendelkezésre. E vizsgarész mérési-értékelési eljárásai jelentős mértékben különböznek a többi vizsgarészétől: 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 beszédképesség-szintje, az értékelési szem-

pontok és skálák jellemzői, az értékelők szigorúsága, a feladatok minősége. E tényezők közül az OH adatbázisaiban csak az értékelés eredményéről szerezhettünk információt. Az adatbázisokban egyértelműen azonosítható értékelési hibát találunk: ennél a vizsgarésznél több esetben félpont szerepel, amely sem az értékelési skálák alkalmazásakor, sem a vizsgapontok átszámításakor (a német nyelvi szóbeli érettség) nem lehetne. A többi készséghez képest magasabb teljesítményátlagokból (41. és 42. táblázat) és ebből az értékelési hibából arra következtetünk, hogy a tanárok az értékelési eljárásokat néhány esetben „kreatívan” alkalmazzák, vagyis feltételezzük a korábbi vizsgára jellemző értékelési gyakorlatot, melyben a vizsgázók írásbeli érettség elérésekor a szóbeli pontszámok kialakításakor, amely így nem hozott összehasonlítható eredményeket. Ezek a körülmények azt mutatják, hogy meg kellene vizsgálni a tanárok által készített mérőfeladatok összhangját a vizsgakövetelményekkel, továbbá elengedhetetlen az értékelési szempontok és skálák empirikus validálása. Ez utóbbira mutatunk példát a 4.3. fejezetben az íráskészség vizsgarész esetében.

A középszinten érettségizők teljesítményében megjelenő különbségeket először iskolatípusok, majd képzési típusok és végül nemek szerinti bontásban mutatjuk be. Elemezzük továbbá az iskolatípusok és a nemek szerint kimutatható különbségek hatását az egyes vizsgarészekben és a teljes vizsgán elért teljesítményre. Ez utóbbi elemzés során az eta-négyzetet számoltuk ki (erről lásd a 3.3.1. fejezetet).

Az érettségizők teljesítményének átlagát és szórását iskolatípus szerinti bontásban a 10. melléklet tartalmazza. A gimnáziumban érettségizettek szignifikánsan jobb teljesítményt értek el, mint a szakközépiskolában végző tanulók. Angolból és németből a legkisebb különbség a két iskolatípusban érettségizők között a hallott szöveg értése vizsgarész átlagos teljesítményében van, amelyet a korábbi vizsgán semmilyen formában nem mértek. Az olvasás és a beszéd-készség esetében angolból átlagosan 17–18%-pontnyi, németből 19–20%-pontnyi különbség van, míg az angolnál a nyelvhelyesség vizsgarészben 21%-pont. Jelentős különbség adódik az íráskészség vizsgarészben is (angolból 25%-pont, németből 28%-pont). A teljes vizsgán az iskolatípusok szerinti különbségek mértéke angolból a 2005. évi érettség 13%-pont volt, a 2009-es érettség 20%-pont; németből 14%-pont volt, amely 2009-re 21%-pontra növekedett.

Az iskolatípusok közötti különbségek nyelvi teljesítményre gyakorolt hatásának mértékét mutató eta-négyzet értékek mindkét nyelvből mindhárom vizsgaévben nőnek, kivéve németből a hallott szöveg értése vizsgarészt (43. táblázat). Ezek az értékek németből magasabbak, mint angolból. Eredményeink arra utalnak, hogy a vizsga növeli a két iskolatípus között egyébként is meglévő különbségeket.

43. táblázat. Az iskolatípusok közötti különbségek hatása a középszintű vizsgán elért teljesítményekre angolból és németből az eta-négyzet (%) alapján

Vizsgarész	Angol			Német		
	2007	2008	2009	2007	2008	2009
Olvasott szöveg értése	11,50	17,53	18,10	15,28	20,00	22,02
Nyelvhelyesség	14,71	19,18	17,71	14,10	19,91	21,09
Hallott szöveg értése	10,67	14,66	15,16	11,94	19,13	16,75
Íráskészség	13,58	16,28	16,88	15,86	17,25	18,30
Beszéd-készség	9,22	11,47	12,74	9,91	12,27	12,45
<i>Teljes vizsga</i>	16,29	20,33	20,58	18,25	22,53	23,49

A 2009. évi NYAT-vizsgálat alapján (15. és 16. táblázat) az adatbázisban megvizsgáltuk, hogy a két iskolatípusban a vizsgázók hány százaléka éri el a B1 szintet (11. melléklet) a 2008. évi érettség. A szakközépiskolában angolból és németből az olva-

sott szöveg értése teszten a vizsgázók 32%-a, a hallott szöveg értése vizsgarészen angolból a szakközépiskolában érettségizők 30%-a, németből 22%-a, míg az íráskészségnél jóval kevesebben érik el a B1 szintet a szakközépiskolában (angolból 16%, németből 14%). A gimnáziumban tanulók a két szövegértést mérő vizsgarészben az olvasásnál 70%-ban, a hallásnál 60%-ban, míg az íráskészségnél 40%-ban érik el a B1 szintet. Mindez arra utal, hogy indokolt e két vizsgaszint mérése a középszintű vizsgán.

A továbbiakban röviden a felnőttoktatás két típusa közötti különbségek mértékét elemezzük a nyelvi készségek fejlettségében (az adatokat a 12. melléklet tartalmazza), mivel a felnőttoktatást nappali és esti/levelező képzésben befejező érettségizők aránya közel azonos (38. táblázat). Mindkét képzéstípus eredményessége szignifikánsan alacsonyabb a nappali tagozatosoknál. A felnőttoktatásban levelező/esti képzésben érettségizők szignifikánsan jobb átlageredményeket értek el a felnőttoktatás nappali képzésében érettségizőknél minden vizsgarészben és a teljes vizsgán. A különbségek mértékét jellemző t-értékek angolból magasabbak, mint németből (44. táblázat). Ezek az értékek mindkét nyelvnél mindhárom évben nem változnak jelentősen, kivéve a 2008. évi német érettségi vizsgán mérhető átlagteljesítmény-különbséget. Ennek feltehető oka, hogy a feladatsor nehézsége nem állandó.

44. táblázat. A középszintű vizsga készségenkénti különbségek mértéke angolból és németből a felnőttoktatásban tanulók szerint a t-értékek alapján

Vizsgarész	Angol			Német		
	2007	2008	2009	2007	2008	2009
Olvasott szöveg értése	16,29	13,76	17,19	9,14	10,34	10,41
Nyelvhelyesség	14,82	12,63	15,18	7,11	6,85	7,67
Hallott szöveg értése	11,18	10,41	13,22	5,92	12,39	9,99
Íráskészség	13,63	13,95	14,09	8,44	11,14	8,73
Beszéd-készség	11,27	10,45	11,96	9,63	7,33	9,47
<i>Teljes vizsga</i>	17,30	15,88	17,67	10,36	12,37	11,82

Megjegyzés: minden t-érték szignifikáns $p < 0,001$ szinten.

Angolból a legkisebb (átlagosan kb. 8%-pontnyi) különbséget a hallott szöveg értése feladatsornál találjuk, az olvasott szöveg értése és a nyelvhelyesség, valamint a beszéd-készség vizsgarészen belül a két felnőttoktatási képzéstípus közötti különbség átlagosan 11–12%-pontnyi, míg a legnagyobb az íráskészség esetében (16%-pont). Német nyelvnél hasonló tendencia figyelhető meg: a legkisebbek a különbségek a nyelvhelyesség (6%-pont), az olvasott és a hallott szöveg értése tesztben (9%-pont) és a beszéd-készség vizsgarészen (11%-pont) és a legnagyobb az íráskészség esetében (13%-pont). Az OH adatbázisainak elemzése és a 2009. évi NYAT-vizsgálat eredményeit összevetve (11. melléklet) a felnőttoktatás nappali képzésében a vizsgázók kb. 20%-a éri el az olvasási és a hallási készségnél a B1 szintet, míg az íráskészség esetében csupán 5% mindkét nyelvből. A felnőttoktatás esti/levelező képzésében résztvevők a szakközépiskolában tanulók arányaihoz hasonló eredményt mutatnak. A felnőttoktatás két képzéstípusa közötti különbségek okainak feltárásához további kutatásokra van szükség.

A nemek szerinti különbségek vizsgálata a középszintű érettségiben azt mutatja, hogy németből a lányok minden vizsgarészben szignifikánsan jobban teljesítettek a fiúknál (13. melléklet). A legjelentősebb különbségek a két produktív készségnél: az íráskészségnél (átlagosan 11%-pont) és a beszéd-készségnél (7%-pont) vannak, míg az olvasott szöveg értésénél és a nyelvhelyességnél alacsonyabb (átlagosan 4%-pont) és elhanyagolható a hallott szöveg értése vizsgarészben (átlagosan 2%-pontnyi a különbség a fiúk és a lányok teljesítményében). Angolból a 2008. és a 2009. évi érettségiben a fiúk átlagtel-

jesítménye a receptív készségeket mérő vizsgarészekben belül szignifikánsan jobb, de a lányok szignifikánsan jobban teljesítettek a produktív készségeknél; a nemek közötti különbségek a teljes vizsgán is szignifikánsak. Az eredményekből megállapítható, hogy míg angolból középszinten kevésbé jellemző a nemek közötti különbség, addig németből jelentősebb.

A nemek közötti különbségek az összteljesítményt angol nyelvből nem határozzák meg jelentősen, kivéve az íráskészség és a hallott szöveg értése teszten elért teljesítményeket (45. táblázat). Németből az összteljesítményt a nemek közötti különbségek az angolnál nagyobb arányban determinálják, kivéve a hallott szöveg értése vizsgarészt. A német nyelvi érettségi íráskészség vizsgarészen az eta-négyzet értékei átlagosan 3%. Úgy tűnik, hogy a vizsga jól kezeli a nemek közötti különbséget: angolból egyáltalán nincs hatása a teljesítményre, míg németből a két produktív készségnél megjelenő különbségeket feltehetően nem a vizsga hozza létre, vagyis nem mondhatjuk azt, hogy a vizsga a lányokat hozná kedvezőbb helyzetbe. Ennek oka, hogy a hazai vizsgálatok azt mutatják, hogy a nemek közötti különbségek jelentősek a produktív készségeknél (2.2. fejezet).

45. táblázat. A nemek közötti különbségek hatása a középszintű vizsgán elért teljesítményekre angolból és németből az eta-négyzet (%) alapján

Vizsgarész	Angol			Német		
	2007	2008	2009	2007	2008	2009
Olvasott szöveg értése	0,19	0,11	0,33	0,78	0,53	0,93
Nyelvhelyesség	0,01	0,00	0,52	0,45	0,84	1,64
Hallott szöveg értése	0,47	0,55	1,01	0,23	0,22	0,10
Íráskészség	0,69	0,43	0,52	2,50	2,68	2,94
Beszéd-készség	0,28	0,29	0,35	1,86	1,77	1,84
<i>Teljes vizsga</i>	0,03	0,02	0,00	1,71	1,59	1,96

4.1.3. Teljesítmények és a teljesítményt meghatározó változók hatása az emelt szintű érettségien

Az emelt szintű idegen nyelvi érettségi vizsgát a középszinthez hasonló módon elemezzük: először megvizsgáljuk az átlagteljesítményeket az egyes vizsgaéveken belül, majd a vizsgaévek közötti tendenciákra következtetünk. Ezt követően az átlagteljesítményeket iskolatípusok és nemek szerinti bontásban mutatjuk be és elemezzük e különbségek hatását a teljesítményekre. A felnőttoktatás eredményességét az emelt szintű vizsgán nem elemezzük, mert a vizsgázók száma nagyon alacsony (38. táblázat).

Angolból az emelt szintű vizsgán elért teljesítmények alapján egységes kép bontakozik ki (46. táblázat): mindhárom évben a legkönnyebben teljesíthető vizsgarész a hallott szöveg értése és az íráskészség. Ezeket követi a beszéd-készség és a legnehezebb vizsgarész az olvasott szöveg értése és a nyelvhelyesség. A három év eredményeinek összehasonlítása alapján a 2009. évi nyelvhelyesség és hallott szöveg értése feladatlap nehezebbnek bizonyult a korábbi évekhez képest. A vizsgázói populáció heterogén a beszéd-készség vizsgarésznél (a relatív szórás a három év átlagában 34%), a teljesítmények folyamatosan emelkednek. E vizsgarész teljesítménye szignifikánsan különbözik a teljes vizsgáétól, de nem olyan mértékben, mint középszinten.

46. táblázat. Az emelt szintű vizsga eredményei készségenként és a teljes vizsgán angolból (%)

Vizsgarész	2007			2008			2009		
	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értése	67,36	21,89	0,36	69,30	16,56	0,27	68,00	21,11	0,33
Nyelvhelyesség	67,27	19,64	0,32	67,72	17,70	0,29	61,65	17,49	0,27
Hallott szöveg értése	78,27	13,01	0,22	76,76	15,74	0,26	71,99	18,76	0,29
Íráskészség	74,84	18,12	0,30	74,07	18,54	0,31	76,48	16,86	0,27
Beszédkésztség	68,78	24,26	0,40	69,99	24,25	0,40	71,02	23,34	0,37
Teljes vizsga	70,98	16,13	0,27	71,24	15,39	0,25	69,50	16,45	0,26

Németből mindhárom évben az olvasott szöveg értése vizsgarész volt a legkönnyebb (47. táblázat), 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. Ennél a készségnél növekszik a vizsgázók átlagteljesítménye és ezzel együtt csökken a szórás, a feladatlap mindhárom évben túl könnyű az érettségizőknek. Lényeges megvizsgálni, hogy az emelt szintű feladatok hogyan viszonyulnak a vizsgakövetelményekben rögzített készség szintekhez és a vizsgázók nyelvtudás-szintjéhez, illetve mi jellemzi a felkészítési folyamatot. Hasonló tendencia figyelhető meg a hallott szöveg értése feladatsornál. A nyelvhelyesség teszt a vizsga legnehezebb része a 2008. és 2009. évi eredmények alapján. A két produktív készséget mérő vizsgarészben elért teljesítmények az évek múlásával nőnek. Ugyanakkor ezekben a vizsgarészekben heterogén a populáció, jelentős különbségek lehetnek a vizsgázók között különböző háttérváltozók szerint.

47. táblázat. Az emelt szintű vizsga eredményei készségenként és a teljes vizsgán németből (%)

Vizsgarész	2007			2008			2009		
	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értése	77,68	21,39	0,50	81,77	13,68	0,33	88,78	10,03	0,23
Nyelvhelyesség	68,90	17,54	0,41	65,11	18,71	0,46	64,25	17,23	0,39
Hallott szöveg értése	68,34	17,38	0,41	74,31	19,11	0,47	79,20	14,26	0,32
Íráskészség	64,98	25,76	0,61	74,38	21,17	0,52	70,70	22,55	0,51
Beszédkésztség	69,10	26,19	0,62	73,36	22,68	0,55	73,60	23,34	0,53
Teljes vizsga	69,47	18,09	0,43	73,47	15,76	0,38	74,98	14,46	0,33

Az iskolatípusok szerinti különbségek mértéke az angol és a német emelt szintű vizsgán nagyon hasonlóan alakul. Mindkét nyelvre jellemző, hogy a gimnáziumban végzetek szignifikánsan jobb teljesítményt értek el, mint a szakközépiskolában érettségizők (14. melléklet). E különbségek a beszédkésztségénél jelentősebbek (mindkét nyelvből átlagosan 10%-pont), a szórásértékek is magasabbak a többi vizsgarészhez képest. Míg angolból a különbségek mértéke a három év viszonylatában közel azonos, addig németből vannak ingadozások. Feltételezzük, hogy ez a feladatsorok eltérő nehézségéből is adódhat. Az iskolatípusok közötti különbségek teljesítményre gyakorolt hatása (48. táblázat) nem különbözik jelentősen a két nyelv között. Az eta-négyzet értékek a középszintű vizsgához képest (43. táblázat) kisebbek. A teljes vizsgán elért összteljesítményeket mindkét nyelvből kb. 5%-ban határozza meg az iskolatípusok közötti kü-

lönbség. A 2009. évi NYAT-vizsgálat eredményei alapján a két iskolatípusban és a nappali képzésben, valamint a jogviszony nélkül vizsgázók nagyon hasonló arányban érték el a B2 szintet (11. melléklet). A középszinttel ellentétben itt nem találunk jelentős különbségeket az egyes vizsgázói csoportok szerint.

A nemek szerinti különbségek a receptív készségeknél és a nyelvhelyesség vizsgarészen az emelt szintű angol nyelvi érettségin jelentősebbek, mint németből: három év átlagában a fiúk 3–4%-ponttal jobban teljesítettek a lányoknál (15. melléklet). E különbségek angolból a teljesítményt 1–2%-ban határozzák meg. Az íráskészségnél a lányok szignifikánsan jobban teljesítettek a fiúknál elsősorban németből (átlagosan 5%-pontnyi a különbség). A beszéd-készség vizsgarésznél a különbségek nem szignifikánsak angolból, németből a lányok szignifikánsan ($p < 0,05$ szinten) jobb teljesítményt értek el a szóbeli vizsgarészen a fiúknál, így a nemek szerinti különbségek hatása ennél a készségnél németből kicsit magasabb, mint angolból. A teljes vizsgán angolból és németből a receptív és a produktív készségeknél lévő különbségek kiegyenlítik egymást.

48. táblázat. Az iskolatípusok és nemek közötti különbségek hatása az emelt szintű vizsgán elért teljesítményekre angolból és németből az eta-négyzet (%) alapján

Vizsgarész	Az iskolatípusok közötti különbségek hatása a teljesítményre					
	Angol			Német		
	2007	2008	2009	2007	2008	2009
Olvasott szöveg értése	3,21	3,32	3,19	3,18	1,53	4,13
Nyelvhelyesség	3,72	3,64	4,18	3,90	2,48	5,52
Hallott szöveg értése	2,25	2,18	3,48	2,99	1,40	3,09
Íráskészség	3,27	2,81	2,88	4,07	0,76	1,99
Beszéd-készség	3,88	3,64	3,56	4,51	1,77	6,00
<i>Teljes vizsga</i>	4,88	4,62	4,87	5,47	2,26	5,91

Vizsgarész	A nemek közötti különbségek hatása a teljesítményre					
	Angol			Német		
	2007	2008	2009	2007	2008	2009
Olvasott szöveg értése	1,14	0,31	1,77	0,13	0,04	0,31
Nyelvhelyesség	0,55	0,24	0,28	0,30	0,03	0,06
Hallott szöveg értése	1,68	1,49	1,71	0,14	0,54	0,01
Íráskészség	0,83	1,85	0,37	1,02	1,27	1,12
Beszéd-készség	0,06	0,04	0,04	0,23	0,24	0,26
<i>Teljes vizsga</i>	0,16	0,01	0,46	0,03	0,04	0,21

4.1.4. Az idegen nyelvi érettségi vizsga jelenlegi működésének jellemzői, vizsgálandó területei

Az OH adatbázisait a vizsgázói összetétel változásai és a vizsgázók eredményei alapján elemeztük, amellyel a vizsgafejlesztési folyamatra és az igények és célok változásaira irányuló visszacsatolás biztosítható. A továbbiakban vizsgálatunk alapján jellemezzük a közép- és emelt szintű vizsga jelenlegi működését, összegezzük az idegen nyelvi érettségiről szerzett tapasztalatainkat és teszünk javaslatokat a vizsgálandó célterületekre.

A vizsgázói összetétel változásai alapján szükséges megvizsgálni az előrehozott vizsga választásának okait és a nyelvtanulási folyamatra gyakorolt hatását. Mindezek mellett nyomon kellene követni a felnőttoktatást befejező érettségizők eredményeit, és az alacsony teljesítményt meghatározó tényezőket.

Az egyes vizsgarészekben és a teljes vizsgán elért eredmények alapján a receptív készséget mérő vizsgarészekben mindkét szinten és nyelvből, de elsősorban a német nyelvi érettségivel feltételezzük, hogy változik a feladatsorok nehézsége. Így feladat- és itemszintű elemzésekre van szükség, melyekkel meg kell határozni a feladatlapon empirikus nehézségi szintjét és egybevetni a feladatkészítők által megállapított nehézségi sorrenddel és a vizsgázók képességszintjével. Az íráskészség vizsgarészben mindkét nyelvből és szinten jellemző az érettségizők teljesítményének javulása, így szükséges megvizsgálni a performancia tesztek jellemzőit, az értékelési eljárások minőségét, valamint, hogy ezek mennyiben vannak összhangban a vizsgázók teljesítményszintjeivel. Mindezek mellett a kutatásoknak további lényeges eleme a vizsga nyelvoktatási folyamatra gyakorolt hatásának megismerése. Ebben főleg azt kell elemezni, hogy a nyelvtanítás mennyiben mozdult el a készségfejlesztés irányába és melyek a vizsgára felkészítő gyakorlat jellemzői. Lényeges kutatási feladat a közép- és emelt szintű szóbeli vizsgarészben alkalmazott feladatsorok elemzése és az értékelés alapjául szolgáló skálák empirikus validálása.

A teljes vizsgán az egyes háttérváltozók alapján iskolatípusok szerint jelentős különbségek vannak a tanulók teljesítményében. Ezek a különbségek középszinten átlagosan 20%-ban, emelt szinten 5%-ban determinálják a teljes vizsgán elért teljesítményeket. A nemek szerinti különbségek hatása angolból és németből emelt szinten elhanyagolható, míg németből középszinten 2%. A további vizsgálatoknak az érettségizők eredményeit meghatározó kognitív és affektív tanulói tényezők, valamint az iskolán belüli és iskolán kívüli változók hatásának feltárására kell irányulniuk.

Az OH adatbázisának elemzése és a 2009. évi NYAT-vizsgálat eredménye szerint középszinten szükség van a két nyelvi szint mérésére, mivel a szakközépiskolát és a felnőttoktatást befejezők közül kevesen érik el a nem túl magas B1 szintet. Az emelt szintű vizsgán egységesen azt találtuk, hogy a vizsgázók kb. 75–80%-a éri el a két receptív készség esetében a B2 szintet, míg az íráskészségnél alacsonyabb ez az arány. Ez azt jelezheti, hogy a vizsgázók nyelvtudása a receptív készségeknél a B2 követelményszintnél magasabb szinten van. Ezek az információk további vizsgafejlesztési igényeket és újabb célokat jelölhetnek ki.

E fejezet további részében a receptív készségeket mérő vizsgarészekben belül végzünk feladat- és itemszintű elemzéseket, majd az íráskészség esetében megvizsgáljuk a mérési-értékelési eljárások minőségi mutatóit. Mindkét vizsgálattal célunk, hogy a vizsgafejlesztési folyamat számára adjunk visszacsatolást.

4.2. Az idegen nyelvi érettségi belső és külső validitása

Ebben az alfejezetben egy IRT-alapú nyelvi feladatbank létrehozásának szükségességét, lehetőségeit, problémáit és módszertani alapkérdéseit vizsgáljuk a receptív készségeket mérő vizsgarészek alapján. Az első részben elemezzük a 2006. évi középszintű német nyelvi érettségi olvasott és hallott szöveg értése és a nyelvhelyesség vizsgafeladatok eredményeit. A második részben a 2005. és 2006. évi közép- és emelt szintű érettségi nyelvhelyesség vizsgarész elemzésének példája és európai feladatbank-építési projektek eredményei alapján teszünk javaslatot arra, hogyan építhető ki egy, a KER (2002) szintrendszerére épülő nyelvi feladatbank.

4.2.1. A német nyelvi érettségi feladatok adatbázisának elemzése

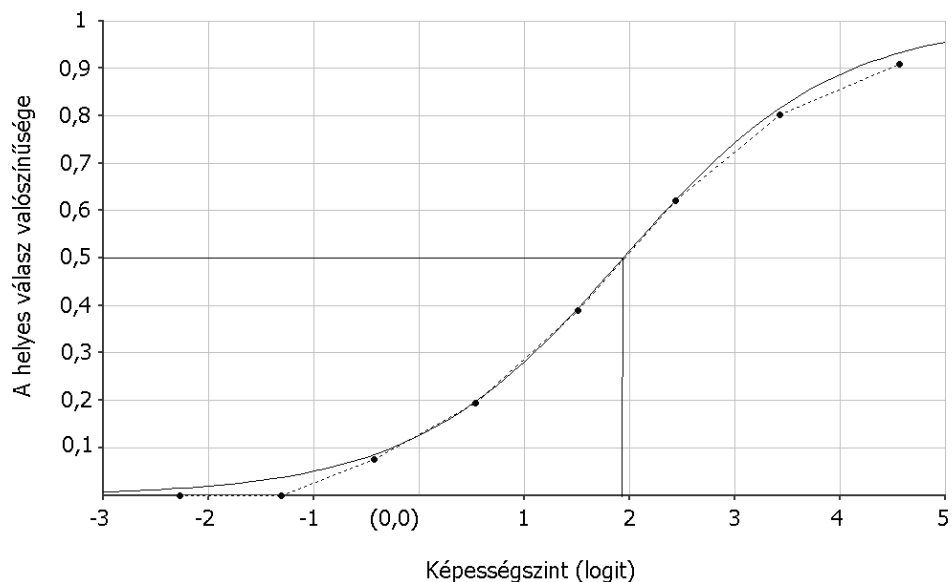
A 2006. évi középszintű érettségi receptív készségeket mérő vizsgarészek elemzésében először az egyes ítemek, feladatok és feladatsorok nehézségi indexét vizsgáljuk, majd

felállítunk a vizsgarészekben belül egy minta- és tesztfüggetlen, objektív itemnehézségi sorrendet, melyet egyrészt egybevetünk a feladatkészítők által meghatározott nehézségi sorrenddel, másrészt a mintába került vizsgázók képességszintjével. Végül az egyes itemek megbízhatósági mutatóit elemezzük a tesztfejlesztés lehetőségeinek bemutatása céljából.

Az itemek nehézségi indexe

Rasch a modell megalkotása során abból indult ki, hogy „a magasabb képességszintű személy nagyobb valószínűség mellett old meg bármely típusú itemet, mint a többi személy és hasonlóan egy item akkor nehezebb, mint a másik, ha bárki nagyobb valószínűséggel oldja meg a másik itemet, mint azt” (*Rasch*, 1960. 117. o., idézi *Molnár*, 2006. 106. o.). Ennek az elgondolásnak megfelelően a *Rasch*-modell segítségével olyan item-jelleggörbéket rajzolhatunk, amelyek egy itemre adott helyes válaszok valószínűségét különböző képességszintek mellett írják le és megadják, hogy az egyes képességszintű diákok milyen valószínűséggel válaszolják meg helyesen az adott itemet.

A 12. ábra példaként egy jól működő item karakterisztikus görbét mutatja. A vízszintes tengelyen a képességszinteket, a függőleges tengelyen az ezekhez tartozó megoldási valószínűségeket (0 és 1 közötti érték) ábrázoltuk. A 12. ábráról leolvashatjuk, hogy mely képességszintű vizsgázók oldják meg nagyobb valószínűséggel az itemet, valamint az item mennyire különíti el egymástól a különböző képességűeket. A helyes válasznak még igen alacsony képességszint értékeknél is van valószínűsége, és a tévedésnek is van valószínűsége még az igen magas képességszint értékek mellett is (*Verhelst*, 2004). Az itemek nehézségét az a képességszint reprezentálja, ahol a helyes megoldás valószínűsége 50 százalék. A 12. ábrán jelöltük ezt az értéket (1,94). A képességszintek és az itemnehézségi paraméterek közös skálájának egységét logitnak nevezzük.



12. ábra

A hallott szöveg értése feladatsor 15. itemének jelleggörbéje

A továbbiakban az egyes feladatok itemeinek nehézségi indexét jellemezzük, amihez a 49. táblázatban zárójelben megadjuk a jellemzett feladatok típusait is. Mindhárom feladatlapot kitöltötték a vizsgázók, a teszteket egy adatbázisban, egy skálára kalibrálva elemezzük, ezáltal közvetlenül összehasonlítható a feladatok itemeinek nehézsége. Az alsó skála logitérték a feladat legkönnyebb, míg a felső logitérték a legnehezebb itemet jelöli. Az adatok további jellemzője, hogy intervallumskálán vannak, így kiszámolható

az egyes feladatok átlagos nehézségi szintje, skálaterjedelme és szórása. A középérték logitok jelölik az egyes feladatok átlagos nehézségét. A feladatokon belüli itemek jellemzőit az 16., 17. és 18. melléklet tartalmazza.

49. táblázat. A feladatlapon nehézségi indexének jellemzői

Feladatlapon	Feladat	Itemek száma	Skála logitok		Középérték logitok	Skála terjedelme	Szórás
			Alsó	Felső			
Olvasott szöveg értéke	1. (táblázat kitöltése)	8	-2,78	0,41	-1,17	3,19	0,96
	2. (interjúkérdések párosítása válasszal)	7	-0,14	2,91	1,08	3,05	1,11
	3. (rövid választ igénylő nyitott kérdések)	10	-1,60	1,41	0,18	3,01	0,99
Hallott szöveg értéke	1. (igaz/hamis állítás)	7	-2,54	0,79	-0,97	3,33	1,03
	2. (rövid választ igénylő nyitott kérdések)	7	-2,48	2,52	0,03	5,00	2,04
	3. (hiányos mondatok kiegészítése)	10	-1,59	1,95	0,66	3,54	1,36
Nyelv-helyesség	1. (mondat alapú feleletválasztás)	7	-2,20	1,18	-0,60	3,38	1,06
	2. (szöveg alapú feleletválasztás)	8	-1,49	0,93	-0,34	2,42	0,81
	3. (igék ragozott alakjainak beillesztése)	9	-1,11	1,07	0,13	2,18	0,82
	4. (szöveg kiegészítése elöljárószavakkal)	6	0,04	2,34	1,22	2,38	0,77

A 49. táblázat bemutatja, hogy a feladatkészítők által megállapított nehézségi sorrend mennyire esik egybe a feladatok empirikus nehézségének növekedésével⁹. Ez többnyire akkor teljesül, ha a feladatok itemeinek alsó és középérték logitértékei egyre növekednek. Ennek a kritériumnak a hallott szöveg értéke és a nyelvhelyesség feladatsor felel meg leginkább, tehát a feladatkészítők megfelelően határozták meg ezekben a tesztekben a feladatok nehézségi szintjét. Ugyanakkor ezek az adatok azt is jelzik, hogy az olvasott szöveg értéke feladatsor második feladata a harmadiknál nehezebb, így nem teljesül az a szándék, hogy a harmadik feladat legyen a legnehezebb. Ennek oka a második feladat típusában keresendő. Ebben a feladatban a vizsgázóknak egy interjú válaszaihoz kellett az előre megadott kérdéseket hozzárendelni. A feladat itemnehézségi indexeinek vizsgálata alapján két kiugróan nehéz (12. és 13.) itemet azonosíthatunk (16. melléklet), amelyek elemzése alapján (Einhorn, 2007a) megállapítható, hogy a kérdésre adott válasz tartalmilag nem kötődik eléggé a kérdéshez, így ezeket az itemeket ki kell hagyni egy tesztfejlesztés során.

A 49. táblázat adatai a hallott szöveg értéke feladatsor második feladatának belső instabilitására is utalnak, mivel az itemek nehézségi indexei nagy szóródást mutatnak, a skálaterjedelem 5 logitegységnyi. Ennek feltételezhető oka, hogy a rövid szöveghez túl sok item készült és a két legnehezebb (8. és 12.) item megoldásakor számokat kellett a vizsgázóknak beírni, amelyeknek megértése általában nehézséget okoz. Mindkét feladatnál a zárójelben megadott itemek értéke a 17. mellékletben található és elhelyezkedését a 13. ábrán tanulmányozhatjuk.

A nyelvhelyesség feladatsor első feladatában a legnehezebb item nehézségi indexe magasabb, mint a második és harmadik feladat legnehezebb iteme (18. melléklet). Ez a 3. item, amelynek viselkedését a későbbiekben részletesen elemezzük (lásd 14. ábra).

Az itemek nehézségi szintjeinek és a vizsgázók képességszintjeinek összekapcsolása

A Rasch-modell segítségével a 49. táblázatban jellemzett feladatok itemeinek nehézségét és a mintába került diákok képességszintjét a minta-, illetve itemtérkép (*map of persons' ability/ items' difficulty map*) segítségével közös képességskálán ábrázolhatjuk, így össze tudjuk hasonlítani a diákok képességszintjét a feladatlapon szereplő itemek 50

⁹ A tesztfüzeteket úgy állítják össze, hogy az első feladat legyen a legkönnyebb az utolsó pedig a legnehezebb.

százalékos valószínűséggel történő megoldásához szükséges képességszintekkel. Mivel a Rasch-moddal elemző szoftverek az adatokat intervallumskálára transzformálják, a térképről leolvasható, hogy az adott item mennyivel könnyebb-nehezebb, illetve a vizsgált személy mennyivel magasabb-alacsonyabb képességű. A térkép az itemnehézségi indexeket és a képességparamétereket közös logitskálán ábrázolja, amely nem határozza meg a képességszintek és a nehézségi indexek abszolút helyét, hanem felállítja a relatív távolságokat a képességszinteken és a nehézségi indexeken belül, továbbá a képességszintek és a nehézségi indexek között.

A minta- és itemtérképek többféleképpen kirajzolhatók (lásd például *Molnár, 2003, 2005, 2006; Molnár és Józsa, 2005; Szabó, 2008*), esetünkben feladatlaponként vagy együtt, több dimenzióban ábrázolhatjuk ezeket. Mindhárom feladatlapot ugyanazon vizsgázók töltötték ki, célunk az eredmények összehasonlítása, így a 13. ábrán az olvasott, a hallott szöveg értése, valamint a nyelvhelyesség feladatsor minta- és itemtérképét közös skálára kalibrálva, három dimenzióban mutatjuk.

A 13. ábrán minden egyes 'x' hét tanulót reprezentál. Fontos utalni arra, hogy az ábrázolt 'x'-eken kívül is található vizsgázók a képességskála magasabb, illetve alacsonyabb tartományában, csak hétnél kevesebben vannak, így az ábra nem tartalmazza az adatokat. Az elemzésben alkalmazott ConQuest program egy modellált populáció képességszintjének átlagát veszi nullának, ezért a negatív számok átlag alatti képességet jelölnek. A 13. ábra alapján a vizsgált személyek képessége mindhárom vizsgarészben rendkívül tág határok között mozog, jelentős különbségek vannak az egyes érettségizők között. Ez a megállapítás nem meglepő, hiszen a mintában nagyon eltérő nyelvtudás-szintű vizsgázók szerepelnek. A három feladatlap képességeloszlását összehasonlítva a hallott és az olvasott szöveg értése feladatsor képességparamétereinek eloszlása hasonló képet mutat.

A közös logitskálán ábrázolt feladatlapok minta- és itemtérképe alapján megállapítható, hogy az adott teszt itemeinek nehézsége mennyire felel meg a vizsgázók nyelvi szintjének, az itemnehézségi indexek mennyire jól fedik le a vizsgázók nyelvi szintjei által meghatározott képességskála-intervallumot. A mintához jól illesztett tesztnek a személyek képességparamétereit jelző 'x'-ek és az itemeket reprezentáló számok egymással párhuzamosan futnak. Ennek a kritériumnak a három feladatlap közül a nyelvhelyesség teszt felel meg a leginkább, hiszen a feladatok itemei jól kiegészítik egymást és így többnyire megfelelően fedik le a vizsgázók képességszintjét. Ugyanakkor az olvasott szöveg értése feladatlapon három helyen találunk olyan itemcsoportosulást, amelyek azonos képességszintet mérnek, így az itemek nem fedik le a teljes képességskála-intervallumot. A két kiugróan nehéz (12. és 13.) itemet a 49. táblázatnál elemeztük. A hallott szöveg értése feladatsorra egyrészt jellemző, hogy az átlagos képességszintű diákok képességszintjét (logitérték=0) mérő itemek hiányoznak a tesztből, másrészt két, egymástól jól elkülöníthető itemcsoportot találunk, amelyek egyik része az átlagos képességszint alatt, a másik az e fölötti képességszintű vizsgázókat méri. A három feladatlap közös jellemzője, hogy van egy magas képességszint, amit a tesztek nem mérnek.

logit	Olvasott szöveg értése		Hallott szöveg értése		Nyelvhelyesség	
	személy	item	személy	item	személy	item
5						
4	X		X			
3	XXXX		XXXX			
2	XXXXXXXX		XXXXXXXX			
1	XXXXXXXX		XXXXXXXX			
0	XXX		XXX			
-1	X		X			
-2						
-3						

13. ábra

Az olvasott és hallott szöveg értése, és a nyelvhelyesség feladatlap minta- és itemtérképe

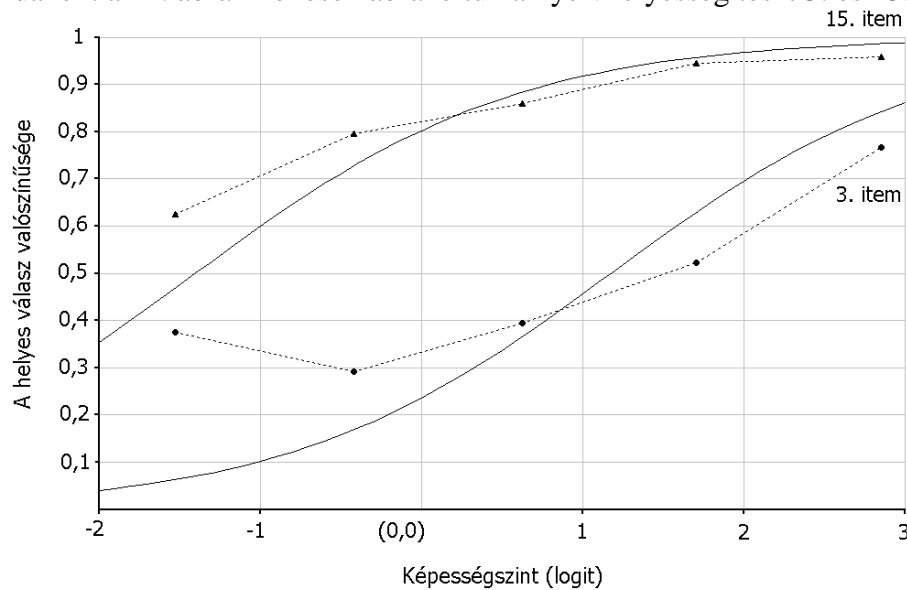
Megjegyzés: A ConQuest program a többdimenziós ábrázolás során egymás mellé helyezi először a képességparamétereket majd az itemeket. Ezt az ábrát szerkesztettük, mivel tesztenként elkülönítve vizuálisan jobban összehasonlítható a minta- és itemtérkép.

Az itemek megbízhatósági mutatói

Az itemek megbízhatósági mutatói megadják, hogyan lehetne a feladatlapokat továbbfejleszteni, illetve azt is, hogy melyik itemet lehet elhagyni vagy módosítani. A Rasch-moddal történő tesztfejlesztés során meg kell vizsgálni a feladatok itemeinek

diszkriminációs indexét, modellilleszkedését, és a nem megfelelően működő itemek karakterisztikus görbéjét.

A 13. ábra alapján az olvasott és a hallott szöveg értése tesztben öt, míg a nyelvhelyesség feladatsorban három olyan item található, amelyik logitértéke alacsonyabb, mint az összes diák képességparamétere. Ez arra utal, hogy annak a valószínűsége, hogy ezeket az itemeket az adott mintában valaki megoldja, nagyobb, mint 50 százalék, továbbá magas annak a valószínűsége, hogy az átlagos képességűek (logitérték=0) megoldják ezeket az itemeket. Ha ezeknek az itemeknek a diszkriminációs indexét megvizsgáljuk (16., 17. és 18. melléklet), akkor többnyire alacsony értékeket találunk. Ezek az itemek nem jól különítik el egymástól a mintába bekerült jobb és gyengébb képességű vizsgázókat. Alacsony diszkriminációs indexű itemet az átlagos képességszint fölött is találunk. Példaként a 14. ábrán közösen ábrázoltuk a nyelvhelyesség teszt 3. és 15. itemét.



14. ábra

A nyelvhelyesség feladatlap 3. és 15. itemének jelleggörbéje

Ez a két item csak a nehézségi indexben különbözik, a diszkriminációs indexük egyaránt alacsony (0,14). Egyik esetben sem rajzolódik ki a teljes logisztikus görbe, mint a 12. ábrán. Az alacsony diszkriminációs indexű nehéz item (3.) azt jelzi, hogy az itemet kevesen, de az egész teszten az átlagnál jobb és gyengébb teljesítményt nyújtó vizsgázók közel azonos valószínűséggel oldják meg. A 14. ábrán az empirikus item karakterisztikus görbe (szaggatott vonal) laposabb, mint az elméleti görbe (folytonos vonal), tehát az adatok nem illeszkednek az elvárt, előre jelzett modellhez.

Az itemek működése és a feladatok továbbfejlesztése során fontos megvizsgálni, hogy az adatok mennyire követik a modellt. Az illeszkedésvizsgálat arról ad információt, hogy mennyire illeszkedik az adott item a többi közé, így rávilágít a problémás itemekre. A három feladatlapon nyolc nem illeszkedő itemet találtunk, amelyek paramétereit az 50. táblázat mutatja. Az infit (súlyozott) paraméter a vizsgázók képességparamétereire és az itemek nehézségéhez közel levő jellemző válaszokra vonatkozik (David, 2006). Az adatok illeszkedésekor az infit paraméter értéke 1-nél van és ekkor az itemek diszkrimináló ereje közel azonos. Az elfogadható értékek sávja függ a minta elemszámától, nem határozható meg előre az elfogadási sáv, ezért ezt a vizsgálat céljától és alkalmazásától függően kell mérlegelni (Wu, 2006b, idézi Molnár, 2006. 112. o.; Linacre, 2002, idézi Eckes 2003. 54. o.). Irányelvként megállapítható, hogy például egy 2000 fős mintánál az elfogadható értékek 0,94 és 1,06 között ingadozhatnak (lásd pél-

dául *Molnár*, 2006 vizsgálatában). Az 1100 fős minta esetében az 1,08 fölötti infit paraméterrel rendelkező itemeket tekintettük nem illeszkedőnek.

50. táblázat. A nem illeszkedő itemek paramétereinek jellemzői

Feladat	Itemszám	Nehézség	Infit paraméter	Diszkrimináció
Olvasott szöveg értése 3. feladat	21	0,54	1,32	0,13
Hallott szöveg értése 1. feladat	3	-0,95	1,09	0,28
	6	-0,45	1,18	0,19
Nyelvhelyesség 1. feladat	2	-1,51	1,11	0,11
	3	1,18	1,17	0,14
Nyelvhelyesség 3. feladat	20	-0,60	1,09	0,25
	21	0,34	1,09	0,31
Nyelvhelyesség 4. feladat	26	0,81	1,11	0,22

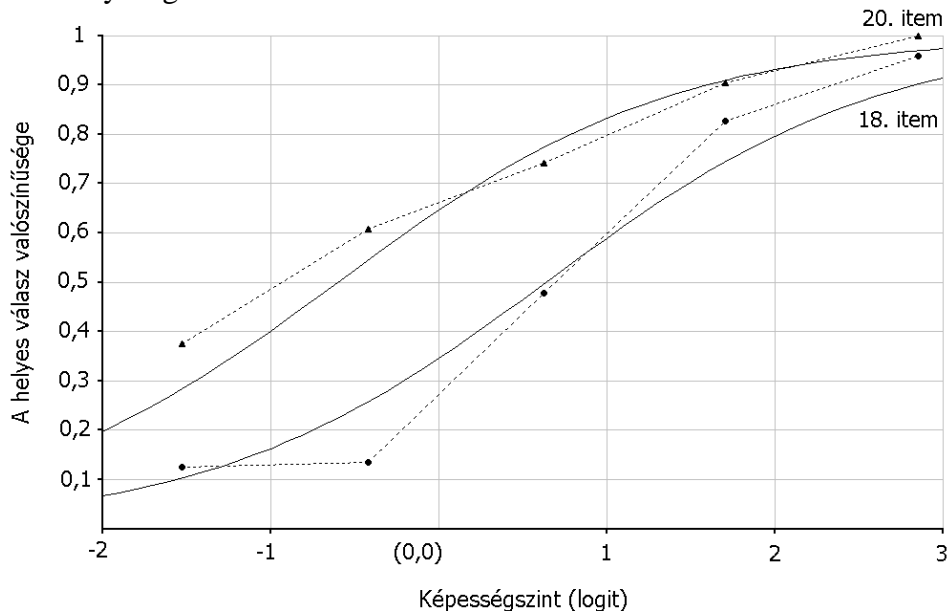
A 13. ábra és az 50. táblázat alapján jól látszanak a tesztfejlesztés lehetőségei. Például az olvasott szöveg értése feladatlap minta- és itemtérképe (13. ábra) mutatja, hogy a harmadik feladathoz tartozó 20., 21. és 22. item azonos képességszintet mér. Mivel célunk az, hogy a képességparamétereket minél jobban lefedje a teszt, indokolt a nem illeszkedő 21. item cseréje.

A nem illeszkedő itemek viselkedését a feladat önálló, elkülönített elemzésével vizsgálhatjuk (e vizsgálat szükségességéről lásd *Nikolov, Pércsich és Szabó*, 2000), amelyhez szükséges a feladatok ismerete is. Ez már túlmutat a fejezet célkitűzésein, ezért csak néhány szempontra térünk ki. A hallott szöveg értése feladatsor első feladatának külön elemzése semmilyen rendellenességet nem mutatott, nem találtunk nem illeszkedő itemeket. A feladat egyedül jól funkcionál, de a tesztben lévő két feladattal együtt már kevésbé. Ennek feltehetően az az oka, hogy a feladat inkább globális szövegértést mér, a vizsgázónak a feladattípusból adódóan (igaz-hamis állítás eldöntése) kevesebb tevékenysége van, magas a találgatás lehetősége, míg a másik kettő feladatban egyszerre kell figyelnie a hallott szövegre, valamint a kérdésekre (2. feladat), vagy a kiegészítendő mondatokra (3. feladat) és közben még írnia is kell.

A nyelvhelyesség feladatsor első feladatának elemzése sok problémára utal (vö. 13. ábra, 49. táblázat és 50. táblázat). Ezt a feladatot össze tudjuk hasonlítani a második feladattal, ami szintén feleletválasztó, csak nem mondat-, hanem szövegalapú. Közös a két feladatban, hogy mindkettő az alacsonyabb (A2) szinthez készült, különbség viszont, hogy a szövegalapú feleletválasztó feladat itemeinek megbízhatósági mutatói jobbak. Az első feladat problémája, hogy a kontextus hiánya alapvetően befolyásolja az itemek megoldottságát és működését. Vizsgálatunk is alátámasztja a mondatalapú feleletválasztó feladat törlésére tett javaslatot a választható feladattípusok közül (*Einhorn*, 2007a).

A nyelvhelyesség teszt harmadik feladatának elkülönített elemzése során az 50. táblázatban jelölt két, nem illeszkedő item mellett négy olyan itemet is találtunk, amelyek infit paraméterei 0,9 körüli értékűek. A továbbiakban példaként ebből a feladatból két item (18. és 20.) karakterisztikus görbéjét ábrázoljuk közösen (15. ábra). A 18. item jelleggörbéje az elvártnál meredekebb, és azt mutatja, hogy egy bizonyos képességszintig alacsony annak a valószínűsége, hogy egy személy helyesen válaszolja meg az itemet, majd megnő az item megoldásának valószínűsége, és a magasabb képességszinten mindenki jól válaszol. Ez azt jelenti, hogy az item túlilleszkedik, mivel diszkriminációs indexe magas (0,54), ám ez az itemműködés kevésbé jelent problémát (*Dávid*, 2006; *Nikolov, Pércsich és Szabó*, 2000). A 20. item jelleggörbéje a 3. itemhez (14. áb-

ra) hasonlóan az elvártnál laposabb, tehát az item nem illeszkedik a többi közé. Ez az itemműködés azért problematikus, mert azt jelzi, hogy nagy számban vált ki váratlan válaszokat a vizsgált személyektől. Ebben a feladatban a vizsgázóknak főnévi igeneveket kellett megfelelő igealakban behelyezni a szövegbe. A vizsgázók elemzett megoldásai (Einhorn, 2007a) tanítási-tanulási problémát feltételeznek, mert az érettségizőknek gyakran nem az okozott problémát, hogy a megfelelő igeidőt válasszák, hanem nem ismerték a helyes igealakokat.



15. ábra

A nyelvhelyesség feladatlap 18. és 20. itemének jelleggörbéje

Megjegyzés: A két item diszkriminációs indexe eltérő (0,54 és 0,25). A Rasch-modell nem határozza meg a diszkrimináció abszolút nagyságát, így a modell képileg az itemeket azonos meredekségű logisztikus görbével ábrázolja. A jobban diszkrimináló itemek azonban jobban széthúzzák a képességparaméter-értékeket (lásd például Molnár, 2006).

4.2.2. Az adatbázisok összekapcsolásának, a feladatbank kiépítésének problémái

A minta- és itemtérkép, valamint az itemek megbízhatóságának vizsgálata alapján számos információ szerezhető az egyes feladatok itemeinek nehézségéről, a fejlesztés lehetőségeiről, azonban több kérdés is felmerül, amelyre nem kapunk választ.

Az egyik kérdés a középszintű érettségi szintjeit érinti. A középszintű vizsga a Referenciakeret két szintjét méri, azonban a minta- és itemtérképről (13. ábra) nem tudjuk megállapítani, hogy mely vizsgázó tudása minősül A2, mely érettségiző tudása B1 szintűnek, továbbá, mely itemek nehézségi indexét lehet az egyik, illetve másik szinthez sorolni. A 49. táblázatban jellemzett feladatok itemnehézségi indexei, valamint a minta- és itemtérkép (13. ábra) mutatja, hogy a könnyebb feladatok általában az átlagos képességszint alatt helyezkednek el, de azt nem tudjuk, hol kell meghúzni a határokat, nem különül el egymástól a kétféle tudásszint.

A további kérdések tárgyalásánál a nyelvhelyesség feladatsor példáján mutatjuk be a vizsgaszintek empirikus vizsgálatának szükségességét. A 16. a, b és c ábrán egymás mellé vetítve mutatjuk a 2005. évi, az előzőekben részletesen elemezett 2006. évi középszintű, valamint a 2005. évi emelt szintű német érettségi nyelvhelyesség feladatlap minta- és itemtérképét. Az adatokat nem skálázhattuk együtt, mint a 13. ábrán, mivel nincs olyan személy vagy item, ami összekötné az adatbázisokat, így az összehasonlítás

mális szinteltolódás esetén is jelentősen eltérő nehézségű feladatsor jöhet létre. Bár közvetlenül nem hasonlítható össze a két feladatlap minta- és itemtérképe, az azonban megállapítható, hogy a személyek képességparaméterei hasonló eloszlásúak a két évben. A 2006. évi német nyelvű érettségi esetében ennél a feladatlapnál az itemek jobban fedik le a vizsgázók képességparaméterei által meghatározott képességskála-intervallumot, mivel a 16. a ábrán két helyen találunk négy olyan itemet, amelyek azonos képességszintet mérnek. A tesztfejlesztés során erre a problémára megoldás lehet az alacsony megbízhatósági mutatóval rendelkező itemek cseréje. A 2005. évi emelt szintű nyelvhelyesség feladatlap minta- és itemtérképéről (16. c ábra) megállapítható, hogy a feladatlap a vizsgázóknak túl könnyű volt. Az itemnehézségi indexek fele helyezkedik el a 0 logitegység alatt.

A 16. a, b és c ábra alapján felmerül az a kérdés, hogy a középszinten magasabb teljesítményt elérők milyen valószínűséggel oldanak meg az emelt szintű feladatokat, vagyis hogyan viszonyul egymáshoz a közép- és az emelt szintű feladatsorok nehézségi szintje? A 16. c ábra egy másik problémára is utal. A magasabb képességtartományban az emelt szintű feladatlap nem működik megfelelően.

További kérdésként fogalmazható meg, hogy valóban B2 szintet mérnek-e az emelt szintű feladatok? Amennyiben a feladatok B2 szintűek, akkor indokolt lehet a vizsgakövetelmények módosítása és a feladatok nehézségi szintjének emelése C1 szintre. Amennyiben a feladatok a B2 szint alatti szintet mérnek, akkor szükséges a vizsgafeladatokat a követelményekben megadott szinthez igazítani. Ezek a hipotézisek is jelzik, hogy az emelt szintű feladatsornál is szükséges megvizsgálni a Referenciakerethez történő illeszkedést.

A Rasch-modell segítségével a fenti kérdésekre tudunk válaszolni, mivel a modell egyik fontos tulajdonsága – amit a PISA-vizsgálatokban (OECD, 2005) is alkalmaznak –, hogy ha ismerjük egy vizsgázó képességszintjét, akkor ez alapján megállapítható, hogy milyen valószínűséggel válaszol meg egy olyan itemet, amelynek nehézségi indexe értelmezhető a közös képességskálán, anélkül, hogy a diáknak a valóságban meg kellene oldania azt. A Rasch-modell a feladatokat összekötő horgonyitemek (*anchor items*) segítségével lehetőséget teremt a feladatsorok együttes elemzésére, a feladatok egy skálára hozására, tehát közvetlenül összehasonlíthatóvá válik az item nehézsége és a diákok képességszintje (a horgonyitemek alkalmazásáról lásd magyar nyelven Molnár, 2003, 2006 és Szabó, 2009 tanulmányát). A feladatok Referenciakerethez történő illeszkedésének vizsgálatához referenciaitemeket tartalmazó feladatokra van szükség. Ebben a folyamatban felhasználhatók a DIALANG (lásd például Alderson és Huhta, 2005) és az EBAFLS-projekt (CITO, 2008) eredményei. Ugyanakkor ezeknek az alkalmazásával számos módszertani kérdés is felmerül (lásd részletesebben: 1.2.2. fejezetet).

A továbbiakban bemutatjuk, hogy hogyan lehetne létrehozni a német nyelvi érettségi feladatsoraiból a vizsgált adatbázisok alapján feladatbankot. Az EBAFLS-projekt ajánlásainak megfelelően a legfontosabb lépések a következők (CITO, 2008):

I. Az érettségi vizsgafeladatok ekvivalenciájának, nehézségi szintjének vizsgálata a Referenciakeret alapján:

1. Az adatbázisban lévő feladatok nehézségi szintjének vizsgálata.
2. Érettségi és referenciafeladatok kiválasztása.
3. Tesztfüzetek kialakítása.
4. A tesztfüzetek bemérése.
5. Adatok bevitele, adatbázis létrehozása.
6. IRT-alapú itemelemzés és -szintezés.

II. Az érettségi feladatokból feladatbank létrehozása:

7. Feladatok kiválasztása.

8. A feladatok jellemzőinek leírása (például a Dutch grid-projekt alapján¹⁰ Alderson és mtsai, 2004).

9. Ajánlások megfogalmazása az érettségi vizsgafeladatok fejlesztéséhez.

Az első fázisban, az idegen nyelvi érettségi vizsgaszintjeinek vizsgálatában a különböző vizsgaidőszakokban keletkező feladatokból horgonyitemek segítségével és újabb felmérésben történő bemérésével meg tudjuk teremteni a különböző vizsgafeladatok itemszintű adatbázisai közötti kapcsolatot, és a Rasch-modell alkalmazásával egy adatbázisban tudjuk elemezni az adatokat. Ennek a kutatásnak a célja az azonos nyelvi konstruktumot mérő, különböző vizsgaidőszakokban keletkezett feladatsorok nehézségének, állandóságának és ekvivalenciájának vizsgálata. A Referenciakeret szintjeihez való illeszkedést az európai projektekből származó referenciafeladatokkal lehetne ellenőrizni. Egy ilyen kutatás lebonyolítása és az eredmények értelmezése során meg tudnánk állapítani, hogy az idegen nyelvi érettségi vizsgafadatainak két szintje mennyire igazodik a vizsgázók nyelvtudásának reális szintjéhez. A referenciafeladatok alkalmazásával választ kaphatunk arra a kérdésre is, hogy a vizsgafeladatok szintje hogyan illeszkedik a Referenciakeretben megadott nyelvtudási szintekhez. A kutatás eredménye hozzájárulhat annak ellenőrzéséhez, hogy valóban megfelelőek-e a szintekről alkotott elképzelések.

A második fázisban az érettségi vizsgafeladatokból feladatbank építhető ki. A feladatbankba bekerülő érettségi feladatoknak a Referenciakeret deskriptorai alapján tartalmaznia kell az itemek pszichometriai jellemzőit, tartalmát és formáját. Ebben a fázisban felhasználhatók a Dutch grid-projekt eredményei (Alderson és mtsai, 2004), amelyekkel a feladatokhoz tartozó szövegek és itemek jellemezhetők és leírhatók a Referenciakeret alapján. A létrejövő feladatbank viszonyítási pontként funkcionálhat a további vizsgafeladatok fejlesztésekor. A kutatás hozzájárulhat annak, a szakirodalomban (például Einhorn, 2007b, 2008b; Nikolov, 2007) gyakran megfogalmazott kérdésnek a tisztázásához, hogy az idegen nyelvi érettségi vizsga két szintje hogyan viszonyul egymáshoz és a Referenciakeret szintrendszeréhez. E vizsgálat segítheti azt a folyamatot, amelynek eredményeként egységes skálán tudjuk elhelyezni a diákok teljesítményét.

4.3. Az íráskészség mérési-értékelési eljárásainak minőségi mutatói

A receptív készségek mellett az íráskészség esetében is megvizsgáljuk a vizsga jelenlegi működését és adunk visszajelzést a vizsgafejlesztés számára. Ebben a fejezetben először egy adatbázis-elemzés eredményeit mutatjuk be, melynek során a 2006. évi emelt szintű német érettségi íráskészség vizsgarész pontszámait vizsgáltuk. Először az egyes értékelési szempontok belső összefüggésrendszerének feltárását végezzük el, majd a parciális kredit modellel az egyes értékelési szempontok nehézségéhez tartozó paramétereket elemezzük. Ezt követően azt vizsgáljuk, hogy az egyes skálapontok mennyiben fedik le az érettségizők íráskészsége alapján meghatározott teljesítményszintet, majd az értékelési skálák működését jellemző paraméterek jellemzőit mutatjuk be. Ezután összegezzük a feltárt paramétereket és bemutatjuk, hogy ezek hogyan használhatók fel a vizsgafejlesztésben és a további elemzésekben. Ehhez példaként a 2006. évi középszintű német érettségi vizsgán alkalmazott íráskészséghez tartozó mérési-értékelési eljárásokat választottuk.

¹⁰ A Dutch Grid a <http://www.lancs.ac.uk/fss/projects/grid/> honlapon érhető el.

4.3.1. Az emelt szintű érettségi íráskészség vizsgarész mérési-értékelési eljárásai

Az értékelési szempontok belső összefüggései

Az egyes értékelési szempontok belső összefüggéseinek feltárásához az íráskészség vizsgarészen elért összpontszámra, mint függő változóra számított regresszióanalízist végeztünk. Ennek célja annak megállapítása, hogy milyen mértékű az egyes értékelési szempontok súlya az összpontszámban. Az eredményeket az 51. és az 52. táblázat tartalmazza.

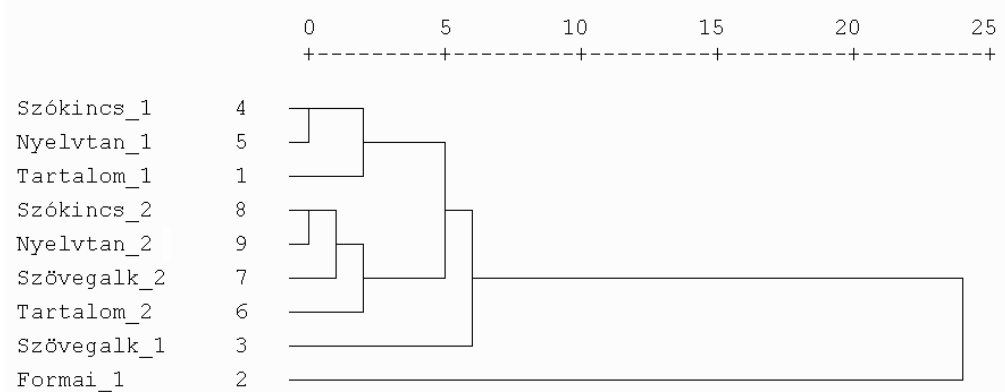
51. táblázat. Az első feladat összpontszámra mint függő változóra számított regresszióanalízis eredménye

Értékelési szempont	r	β	$r*\beta*100$	Szign.
Tartalom	0,86	0,32	27,6	$p<0,001$
Formai jegyek és hangnem	0,47	0,16	7,8	$p<0,001$
Szövegalkotás	0,85	0,13	11,4	$p<0,001$
Szókincs, kifejezésmód	0,91	0,28	25,7	$p<0,001$
Nyelvtan, helyesírás	0,90	0,29	26,2	$p<0,001$
Megmagyarázott variancia (%)	98,6			

52. táblázat. A második feladat összpontszámára mint függő változóra számított regresszióanalízis eredménye

Értékelési szempont	r	β	$r*\beta*100$	Szign.
Tartalom	0,88	0,30	26,2	$p<0,001$
Szövegalkotás	0,93	0,23	21,7	$p<0,001$
Szókincs, kifejezésmód	0,93	0,28	25,7	$p<0,001$
Nyelvtan, helyesírás	0,91	0,28	25,3	$p<0,001$
Megmagyarázott variancia (%)	98,9			

Az egyes értékelési szempontokkal az összpontszám 99%-ban magyarázható, amelyet mindkét feladatban leginkább a tartalmi és a nyelvi komponensek határoznak meg. A szövegalkotás mindkét feladatban alacsonyabb magyarázóerővel rendelkezik. A formai jegyek és hangnem szempont magyarázóereje az első feladatban nagyon alacsony, az összpontszámmal való korreláció értéke is közepes ($r=0,47$; $p=0,001$). Az első feladat megbízhatósága (19. táblázat) 0,91-re növelhető, ha ezt az értékelési szempontot törölnénk. A klaszteranalízis alapján (17. ábra) ez a szempont kapcsolódik legtávolabbról az első feladat többi értékelési szempontjához.



17. ábra

Az értékelési szempontok összefüggése

A második feladat értékelési szempontjai együtt mozognak, mindegyik szempont jelentősen hozzájárul ahhoz, hogy megbízható képet kapjunk a vizsgázók feladatban mért íráskészségéről. Az értékelési szempontok eltérő működésében számos tényező együttesen jelenhet meg. Ide tartozhat az egyes szempontok eltérő nehézsége, skálaszerkezete, nem megfelelő működése és használata. A továbbiakban ezeket a tényezőket a parciális kredit modell segítségével egymástól elkülönítve vizsgáljuk.

Az értékelési szempontok nehézsége

A parciális kredit modellel vizsgált adatbázis eredményeként az 53. táblázat az értékelési szempontok nehézségi indexét és modellilleszkedését mutatja. A táblázatban lévő logitérték a vizsgázók íráskészség-szintjének és az egyes értékelési szempontokhoz tartozó nehézségi paraméterek közös skálájának egységét jelöli. A középérték logitokkal az értékelési szempontok átlagos nehézségét határoztuk meg a vizsgázók íráskészség-szintje alapján. A standard hiba alacsony, így a minta megfelelő információmennyiséget biztosít az értékelési szempontok nehézségének megállapításához.

53. táblázat. Az értékelési szempontok nehézségi indexe és modellilleszkedése

<i>Feladat</i>	<i>Értékelési szempont</i>	<i>Középérték logitok</i>	<i>Standard hiba</i>	<i>Infit paraméter</i>
1.	Tartalom	-0,33	0,03	1,3
	Formai jegyek és hangnem	1,28	0,03	1,7
	Szövegalkotás	-0,82	0,03	0,8
	Szókincs, kifejezésmód	-0,24	0,03	0,9
	Nyelvtan, helyesírás	-0,07	0,03	1,0
2.	Tartalom	0,06	0,03	1,3
	Szövegalkotás	-0,20	0,03	0,8
	Szókincs, kifejezésmód	-0,02	0,03	0,8
	Nyelvtan, helyesírás	0,35	0,09	0,8

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. 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. Az íráskészség azonos komponenseit mérő szempontok közötti távolság a szókincs, kifejezésmódnál a legkisebb (0,2), azonos a tartalom és a nyelvtan, helyesírás szempontnál (0,4) és a legnagyobb a szövegalkotásnál (0,6). Az itt megjelenő nagyobb különbség feltehetően az eltérő skálaszerkezettel (2. melléklet) 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. *Park* (2004) kutatásában a vizsgálatunkhoz hasonló elemszámú mintát használt és 0,8 és 1,2 között határozta meg ezt a sávot. Az adatok értelmezésekor ezt vesszük alapul.

Az 53. táblázatban az első feladatnál két nem illeszkedő szempontot találunk, míg a második feladatban egyet. Az infit paraméter értéke 1,2 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: a két tartalmi szempont, valamint a formai jegyek és hangnem nem illeszkedik. Ezeket az íráskészség kommunikatív komponenseiként határoztuk meg (11. táblázat). Feltételezhetjük, hogy ezek a szempontok függenek az adott feladat megértésétől, a kommunikatív keret felismerésétől, ezért a

vizsgálókból nagy arányban válhatnak ki következetlen válaszokat, de az is elképzelhető, hogy az értékelési skálák nem működnek megfelelően, vagy az értékelők nem használják következetesen az egyes skálapontokat. Ezeket a feltételezéseket a következő alfejezetben részletesebben elemezzük.

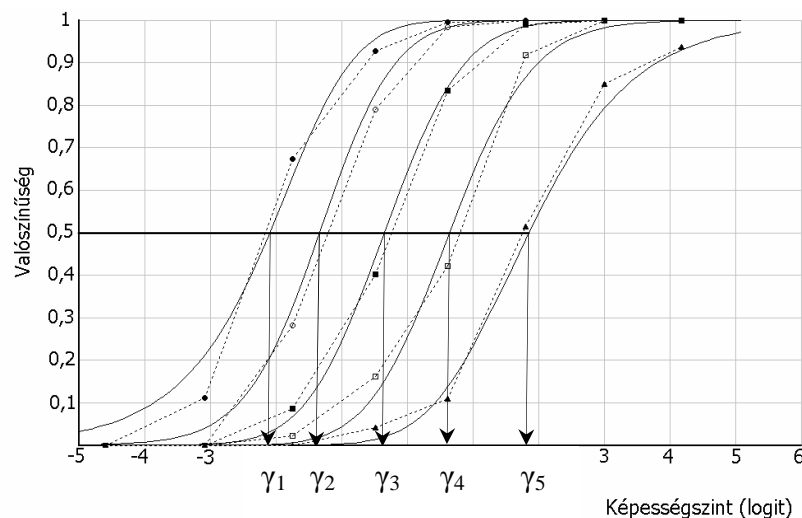
A 0,8 infit paraméter alatt lévő szempontok túlilleszkednek, amely arra utal, hogy az értékelők nem használják azonosan az összes skálafokot, a pontszámok emelkedése nem követi az íráskészség-szint növekedését. Ennyire szélsőséges skálahasználatot az adatok nem mutatnak, de több értékelési szempont 0,8 infit paraméter értékkel rendelkezik, érdemes tehát megvizsgálni részletesen, hogy az értékelők hogyan használják az egyes értékelési skálához tartozó pontszámokat.

Az értékelési skálapontok jellemzői

Az értékelési skálapontok akkor működnek az elvártak megfelelően, ha az íráskészség-szint növekedésével együtt jelentősen és arányosan nő az értékelési szempont-hoz tartozó pontszám elérésének valószínűsége. Ez azt mutatja, hogy a magasabb pontszámot elérő vizsgáló valóban magasabb készség-szinttel rendelkezik, mint az a vizsgáló, aki alacsonyabb pontszámot ér el. Ebből következik, hogy az értékelési skálák alkalmazása megegyezik azzal a szándékkal, amiért készítették őket.

Az értékelési skálákhoz tartozó pontok megfelelő elrendezésének és világos elkülöníthetőségének vizsgálatához az egyes szempontokhoz tartozó ún. thurstoni küszöbértékeket kell megvizsgálni. Ez a mutató megadja, hogy milyen készség-szint szükséges ahhoz, hogy valaki 50%-os valószínűséggel elérjen egy adott pontszámot, így ez az adott értékelési szempont-hoz tartozó nehézségi index értelmezésében játszik szerepet (Wu, 2006a, idézi Molnár, 2008). Ez a mutató hasonlít a dichotóm itemek nehézségi indexéhez (lásd például 12. ábra), de ebben az esetben minden szempont minden pontszámához rajzolhatunk itemkarakterisztikus görbéket.

A 18. ábra példaként az első feladat szókincs, kifejezésmód értékelési szempontnál mutatja az egyes pontszámok thurstoni küszöbértékét. Az ábrán jelöltük ezeket az értékeket: $\gamma_1=-2,09$; $\gamma_2=-1,33$; $\gamma_3=-0,34$; $\gamma_4=0,65$; $\gamma_5=1,85$. Az egyes pontszámokhoz tartozó készség-szint egyre növekszik, ahogyan a pontszámok emelkednek. A thurstoni küszöbértékek további tulajdonsága, hogy a logitskálát olyan intervallumokra osztja fel, amelyen értelmezhetők az értékelési skálához tartozó pontszámok. Ez a konkrét példában arányos, γ_1 és γ_2 pontszám közötti távolság 0,8; míg a többi esetében 1 logitérték.



18. ábra

Az első feladat szókincs, kifejezésmód thurstoni küszöbértéke és modellilleszkedése

A 18. ábra azt is mutatja, hogy az egyes pontszámokhoz tartozó empirikus karakterisztikus görbék (szaggatott vonal) mennyire illeszkednek az adatok alapján elvárt és előre jelzett elméleti görbékhez (folytonos vonal). Ezt a modellilleszkedést az outfit paraméterekkel jellemezhetjük számszerűen (Engelhard és Myford, 2003; Park, 2004), amely az értékelési skálapontok működésének vizsgálatára szolgál. A parciális kredit modellel elemző ConQuest program minden értékelési skálapontnál kiszámolja ezt az értéket. Ha az adatok megfelelően illeszkednek, akkor a paraméter értéke közel van 1-hez. A 18. ábrán lévő értékelési szempontokhoz tartozó pontszámoknál ezek az értékek 0,4 és 1,5 között vannak (54. táblázat).

Az 54. és 55. táblázat tartalmazza az értékelési szempontoknál és pontszámoknál a thurstoni küszöbértékeket és a hozzájuk tartozó outfit paramétereket. Az értékelési szempontokhoz tartozó küszöbértékek az íráskészség-szint növekedésével kivétel nélkül emelkednek. Minden értékelési szempont megfelelően működik: minél magasabb készség-szinttel rendelkeznek a vizsgázók, annál valószínűbb, hogy magasabb pontszámot érnek el. Ebből következik, hogy az értékelési skálák alkalmazása megegyezik azzal a szándékkal, amiért készítették ezeket. (Az egyes értékelési szempontok jellemzőit a 19. ábra alapján értelmezzük.)

54. táblázat. Az értékelési szempontok thurstoni küszöbértéke (γ) és modellilleszkedése (outfit) az első feladatban

Pont	Tartalom		Formai jegyek és hangnem		Szövegalkotás		Szókincs, kifejezőmód		Nyelvtan, helyesírás	
	γ	Outfit	γ	Outfit	γ	Outfit	γ	Outfit	γ	Outfit
1	-2,14	0,5	0,60	1,1	-2,41	0,8	-2,09	0,4	-2,00	0,5
2	-1,37	1,3	1,95	3,4	-0,96	0,8	-1,33	0,7	-1,29	0,9
3	0,09	1,0	-	-	0,91	0,7	-0,34	0,9	-0,32	1,1
4	0,36	1,1	-	-	-	-	0,65	0,9	0,89	0,9
5	1,34	2,0	-	-	-	-	1,85	1,5	2,36	2,4

55. táblázat. Az értékelési szempontok thurstoni küszöbértéke (γ) és modellilleszkedése (outfit) a második feladatban

Pont	Tartalom		Szövegalkotás		Szókincs, kifejezőmód		Nyelvtan, helyesírás	
	γ	Outfit	γ	Outfit	γ	Outfit	γ	Outfit
1	-1,63	1,2	-1,59	0,4	-1,58	0,4	-1,51	0,6
2	-0,86	1,1	-1,05	0,6	-0,93	0,6	-0,70	0,7
3	0,23	1,3	-0,53	0,7	-0,21	0,8	0,01	0,8
4	0,71	1,7	0,44	1,0	0,68	0,9	1,14	0,9
5	1,82	2,0	1,70	0,8	1,89	0,8	2,77	3,5

Az outfit paraméterek jellemzik az adatok modellilleszkedését, azaz a megfigyelt és az elvárt adatok illeszkedését az egyes értékelési pontszámoknál. Ideális esetben ez az érték 1,0-hez van közel. Ha ez az érték 2,0-nél nagyobb, akkor azt jelölheti, hogy az adott pontszámhoz tartozó leírás nem működik megfelelően, az értékelők a különböző képességszintű vizsgázók esetén következtelenül használják az adott pontszámot (Engelhard és Myford, 2003). A formai jegyek és hangnem szemponton kívül mindkét feladatban a nyelvtan, helyesírás szempontnál találtunk 2,0-nél magasabb outfit paramétereket az 5. pontnál. Ennek oka, hogy a 4–5 pont közötti távolság aránytalanul megnő: az első feladatban 1,47 a másodikban 1,63 logitegység. Ez jelezheti, hogy az ötpontos

deskriptort az értékelők ennél a szempontnál nem használják következetesen, de utalhat arra is, hogy ebben az esetben további skálafokra lenne szükség, hogy jobban leírható legyen a 4 és 5 pont között lévő vizsgázók íráskészség-szintje. Ennek megállapításához azonban már szükséges annak ismerete is, hogy egyrészt az értékelők mennyire szigorúan értékelték, másrészt pedig, hogy melyek a vizsgázók íráskészség-paramétereinek jellemzői. Adataink csak a második kérdés vizsgálatára adtak lehetőséget, mivel az adatbázisban nem volt arra vonatkozó adat, hogy egy értékelő hány vizsgázó írásművét értékelté.

A vizsgázók íráskészség szintje és az értékelési skálapontok összekapcsolása

A parciális kredit modell segítségével közös skálán tudjuk ábrázolni a vizsgált személyek íráskészség-szintjét és az 54. és 55. táblázatban jellemzett skálapontok thurstoni küszöbértékét. A minta- és skálapont-térkép (19. ábra) első oszlopában a közös logitskálán kifejezett íráskészség feladatokon elért eredményeket egy képességskálán helyezi el a ConQuest program. A második oszlop a vizsgázók íráskészség-paramétereit, míg a többi oszlop az egyes értékelési szempontokhoz tartozó thurstoni küszöbértékeket jeleníti meg. A parciális kredit modellel elemző szoftverek az adatokat intervallumskálára transzformálják, így a logitskála a készségi szinteken és a thurstoni küszöbértékeken belüli és ezek közötti relatív távolságokat mutatja. A továbbiakban először a minta íráskészség-paramétereinek eloszlását, majd az értékelési skálapontok jellemzőit, végül a kettőt együtt elemezzük.

A mintatérképen minden 'x' öt vizsgázót reprezentál, ugyanakkor az ábrázolt 'x'-eken kívül is vannak még vizsgázók a logitskála magasabb-alacsonyabb tartományában, csak ötnél kevesebben vannak, így a 19. ábra nem tartalmazza az adataikat. A logitskálán megadott negatív számok átlag alatti íráskészség-szintet, a 0 az átlagos, míg a pozitív számok az átlag feletti szintet jelölik. A mintatérképről leolvasható a vizsgázók íráskészség paramétereinek jellemzői. A vizsgázók íráskészsége tág határok között mozog, 6 logitegységnyi a skála. A minta eloszlása két maximummal rendelkezik: a -0,2 és az 1-es logitértéknél, így bimodális eloszlás jellemző. Az érettségizők többsége, a minta kb. 70%-a helyezkedik el az átlagos készségi szint felett.

A skálapont-térképen az egyes értékelési szempontokhoz tartozó pontszám thurstoni küszöbértékét jelöltük, amelyet az a készségi szint reprezentál, ahol az adott pontszám elérésének valószínűsége 50%. A pontos értékeket az 54. és 55. táblázat tartalmazza. A továbbiakban az egyes szempontok jellemzőit elemezzük.

A két feladatban a tartalom szempontokhoz tartozó pontszámok (19. ábra: 1. és 6. szempont) nagyon hasonlóan működnek. Jellemző, hogy az 1 és 2, továbbá a 3 és 4 pont közötti távolság kisebb, mint a 2 és 3, valamint a 4 és 5 pont közötti távolság¹¹. Ez feltehetően az értékelési skála szerkezetéből adódik, hiszen az 1–2, 3–4 pontszámokhoz azonos leírás tartozik, míg a 2–3 és a 4–5 pontszámhoz különböző szintleírások (a szövegalkotás szempontokhoz hasonlóan, lásd a 2. mellékletet). Az empirikus adatok mutatják az értékelési skála szerkezetét: az egyes pontszámok nem azonos lépésköze helyezkednek el egymástól.

¹¹ Az 1–2 pont esetén a képességi szintbeli távolság 0,8; a 3–4 pont esetén 0,3 az első és 0,5 a második feladatban. A 2–3 pont esetén a γ_2 és γ_3 közötti távolság 1,5 és 1,1; a 4–5 pont esetén pedig 1 logitegység.

logit- skála	személy	1. feladat					2. feladat				
		1	2	3	4	5	6	7	8	9	
6											1: Tartalom 2: Formai jegyek, hangnem 3: Szövegalkotás 4: Szókincs, kifejezésmód 5: Nyelvtan, helyesírás
5											6: Tartalom 7: Szövegalkotás 8: Szókincs, kifejezésmód 9: Nyelvtan, helyesírás
4	X X X X XX										
3	XXX XXX XXX XXX										9.5
2	XXXXX XXXXX XXXXXX XXXXXXXX XXXXXXXX XXXXXXXX	2.2		4.5		6.5		8.5			
1	XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX	1.5									9.4
0	XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX XXXXX XXXXX XXXXX	1.3									9.3
-1	XXXX XXX XXXX XXX XX XX X		3.2								9.2
-2	X X X	1.1		4.1	5.1						
-3			3.1								
-4											
-5											

19. ábra

Az emelt szintű íráskészség feladatlap minta- és skálapont-térképe

Megjegyzés: A ConQuest program az eredmények ábrázolásakor az értékelési szempontokat jellemző küszöbértékeket egymás mellé helyezi. Ezt az ábrát szerkesztettük, hogy az egyes értékelési skálapontok thurstoni küszöbértéke vizuálisan jobban elkülöníthető legyen.

A szövegalkotásnál (19. ábra: 3. és 7. szempont) a két feladathoz különböző skálaszerkezet tartozik. Az első feladatban minden pontszámhoz azonos leírás tartozik, az

adatok is azt mutatják, hogy a thurstoni küszöbértékek közel azonos intervallumokra vannak (1,5 és 1,9 logit-egységnyi különbség van az egyes pontszámok között). A második feladatban viszont az 1–2 és a 2–3 közötti távolság közel azonos (0,5 logitegység), majd egyre növekszik: a 3–4 közötti távolság 1 logitegység, a 4–5 között pedig 1,3, így ennél a szempontnál az empirikus adatok nem követik a javítási-értékelési útmutatóban megadott skála szerkezetét (2. melléklet). Ennek lehetséges okait a következő alfejezetben részletezzük.

A szókincs, kifejezésmód és nyelvtan, helyesírás esetében (19. ábra: 4., 5., 8. és 9. szempont) az egyes pontszámok nagyon hasonlóan működnek. Mindkét feladatra jellemző, hogy az 1, 2, 3 és 4 pont közötti thurstoni küszöbértékek közel azonosak és a köztük lévő távolság is arányos, míg a 4–5 pont között aránytalanul megnő és ezek modellilleszkedése sem megfelelő (lásd 54. és 55. táblázat).

Ha a közös logitskálán kifejezett vizsgázók íráskészség-paramétereit és az egyes értékelési szempontokhoz tartozó thurstoni küszöbértékeket együtt vizsgáljuk, akkor a minta- és skálapont-térképről leolvashatjuk, hogy az egyes értékelési szempontokhoz tartozó thurstoni küszöbértékek mennyire jól fedik le a vizsgázók készség szintjei által meghatározott intervallumot. A mintához jól illesztett tesztnél a készségparamétereket jelző 'x'-ek és az értékelési skálapontokat jellemző thurstoni 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. 37%-át az értékelési skálákhoz tartozó maximális pontszámot leíró szintleírások jellemzik, amelyek nem fedik le teljesen a vizsgázók íráskészség-szintjét. Az 1–2 pontot leíró deskriptor a minta 15%-ára jellemző. A minta többségének szintjét a 3–4 pontot tartalmazó szintleírások jellemzik. Ezek az értékek arra utalhatnak, hogy az alacsonyabb pontszámhoz tartozó deskriptorok kevésbé jellemzők az érettségizőkre, míg az ötpontos esetében további differenciálás lenne indokolt.

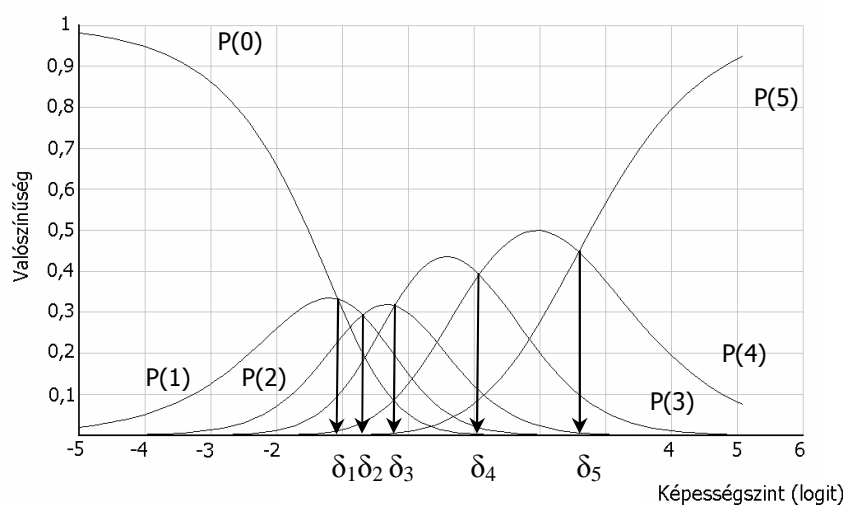
Az értékelési skálák működésének mutatói

Az előzőekben az egyes értékelési szempontok nehézségét és a vizsgázók íráskészségét 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.

A kérdések vizsgálatához a parciális kredit modellel felrajzolható karakterisztikus görbéket vizsgáljuk. Ezek különböző íráskészség-szintek mellett annak valószínűségét mutatják, hogy a vizsgázó teljesítményét az adott értékkategóriába sorolják az értékelők. A továbbiakban a második feladat három értékelési szempontjához tartozó karakterisztikus görbéket elemezzük, amelyek tipikusnak mondhatók. A 19., 20. és 21. mellékletben található a többi értékelési szemponthoz tartozó skálák karakterisztikus görbéi.

A 20. ábra a nyelvtan, helyesírás szempontjához tartozó értékelési skála használatának eredményét mutatja. A legfelső görbe a 0 pontba tartozás valószínűségét jelöli, amely azt mutatja, hogy a készség szint növekedésével hogyan csökken a 0 pontszám elérésének valószínűsége. Az 1 pontszám valószínűségi görbéje azt jelöli, hogy a készség szint növekedésével hogyan növekszik, majd csökken az 1-es pontszám valószínűsége. Ugyanezt látjuk a többi pontszámhoz tartozó görbénél is. Az egyes görbék metszéspontját $\delta_1, \delta_2 \dots \delta_5$ paraméterekkel jelöltük. A δ_1 paraméter például azt a készség szintet jelöli, ahol azonos annak valószínűsége, hogy a vizsgázók 0 vagy 1 pontot érnek el. A többi delta paraméter azokat a készség szinteket jelöli, ahol azonos annak a valószínűsége, hogy a vizsgázó eredménye a k-adik, vagy a k+1-dik pontszámba so-

rolható. Az öt δ érték (-1,06; -0,68; -0,19; 1,09; 2,59) emelkedik a készpénz szint növekedésével.



20. ábra

A második feladat nyelvtan, helyesírás értékelési szempont karakterisztikus görbéje

A paraméterek közül a δ_1 , δ_2 és δ_3 egymáshoz közelebb van, mint a δ_4 és δ_5 , amely arra utal, hogy kevés azon diákok száma, akik 1 vagy 2 pontot kapnak és meghatározóvá válik az a vizsgázói csoport, akik 3, 4 és 5 pontot érnek el (20. ábra). A δ értéke attól függ, hogy az egyes pontszám-kategóriában hány tanuló van, viszont az elemzésbe bevont minta vizsgapontszám eloszlása megegyezik a vizsgázói populáció teljesítmény-eloszlásával (18. táblázat).

A többi nyelvi komponens mérő szempontok δ paramétereit az 56. táblázat tartalmazza és a 19. mellékletben található a karakterisztikus görbék. Ezek azt mutatják, hogy a δ_1 és δ_2 értékek közel azonosak, majd a δ_3 , δ_4 és δ_5 paraméterek azonos távolságra vannak. Ez azt jelzi, hogy a 0–1 és az 1–2 pontba tartozás valószínűsége nagyon alacsony és meghatározóvá válnak azok az érettségizők, akik 0, 2 és 5 pont közötti eredményeket érnek el, azaz a 6 pontszámból álló skála 5-fokozatúként működik. Ez azt jelezheti, hogy az 1–2 pontos deskriptor megfogalmazása nem megfelelő, vagy nincs szükség rá, mert nem jellemző a vizsgázókra, vagy pedig az értékelők valamilyen okból nem akarják alkalmazni ezeket a skálapontokat. Ennek feltehetően az az oka, hogy a vizsgázókra kevésbé jellemző az 1 pont elérése (lásd 19. ábra).

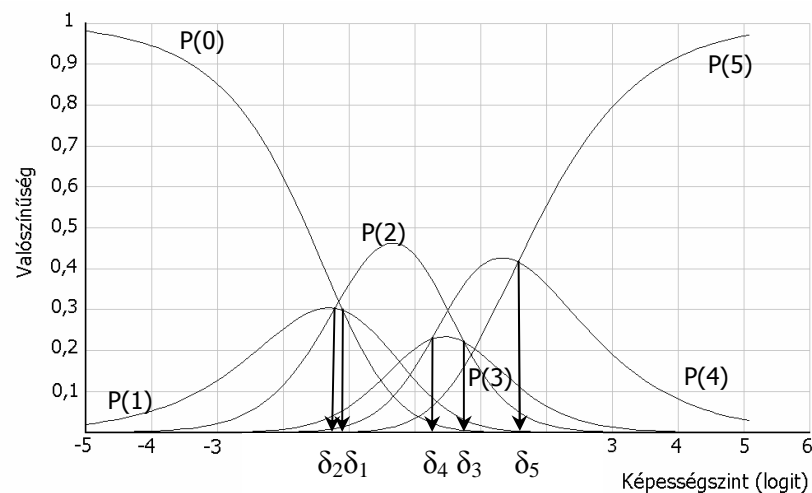
56. táblázat. Az értékelési szempontokhoz tartozó skálák minőségi mutatói

Feladat	Értékelési szempont	δ_1	δ_2	δ_3	δ_4	δ_5
1.	Tartalom	-1,56	-1,86	1,18	-0,51	1,08
	Formai jegyek és hangnem	0,91	1,65	-	-	-
	Szövegalkotás	-2,18	-1,06	0,77	-	-
	Szókincs, kifejezésmód	-1,55	-1,57	-0,30	0,63	1,57
	Nyelvtan, helyesírás	-1,44	-1,55	-0,35	0,87	2,15
2.	Tartalom	-1,09	-1,21	0,76	0,27	1,57
	Szövegalkotás	-0,91	-1,13	-0,82	0,45	1,43
	Szókincs, kifejezésmód	-0,99	-1,09	-0,26	0,62	1,62
	Nyelvtan, helyesírás	-1,06	-0,68	-0,19	1,09	2,59

A szövegalkotás szempontozott emelt szinten kétféle skálaelrendezés tartozik (2. melléklet). Ugyanakkor a két karakterisztikus görbe (20. melléklet) hasonló működést

mutat. Az első feladathoz tartozó skálán a δ értékek folyamatosan növekednek (56. táblázat), a második feladathoz tartozó szövegalkotást jellemző paraméterek a 20. ábrához hasonló képet mutatnak, csak szélsőségesebb a skálahasználat. Ennek oka, hogy a paraméterek felcserélődnek: $\delta_2 < \delta_1 < \delta_4 < \delta_3 < \delta_5$, ahol δ_1 , δ_2 és δ_3 paraméterek értéke közel azonos (56. táblázat). Ez a skálaműködés azt jelzi, hogy a -1-es logitértékig a 0 pontszám, majd a 0,45 logitértéktől a 3 pont elérése valószínű. Ennek az értékelési skálának a diszkriminációs indexe is magas (0,9), így ez a skála egy túldiszkrimináló dichotóm item jelleggörbéjéhez hasonlít (lásd például 15. ábra). 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.

A tartalom értékelési szemponthoz tartozó értékelési skála (21. ábra és 21. melléklet) mindkét feladatban hasonlóan következtelenül működött. Jellemző az egyes paraméterek felcserélődése: $\delta_2 < \delta_1 < \delta_4 < \delta_3 < \delta_5$, ahol a δ_1 és δ_2 értéke hasonló és a δ_4 közel van a δ_3 -hoz (56. táblázat). 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, de a szövegalkotás szemponttól eltérően itt azt tapasztaljuk, hogy mind a négy deskriptort használják az értékelők, csak az egyes pontszámok valószínűsége nem azonos. Feltehető az a vizsgáztatói gyakorlat, hogy az azonos deskriptorokat leíró pontszámoknál a magasabbat adják a vizsgázónak.



21. ábra

A második feladat tartalom értékelési szempont karakterisztikus görbéje

A formai jegyek és hangnem szempontnál egy dichotóm item jelleggörbéje rajzolódik ki (21. melléklet). A 0,91 és 1,65 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 címzetthez való viszonyoknak, vagy pedig nem felelnek meg. Feltehetően ez a skálaműködés az oka annak, hogy e szempont nem illeszkedik a többi közé (53. táblázat), a legtávolabbról kapcsolódik a többi szemponthoz (17. ábra) és a legalacsonyabb magyarázóerővel rendelkezik (51. táblázat).

A feltárt paraméterek jellemzői és értelmezése

A parciális kredit modellel végzett elemzés során számos paramétert definiáltunk, amelyek segítségével a vizsgázók íráskészségének jellemzőit, az értékelési szempontok nehézségét és az egyes értékelési skálák működését elemezhetjük. Eredményeink szerint az értékelési szempontokat jellemezhetjük (1) az átlagos nehézségi szintjük és (2)

modellilleszkedésüket az *infit* paraméterek alapján. A nem illeszkedő értékelési szempontok azt jelezhetik, hogy az íráskészség más dimenzióit mérik és így működésük eltérő (például a kommunikatív komponensek az első és a második feladatban). Az egyes értékelési skálapontok működését (3) a *thurstoni* küszöbértékek jellemzik, amelyek akkor megfelelőek, ha a készségszint növekedésével arányosan emelkednek. Ehhez a mutatóhoz tartozik (4) az *outfit* modellilleszkedés-paraméter, amelynél a 2,0 fölötti érték azt jelzi, hogy az adott skálapont nem működik megfelelően. Ennek oka például az lehet, hogy nagyon különböző készségszintű vizsgázók esetében is azonos pontszámot adnak az értékelők (például a második feladat nyelvtan, helyesírás szempontnál). A vizsgázók íráskészség-szintjét és az értékelési szempontok *thurstoni* küszöbértékét (5) minta- és skálapont-térképen ábrázolhatjuk. Ez azt mutatja, hogy az egyes értékelési szempontok mennyire fedik le a vizsgázók íráskészség-szintje alapján meghatározott intervallumot. Ez alapján azonosítható egy magasabb szint, amelyet az egyes skálapontok küszöbértékei nem fednek le. Az értékelési skálahasználatot (6) karakterisztikus görbékkel jellemezhetjük, amelyek mutatói (7) a *delta*-paraméterek. Ezek felcserélődése és egymástól való távolságuk jellemezhetik az értékelési skálák redundáns működését (például a második feladatban a szövegalkotásnál), valamint azt, hogy az értékelési skála bizonyos pontszámait az értékelők nem használják (például a tartalom szempontoknál).

Eredményeink szerint a vizsgán alkalmazott értékelési eljárások magas megbízhatóságot eredményeznek, ugyanakkor vizsgálatunk az egyes értékelési szempontok (például formai jegyek és hangnem) diszfunkcióit, az értékelési skálák nem teljesen megfelelő működését és használatát is mutatják. Az egyes értékelési skálák módosítására a bemutatott eredmények alapján elhamarkodott lenne javaslatokat tenni, hiszen szükséges a fenti paraméterek mentén több évi érettségi eredményeinek ismerete. Az azonban látható, hogy a teljesítmények javulásával (lásd 4.1.3. fejezet) az értékelési szempontokhoz tartozó alacsonyabb pontszámok inflálódhatnak, azaz egy idő elteltével az alacsonyabb szintleírás feltehetően nem lesz jellemző az érettségizőkre. Ennek oka, hogy a feladatokat és az értékelési rendszert a diákok a vizsgára való felkészülés során megismerik, amely jelentősen befolyásolja a teljesítményüket.

A vizsgálatunk jelentősége, hogy a feltárt paraméterek mentén a közép- és emelt szinten megjelenő tendenciák leírhatók és a különböző évekből származó vizsgaeredmények összehasonlíthatók és változásaik nyomon követhetők. A továbbiakban a feltárt paraméterek használhatóságát mutatjuk be a 2006. évi középszintű érettségi vizsga íráskészség részének elemzésével.

4.3.2. A középszintű érettségi íráskészség vizsgarész mérési-értékelési eljárásai

A középszintű érettségi íráskészség vizsgarész mérési-értékelési eljárásai az emelt szintű vizsgához képest más jellegű kutatási kérdéseket is felvetnek. A középszintű érettségin a feladatoknak a *KER* (2002) két nyelvi szintjét, az A2–B1 szintet kell lefedniük, így a feladatoknak az emelt szinthez képest egy nagyobb képességtartományt kell átfolgni. A feladatok és az értékelési szempontok előzetes kipróbálása, nehézségi szintjük vizsgálata nem történt meg, így ebben a részben a 4.2. fejezethez hasonlóan arra koncentrálunk, hogy a Referenciakeret két nyelvi szintje mennyire különül el a mérési-értékelési eljárások alkalmazásában.

Vizsgálatunk célja, hogy – a receptív készségekhez hasonlóan – az előző alfejezetben részletesen, példákkal illusztrált minőségi mutatókat a középszintű német érettségin is megvizsgáljuk. Az értékelési szempontok nehézségénél így azt elemezzük, hogy a feladatkészítők által megállapított nehézség mennyire van összhangban az empirikus

adatokkal. A vizsgázók íráskészségével kapcsolatban azt vizsgáljuk, hogy az értékelési szempontokhoz tartozó thurstoni küszöbértékek mennyire fedik le a vizsgázók íráskészség-szintje alapján meghatározott intervallumot. Vizsgálatunkban a 2006. évi középszintű német érettségi vizsgafeladatokhoz tartozó adatbázist (az adatbázis jellemzőiről lásd a 3.3.1. fejezetet) alkalmaztuk. A középszintű vizsgán lévő teljesítmény-eloszlásról nincs adatunk, így nem vizsgáltuk az egyes értékelési skálák működését, mert ezek megítélése függ attól, hogy hány vizsgázó ér el egy adott pontszámot.

A 2006. évi érettségi íráskészség feladatainak megbízhatósága a többi nyelvi készséghez képest magas (17. táblázat): az első feladatban mért Cronbach- α érték 0,86, a második feladaté 0,93, a teljes teszté 0,92. A továbbiakban csak a parciális kredit modellel végzett elemzés eredményeit mutatjuk be, ugyanezen adatbázis vizsgálata alapján az egyes szempontoknál elért eredményeket *Einhorn* (2007a) tanulmányában olvashatjuk.

Az 57. táblázat az értékelési szempontok átlagos nehézségi szintjét a készség-szint alapján logitértékben kifejezve adja meg. A standard hiba alacsony, azaz a minta elemszáma megfelelő információmennyiséget ad az átlagos nehézségi szint meghatározásához. A táblázat az adatok illeszkedését mutató infit paramétereket is tartalmazza.

57. táblázat. Az értékelési szempontok nehézségi indexe és modellilleszkedése

Feladat	Értékelési szempont	Középérték logitok	Standard hiba	Infit paraméter
1.	Tartalom és szöveghossz	-0,74	0,03	1,7
	Nyelvhelyesség	-0,05	0,03	1,3
2.	Tartalom	0,03	0,03	0,9
	Formai jegyek, hangnem	0,13	0,03	1,5
	Szövegalkotás	-0,04	0,03	0,7
	Szókincs, kifejezőmód	0,05	0,03	0,7
	Nyelvtan, helyesírás	0,62	0,07	0,8

Az első feladathoz tartozó értékelési szempontok teljesítésének nehézsége a modellált populáció átlagos készség-szintje (0 logitérték) alatt helyezkedik el, a második feladat értékelési szempontjainak átlagos nehézsége az átlagos szintnél és a nyelvtan, helyesírás szempont az átlag fölötti szinten helyezhető el. Mindkét feladatban a legszigorúbban a nyelvi komponenseket értékelték. Az azonos szempontok közötti távolság a tartalomnál 0,8; míg a nyelvhelyességnél 0,7, így a két feladat feltehetően különböző nyelvi szintet mér a feladatkészítők szándékainak megfelelően.

Az első feladat két szempontja, valamint a formai jegyek és hangnem nem illeszkedik, mivel az infit paraméter érték 1,2 fölött van. 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 az emelt szintű feladatokhoz hasonlóan.

Két értékelési szempontnál utalnak az adatok túlilleszkedésre (az infit paraméter értékek 0,8 alatt vannak). Az önálló elemzés során is 0,7 körüli értéket kaptunk a szövegalkotás és a szókincs, kifejezőmód szempontnál. Mindkét esetben a diszkriminációs index magas (0,94). 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 vizsgál-

zónál 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.

A 22. ábra mutatja a középszintű íráskészség vizsgarész minta- és skálapont-térképét, amelyen a vizsgázók íráskészség-szintjét és az értékelési szempontokhoz tartozó thurstoni küszöbértékeket helyeztük el közös logitskálán. Ennek jellemzőit a 19. ábránál mutattuk be.

A mintatérkép alapján megállapítható, hogy nagy különbség van a vizsgázók íráskészség-szintje között, a logitskála 8 egységnyi, így indokolt a vizsgán a két nyelvi szint mérése. A minta közel 70%-a az átlagos készségszint felett helyezkedik el.

A skálapont-térképen elhelyezett értékelési szempontokhoz tartozó thurstoni küszöbértékek az íráskészség-szint növekedésével kivétel nélkül emelkednek. Minden értékelési szempont megfelelően működik: minél magasabb készségszinttel rendelkeznek a vizsgázók, annál valószínűbb, hogy magasabb pontszámot érnek el.

Az első feladathoz tartozó értékelési skálapontok többsége (22. ábra: 1. és 2. szempont) az átlagos készségszint alatt helyezkedik el. Ugyanakkor a nyelvhelyességnél a 4. és 5. pont közötti távolság aránytalanul nagy. Az outfit paraméter még elfogadható (1,9) a két pont esetében, így nem indokolt újabb deskriptor alkalmazása. Egy korábbi, kis-mintás vizsgálat tapasztalatai azt mutatták (Einhorn, 2007a), 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 az összes szempontnál az 1, 2 és 3 pontszám elérése az átlagos íráskészség-szint alatt jellemző. A kommunikatív komponenseket mérő szempontoknál a készségszint növekedésénél viszonylag arányosan emelkedik az adott pontszám küszöbértéke. A szövegalkotás és a szókincs, kifejezésmód esetében az 1–2 pont közötti távolság kicsi (0,4) a 2–3 pont közötti távolsághoz (0,6) hasonlóan, a 3–4 pont közötti távolság a szövegalkotásnál 0,8, a szókincs, kifejezésmódnál 1 logitegység; a 4–5 pont között pedig 1,4. Ezek az adatok arra utalnak, hogy a 4. és az 5. pontszám közé egy, a diákok készségszintjét jobban leíró deskriptorra lenne szükség. A nyelvhelyesség esetében az 1–2 és a 2–3 közötti távolság közel azonos (0,6 és 0,8 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 2-es logitértékig többnyire megfelelően fedik le az egyes értékelési szempontokhoz tartozó pontszámok thurstoni küszöbértékei a vizsgázók íráskészség-szintje alapján meghatározott intervallumot. Ugyanakkor a minta kb. 30%-ának jellemzésére az értékelési skálák deskriptorai kevésbé alkalmasak.

logit- skála	személy	1. f.		2. feladat					
		1	2	3	4	5	6	7	
									1: Tartalom és szöveghossz 2: Nyelvhelyesség
									3: Tartalom 4: Formai jegyek és hangnem 5: Szövegalkotás 6: Szókincs, kifejezőmód 7: Nyelvtan, helyesírás
7									
6									
5	X								
	X								
	X								
	X								
4	XX								
	XXX								
	XXX							7.5	
	XXXX								
3	XXXX								
	XXXXX								
	XXXXX								
	XXXXX		2.5						
	XXXXXX						6.5		
2	XXXXXXXX			3.5		5.5			
	XXXXXXXX							7.4	
	XXXXXXXXXX								
	XXXXXXXXXX				4.3				
1	XXXXXX								
	XXXXXXXX	1.5	2.4						
	XXXXXXXX			3.4			6.4		
	XXXXXX					5.4			
	XXXXXXXXXX				4.2			7.3	
0	XXXXXX			3.3					
	XXXXXX	1.4	2.3			5.3	6.3		
	XXX								
	XXX	1.3		3.2	4.1			7.2	
-1	XXXX					5.2	6.2		
	XX	1.2	2.2			5.1	6.1		
	XX			3.1				7.1	
	XX								
-2	X	1.1							
	X		2.1						
	X								
	X								
-3									
-4									
-5									

22. ábra

A középszintű íráskészség feladatlap minta- és skálapont-térképe

Megjegyzés: A minta- és itemtérképen minden 'x' hét vizsgázót jelöl.

A 2006. évi középszintű érettségi íráskészséget mérő vizsgarészében megállapítottuk, hogy a Referenciakeret két nyelvi szintje a vizsgán alkalmazott értékelési eljárásokban is elkülönül: az infit paraméterek alapján kimutattuk, hogy az első és második feladat az íráskészség különböző dimenzióit méri. Az egyes értékelési skálapontok középszinten sem fedik le a vizsgázók íráskészsége alapján meghatározott intervallumot. Vizsgálatunk érdekessége, hogy feltehetően a tanárok a nyelvtan, helyesírás szempontot kiemelve értékelik a vizsgázók írásműveit.

A fejezetben bemutatott vizsgálatunk hazai kontextusban több szempontból tekinthető újszerűnek. Egyrészt a parciális kredit modell esetén a hazai szakirodalomból a parametrizálási lehetőségeket (*Molnár, 2008*) és egy olvasási attitűdkérdőív adatainak elemzését (*Kontra, 2009*) ismerjük, az íráskészségnél viszont nem jellemző az alkalmazása. Másrészt a bemutatott minőségi mutatók olyan vizsgálatokban is értelmezhetők, alkalmazhatók és elemezhetők, ahol az értékelés analitikus skálákkal történik. Harmadrészt pedig az idegen nyelvi érettségiről főként kvalitatív elemzések és a nagymintás vizsgálatokból származó leíró statisztikai elemzések eredményeit ismerjük, így kvantitatív vizsgálatunk segítheti a vizsgafejlesztési folyamatot. A vizsgálatunkban feltárt paraméterek mentén a különböző évekből származó vizsgaeredmények összehasonlíthatók és változásai nyomon követhetők.

5. AZ IDEGEN NYELVI ÉRETTSÉGI ELEMZÉSE A TANÁROK VISSZAJELZÉSE ALAPJÁN

A fejezetben az értekezés 3. fejezetében kijelölt kutatási területekből a negyediket, a tanárok vizsgával kapcsolatos (1) nézeteit és attitűdjeit elemezzük, (2) a vizsga nyelvoktatási folyamatra és (3) osztálytermi mérés-értékelésre kifejtett hatását vizsgáljuk a tanárok nézetei alapján, majd pedig (4) a vizsgával kapcsolatos nézetek és attitűdök befolyásoló szerepét a nyelvoktatás folyamatára és mérési-értékelési eljárásaira. Ebben a fejezetben egy kötelező vizsgatárgy, az idegen nyelvi érettségi nyelvoktatási folyamatra gyakorolt hatását feltáró vizsgálatunk eredményeit mutatjuk be. A nyelvoktatás sajátos módszertannal, mérési-értékelési jellemzőkkel rendelkezik, így az empirikus kutatásunk eredményei tartalmi kötöttségűek, területspecifikusak. A bemutatandó vizsgálat eszközei és eredményei valószínűleg általánosíthatók a nyelvtudást mérő vizsgák hatásaira, de a kutatásunk eredményei további érettségi tárgyakra csak korlátozottan lehetnek érvényesek és nagyon korlátozottan más értékelési rendszerek, például a kompetenciamérések hatásmechanizmusaira. A vizsgálatot a pedagógiai kutatómódszertanban megszokott eljárások szerint terveztük meg, bonyolítottuk le és elemezzük az eredményeket. A kérdőívünk validitásáért nagy hangsúlyt fektettünk (lásd 3.3.2. fejezet), a kiválasztott minta reprezentatívnak tekinthető a tanított idegen nyelv és iskolatípus szerint (21. táblázat), az eredményeket statisztikai eljárásokkal dolgoztuk fel. Ugyanakkor hangsúlyozni kell, hogy az ebben a fejezetben bemutatandó eredmények tartalomfüggők és általánosíthatóságuk az empirikus vizsgálatunk tárgya és korlátai miatt nem lehetséges.

A tanári kérdőívben ötfokozatú Likert-skálát használtunk, melyet azért választottunk, hogy a tanárok a vizsgához kapcsolódó semleges álláspontjukat is kifejezhessék és így a közömbös megkérdeszettek nem kényszerítettük döntésre. Minden esetben kiszámoltuk az adott kérdőív-tétel átlagos támogatottságát. Több kérdőív-tételt a nemzetközi teszthatást vizsgáló felmérésekből ismert és publikált tanári kérdőívekből vettünk át és alakítottuk saját kontextusunkra (lásd 3.3.2. fejezet és 3. melléklet). Ezzel lehetőség nyílik arra, hogy a tanárok válaszait viszonyítsuk az ott kapott eredményekhez. Természetesen nem lehet a hasonlóságokból és a különbségekből messzemenő következtetéseket levonni. A teszthatást vizsgáló nemzetközi empirikus kutatások tanulságai szerint (1.3.3. fejezet) oktatási kontextusonként jelentős különbségek lehetnek a hatás mértékében, irányában és jellegében. A felhasznált kérdőívek eredményei viszont az adatelemzésben támpontot nyújthatnak.

5.1. A tanárok vizsgával kapcsolatos nézetei és attitűdjei

A vizsgára vonatkozó tanári nézetek és attitűdök kérdőív-rész elemzésében három faktor jött létre: (1) a tanárok nézetei a vizsga jelentőségéről, (2) a vizsga következményeivel kapcsolatos attitűdök, (3) a tanórai felkészüléssel kapcsolatos attitűdök, melyeket a vizsga befolyásolhat (34. táblázat). A továbbiakban először a faktorok alapján jellemezzük az egyes kérdőív-tételek átlagos támogatottságát, elemezzük a kutatási kérdéseknek megfelelően a háttérváltozókkal való kapcsolatát, majd pedig azokat a kérdőív-tételeket is vizsgáljuk egyenként, melyekből a teljes kérdőív-re végzett faktoranalízis alapján nem hoztunk létre összevont változót.

Az 58. táblázat „a tanárok nézetei a vizsga jelentőségéről” faktorba sorolt állítások alapstatisztikai adatait tartalmazza. A kérdőív-tételeket az átlagos támogatottság mértéke szerint csökkenő sorrendben helyeztük el, és megadtuk az átlagot, a szórást és az átlag

standard hibáját, amely minden esetben 0,10 alatt van. Ezen kívül pedig páros t-próbával meghatároztuk az egyes állítások átlagos támogatottsága közötti különbséget.

Az első faktorba sorolható változók közül a nyelvtanárok leginkább azt az állítást fogadták el, hogy a vizsga illeszkedik a *Közös Európai Referenciakerethez* (2002). Ugyanakkor ezen állítás igazságtartalmának vizsgálata, az érettségi szintjeinek a *KER*-hez (2002) való illesztése a nyelvtudásmérés egyik legfontosabb kutatási feladata, hiszen jelenleg nem rendelkezünk empirikus adatokkal erről (2.3.4. fejezet). A vizsgakövetelmények és a vizsgaleírás alapján a tanárok megfelelőnek találják a vizsga illeszkedését a Referenciakerethez. A tanárok úgy érzik, hogy a kétszintű érettségi a korábbi vizsgánál jobban segíti a javítói és értékelői munkájukat, hasonló módon értettek egyet azzal, hogy a vizsga elősegíti a nyelvi készségek fejlesztését a nyelvvoktatásban, a korábbi vizsgánál jobban méri a diákok nyelvtudását és a NAT-ban deklarált célokat. A vizsga presztízsével és hasznosságával kapcsolatos kérdőívtelemek (1., 3. és 7.) megítélésében a tanárok kissé kritikusabbak voltak. A kétszintű nyelvi érettségi tartalmi és módszertani modernizációjának elindításával kapcsolatos állítást többnyire elfogadták a válaszadók. Hasonló mértékben értettek egyet azzal, hogy a vizsga a korábbi érettséginél nagyobb jogosítványt nyújt és ezektől szignifikánsan elkülönülve legkevésbé azzal értettek egyet, hogy az érettségi a korábbi vizsgánál jobban szolgálja a felsőoktatás igényeit. Ebben az esetben a vizsgának közepes hatása van a tanárok nézeteire. Ez azt jelezheti, hogy a tanárok szerint a vizsga kevéssé teljesíti a felvételi funkcióját. Ez összhangban van a 2.3.4. fejezetben bemutatott szakirodalmi összegzéssel és a 4.1.2. és a 4.1.3. fejezet empirikus eredményeivel. A tanárok a saját tanítási gyakorlatuk minőségére az érettségi eredményekből közepes mértékben következtetnek.

58. táblázat. A tanárok nézetei a vizsga jelentőségéről

Nr.	Kérdőívtelem	Átlag	Szórás	Std. hiba
<i>A kétszintű idegen nyelvi érettségi vizsga...</i>				
4.	illeszkedik a Közös Európai Referenciakeret szintrendszeréhez.	4,16	0,94	0,08
8.	a korábbi vizsgánál jobban segíti a javítói és értékelői munkámat.	3,90*	1,04	0,09
6.	elősegíti a nyelvi készségek fejlesztését a nyelvvoktatásban.	3,89	1,12	0,09
2.	a diákok nyelvtudását a korábbi vizsgánál jobban méri.	3,83	1,18	0,10
5.	a Nemzeti alaptantervben deklarált célokat a korábbi vizsgánál jobban méri.	3,81	1,04	0,09
7.	a nyelvvoktatásban tartalmi és módszertani modernizációt indított el.	3,68	1,12	0,09
1.	a diákoknak a korábbi vizsgánál nagyobb jogosítványokat nyújt.	3,52	1,24	0,10
3.	a korábbi vizsgánál jobban szolgálja a felsőoktatás igényeit.	3,33*	1,23	0,10
14.	megbízható visszajelzést ad a tanításom minőségéről.	2,95*	1,15	0,09
<i>A tanárok nézetei a vizsga jelentőségéről</i>		3,69	0,83	0,07

* Szignifikánsan különbözik az előző állítástól $p < 0,05$ szinten.

A faktoron belüli állításoknál megvizsgáltuk, hogy a leggyakrabban tanított két idegen nyelv esetében találunk-e különbségeket. Az angolt tanító tanárok (N=79) többnyire egyetértettek (3,99) azzal, hogy a kétszintű nyelvi érettségi a diákok nyelvtudását a korábbi vizsgánál jobban méri, míg a némettanárok (N=59) ezzel az állítással csak közepes mértékben (3,54) értettek egyet (a különbség $p < 0,05$ szinten szignifikáns). A többi változónál és a teljes faktor esetében nem találtunk szignifikáns különbséget. Összességében azonban a nyelvtanárok többnyire egyetértettek a felsorolt állításokkal, vagyis a vizsgareform eredményeként egy olyan érettségi jött létre, melyet a tanárok elfogadhatónak és a korábbi vizsgához képest megbízhatóbbnak tartanak. Ez összhangban van a

vizsgafejlesztők szándékaival. Cheng (1998), Al-Jamel és Ghadi (2008) vizsgálatában hasonló kép bontakozott ki (3. melléklet): összességében ezekben a kutatásokban is közepes vagy erős hatásról nyilatkoztak a megkérdezett tanárok a módosított vagy megreformált vizsgáról (bár az egyes állításoknál jelentősebb különbségek is vannak, például a 2. kérdőív-tételnél). Ez megerősíti az elméleti modellben (11. ábra) az igényeknek és a céloknak megfelelő vizsgafejlesztési folyamat jelentőségét, amelynek figyelembevételével olyan vizsga jöhet létre, melyet a tanítási-tanulási folyamatban részt vevők megfelelőnek tartanak.

A vizsgával kapcsolatos nézetek formálódásában a tanárok tapasztalatai több forrásból származhatnak: (1) melyik vizsgaszintre készítik fel diákjaikat, (2) melyik vizsgaszinten érettségizettek eddig és (3) vettek-e részt érettségi képzésen. E változók mentén a közép- és emelt szintre való felkészítés egy kérdőív-tételnél mutatott szignifikáns különbséget. Azok a tanárok, akik emelt szintre készítik fel a diákokat többnyire egyetértettek (3,82) azzal, hogy az érettségi a diákoknak a korábbi vizsgánál nagyobb jogosítványokat nyújt, míg a csak középszintre felkészítő kollégáik közepes mértékben értettek egyet ezzel (3,29; a különbség $p < 0,01$ szinten szignifikáns). Ennek feltehető oka, hogy az emelt szintű érettségire készülők célja a nyelvvizsga-bizonyítvány és/vagy a nyelvszakra való továbbtanulás (erre következtethetünk a 4.1. fejezet eredményeiből), így a tanáraik pozitívabbnak ítélik meg a vizsga által felkínált jogosítványokat.

A tanárok érettségivel kapcsolatos nézeteinek alakulását a vizsgáztatási tapasztalatok jelentősen meghatározzák (59. táblázat).

59. táblázat. A vizsgáztatási tapasztalat szerinti különbségek a tanárok vizsgával kapcsolatos nézeteiben

Nr.	Kérdőív-tétel	Csak középszinten vizsgáztatott (N=91)			Közép- és emelt szinten vizsgáztatott (N=51)			Kül. szign.
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	
<i>A kétszintű idegen nyelvi érettségi vizsga...</i>								
1.	a diákoknak a korábbi vizsgánál nagyobb jogosítványokat nyújt.	3,32	1,24	0,13	3,86	1,20	0,17	$p < 0,05$
2.	a diákok nyelvtudását a korábbi vizsgánál jobban méri.	3,63	1,29	0,14	4,24	0,81	0,11	$p < 0,01$
3.	a korábbi vizsgánál jobban szolgálja a felsőoktatás igényeit.	3,12	1,26	0,13	3,63	1,13	0,16	$p < 0,05$
4.	illeszkedik a Közös Európai Referenciakeret szintrendszeréhez.	4,02	0,99	0,10	4,44	0,81	0,11	$p < 0,05$
5.	a Nemzeti alaptantervben deklarált célokat a korábbi vizsgánál jobban méri.	3,58	1,03	0,11	4,27	0,85	0,12	$p < 0,001$
6.	elősegíti a nyelvi készségek fejlesztését a nyelvoktatásban.	3,62	1,19	0,12	4,39	0,83	0,12	$p < 0,001$
7.	a nyelvoktatásban tartalmi és módszertani modernizációt indított el.	3,51	1,14	0,12	3,98	1,03	0,14	$p < 0,05$
8.	a korábbi vizsgánál jobban segíti a javítói és értékelői munkámat.	3,73	1,10	0,11	4,24	0,86	0,12	$p < 0,05$
14.	megbízható visszajelzést ad a tanításom minőségéről.	2,90	1,16	0,12	3,14	1,13	0,16	n.s.
<i>A tanárok nézetei a vizsga jelentőségéről</i>		3,50	0,85	0,09	4,04	0,67	0,09	$p < 0,001$

A 14. állítást kivéve, minden esetben pozitívabban ítélték meg a vizsgát azok a tanárok, akik a középszint mellett emelt szinten is vizsgáztatnak. A vizsga filozófiájához tartozó két állításnál (5., 6.) és a teljes faktorban a különbség mértéke $p < 0,001$ szinten szignifikáns. Mindez arra utal, hogy azok a tanárok, akik a vizsgáról több tapasztalattal rendelkeznek, összességében elfogadhatóbbnak, és mérési-értékelési eljárásait megbízhatóbbnak tartják.

Az érettségi képzésen való részvétel a 2. és 4. kérdőív-tételnél (58. táblázat) okozott szignifikáns ($p < 0,05$) véleménykülönbséget. Ennek feltehető oka, hogy a vizsgáztatói képzéseken kiemelt elem, hogy a tanárok megismerjék a vizsga mérési-értékelési eljárásait és ez tükröződhet a válaszokban. Egyúttal ez arra is utal, hogy a vizsgáztatói képzések segíthetik a vizsga filozófiájának megértését, megismertetését és elfogadtatását.

A tanárok a második faktorba sorolható állításokkal közepes mértékben értettek egyet, de a páros t-próba eredményei alapján a változók átlagukban szignifikánsan különböznek egymástól (60. táblázat). E kérdőív-tételeket a nemzetközi vizsgálatokból vettük át (Chen, 2002; Al-Jamal és Ghadi, 2008) és alakítottuk saját kontextusunkra (3. melléklet). A 12. állítás átlagos támogatottsága összehasonlítható Al-Jamal és Ghadi (2008) vizsgálatának eredményével, melyhez hasonlóan e kijelentéssel a tanárok közepes mértékben értettek egyet. Chen (2002) szerint azok a tanárok, akik jelentős külső elvárásokat érzékeltek, nagyobb arányban érezték magukat kínosan és felelősnek, ha a diákjaik nem sikeresek a vizsgán és a tanításuk minőségét ez alapján ítélték meg. A kutató e változók, valamint a tanulói elvárások mértéke, az iskolatípus és a nyelvvizsgára felkészülő osztály típusa között szoros összefüggéseket talált. A teljes faktor esetében azonban nincsenek szignifikáns összefüggések, de a 15. állítás alacsony, $p < 0,05$ szintű szignifikáns korrelációt mutat a nyelvtanulók elvárásainak a tanítási-tanulási folyamatra gyakorolt hatásával (0,21), a vizsgaszintre való felkészítéssel (0,20) és a vizsgára felkészítő csoport típusával (0,23).

60. táblázat. A vizsga következményeivel kapcsolatos attitűdök

Nr.	Kérdőív-tétel	Átlag	Szórás	Std. hiba
15.	Felelősnek érzem magam, ha a tanulók nem sikeresek az érettségi vizsgán.	3,40	1,09	0,09
13.	Az érettségi vizsgaeredmények befolyásolják, hogy a tanulók és a kollégák engem jó tanárnak tartanak-e.	3,04*	1,28	0,10
12.	Kínosan érzem magam, ha az érettségim a diákjaim rosszabbul teljesítenek, mint a kollégám diákjai.	2,59*	1,26	0,10
<i>A vizsga következményeivel kapcsolatos attitűdök</i>		3,00	0,94	0,08

* Szignifikánsan különbözik az előző állítástól $p < 0,05$ szinten.

A tanórára való felkészüléssel kapcsolatos attitűdökhöz tartozó 18. állítással (Szükségesnek érzem a tankönyvi feladatok kiegészítését az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatokkal.) többnyire egyetértettek a válaszadók (átlag: 4,44; szórás: 0,87, std. hiba: 0,07). A magas érték felhívja a figyelmet annak fontosságára, hogy a felhasznált tananyagok kritikus szemlélete alapfeltétel a vizsgára való sikeres felkészülésben, hiszen a mérőfeladatoknak alapvetően más a funkciójuk, mint a gyakorlófeladatoknak. A továbbiakban érdemes megvizsgálni, hogy a tanárok hogyan látják a vizsgafeladatok tanítási-tanulási folyamatban való felhasználásának céljait és milyen felkészítő feladatgyűjteményeket használnak. Ezt az 5.2. fejezetben tárgyaljuk. A mérőfeladatok tanítási-tanulási folyamatban való megjelenését fontosnak érzik a tanárok. A 11. állítással (Mióta bevezették a kétszintű érettségit, több időre van szükségem az óráimra való felkészüléshez) a tanárok közepes mértékben értettek egyet (átlag: 2,73, std. hiba: 0,10), de a magas szórás (1,27) és a relatív szórás (47%) eltérő válaszokra utal. E

faktorba tartó változók szignifikánsan különböznek aszerint, hogy a tanárok váltottak-e tankönyvet az új érettségi hatására (61. táblázat). Ez arra utal, hogy az érettségi hatására végbement tankönyvváltás befolyásolhatja a tanórai tervezési folyamatot: a tankönyvi feladatok kiegészítését az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatokkal és az órákra való felkészülést.

61. táblázat. A tankönyvváltás szerinti különbségek a tanórára való felkészüléssel kapcsolatos attitűdökben

Nr.	Kérdőívtétel	Váltott tankönyvet az érettségi hatására (N=109)			Nem váltott tankönyvet (N=40)			Kül. szign.
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	
11.	Mióta bevezették a kétszintű érettségit, több időre van szükségem az óráimra való felkészüléshez.	2,91	1,30	0,12	2,25	1,06	0,17	p<0,01
18.	Szükségesnek érzem a tankönyvi feladatok kiegészítését az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatokkal.	4,53	0,82	0,08	4,20	0,97	0,15	p<0,05
	<i>A tanórára való felkészüléssel kapcsolatos attitűdök</i>	3,72	0,90	0,09	3,23	0,69	0,11	p<0,01

Mindezek mellett a 18. kérdőívtételnél a tanított évek száma szerint is szignifikáns különbségek vannak ($F=4,07$; $p<0,01$). A Tukey-b-próba alapján a 20–29 éve tanító pedagógusok többnyire egyetértettek azzal, hogy szükséges a tankönyvi feladatok kiegészítése (4,00), míg a többi csoportba tartozó, 10–19 (4,49), 30–39 (4,61) és 1–9 (4,63) éve tanító tanárok teljes mértékben értettek egyet ezzel a kijelentéssel. Feltételezhető, hogy a 20–29 éve pályán levők elsősorban a nyelvi készségek fejlesztésére helyezhetik a hangsúlyt és kevésbé arra, hogy mérőfeladatokkal gyakoroltassanak, míg a pálya elején járók és az idősebb kollégáik rövid távú célként valószínűleg inkább mérőfeladatokat használnak. Ugyanakkor ennek elemzésére más adatgyűjtő eszközök is szükségesek.

A három faktor közötti összefüggést a 22. melléklet tartalmazza, mely szerint a két, attitűdöket tartalmazó faktor között szignifikáns korreláció van (0,38), míg a nézetek és az attitűdök között nincs szignifikáns kapcsolat. Ez arra utal, hogy a tanárok gondolkodását meghatározó kognitív és affektív változók egymástól elkülönülnek, míg az attitűdállításokat tartalmazó faktorok között szoros kapcsolat van. A tanórai tervezés és a vizsga következményei egymással összefüggnek és kölcsönösen hatnak egymásra.

A harmadik faktor értelmezéséhez a vizsgára felkészítő tananyagok megítélésére is kértük a tanárokat. Válaszaikból az derül ki, hogy alkalmasnak találják a nyelvoktatási folyamatban használt tankönyveket a készségfejlesztés szempontjából (átlag: 4,09; szórás: 0,93, std. hiba: 0,08) és megfelelőnek ítélik az érettségire felkészítő feladatgyűjteményeket a vizsgaszituáció szimulálása tekintetében (átlag: 4,19; szórás: 0,84, std. hiba: 0,07). Az eredmények értelmezéséhez arra kértük a nyelvtanárokat, jelöljék, hogy (1) váltottak-e tankönyvet az új érettségi hatására, (2) használnak-e a nyelvi érettségire felkészítő feladatgyűjteményt; (3) melyik tankönyvet és (4) feladatgyűjteményt használják leggyakrabban az érettségire való felkészülés során.

A megkérdezettek 73,2%-a váltott tankönyvet az érettségi hatására és 90,6%-uk használ a nyelvi érettségire felkészítő feladatgyűjteményt. A leggyakrabban használt tankönyvek és feladatgyűjtemények szerzőjét és címét ábécésorrendben, nyelvenként és az említés gyakorisága alapján a 23. melléklet tartalmazza. A nyelvoktatási folyamatban

használt tankönyvek többnyire meghatározzák a nyelvoktatási folyamatot (erről lásd a későbbiekben elemzendő 62. táblázatot). A tankönyvpiacon jelentős számban jelentek meg az érettségi feladatokra hasonlító mérőfeladatokat tartalmazó feladatgyűjtemények és ezeket a tanárok nagyon magas arányban használják.

A tankönyvekkel és a feladatgyűjteményekkel kapcsolatban gyűjtött információk segíthetik az elméleti modellben szereplő 5. visszacsatolási kör (11. ábra) vizsgálatát, mellyel lehetőség nyílik annak ellenőrzésére, hogy a vizsga hatására létrejövő tankönyvek és feladatgyűjtemények mennyiben vannak összhangban a tesztkészítők szándékával. Erre a célra alkalmazható például a nemzetközi szakirodalomból ismert és empirikusan validált tananyagokat vizsgáló eszköz (*Saville és Hawkey, 2004*). Ezzel például megvizsgálható, hogy a tankönyvek mennyiben alkalmasak a használható nyelvtudás fejlesztésére, az érettségire való felkészülésre: tartalmaznak-e elegendő autentikus szöveget, változatos feladattípusokat és lehetővé teszik-e a változatos munkaformákat. Ennek elemzése meghaladná az értekezés kereteit, így erre nem vállalkoztunk, de a 23. mellékletben megadott lista kiindulási alap lehet ehhez.

A vizsga megítélésével kapcsolatban két állítás értékelésére kértük a nyelvtanárokat, amely fontos visszajelzés lehet a vizsgafejlesztés számára. A vizsga továbbfejlesztése szempontjából jelent meg az az elképzelés, hogy a nyelvhelyesség külön mérésére nincs szükség, több kommunikatív szemléletű nyelvvizsgán sem alkalmazzák ezt a vizsgarészt. Ugyanakkor a tanárok többnyire egyet értettek azzal, hogy szükséges a külön nyelvhelyesség feladatsor (átlag: 4,46; szórás: 0,88, std. hiba: 0,07), és a vizsga mérési gyakorlata szempontjából fontosnak tartják. A 16. kérdőívvel azt vizsgáltuk, hogy az emelt szintű érettségien jelenjen-e meg a közvetítési készség mérése vagy sem. A nyelvtudás-mérésben e készség egyre inkább előtérbe kerül, a négy alapkészség mellett az ötödiknek tekintik. A KER is külön foglalkozik a közvetítési stratégiákkal, *Dévény (2008)* empirikus vizsgálatai pedig azt mutatják, hogy ez a készség jól elkülönül a többitől. Bár a kétnyelvű vizsgák ellenzői a fordítási/közvetítési feladat negatív hatásával és a feladat értékelésének problémáival érvelnek. Az emelt szintű vizsgán a közvetítési készség mérése azért lenne releváns, mert a nyelvszakos felsőoktatásban fontos szerepe van ennek a készségnek (lásd például *Szénich (2007)* vizsgálatát). A tanárok a közvetítési készség mérésével közepes mértékben (átlag: 3,21, std. hiba: 0,11) értettek egyet, de a szórás (1,31) és relatív szórás (41%) értékek magasak, így eltérő válaszok figyelhetők meg. Eredményeink szerint azok a nyelvtanárok, akik csak középszinten vizsgáztatnak (N=87) fontosabbnak tartják a közvetítési készség mérését (átlag: 3,44), mint azok a kollégáik (N=49), akik az emelt szintű vizsga követelményeit, a feladatokat és a tanulók teljesítményeit saját tapasztalataikból is ismerik, mert mindkét szinten vizsgáztatnak (átlag: 2,88). A véleménykülönbség $p < 0,01$ szinten szignifikáns. Adataink ebből a szempontból nem tekinthetők reprezentatívnak, de arra utalnak, hogy e készség mérésének bevezetése alapos körültekintést, mérlegelést igényel.

5.2. Az idegen nyelvi érettségi hatása a nyelvoktatási folyamatra a tanárok nézetei alapján

A kétszintű nyelvi érettségi nyelvoktatási folyamatra gyakorolt hatását a kérdőív két részével vizsgáltuk. Az elsőben arra kértük a nyelvtanárokat, hogy értékeljék a nyelvoktatási gyakorlatot befolyásoló tényezőket, a második részben a nyelvoktatási folyamatra vonatkozóan olyan területeket határoztunk meg, amelyre az érettségi vizsgának hatása lehet (lásd 26. táblázat). A továbbiakban a két kérdőív rész alapján gyűjtött információkat elemezzük.

5.2.1. A tanárok nyelvoktatási folyamatát befolyásoló tényezők

A kérdőív részenként végzett faktoranalízis alapján (30. táblázat) két faktor jött létre: (1) a nyelvtanítási folyamatot befolyásoló külső, tárgyi tényezők (23–25. állítás); (2) a nyelvtanítási folyamatot befolyásoló tanári tényezők (20–22, 26. állítás). Ugyanakkor ez a kérdőív rész jelentősen elkülönül a kérdőív többi dimenzióitól és feltehetően ez az oka annak, hogy a teljes kérdőívre végzett faktoranalízisből kimaradt (34. táblázat), így ebből a részből nem hoztunk létre összevont változót.

A tanárok adott faktoron belüli válaszait átlagolva rangsorolhatjuk a két látens változót aszerint, hogy mennyiben határozzák meg a nyelvoktatási folyamatot (62. táblázat). Mindkét tényező hatásával többnyire egyetértettek a nyelvtanárok, de ezek között statisztikailag szignifikáns különbség mutatható ki (a páros t-próba eredménye: $t=2,18$; $p<0,05$). A külső, főleg a nyelvi érettségre vonatkozó tényezőknek a tanárok válaszai alapján valamivel nagyobb a hatása a nyelvoktatási folyamatban, mint a tanári tényezőknek.

A külső, tárgyi tényezőkhez tartozó változók elemzésével arra is választ kaphatunk, hogy az érettségi vizsgakövetelményeknek vagy a korábbi vizsgafeladatoknak van-e nagyobb hatása. Ez a kérdés azért releváns, mert a korábbi érettségre való felkészülés során elterjedt volt az a gyakorlat, amelyben a tanárok a vizsgával kapcsolatban főleg a mintafeladatokból és a korábbi vizsgafeladatokból tájékozódtak és nem a vizsgakövetelményekből. A páros t-próba alapján a vizsgakövetelmények szignifikánsan nagyobb ($p<0,01$) hatást gyakorolnak a nyelvtanítási folyamatra, mint a vizsgafeladatok. Ennek feltehető oka, hogy a vizsgakövetelmények a korábbi vizsgánál jóval részletesebben és pontosabban adják meg a készségekre lebontott követelményeket, a szövegek jellemzőit, a témaköröket, a kommunikációs helyzeteket és szándékokat, továbbá a nyelvtani szerkezeteket, valamint a szókincsre vonatkozó információkat (ezek elemzéséről lásd a 2.3.1. fejezetet), amelyek elsősorban a tervezési munkát segíthetik. Ugyanakkor ennek vizsgálatára a kérdőívünk nem alkalmas. Érdeemes a továbbiakban megvizsgálni, hogy a tanárok milyen céllal és hogyan használják a vizsgakövetelményeket.

62. táblázat. A nyelvtanítási gyakorlatot befolyásoló tényezők jellemzői

Nr.	Kérdőív-tétel	Átlag	Szórás	Std. hiba
	<i>A nyelvoktatási folyamatot befolyásolják...</i>			
22.	a korábbi nyelvoktatási tapasztalataim.	4,62	0,58	0,05
24.	a kétszintű nyelvi érettségi vizsgakövetelményei.	4,28*	0,75	0,06
23.	az aktuálisan alkalmazott nyelvkönyv.	3,96*	0,85	0,07
25.	a kétszintű nyelvi érettségi korábbi vizsgafeladatai.	3,93	1,01	0,08
26.	a nyelvtanulók elvárásai.	3,92	0,89	0,07
21.	a módszertani továbbképzésen tanultak.	3,80	1,10	0,09
20.	a felsőoktatásban, a nyelvtanárképzésben tanultak.	3,35*	1,12	0,09
	<i>Nyelvoktatási folyamatot befolyásoló külső, tárgyi tényezők</i>	4,06	0,67	0,05
	<i>Nyelvoktatási folyamatot befolyásoló tanári tényezők</i>	3,92*	0,59	0,06

* Szignifikánsan különbözik az előző állítástól $p<0,05$ szinten.

Cheng (1998) és Caine (2005) vizsgálatával összhangban (3. melléklet) a tanárok tapasztalatai határozzák meg leginkább a nyelvoktatási folyamatot és csak közepes mértékben a felsőoktatásban és a nyelvtanárképzésben tanultak. Ugyanakkor a két kutatáshoz képest a nyelvkönyveknek nagyobb szerepe van a felmért tanárok körében. A nyelvtanulók elvárásai a gimnáziumban tanító pedagógusok (átlag: 4,10) oktatási folyamatát szignifikánsan ($p<0,01$) nagyobb mértékben befolyásolja, mint a szakközépis-

kolában tanítóké (átlag: 3,64). Ehhez hasonlóan azok a nyelvtanárok, akik emelt szintre készítik fel a diákokat (átlag: 4,10) szignifikánsan nagyobb mértékű ($p < 0,01$) nyelvtanulói elvárásból adódó hatásról nyilatkoztak, mint azok a kollégáik, akik csak középszintre készítik fel a diákokat (átlag: 3,71). Ez feltehetően azzal magyarázható, hogy a gimnáziumból többen tesznek emelt szintű érettségit (lásd 39. táblázat) és a vizsgán való sikeres tanulói szereplés célja és az ezzel kapcsolatos elvárások meghatározzák a tanárok munkáját.

5.2.2. A tanárok nyelvoktatási folyamatának jellemzői az érettségi alapján

A kérdőív negyedik részében az érettségi jellemzői alapján fogalmaztunk meg állításokat a nyelvoktatási gyakorlatra vonatkozóan. Minden állításhoz két azonos szerkezetű, de tartalmában különböző skála tartozik (31. táblázat): az egyik skálán azt jelölték a tanárok, hogy nézeteik szerint mennyire jellemző az adott nyelvoktatási gyakorlat, míg a másik skálán azt, hogy mennyire jellemző abban az érettségi hatása. A két skálában lévő faktorok jellemzőit a 34. táblázat, míg e kérdőív rész reliabilitását a 35. táblázat tartalmazza. Az egyes állításoknál mindkét skálán kiszámoltuk az átlagos támogatottság mértékét és az állításokhoz tartozó két skálán mért értékek között korrelációt számoltunk (63. táblázat). Ezzel az elemzéssel azt vizsgáljuk, hogy a tanárok nyelvoktatási gyakorlata és az érettségi észlelt hatása között milyen összefüggéseket találunk.

63. táblázat. A nyelvoktatási gyakorlat jellemzői és az érettségi hatása

Nr.	Kérdőív-tétel	Mennyire jellemző?			Érettségi hatása			Korreláció
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	
27.	A hallott szöveg értésének fejlesztése nagy szerephez jut az óráimon.	4,09	0,77	0,06	2,99	1,10	0,09	n.s.
28.	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.	3,32	0,95	0,08	2,39	1,21	0,10	n.s.
29.	A nyelvrákon a célnyelvet használom.	4,18	0,73	0,06	2,47	1,42	0,12	n.s.
30.	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.	4,38	0,63	0,05	2,87	1,24	0,10	n.s.
31.	A nyelvrákon autentikus szövegekkel dolgozom.	3,97	0,86	0,07	2,82	1,23	0,10	n.s.
32.	A diákoknak fordítási feladatokat adok.	2,57	1,06	0,09	2,17	1,29	0,11	n.s.
33.	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.	3,78	0,95	0,08	2,62	1,15	0,09	0,28*
34.	A nyelvrákat a nyelvi szituációk gyakorlására építem.	3,42	0,75	0,06	2,76	1,02	0,08	0,28**
35.	Változatos munkaformákat alkalmazok.	4,13	0,73	0,06	2,52	1,17	0,10	n.s.
36.	Változatos, szövegközpontú feladattípusokat használlok.	3,87	0,76	0,06	2,89	1,18	0,10	0,28**
37.	A produktív készségek mérésénél értékelési szempontokat és skálákat használlok.	3,60	1,12	0,09	3,56	1,20	0,10	0,49**

Jelölések: * $p < 0,01$; ** $p < 0,001$; n.s: nem szignifikáns.

A továbbiakban először az első skála adatait elemezzük, mellyel a tanárok nyelvoktatási gyakorlatát jellemezzük. Ezt követően a két skála közötti kapcsolatokat és a második skálán mért eredményekben megjelenő különbségeket vizsgáljuk különböző háttérváltozók szerint. E kérdőív rész harmadik elemzési szempontjaként a létrejött faktorok (34. táblázat) közötti összefüggéseket (22. melléklet) vizsgáljuk.

A nyelvoktatási gyakorlat jellemzői

Az első skálán leginkább azzal értettek egyet a nyelvtanárok, hogy a tanulókat folyamatosan, a nyelvi készségek használatára készítik fel, amely a készségfejlesztő kommunikatív szemléletű nyelvoktatás egyik legfontosabb alapelve. Ugyanakkor ennek némiképp ellentmond, hogy a tanárok a 18. állításra adott válaszukban teljes mértékben egyetértettek (átlag: 4,44) azzal, hogy szükséges a tankönyv feladatainak kiegészítése az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatokkal, amely inkább megnehezíti a készségfejlesztést, mint elősegíti azt. Ennek oka a fejlesztő- és a mérőfeladat közti különbségekben rejlik, mert a mérőfeladatok a készség szintjét rögzítik, így kevésbé alkalmasak a fejlesztésre. Ez az ellentmondás felhívja a figyelmet az alkalmazás célját és módját feltáró vizsgálatok szükségességére.

A nyelvtanárok többnyire jellemzőnek tartják saját nyelvoktatási gyakorlatukra a célnyelv használatát, a változatos munkaformák alkalmazását, az autentikus szövegek használatát, a hallott szöveg értésének fejlesztését. Ugyanakkor az eredmények azt is mutatják, hogy a nyelvórákon a nyelvi szituációk gyakoroltatása és különösen az adott csoporton belüli differenciált nyelvoktatás csak közepes mértékben jellemző. Ezekről az állításokról szignifikánsan elkülönülve kevésbé jellemző a fordítási feladatok használata, de a magas szórás és relatív szórásértékek (41%) eltérő válaszokat jelölnek. Az első skálán szerzett adatok értelmezésében mindenféleképp utalni kell arra, hogy ezekből kevésbé lehet megbízható módon következtetni a nyelvoktatási gyakorlatra (ahogy erre utaltunk a kérdőív reliabilitásának elemzésekor: 35. táblázat), mert a nyelvtanárok ismerik a kommunikatív nyelvoktatás alapelveit, amelyekhez a válaszaikban igazodnak.

A tanárok nyelvoktatási folyamatokról kialakított nézetei kapcsolatban állhatnak a környezetre vonatkozó tényezők közül az iskolatípussal, az érettségire felkészítő csoport jellegével, a vizsgaszintre való felkészítéssel; a tanárookra vonatkozó tényezők közül a képzéseken való részvétellel és a tanárok végzettségével. A továbbiakban e tényezők szerinti különbségeket elemezzük a nyelvoktatási gyakorlatra jellemző faktorok szerint. Az első skálán mért változókból a teljes kérdőívre vonatkozó faktoranalízis során három faktor jött létre: (1) az eredményes tanítási módszerekről, (2) a tanítási gyakorlatról és (3) az oktatás tartalmáról kialakított tanári nézetek. E három faktorba tartozó változókat (34. táblázat) az érettségi jellemzői alapján határoztuk meg (27. táblázat).

A 64. táblázat azokat a változókat tartalmazza, amelyek szerint a kétmintás t-próba alapján különbséget találtunk a három faktor valamelyikében. Az egyes változók szerinti csoportok elemszámát is megadtuk, a különbségeket a szignifikanciaszint jelölésével jellemeztük. Eredményeink szerint a környezetre vonatkozó változók közül az iskolatípus szerint találunk két faktor esetében szignifikáns különbséget, továbbá a tanárok aktuális tanítási gyakorlatát befolyásolja, hogy milyen vizsgaszintre készítik fel a diákokat és vettek-e részt érettségi vagy nyelvvizsgáztatói képzéseken. Mindezek arra hívják fel a figyelmet, hogy a különböző vizsgával kapcsolatos továbbképzések meghatározhatják a tanárok nyelvoktatásról kialakított nézeteit és feltételezhetően a gyakorlatukat is, de ezt közvetlenül nem elemeztük. E három faktor esetében az angol és a német nyelvénél nincs szignifikáns különbség.

64. táblázat. A környezetre és a tanárookra vonatkozó tényezők szerinti különbségek a nyelvoktatási gyakorlatot jellemző faktorok megítélésében

Változó	Csoport	Vizsgált faktor		
		Nyelvoktatás módszerei	Tanítási gyakorlat	Nyelvoktatás tartalma
Iskolatípus	1. Gimnázium (N=91)	1>2*	n.s.	1>2*
	2. Szakközépiskola (N=58)			
Érettségi szintre való felkészítés	1. Középszintre (N=75)	1<2*	1<2**	n.s.
	2. Közép- és emelt szintre (N=69)			
Részvétel érettségi képzésen	1. Részt vett (N=99)	n.s.	1>2*	n.s.
	2. Nem vett részt (N=49)			
Részvétel nyelvvizsgázatói képzésen	1. Részt vett (N=60)	n.s.	1>2**	1>2*
	2. Nem vett részt (N=88)			

Jelölések: 1.: első csoport; 2.: második csoport; >/< szignifikánsan nagyobb/kisebb érték * $p<0,05$; ** $p<0,01$; n.s.: nem szignifikáns.

Az érettségi hatása a nyelvoktatási gyakorlatra a tanárok nézetei alapján

A vizsga tanítási-tanulási folyamatra gyakorolt hatásának megítélésében a kérdőívünk megbízhatóan mér (35. táblázat), a faktoranalízis alapján ezek az állítások jól elkülönülnek az előző skálán mért adatoktól (31. táblázat). A 63. táblázat alapján az oktatás tartalmát jellemző 33. és 36. állítás, valamint a tanítási gyakorlathoz tartozó 34. kérdőívétel esetében találunk szignifikáns összefüggést (a 37. állításhoz hasonlóan, melyet később elemzünk). Mindez arra utal, hogy elsősorban az oktatás tartalma és az érettségi hatása között van kapcsolat, míg a módszereket jellemző állítások esetében nincs szignifikáns összefüggés. Az érettségi hatását jellemző skálán mért átlagos támogatottság mértékét megvizsgálva megállapítható, hogy a vizsgának közepes hatása van az oktatás tartalmára (például 27., 31., 33., 36. állítás), továbbá olyan, a nyelvoktatási gyakorlatra jellemző területekre (például 30., 34.), amelyek a készségorientált nyelvoktatást jellemzik. A nemzetközi teszthatást vizsgáló kutatásokkal összhangban (például Alderson és Wall, 1993; Cheng, 1998; Shohamy, Donitsa-Schmit és Ferman, 1996) a tanárok szerint a vizsga hatása kevésbé jellemző a tanítási módszerekre (28., 29., 35.), és meghatározóbb az oktatás tartalmára. Az érettségi tanítási módszerekre gyakorolt hatását jellemző állításoknál a szórásértékek nagyon magasak (a relatív szórás értékei 46%-tól 57%-ig terjednek), azaz a minta heterogén a vizsga hatásának megítélésében, így feltételezhető, hogy különböző háttérváltozók mentén azonosíthatók olyan csoportok, ahol a vizsga hatása erősebb.

Az iskolatípusok szerinti különbségek a 27b és 29b állításoknál mutathatók ki. A hallási készség fejlesztése (27b) általában eléggé elhanyagolt területe a nyelvoktatásnak. E készség fejlettségét a korábbi vizsgán nem mérték (1. melléklet). A próbaérettségiken és a 2005. és 2006. évi középszintű vizsgán ezen a területen rosszul teljesítettek a diákok (a nyelvhelyesség mellett) különösen angolból (14. táblázat), amely jelezte a fejlesztés szükségességét. A 2007. évi érettségi vizsgát követően a diákok jobban teljesítettek, a legkisebb különbség a két iskolatípusban érettségizők között e készségnél van (43. táblázat). Ennek okára a tanári kérdőíves vizsgálatunkból következtethetünk. Eredményeink szerint a szakközépiskolában tanító pedagógusok (átlag: 3,24, szórás: 1,08) szignifikánsan ($p<0,05$) nagyobb hatásról nyilatkoztak, mint a gimnáziumban tanítók (átlag: 2,84, szórás: 1,09). Hasonló tendencia figyelhető meg az érettségi hatásának mértékénél a célnyelv használatában is (29b állítás). A szakközépiskolában tanító nyelvtanárok (átlag: 2,77, szórás: 1,56) körében az érettségi hatása nagyobb ezen a terü-

leten, míg a gimnáziumban tanító nyelvtanároknál az érettségi hatása kevésbé jellemző (átlag: 2,29, szórás: 1,29). A vizsgált minták varianciája szignifikánsan különbözik egymástól, az $F=8,77$; $p<0,01$, a Welch-próba alapján $t=1,97$; $p<0,05$).

A 29b állításnál a csak középszinten vizsgáztató nyelvtanárok is szignifikánsan nagyobb hatást jelöltek válaszaikban (átlag: 2,70, szórás: 1,47), mint az emelt szinten is vizsgáztató kollégáik (átlag: 2,16, szórás: 1,21) ($F=10,48$; $p<0,01$; Welch-próba alapján $t=2,33$; $p<0,05$). Ehhez hasonlóan a csak középszintre felkészítő tanárok nagyobb hatásról nyilatkoztak (átlag: 2,73, szórás: 1,45), mint azok, akik a középszint mellett emelt szintre is felkészítenek (átlag: 2,22, szórás: 1,36). Ezen kívül pedig a tanárok válaszai a nyelvi csoport jellege (65. táblázat), a tanított évek száma (66. táblázat) és a vizsgáztatói tapasztalat (67. táblázat) szerint is különböznek ennél az állításnál. A továbbiakban e különbségeket elemezzük.

A készségfejlesztést a célnyelv és a szövegközpontú feladattípusok használatával megvalósító nyelvoktatási gyakorlatra az érettséginek közepes hatása van a kizárólag normál óraszámú csoportban tanító pedagógusok válasza alapján (65. táblázat). Az emelt óraszámú és kéttannyelvű csoportban tanító pedagógusok kisebb hatásról számoltak be. Összességében az eredmények arra utalnak, hogy a magasabb óraszámú és szintű nyelvi képzésekben az érettségi hatása nem jellemző, vagyis már korábban megtörténhetett a kommunikatív szemlélet érvényesítése. Ugyanakkor a kommunikatív szemléletű vizsgának valamivel nagyobb a hatása az alacsonyabb óraszámú csoportban felkészítő tanárok nézeteire.

65. táblázat. Az érettségire felkészítő csoport szerinti különbségek a vizsga nyelvoktatási folyamatra gyakorolt hatásában a tanárok nézetei alapján

Nr.	Kérdőívtétel	Kategória	N	Szignifikáns különbséget mutató csoportok		F érték
				1	2	
29b	A nyelvórákon a célnyelvet használom.	kéttannyelvű csoport	15	1,67		6,86
		csak emelt óraszámú csoport	21	1,81		
		normál és emelt óraszámú csoport	42	2,29	2,29	
		csak normál óraszámú csoport	70		2,96	
30b	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.	csak emelt óraszámú csoport	21	2,29		4,05
		kéttannyelvű csoport	15	2,40	2,40	
		normál és emelt óraszámú csoport	42	2,79	2,79	
		csak normál óraszámú csoport	71		3,18	
36b	Változatos, szövegközpontú feladattípusokat használlok.	csak emelt óraszámú csoport	21	2,19		3,11
		kéttannyelvű csoport	15	2,80	2,80	
		normál és emelt óraszámú csoport	42		3,02	
		csak normál óraszámú csoport	71		3,03	
	<i>A vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézetek</i>	csak emelt óraszámú csoport	21	2,21		4,61
		kéttannyelvű csoport	15	2,41	2,41	
		normál és emelt óraszámú csoport	41	2,67	2,67	
		csak normál óraszámú csoport	70		2,94	

Megjegyzés: Minden F-érték szignifikáns $p<0,05$ szinten.

A tanított évek száma szerinti különbségeket több területen is azonosítottuk (66. táblázat). Eredményeink szerint a vizsga hatása kevésbé jellemző a 20–29 éve pályán lévő pedagógusok körében és meghatározóbb az idősebb és hosszabb ideje pályán lévő

pedagógusoknál, valamint a pályájuk elején járó tanároknál. Ugyanakkor adataink csak tájékoztató jellegűek. A minta reprezentatív a tanított idegen nyelv és az iskolatípus szerint (21. táblázat), de elképzelhető, hogy e változó mentén képzett csoportok nem reprezentálják megfelelően a populációt. Ezek az adatok további vizsgálatok szükségességére utalnak más adatgyűjtő eszközök és módszerek kombinálásával.

66. táblázat. A tanított évek száma szerinti különbségek az érettségi nyelvoktatási folyamatra gyakorolt hatásában a tanárok nézetei alapján

Nr.	Kérdőív-tétel	Pedagóguspályán eltöltött évek száma	N	Szignifikáns különbséget mutató csoportok		F érték
				1	2	
28b	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat	10–19	42	2,17		2,78
		20–29	33	2,21		
		1–9	40	2,35	2,35	
		30–39	32		2,91	
29b	A nyelvórákon a célnyelvet használom.	20–29	33	2,12		3,40
		10–19	42	2,29		
		1–9	40	2,43	2,43	
		30–39	33		3,12	
30b	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.	20–29	33	2,36		4,56
		10–19	43	2,70	2,70	
		1–9	40	3,03	3,03	
		30–39	33		3,39	
31b	A nyelvórákon autentikus szövegekkel dolgozom.	20–29	33	2,42		3,34
		10–19	43	2,63	2,63	
		1–9	40	2,98	2,98	
		30–39	33		3,27	
33b	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.	20–29	33	2,24		3,01
		10–19	43	2,53	2,53	
		1–9	40	2,68	2,68	
		30–39	33		3,06	
36b	Változatos, szövegközpontú feladattípusokat használlok.	20–29	33	2,45		2,84
		10–19	43	2,79	2,79	
		30–39	33	3,06	3,06	
		1–9	40		3,20	
	<i>A vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézetek</i>	20–29	33	2,41		4,05
		10–19	42	2,56		
		1–9	40	2,78	2,78	
		30–39	32		3,10	

Megjegyzés: Minden F-érték szignifikáns $p < 0,05$ szinten.

A vizsgáztatói tapasztalat szerinti különbségek (67. táblázat) abból adódhatnak, hogy a vizsga a készségek mérésével azt jelzi a tanároknak, hogy szükséges a készségfejlesztés a célnyelv használatával és ez megjelenhet a tanári gyakorlatban is. Annak oka, hogy a vizsga közepes hatást fejt ki azokra a tanároknak, akik csak középszinten vizsgáztatnak, abban keresendő, hogy a középszintű vizsgára többnyire az alacsonyabb nyelvtudás-szintű érettségizők készülnek, míg az emelt szintre való felkészülés során e nélkül is

kiemelten fontos a készségfejlesztés a célnyelv használatával. Az értékelési szempontok és skálák használatára vonatkozó állítás elemzésére a későbbiekben térünk ki.

67. táblázat. A vizsgáztatói tapasztalat szerinti különbségek az érettségi nyelvoktatási folyamatra gyakorolt hatásában a tanárok nézetei alapján

Nr.	Kérdőívétel	Csak középszinten vizsgáztatott (N=91)			Közép- és emelt szinten is vizsgáztatott (N=51)			Kül. szign.
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	
29b	A nyelvrákon a célnyelvet használom.	2,67	1,49	0,16	2,16	1,21	0,17	p<0,05
30b	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítettem fel.	3,07	1,28	0,13	2,53	1,16	0,16	p<0,05
37b	A produktív készségek mérésénél értékelési szempontokat és skálákat használok.	3,42	1,26	0,13	3,84	1,03	0,14	p<0,05

A kérdőíves felmérésünk azt jelzi, hogy a készségfejlesztésre irányuló kommunikatív nyelvoktatás már a vizsga bevezetése előtt megjelent az iskolákban, a vizsga szemléletében csak követi ezt. Feltételezzük, hogy ez az oka annak, hogy a tanárok nézetei szerint a vizsga csak kevés, vagy közepes mértékű hatást fejt ki a nyelvoktatási folyamatra. Ugyanakkor valamennyivel nagyobb hatásról nyilatkoztak a szakközépiskolában, elsősorban normál óraszámú csoportban tanító, fiatalabb és idősebb tanárok, akik elsősorban a középszintű vizsgára készítik fel a diákjaikat.

A vizsgának legkevésbé van arra hatása, hogy a tanárok fordítási feladatokat adnak. A tanárok 42%-a nyilatkozott úgy, hogy az érettségi hatása egyáltalán nem jellemző erre a gyakorlatra. Ennek oka, hogy a fordítás semmilyen formában nem jelenik meg a vizsgán. Ugyanakkor a német nyelvet tanító tanárok gyakorlatában szignifikánsan ($p<0,05$) nagyobb mértékben jellemző a fordítási feladatok használata.

A vizsga közvetlen és erős hatását egy területen tudjuk azonosítani. Úgy tűnik, hogy a vizsga hatása többnyire jellemző arra, hogy a tanárok a produktív készségek mérésénél értékelési szempontokat és skálákat használnak (63. táblázat). Ennél az állításnál a két skála között szoros korreláció van (0,49, $p<0,001$), a faktoranalízis során közös faktorba kerültek (8. és 9. melléklet, 31. táblázat). E két skálán mért állítás megítélésében a minta homogénebb (relatív szórás: 34%). Eredményeink szerint a vizsgáztatói tapasztalat mellett (67. táblázat) az érettségi képzésen való részvételnek is jelentős hatása van erre a gyakorlatra. Azok a tanárok, akik részt vettek az érettségre vonatkozó vizsgáztatói képzésen szignifikánsan nagyobb ($p<0,01$) hatásról nyilatkoztak (átlag: 3,75, szórás: 1,10), mint azok a kollégáik, akik nem vettek részt érettségi képzésen (átlag: 3,14, szórás: 1,31). Mindez azt jelzi, hogy a vizsgához kapcsolódó képzéseknek jelentős hatása lehet az iskolai gyakorlatra is, vagyis a vizsgán alkalmazott értékelési szemlélet megjelenik a napi tanári munkában, amely több okból tekinthető pozitív fejleménynek. Korábban, a vizsgakoncepció kialakításában a tanárok részéről jelentős ellenállást lehetett megfigyelni az értékelési skálák alkalmazásával kapcsolatban (Einhorn, 2007b), hiszen ezek többletmunkát jelentenek, de hosszabb idő elteltével javíthatják az értékelés minőségét. Az értékelési szempontok és skálák használata pedig kritériumorientált értékelést is lehetővé tesz, hiszen a készséget felépítő komponensek fejlettségéről is szerezhető információ, amely elősegítheti a nyelvoktatási folyamatot. Szükséges megvizsgálni, hogy a nyelvtanárok milyen módon alkalmazzák az értékelési skálákat a napi munkájukban. Leginkább az alkalmazás célja fontos: vagyis, hogy kizárólag a vizsgára felkészítő folyamatban mutatják-e be a tanulóknak az értékelési rendszert és benne az egyes

skálákat, vagy a folyamatban is alkalmazzák-e a szóbeli és írásbeli produktumok értékelésénél, vagy felhasználják-e a tanulói önértékelésben.

A nyelvoktatási gyakorlat jellemzői és az abban megjelenő érettségi hatása közötti összefüggések

A 22. melléklet alapján a harmadik kérdőív-részhez tartozó faktorok (lásd 34. táblázat) közül az első skálán mért tanítási módszerek, a tanítási gyakorlat és az oktatás tartalma (a 22. mellékletben az 5., 6. és 7. faktor) között mutatható ki szoros összefüggés. A tanítási módszereket és a tanítási gyakorlatot jelölő 5. és 6. faktort a kommunikatív nyelvoktatás célját jelentő, készségfejlesztő gyakorlat kapcsolja össze, és ez jelenik meg a korreláció értékében (0,61).

A vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézeteket jelölő 4. faktor és az első skálával mért 5., 6. és 7. faktor közül egyikkel sincs szignifikáns korreláció, nem mutat szignifikáns összefüggést a tanárok nézeteivel (1. faktor), de a két attitűdöt jelölő faktorialacsony (0,17 és 0,18) korrelál. Az eredmények alapján megállapítható, hogy az érettségi hatását elsősorban a tanórai felkészüléssel és a vizsga következményeivel kapcsolatos attitűdök befolyásolják kis mértékben, viszont a nézetekkel nem mutatható ki szignifikáns összefüggés.

A tanárok vizsgával kapcsolatos nézeteit jelölő 1. faktor csak a 8. faktoral mutat összefüggést (22. melléklet). A vizsgával kapcsolatos nézetek hatnak a vizsga és a mérési és értékelési gyakorlat közti viszonyra ($\beta=0,23$; $p<0,01$, a megmagyarázott variancia 5,12%).

5.3. Az idegen nyelvi érettségi hatása az osztálytermi mérési-értékelési eljárásokra a tanárok nézetei alapján

A kérdőív két részében kértük arra a tanárokat, hogy a nyelvoktatási folyamatban megjelenő mérési-értékelési eljárásokat jellemezzék és adják meg, milyen céllal alkalmazzák a tanítási-tanulási folyamatban a központilag összeállított vizsgafeladatokat (25. táblázat). A továbbiakban a két kérdőív-részrel gyűjtött adatok elemzését mutatjuk be.

Az idegen nyelvi tesztelés kritériumaival végzett faktoranalízis során három faktor jött létre. Az első faktor tartalmazza a klasszikus tesztelmélet, a második a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézeteket, míg a harmadik faktor a validitás különböző aspektusait jelöli, amely hagyományosan megjelenik a tanítási-tanulási folyamatban (34. táblázat). Ha a tanárok válaszai alapján létrejött faktorokat átlagoljuk, akkor megállapítható a faktorok fontossági sorrendje (68. táblázat). A tanárok a klasszikus tesztelmélet jóságmutatóit tartják a legfontosabbnak, ettől szignifikánsan elkülönülve ($t=14,14$; $p<0,001$) a validitás aspektusait, majd pedig a kommunikatív tesztelés kritériumait. Ezek szignifikánsan különböznek ($t=21,83$; $p<0,001$) a klasszikus jóságmutatók fontosságától a validitás aspektusaihoz hasonlóan ($t=2,76$; $p<0,01$). Ez azt jelzi, hogy a kommunikatív tesztelés kritériumai megjelennek a tanítási-tanulási folyamatban alkalmazott mérőeszközök készítésének folyamatában, de a többi kritériumhoz képest némiképp háttérbe szorulnak.

Az egyes háttérváltozók szerint az érettségi képzésen való részvétel alapján találunk szignifikáns különbséget annak megítélésében, hogy a tanárok mennyiben tartják fontosnak, hogy mérőfeladataik igazodjanak a kétszintű nyelvi érettségi szemléletéhez. Azok a tanárok, akik vettek részt érettségi képzésen (átlag: 4,10, szórás: 0,92) fontosabbnak tartják a 49. állítást, mint azok a tanárok, akik nem vettek részt érettségi képzésen (átlag: 3,65, szórás: 0,99; szignifikanciaszint: $p<0,05$). E képzésen való részvétel

befolyásolja a tanárok mérés-értékeléssel kapcsolatos nézeteit. A 38. állítást a némettanárok szignifikánsan ($p < 0,05$) fontosabbnak találják az angoltanároknál.

68. táblázat. A tanárok nézetei az idegen nyelvi tesztelés kritériumairól

Nr.	Kérdőív-tétel	Átlag	Szórás	Std. hiba
	A saját magam által készített mérőfeladat...			
38.	mérjen megbízhatóan, pontosan.	4,78	0,43	0,04
40.	eredménye segítse elő a nyelvtanuló előrehaladását.	4,65*	0,57	0,05
41.	legyen objektíven értékelhető.	4,60	0,63	0,05
46.	mérje a tananyag elsajátításának mértékét.	4,39*	0,65	0,05
43.	tükrözzön reális, a valóságban is várható nyelvi szituációt.	4,38	0,68	0,06
44.	a tanuló szókincsét célzottan mérje.	4,17*	0,77	0,06
50.	épüljön olyan szövegekre, szituációkra, amelyeket a diákok szívesen oldanak meg.	4,14	0,85	0,07
49.	igazodjon a kétszintű idegen nyelvi érettségi szemléletéhez.	3,93*	0,96	0,08
42.	mérje külön az egyes nyelvi készségeket.	3,87	0,98	0,08
39.	épüljön autentikus szövegekre.	3,85	0,93	0,08
48.	épüljön a Közös Európai Referenciakeret szintrendszerére.	3,48*	0,98	0,08
47.	a tanuló formális nyelvtani ismereteit mérje.	3,33	0,85	0,07
45.	csak a nyelvtudást, az általános műveltséget, háttérismeretet ne mérje.	2,88*	0,92	0,09
	<i>A klasszikus tesztelmélet jóságmutatóinak fontosságáról kialakított nézetek.</i>	4,68	0,39	0,03
	<i>A validitás fontosságáról kialakított nézetek.</i>	3,96*	0,58	0,05
	<i>A kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek.</i>	3,79*	0,47	0,04

* Szignifikánsan különbözik az előző állítástól $p < 0,05$ szinten.

A három faktornál (68. táblázat) feltételezzük, hogy a tanárok nézetei az oktatási folyamatot befolyásoló tényezőkkel szoros kapcsolatot mutatnak. Ezek közül elsősorban a külső, tárgyi tényezőkkel (elsősorban a vizsgafeladatok és vizsgakövetelmények nyelvoktatásra gyakorolt hatása változókkal) várunk szignifikáns összefüggést. Az elemzés során az egyes, nyelvoktatási folyamatot meghatározó tényezők és a nyelvi tesztelés kritériumairól kialakított tanári nézetekhez tartozó három faktor között számoltunk korrelációt. Ennek oka, hogy az oktatási folyamatot befolyásoló tényezőkből a teljes kérdőívre elvégzett faktoranalízis során (34. táblázat) nem hoztunk létre összevont változót. E vizsgálat eredményeként csak a kommunikatív tesztelés kritériumainak fontosságával (a 10. faktorral) találtunk szignifikáns kapcsolatot. Eredményeink szerint szoros összefüggés van a vizsgakövetelmények nyelvoktatásra gyakorolt hatása és a kommunikatív tesztelés kritériumainak fontossága között (0,34), míg a módszertani továbbképzésen tanultak, az érettségi vizsgafeladatai és a nyelvtanulók elvárásai kis mértékben hatnak arra, hogy a tanárok fontosnak tartják-e ezeket a szempontokat a feladatok elkészítésénél. E változókra végzett regresszió-analízis szerint csak a módszertani továbbképzésen tanultaknak és a vizsgakövetelményeknek van szignifikáns hatása a függő változóra, de a legnagyobb megmagyarázott varianciát a vizsgakövetelmények esetében találjuk (69. táblázat). Feltételezzük, hogy a vizsgakövetelményekben lévő alapelvek megismerése hat arra, hogy a tanárok a saját mérőfeladataik elkészítésében igazodni szeretnének a kommunikatív tesztelés kritériumaihoz.

69. táblázat. A kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek mint függő változóra számított regresszióanalízis eredménye

Háttérváltozó	<i>r</i>	β	$r*\beta*100$	Szign.
<i>A nyelvoktatási folyamatot befolyásolják...</i>				
a módszertani továbbképzésen tanultak.	0,26	0,17	4,28	p<0,05
a kétszintű nyelvi érettségi vizsgakövetelményei.	0,34	0,23	7,56	p<0,05
a kétszintű nyelvi érettségi korábbi vizsgafeladatai.	0,26	0,08	2,24	n.s.
a nyelvtanulók elvárásai.	0,23	0,14	3,24	n.s.
<i>Megmagyarázott variancia (%)</i>			17,32	

A vizsgakövetelmények összhangban vannak a KER nyelvtudás-modelljével (2.3.1. fejezet), a követelményeket a Referenciakeret szintjei alapján definiálják és a feladatokhoz felhasználható szövegek jellemzői és a szövegfajták esetében utalnak az autentikuság és az interaktivitás kritériumaira, a receptív, a produktív és az interaktív stratégiák mérésének alapelveire. A dokumentum alapot szolgáltatathat a tanári gyakorlatban alkalmazott mérőfeladatok elkészítéséhez is, melyet a vizsgálatunk eredményei is alátámasztanak. Ez azért lényeges, mert az idegennyelv-tudás mérésére a tanároknak napi munkájuk során szükségük van, ugyanakkor ez eléggé elhanyagolt területe a tanárképzésnek és úgy tűnik, hogy a vizsgához készült részletes dokumentáció támogatja a kommunikatív nyelvi mérés alapelveinek megismertetését, fontosságának felismerését. Ehhez pedig lényegesen hozzájárulnak a módszertani továbbképzések, valamint az érettségi vizsgához kapcsolódó vizsgáztatói képzések.

A vizsgafeladatok tanítási-tanulási folyamatban való alkalmazásának céljait jellemző állításokkal való egyetértés mértéke alapján egy faktor jött létre, azaz a tanárok válaszmintázatai alapján nem osztható két részre a vizsgafeladatok felhasználásának lehetséges valódi funkciói és diszfunkciói (26. és 34. táblázat). A tanárok többnyire egyetértettek azzal, hogy a vizsgafeladatok fő funkciója, hogy felkészítse a diákokat a vizsgán a konkrét teendőkre, ettől szignifikánsan elkülönülve a készségek fejlettségi szintjének megállapítására szolgálnak, és előrejelző funkcióval rendelkeznek. A két diszfunkciót jelölő állítás nem különül el egy másik faktorba, de a nyelvtanárok a többihez képest szignifikánsan kevésbé értettek azzal egyet, hogy a vizsgafeladatok alkalmasak a nyelvi készségek fejlesztésére és a tanulók tanulásra való motiválására (70. táblázat).

70. táblázat. A tanárok nézetei a vizsgafeladatok oktatási folyamatban való alkalmazásának céljairól

Nr.	Kérdőív-tétel	Átlag	Szórás	Std. hiba
<i>A központilag összeállított vizsgafeladatok tanítási-tanulási folyamatban való alkalmazásának célja, hogy...</i>				
53.	a diákokat felkészítse az érettségi vizsgán a konkrét teendőkre.	4,38	0,82	0,07
54.	megállapítsa, az adott készséggel milyen minőségben rendelkezik a diák.	4,20*	0,86	0,07
51.	a diáknak a vizsgán várható teljesítményéről nyújtson információkat.	4,05	0,93	0,08
55.	a diákok nyelvi készségeit fejlessze.	3,69*	1,13	0,09
52.	a diákokat tanulásra motiválja.	3,68	1,01	0,08
<i>A vizsgafeladatok alkalmazásának céljairól kialakított nézetek</i>		4,01	0,70	0,06

* Szignifikánsan különbözik az előző állítástól p<0,05 szinten.

Eredményeink összhangban vannak *Cheng* (1998) vizsgálatával (3. melléklet). Abban a kutatásban is elsősorban a vizsgafeladatok felkészítő szerepét emelték ki a tanárok, míg a készségfejlesztő és motiváló funkciót a legkevésbé. E faktoriala vizsgafeladatok nyelvoktatási folyamatra gyakorolt hatása áll összefüggésben ($r=0,26$; $p<0,01$).

A négy mérésel és értékeléssel kapcsolatos faktor közötti korrelációs értékeket a 22. melléklet tartalmazza. Ezek alapján megállapítható, hogy a klasszikus tesztelmélet és a kommunikatív teszteléshez tartozó faktorok fontossága között mutathatók ki összefüggések, amelyek arra utalnak, hogy a kritériumok egymást támogatják.

A vizsgafeladatok felhasználásának céljaival kapcsolatos nézetek (12. faktor) összefüggéseket mutatnak a kommunikatív tesztelés kritériumainak fontosságával (10. faktor) (0,28) és a klasszikus tesztelmélet kritériumainak (9. faktor) megítélésével. A produktív készségek mérésénél és értékelésénél alkalmazott eljárások (8. faktor) összhangban vannak a kommunikatív tesztelés kritériumaival (10. faktor), melyet azért alkalmaznak a tanárok, hogy objektívan értékelve, megbízható módon segítsék a nyelvtanulók előrehaladását (9. faktor).

A 22. mellékletben lévő további kapcsolatok elemzése alapján a kommunikatív tesztelés jóságmutatóinak fontosságával és az oktatás tartalmával kapcsolatos nézetek (7. faktor) között (0,43), az oktatási módszerek (0,18), valamint a tanítási gyakorlat (0,23) között mutatható ki összefüggés. Ez arra utalhat, hogy a tanárok a kommunikatív nyelvoktatás alapelveihez a megfelelő mérési alapelveket társítják. Az oktatás tartalma a vizsga osztálytermi mérési eljárásai, valamint a klasszikus tesztelmélet jóságmutatóinak fontosságával kapcsolatban kialakított nézetekkel is szignifikáns korrelációt (0,22 és 0,27) mutat. Mindezek mellett a vizsga nyelvoktatási gyakorlatra kifejtett hatása és a kommunikatív tesztelés fontosságával kapcsolatos nézetek között is van kapcsolat (0,33). Ez arra utalhat, hogy az érettségi kommunikatív szemlélete összefüggésben van a vizsga nyelvoktatási folyamatra kifejtett hatásával és a kommunikatív tesztelés alapelveivel is. Ugyanez az összefüggés érvényesül a vizsgafeladatok felhasználásával kapcsolatban (a korreláció értéke: 0,21).

5.4. A vizsgával kapcsolatos nézetek és attitűdök hatása a nyelvoktatás folyamatára, mérési-értékelési eljárásaira

A nemzetközi teszthatást vizsgáló kutatások tanulságai szerint a tanárok vizsgával kapcsolatos nézetei és attitűdjei meghatározó jelentőségűek a tanítási-tanulási folyamatra gyakorolt hatás mértékében (1.3.2. és 1.3.3. fejezet). Ebben a fejezetben így azt elemezzük, hogy a vizsga jelentőségéről kialakított tanári nézetek, a vizsga következményeivel kapcsolatos attitűdök, valamint a tanórai felkészüléssel kapcsolatos attitűdök mennyiben határozzák meg a vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézeteket, a vizsga és az osztálytermi értékelési eljárások közti kapcsolatról és a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézeteket, valamint a vizsgafeladatok oktatási folyamatban való alkalmazásának céljait. A továbbiakban a három nézettel és attitűddel kapcsolatos faktort (5.1. fejezet) háttérváltozóként kezeljük, és azt vizsgáljuk, hogy ezek mennyiben határozzák meg az elemzés során definiált nyelvoktatás folyamatát (5.2. fejezet) és mérési-értékelési eljárásait jellemző faktorokat (5.3. fejezet). Ebben a fejezetben szintetizáljuk az előzőekben részletesen jellemzett faktorokat.

A fejezet további részében először az egyes faktorok közötti összefüggéseket elemezzük a 22. melléklet alapján. Ezt követően regresszió-analízissel vizsgáljuk az érettségire vonatkozó tanári nézetek és attitűdök meghatározó szerepét a nyelvoktatás folyamatára és mérési-értékelési eljárására. A harmadik részben pedig a nemzetközi vizs-

gálatok alapján értelmezzük az eredményeket és megfogalmazzuk a vizsgálat legfontosabb tanulságait.

A tanárok vizsgával kapcsolatos nézetei (1. faktor) az osztálytermi mérési eljárásokhoz hasonló összefüggést mutatnak a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetekkel (0,23) és a vizsgafeladatok tanítási-tanulási folyamatban való felhasználásának céljaival (0,18). A nézetek főleg a mérés-értékelés területéhez kapcsolódó 8., 10. és 12. faktorról vannak összefüggésben, míg ez a vizsga nyelvoktatási folyamatra gyakorolt hatása (4. faktor) esetében nem jellemző. A vizsga következményeivel kapcsolatos attitűdök és a kommunikatív tesztelés jóságmutatóiról, a validitás fontosságáról kialakított nézetek, valamint a vizsgafeladatok felhasználása között mutatható ki szignifikáns kapcsolat (0,18–0,19). Ez azt jelzi, hogy a vizsgaeredményekből levont következtetések összefüggésben vannak az érettségi mérési-értékelési eljárásaira gyakorolt hatással. A tanórai felkészüléssel kapcsolatos attitűdök és a kommunikatív tesztelés jóságmutatóinak fontossága között is találunk szignifikáns összefüggést (0,26), a validitás aspektusaihoz hasonlóan (0,18).

A továbbiakban a vizsgával kapcsolatos attitűdök és nézetek mint háttérváltozók hatását elemezzük a fejezet elején definiált négy függő változóra. Mind a négy elemzésbe ugyanazokat a független változókat vontuk be, így az eredmények közvetlenül összehasonlíthatók. A 71. táblázatban az $r \cdot \beta \cdot 100$ értékeket foglaltuk össze és a modellek elemeinél jelöltük, hogy a független változó hatása a β értékek alapján szignifikáns-e; az egyes modelleket a szignifikanciaszintek jelzésével jellemezzük. A továbbiakban először a négy függő változóra végzett regressziós modelleket elemezzük egyenként, majd a háttérváltozók meghatározó szerepét hasonlítjuk össze a modellekben.

71. táblázat. Regressziós modellek a vizsga nyelvoktatási folyamatra gyakorolt hatásával kapcsolatos nézetekre mint függő változókra

Háttérváltozó	Vizsgált faktor			
	A vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézetek	A vizsga és az osztálytermi értékelési eljárások közti kapcsolatról kialakított nézetek	A kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek	A tanárok nézetei a vizsgafeladatok oktatási folyamatban való felhasználásának céljairól
1. A tanárok nézetei a vizsga jelentőségéről	0,18 n.s.	4,23 p<0,05	5,82 p<0,01	3,31 p<0,05
2. A vizsga következményeivel kapcsolatos attitűdök	2,16 n.s.	-0,01 n.s.	0,23 n.s.	2,88 n.s.
3. A tanórára való felkészüléssel kapcsolatos vizsgához kötődő attitűdök	3,36 p<0,05	2,07 n.s.	3,83 p<0,05	0,22 n.s.
Megmagyarázott variancia (%)	5,70	6,29	9,88	6,40
Szign.	p<0,05	p<0,05	p<0,02	p<0,05

A vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézetekre nincs szignifikáns hatása a tanárok vizsgára vonatkozó nézeteinek (1. faktor) és a vizsga következményeivel kapcsolatos attitűdöknek (2. faktor). A megmagyarázott variancia nagy része a 3. faktorból adódik. Ez arra utal, hogy a tanárok nem a vizsgával kapcsolatos nézeteik és attitűdjeik hatására követik a kommunikatív nyelvoktatás alapelveit. En-

nek feltehető oka, hogy a kommunikatív szemlélet már jóval a vizsga bevezetése előtt megjelent, az érettségi vizsga csak követte ezt. Ebből viszont az következik, hogy vizsgálatunkban nem tudjuk empirikusan igazolni a vizsgával kapcsolatos nézetek és a vizsga következményével kapcsolatos attitűdök meghatározó szerepét.

A vizsga és az osztálytermi értékelési eljárások közti kapcsolatról kialakított nézeteket az 1. faktor határozza meg. Ez azt mutatja, hogy a vizsgán követett értékelési szemlélet megjelenik a napi tanári munkában. Feltételezhető, hogy az értékelési szempontok és skálák használata a vizsga jelentőségének felismerése miatt vált meghatározóvá.

A kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetekre az 1. és a 3. faktornak van szignifikáns hatása. A vizsgával kapcsolatos nézetek befolyásoló szerepe abban jelenhet meg, hogy a vizsgakövetelményekben definiált értékelési szemlélet részeként értékelődött fel e tesztelési forma az osztálytermi mérésekben (lásd erről a 69. táblázat eredményeit). A nézetek mellett a tanóra való felkészüléssel kapcsolatos attitűdöknek is meghatározó szerepe van. Feltételezhető, hogy a kommunikatív tesztelés alapelveinek érvényesítése a tanárok által készített mérőfeladatokban alapos felkészülést, mérlegelést igényel. Ezt az mutatja, hogy 3. faktorba tartozó 11. állítás (*Mióta bevezették a kétszintű érettségit, több időre van szükség az órára való felkészüléshez*) nagy arányban határozza meg ($r=0,25$; $\beta=0,23$; $p<0,01$) a 8. faktor varianciáját.

A tanárok nézetei a vizsgafeladatok oktatási folyamatban való felhasználásának céljairól függő változóra az 1. faktor van szignifikáns hatással. Ugyanakkor ennél a faktornál a vizsga következményeivel kapcsolatos attitűdöknek arányaiban és a többi faktorhoz képest meghatározó szerepe van, de a 2. faktor hatása nem szignifikáns. E faktorba tartozó változók közül a 12. kérdőívtételnek (*Kínosan érzem magam, ha az érettségim a diákjaim rosszabbul teljesítenek, mint a kollégám diákjai*) van szignifikáns hatása ($r=0,12$; $\beta=0,23$; $p<0,02$), vagyis feltételezhető, hogy főleg e negatív hatás elkerülése érdekében alkalmazzák a tanárok a vizsgafeladatokat a tanítási-tanulási folyamatban.

A vizsga jelentőségével kapcsolatos tanári nézetek előzetes elvárásainkkal ellentétben nem fejtenek ki hatást a tanárok nyelvoktatási folyamatára, de meghatározzák a mérési-értékelési eljárásokkal kapcsolatos nézeteket. A nemzetközi szakirodalomban egyetértés van arról, hogy a vizsgával kapcsolatos nézetek elsősorban az oktatás tartalmát befolyásolják, és kisebb mértékű hatást gyakorolnak a módszerekre. Az értékelésről kevés információnk van, ezt a területet kevés vizsgálatban elemzik (lásd 1.3.3. fejezet). Eredményeink szerint a vizsgán követett értékelési szemlélet megjelenik a nyelvoktatási folyamatban: a vizsgán alkalmazott értékelési eljárásokban, a kommunikatív tesztelés fontosságának felértékelődésében, és a vizsgafeladatok felhasználásában.

A vizsga következményeivel kapcsolatos attitűdöknek egyik vizsgált függő változóra sincs szignifikáns hatása. Vizsgálatunk nem támasztja alá a nemzetközi szakirodalomban gyakran meghatározott összefüggést, mely szerint a vizsga következményei gyakran határozzák meg a tanárok tanítási gyakorlatát, valamint, hogy emiatt a tanárok célzottan a vizsgára tanítanak, és használják a vizsgafeladatokat a tanítási-tanulási folyamatban (lásd például *Alderson és Wall, 1993*). *Chen (2002)* kutatásában például 0,37-es korreláció volt a vizsga következményeivel kapcsolatos attitűdök és a vizsga tanítási folyamatra gyakorolt hatása között. Az összehasonlítást azért tesszük meg, mert az ebbe a dimenzióba tartozó kérdéseket *Chen (2002)* kutatásából vettük át (3. melléklet). Ez az eredmény arra utalt, hogy azok a tanárok, akik felelősnek érzik magukat a vizsgaeredmények miatt, nagyobb hatásról számoltak be. A saját vizsgálatunkban ilyen jellegű összefüggést és az oktatási folyamatra gyakorolt hatást nem mutattunk ki. Ez a tesztelést vizsgáló nemzetközi kutatások alapján azzal magyarázható, hogy a visszahatás jellege és mértéke oktatási kontextusonként eltérő mértékű (lásd 1.3.3. fejezet).

A tanórára való felkészüléssel kapcsolatos vizsgáláshoz kötődő attitűdök meghatározó szerepét a nyelvoktatási folyamatban, és a kommunikatív tesztelés kritériumainak fontosságában mutattuk ki. Mindkét esetben meghatározó volt az az állítás, hogy az érettségi hatására több idő szükséges az órákra való felkészülésre. Ez adódhat egyrészt abból, hogy a tanárok többsége váltott tankönyvet az új érettségi hatására és alkalmazza az érettségire felkészítő feladatgyűjteményeket. Másrészt abból is adódhat, hogy a tanárok a tanítási-tanulási folyamatban alkalmazott mérőfeladataik elkészítésekor ügyelnek a kommunikatív tesztelés kritériumainak biztosítására. Harmadrészt pedig az értékelési skálák használata és az értékelés elkészítése többletmunkát is jelenthet.

E fejezetben bemutatott vizsgálatunk több szempontból tekinthető újszerűnek. Egyrészt a hazai szakirodalomban a kommunikatív szemléletű nyelvtudást mérő vizsgák hatásának vizsgálata egyáltalán nem jellemző. Ez a terület jelentősen háttérbe szorult, annak ellenére, hogy ennek nagyon gazdag nemzetközi szakirodalma van, melyben a teszthatás mint a kommunikatív tesztelés fontos minőségi mutatója jelenik meg. Másrészt vizsgálatunkban kimutattuk, hogy a vizsgának hatása van az oktatás tartalmára: a vizsgára felkészítő tananyagok és mérőfeladatok alkalmazására, a vizsgán mért készségek fejlesztésére és megállapítottuk, hogy a vizsgával kapcsolatos nézetek és attitűdök az osztálytermi mérés-értékelés több aspektusára fejtik ki hatásukat. Ezek közé tartozik a vizsga értékelési eljárásainak tanítási gyakorlatban való alkalmazása, a kommunikatív tesztelés kritériumainak felértékelődése a tesztfeladatok elkészítésekor, valamint a vizsgafeladatok funkció szerinti alkalmazása. A vizsga ezen a területen kimutatható erős hatása azzal magyarázható, hogy a vizsgához részletes dokumentáció tartozik, értékelési szemlélete megfelel a kommunikatív tesztelés alapelveinek (lásd 2.3.1. fejezet) és a tanárok aktív résztvevői az értékelési folyamatnak: ismerik a vizsgán alkalmazott értékelési eljárásokat. Így a vizsga osztálytermi mérésekre gyakorolt hatásának mértékét, jellegét és aspektusait is vizsgáltuk, mely a nemzetközi teszthatást vizsgáló empirikus kutatások elhanyagolt területe (lásd 1.3.3. fejezet).

Az elemzésünk korlátai a bevezetőben említettekén kívül sokrétűek: a teszthatás jelenségét csak közvetve, a tanárok nézetein és attitűdjein keresztül elemeztük, de konkrét viselkedésüket nem. Ebből következik, hogy a teszthatás jelenségét és természetét a tanórai munkától eléggé távoli dimenzióban ragadtuk meg. Mindezek mellett nem ötvöztük a kvalitatív és kvantitatív kutatási módszereket: nem készítettünk a teszthatást vizsgáló kutatásokban megszokott tanári interjúkat, nem végeztünk osztálytermi megfigyeléseket. Nem vizsgáltuk a tanulók e témákkal kapcsolatos nézeteit, attitűdjeit, nem elemeztük a vizsga hatására létrejövő tananyagok és tankönyvek minőségét és összhangját a vizsga szemléletével. Elemzéseinkből nem derül ki egyértelműen, hogy a vizsga hogyan fejt ki hatását a tanítási-tanulási folyamatra, a kapott eredmények értelmezésében az okokra sok esetben csak hipotéziseket tudtunk megfogalmazni. Mindezek alapján egyértelmű, hogy további kutatások szükségesek. A további nyelvpedagógiai vizsgálatokhoz alkalmazható lehet a vizsgálatunkban használt elméleti modell (11. ábra), mellyel jól lehet modellezni a teszthatás komplex folyamatát, ez alapján pontosan definiálhatók a vizsgálandó területek és az azok közötti kapcsolatok. Mindezek mellett a felhasznált adatgyűjtő eszközünk validitására nagy hangsúlyt fektettünk, melyet a nemzetközi vizsgálatok alapján készítettünk. A kérdőív az angol és német nyelvet tanító tanárok körében azonosan működik: csak a 2., 32. és 38. állításnál találtunk szignifikáns különbséget, a többi kérdőív-tételnél és a faktoroknál nincs jelentős különbség. A vizsgálatban francia és olasz nyelvtanárok is részt vettek, de számuk nagyon kevés volt. Ugyanakkor feltételezzük, hogy a kérdőívünk alkalmas a többi nyelvi érettségi hatásának vizsgálatára. Az elemzésünk során számos vizsgálandó területet azonosítottunk. Ezek az információk alapjai lehetnek a további empirikus vizsgálatoknak.

6. ÖSSZEGZÉS, KÖVETKEZTETÉSEK ÉS JAVASLATOK

Az értekezés célja az volt, hogy feltárja a nyelvi mérés-értékelés nemzetközi tendenciáit és bemutassa hazai jellemzőit. Mindezek mellett célul tűztük ki e két terület legfontosabb tanulságainak szintetizálását egy elméleti modellben, melyben megjelöltük és indokoltuk az empirikus vizsgálatban elemzett célterületeket. Ezekkel összhangban az idegen nyelvi érettségi működését a vizsgán elért eredmények alapján, míg a nyelvoktatási folyamatra kifejtett hatását a tanárok visszajelzése alapján elemeztük.

E fejezet két részből áll: az elsőben összegezzük a szakirodalmi kutatások legfontosabb tanulságait és az empirikus vizsgálat eredményei alapján adunk választ és reflektálunk a kutatási kérdésekre és a hipotézisekre (3.2. fejezet). A második részben az empirikus vizsgálatok alapjául szolgáló szintetizáló modellben lévő elemek és visszacsatolási folyamatok (11. ábra) alapján fogalmazzuk meg az empirikus vizsgálatainkból levonható következtetéseket, majd teszünk javaslatokat a további vizsgálandó területekre és a vizsgafejlesztési folyamat számára.

Az elméleti és gyakorlati kérdésekkel foglalkozó nemzetközi szakirodalom alapján megállapítható, hogy a kommunikatív tesztelés mára igen gazdag szakterületté vált, amely a pedagógiai értékelés sajátos jellegű önálló területe. A kommunikatív tesztelés a klasszikus tesztelméletre építve határozza meg a kommunikatív tesztek minőségi mutatóit, funkcióit és típusait. A kommunikatív nyelvi kompetenciák jellegéből adódóan azonban a tesztelmélet egészére vonatkozó mérésmetodikai fogalmakat, mérési és értékelési eljárásokat specifikus tartalmakkal tölti fel. Gazdagítja a klasszikus tesztelmélet jóságmutatóit, kiemeli a validitás fontosságát, és újabb típusait, aspektusait is meghatározza. Ezek közé tartozik a tesztfejlesztési folyamatra vonatkozó autentikusság, interaktivitás és a hatékonyság kritériuma, a teszt lebonyolítására és következményeire vonatkozó igazságosság, összehasonlíthatóság, a döntések reprodukálhatósága és a tesztthatóság, amelyek rendkívül fontosak és e kritériumok biztosítása a teszt érvényességének és a megbízhatóságának az előfeltételeiként fogható fel. Szakirodalmi elemzésünk alapján ezek a kritériumok összhangban vannak a kompetenciaértékelés minőségi mutatóival. A kommunikatív tesztek típusai közül kiemelendők a kompetencia és performancia tesztek, és az ezzel szorosan összefüggő közvetett és közvetlen tesztek, amelyek egy-egy skála két végpontját jelenítik meg és a tesztalkalmazóknak a feladata, hogy ezeknek a teszteknek az egyensúlyát megtalálják.

A nyelvtudást mérő vizsgák figyelembe veszik a minőségi mutatókat és alkalmazzák a kommunikatív tesztelési eljárásokat. Ezt a folyamatot a Közös Európai Referenciakeret szintrendszere is elősegíti, amely esetén fontos kérdés, hogy miként tudják a vizsgafejlesztők a tesztfeladatokat, értékelési eljárásokat a Referenciakeret szintjeihez igazítani. E folyamat eredményességének növelésében alkalmazhatók a valószínűségi tesztelméleti modellek.

A kommunikatív tesztelés alapelveit alkalmazó vizsgák működtetésekor fontosak az alapos empirikus fejlesztő kutatások. Ezek hozzájárulnak ahhoz, hogy a vizsgázók nyelvtudásáról pontos adatokat szerezzünk, és ezek alapján megalapozott döntéseket hozzunk. E vizsgákkal szemben az utóbbi évtizedben számos elvárást fogalmaztak meg, melyek közül a legfontosabb, hogy a vizsga összhangban legyen a kommunikatív nyelvoktatás alapelveivel és pozitívan hasson a tanítási-tanulási folyamatra. A nemzetközi szakirodalomban a vizsgák hatásának fogalmi szintű megnevezése nem egységes, ugyanazt a jelenséget többféleképpen, más-más jelentéstartalommal definiálják a kutatók. Ezek szintetizálása alapján a tesztthatóság komplex jelenség, mely közvetlen és közvetett hatást fejt ki a vizsgával kapcsolatban álló személyekre, a tanítási-tanulási folyamatra, a tanulás tartalmára és minőségére; tágabb értelemben véve az eredmények alapján

hozott döntésekkel a közoktatásra és a társadalomra. Az elméleti kutatások során több modell készült, amelyek nem határozzák meg pontosan a folyamatot leíró jellemzőinek leírását és kölcsönhatását az eredményekkel, így nagy igény van egy, a folyamatot szintetizáló modellre. Bár az empirikus vizsgálatok eredményei csak az adott kontextusban értelmezhetők és nehezen általánosíthatók, de megállapítható, hogy a vizsgának a tanítás tartalmára általában nagyobb hatása van, mint a módszerekre, az osztálytermi értékelésre gyakorolt hatásokat azonban kevés kutatásban vizsgálták.

A hazai kontextusban a kétszintű nyelvi érettségi vizsgakövetelményének és vizsgaleírásának alapján a vizsga elhelyezhető a kommunikatív tesztelés rendszerében. Az érettségi törekszik az autentikusság, az interaktivitás, a hatékonyság és az igazságosság kritériumainak megvalósítására. Ugyanakkor a vizsgafunkciók keveredése, a vizsgaminőség ellenőrzésének hiánya miatt kérdéses az összehasonlíthatóság, a döntések reprodukálhatóságának teljesülése. Ugyanígy nem tudjuk megállapítani, hogy a vizsga a közoktatásra milyen jellegű hatást fejt ki. A kommunikatív tesztek típusai közül a vizsgán elsősorban szűk értelemben vett performancia teszteket használnak.

Az empirikus vizsgálatokban az idegen nyelvi érettségit több szempontból, négy területen vizsgáltuk. Egyrészt elemeztük a vizsgán elért készségenkénti teljesítmények változásait és meghatározó tényezőit, másrészt a receptív készséget mérő vizsgarészekkel a vizsga külső és belső validitását a Rasch-moddal, harmadrészt feltártuk az íráskészség vizsgarészben alkalmazott értékelési eljárások minőségét a parciális kredit-moddal és bemutattuk az értékelési skálák empirikus validálásának folyamatát. Negyedik területként elsősorban az angol- és németnyelv-tanárok vizsgával kapcsolatos visszajelzéseit elemeztük. Fő célunk az volt, hogy ezek alapján visszajelzést adjunk a vizsgafejlesztés számára. A továbbiakban e négy területen elért eredményeinket összegezzük a 3.2. fejezetben megfogalmazott kutatási kérdések és hipotézisek alapján.

I. A vizsgán elért készségenkénti teljesítmények alapján megállapítható, hogy (1) középszinten mindkét nyelvből általában az íráskészség és a nyelvhelyesség vizsgarészben alacsonyabbak a teljesítmények, míg a szövegértést mérő teszteken és a szóbeli vizsgán magasabbak. Mindkét nyelvre jellemző, hogy az olvasott és a hallott szöveg értése feladatlapok átlagos megoldottsága ingadozik. Ez arra utal, hogy a feladatsorok teljesítettsége nem állandó, ezért a döntések reprodukálhatósága alacsony. Az emelt szintű vizsgán mindkét nyelvből magasak az átlagteljesítmények, a vizsgázók többsége a nyelvhelyesség vizsgarész kivételével a 80–100%-os teljesítménysávban van, így a vizsga nem tölti be megfelelően a felvételi funkcióját. (2) Az iskolatípusok szerinti különbségek jelentősek középszinten, és minden évben folyamatosan emelkednek; e különbségek teljesítményre gyakorolt hatása is jelentős. (3) A két nyelv esetében hasonló tendenciák figyelhetők meg: a megfogalmazott hipotézisekkel összhangban a gimnáziumban érettségizők szignifikánsan jobban teljesítenek a szakközépiskolában tanulóknál mindkét vizsgaszinten, a nappali tagozaton tanulók szignifikánsan magasabb teljesítményt érnek el a felnőttoktatásban tanulóknál. Mindezek mellett kimutattuk, hogy a középszintű vizsgán a felnőttoktatásban a levelező képzésben tanulók jobban teljesítenek a felnőttoktatás nappali képzésében résztvevőknél. Elvárásainknak megfelelően a produktív készségeknél a lányok teljesítenek jobban a fiúknál, de a nemek közötti különbségek teljesítményre gyakorolt hatása a teljes vizsgán nem jelentős.

II. A receptív készségeket mérő vizsgarészek feladat- és itemszintű vizsgálati a 2006. évi középszintű német érettségien azt mutatták, hogy (4) a feladatkészítők által megállapított nehézségi sorrend többnyire megegyezett az empirikus nehézség növekedésével. (5) Feltételezéseinknek megfelelően a vizsgafeladatok itemnehézségi indexei nem fedik le teljesen a vizsgázók nyelvi készség szintjei által meghatározott intervallumot, de ez nem általánosítható az összes vizsgarészre. A nyelvhelyesség feladatok

itemei többnyire megfelelően fedik le a vizsgázók képességszintjét. Az olvasott és a hallott szöveg értése feladatainál két, jól elkülöníthető itemcsoportot találtunk, amelyek egyik része az átlagos képességszint alatt és felett lévő vizsgázók nyelvtudását méri. Azonosítottunk egy magasabb képességszintet is, amit a feladatsorok nem mérnek. (6) Az egyes készségek minta- és itemtérképeinek elemzése alapján szükséges a feladatsorok ekvivalenciáját vizsgálni, a két vizsgaszint egymáshoz és a KER-hez való viszonyát horgony- és referenciaitemek segítségével egy újabb beméréssel elemezni és ezek alapján egy IRT-alapú feladatbankot kialakítani. Az európai projektek eredményei alkalmazhatók a referenciateladatok kiválasztásában, az érettségi feladataihoz tartozó szövegek és itemek jellemzésében. A feladatbank kiépítésével empirikusan igazolhatóvá válhat a vizsgaszintek KER-hez való illeszkedése, és a feladatbank viszonyítási pontként funkcionálhat a további feladatok fejlesztésekor.

III. A német érettségi íráskészség-vizsgálatának eredményei szerint (7) az értékelési szempontokat jellemezhetjük az átlagos nehézségi szintjük és modellilleszkedésük alapján. A nem illeszkedő értékelési szempontok azt jelezhetik, hogy az íráskészség más dimenzióit méri (például a középszintű vizsga első feladata), vagy működésük eltérő (például a kommunikatív komponensek). A skálapontok működése akkor megfelelő, ha küszöbértékei az íráskészség-szint növekedésével arányosan emelkednek. Ez a vizsgán megvalósult, de mindkét szinten a szövegalkotásnál és a szókinés, kifejezőmódnál aránytalanul. A küszöbértékekhez tartozó modellilleszkedés-paraméter magas értéke azt jelezheti, hogy különböző képességszintű vizsgázóknál azonos pontszámot adnak az értékelők. Ezt emelt szinten elsősorban a nyelvtan, helyesírás értékelési szempontnál a magasabb képességtartományban állapítottuk meg, míg középszinten az értékelők főleg e szempont alapján ítélik meg a vizsgázók írásművét. (8) A minta- és itemtérkép alapján azonosítottunk egy magasabb képességszintet, amelyet az egyes skálapontok küszöbértékei nem fednek le. Ez a megállapítás igazolja a hipotézisünket, vizsgálatunk azt is kimutatták, hogy az egyes értékelési skálák működése szempontként eltérő. (9) Az értékelésiskála-használatot karakterisztikus görbékkel jellemeztük, amelyek a skálák redundáns működését mutatták például a szövegalkotásnál, valamint arra utalnak, hogy a skála bizonyos pontszámait az értékelők nem használják például a tartalom szempontnál.

IV. A tanárok visszajelzéseit vizsgáló kérdőíves felmérésünk eredményei a két leggyakrabban tanított idegen nyelvre, az angolra és a németre általánosíthatók. Eredményeink szerint (10) a vizsgát a tanárok elfogadhatónak és a korábbinál megbízhatóbbnak és érvényesebbnek tartják. Ez a megállapítás megfelel az előzetes elvárásainknak. Ugyanakkor azt mutattuk ki, hogy a vizsgával kapcsolatos nézetek formálódásában elsősorban a tanárok vizsgáztatási tapasztalatai meghatározók. (11) A tanárok nézetei szerint a nyelvoktatási gyakorlatot jelentős mértékben befolyásolják a korábbi nyelvtanítási tapasztalatok és az érettségi vizsgakövetelményei. A vizsgadokumentumokban lévő alapelvek megismerése hatással van arra, hogy a tanárok a mérőfeladataik elkészítésében fontosnak tartják a kommunikatív tesztelés kritériumait. (12) A tesztelést vizsgáló nemzetközi kutatások alapján megfogalmazott hipotézisünkkel összhangban a tanárok szerint a vizsga a nyelvoktatás tartalmára nagyobb hatást fejt ki, mint a módszerekre. Feltételeztük, hogy a vizsga értékelési szemlélete befolyásolja a tanárok mérési-értékelési eljárásokkal kapcsolatos nézeteit. Ezt megerősítve vizsgálatunkban kimutattuk, hogy (13) a tanárok vizsgával kapcsolatos nézetei és attitűdjei az osztálytermi mérés-értékelés több aspektusára fejtik ki hatásukat. Felértékeli a kommunikatív tesztelés kritériumainak fontosságát és (14) a vizsgafeladatok funkció szerinti felhasználását.

Végül, az értekezés zárásaképpen, az elért eredményeket viszonyítjuk az elméleti modellhez és ez alapján fogalmazzuk meg a következtetéseket, meghatározzuk a vizsgálatunk elméleti és gyakorlati jelentőségét, az ebből levonható konzekvenciákat.

Az idegen nyelvi érettségi fejlesztése a kilencvenes években felértékelődő és az ezredfordulón kiemelt szerepűvé váló nyelvtudás-konceptió változásaival összhangban, a társadalmi és kliensi igények és célok figyelembevételével történt. Ennek eredményeként egy olyan vizsga jött létre, amely magas presztízzsel rendelkezik, mivel záróvizsgaként, felvételiként és nyelvvizsgaként is működik. A tanári kérdőíves felmérésünk alapján a nyelvtanárok elfogadják, a korábbinál megbízhatóbbnak tartják a vizsgát (5.1. fejezet). Ez a nemzetközi vizsgálatokkal összhangban megerősíti és általánosíthatóvá teheti az igényeknek és céloknak megfelelő vizsgafejlesztési folyamat fontosságát, a vizsga státuszának jelentőségét a teszthatás elindulásában.

Eredményeink szerint ezek a tényezők meghatározzák a tanárok vizsgával kapcsolatos nézeteit, a tanórai felkészüléssel kapcsolatos attitűdöket, közepes hatásuk van a vizsga következményeivel kapcsolatos attitűdökre. Vizsgálatunkban nem tudtuk empirikusan igazolni e nézetek és attitűdök meghatározó szerepét a tanárok nyelvoktatási folyamatára, de kimutattuk, hogy a vizsga jelentőségére vonatkozó nézetek befolyásolják a tanárok mérés-értékeléssel kapcsolatos vélekedését (5.4. fejezet). Mindezek mellett vizsgálatunk megerősíti, hogy a vizsgával kapcsolatos képzéseknek, a vizsgadokumentációknak nagy szerepe van a vizsga megítélésében és a vizsga észlelt hatásában.

A nyelvi érettségi hatásait azonban tanórai szinten nem elemeztük, a kutatásunk semmit nem mond a tanulók tanulási folyamatainak az érettségi vizsgával való kapcsolatáról, így nem egyértelmű, hogy a vizsga hogyan fejt ki hatását a tanítási-tanulási folyamatra. Ezzel kapcsolatban azonban közvetve két megállapítást tehetünk. Egyrészt a vizsga hatására létrejövő új tananyagokat, érettségi feladatgyűjteményeket a tanárok szükségesnek érzik alkalmazni a tanítási-tanulási folyamatban, vagyis ez meghatározza a tanórai tervezésüket, az oktatás tartalmát. Ez megerősíti a modellben lévő visszajelentési kört (11. ábra). Másrészt a tanulói teljesítmények javulásából (4.1. fejezet) arra következtetünk, hogy megindult a vizsgára való felkészítés, a tanárok a vizsgán előforduló feladattípusok és értékelési eljárások használatára az évek múlásával egyre nagyobb hangsúlyt fektetnek. Ugyanakkor ezek a megállapításaink kizárólag az idegen nyelvi érettségi vizsgákra vonatkoznak és közvetve a nyelvtudást mérő vizsgákra általánosíthatók.

A modell öt visszacsatolási kört tartalmaz, amelyek egyik funkciója a teszthatás folyamatának empirikus vizsgálata. Az 1. visszacsatolási kör vizsgálata (4.2. és 4.3. fejezet) elsősorban a vizsgán alkalmazott mérési-értékelési eljárások és a követelményszintek meghatározására szolgált. Mindkét esetben a valószínűségi tesztelméleti modelleket alkalmaztuk, így az eredményeink módszertanilag általánosíthatók. A követelményszintek empirikus meghatározása, bemérése más vizsgatárgyban is lényeges kutatási feladat: fontos lenne például a matematikai, természettudományi és az olvasás-szövegértési feladatsorok nehézségének összekapcsolása a nemzetközi mérések (például a PISA-vizsgálatok) kompetenciaszintjeivel, a vizsgaszintek közötti tartalmi és mennyiségbeli különbségek meghatározása és feladatbank kiépítése. Ehhez felhasználhatók lehetnek a 4.2. fejezetben bemutatott elemzési eljárások. A 4.3. fejezetben az értékelési szempontok és skálák validálásának folyamatára mutattunk példát. A fejezetben lévő elemzések az idegen nyelvi érettségin az alacsony megbízhatósággal rendelkező beszédképesség esetében is kellő alapot szolgáltatnak. Mindezek mellett a vizsgálatunk módszerei, az értelmezett paraméterek olyan vizsgatárgyakban is alkalmazhatóak, ahol az értékelés analitikus skálákkal történik.

A 2. visszacsatolási körre vonatkozó elemzéseink (4.1. fejezet) középszinten jelentős különbségeket mutattak iskolatípusok és képzési típusok szerint. A felnőttképzésben és a szakközépiskolákban érettségizők nyelvtudása messze elmarad a gimnáziumban tanulókéétól és ez a különbség az évek múlásával növekszik. Emelt szinten a követelményszintek, a feladatsorok minősége és a vizsga felvételi funkciójának teljesítése jelent problémát. Ezek további vizsgafejlesztési igényeket és újabb fejlesztési célokat jelölhetnek ki.

A 3. visszacsatolási kört feltáró kérdőíves felmérésünk (5. fejezet) a vizsgafejlesztés számára felhívja a figyelmet a vizsgával kapcsolatos képzések fontosságára, melyek a vizsga filozófiájának megismertetését szolgálják. Mindezek mellett főleg a produktív készségek mérésével kapcsolatban van szükség további fejlesztésre (például az íráskészség vizsgarészhez értékelési példák a javítókulcsban, felvételek a szóbeli produktumok megítéléséhez). Ezek az értékelési eljárások nemcsak magasabb megbízhatóságát eredményezhetnék, hanem a vizsgához tartozó dokumentumokhoz hasonlóan (lásd 5.3. fejezet) hatással lehetnének a tanítási-tanulási folyamatra is.

Vizsgálatunkban e három visszacsatolási kört közvetlenül elemeztük, de törekedtünk arra, hogy a modellben szereplő másik két körről is szerezzünk információkat. A 4. visszacsatolási kör igazolásához a tanulók vizsgával kapcsolatos véleményét, felkészülési folyamatát egyáltalán nem, de a vizsgaszintek választásának eredményét vizsgáltuk az OH adatbázisai alapján (4.1.1. fejezet). Ebből megállapítottuk, hogy a diákok egyre nagyobb arányban választják az előrehozott vizsgát, amely a nyelvi érettségik esetében ellentmondásos. Szükséges megvizsgálni a vizsgafajták választásának okait és a nyelvtanulási folyamatra gyakorolt hatását.

A tananyagoknak a vizsga alapelveivel és a fejlesztők szándékaival való összhangját jelölő 5. visszacsatolási kört is csak közvetve, a tanári kérdőívvel elemeztük (5.1. fejezet). Ebből megállapítottuk, hogy a tanárok többsége váltott tankönyvet az új érettségi hatására és használ felkészítő feladatgyűjteményeket, így lényeges lenne ezek minőségének elemzése.

Vizsgálataink egyértelműen jelzik a vizsgafejlesztés szükségességét. Lényeges lenne a mérési-értékelési eljárások és a követelményszintek ellenőrzése, felülvizsgálata. Erre azonban az OH adatbázisai nem alkalmasak, így szükségesnek tartjuk a feladat- és itemszintű elemzéseket, és az IRT-alapú feladatbank kiépítését. Ezekkel emelhető a mérés színvonala, a döntések reprodukálhatósága. A további vizsgálatoknak a teljesítmények közötti különbségek okainak feltárására, diagnosztizálására kellene vonatkozniuk. Szükségesek a tantárgy-pedagógiai vizsgálatok is, amelyek egy-egy területre kiterjedően részletesen elemzik a tanárok tanítási gyakorlatát és a tanulók nyelvtanulási folyamatát, és az ebben megjelenő hatásokat. E vizsgálatok megtervezéséhez, a vizsgálandó célterületek kijelöléséhez és az azok közötti kapcsolatok jellemzéséhez jól használhatónak tartjuk az elméleti modellt, melynek relevanciáját az empirikus vizsgálataink többnyire alátámasztják. Ugyanakkor ez a megállapításunk csak korlátozásokkal érvényes: eredményeink elsősorban az idegen nyelvi érettségire vonatkoznak, a modellben lévő területek közül sem elemeztük az összeset, mivel ez meghaladta volna az értekezés kereteit, továbbá ezek a területek csak hosszabbtávon vizsgálhatók. Így szükségesek a további kutatások, amelyek alapján és az érettségi továbbfejlesztésével a vizsga olyan évenkénti, rendszerszintű felméréssé válhat, amely egyrészt visszajelzést nyújt a közoktatás résztvevői számára, másrészt lehetővé teszi a vizsgázók tudásszintjének követését az évek során.

KÖSZÖNETNYILVÁNÍTÁS

Köszönettel tartozom témavezetőmnek, Vidákovich Tibornak, hogy munkámat útmutatásaival és szakmai tanácsaival mindvégig irányította, formálta kutatói szemléletemet, és segítette az értekezés elkészítésének folyamatát. Ugyancsak köszönöm a Neveléstudományi Doktori Iskola vezetőjének, Csapó Benőnek a folyamatos támogatását és ösztönzését a PhD-fokozat megszerzésére.

Az értekezés és a tézisek színvonalának növeléséhez Molnár Gyöngyvér és Hercz Mária járult hozzá, mindkettőjüknek köszönöm a dolgozathoz fűzött észrevételeiket, javaslataikat. Külön köszönöm Molnár Edit Katalinnak a szakmai támogatását és bátorítását, az angol nyelvű téziszfüzet lektorálását és javítását.

Az értekezésben bemutatott adatbázisok másodelemzését Einhorn Ágnes tette lehetővé, köszönöm, hogy a kérdőívhez és a dolgozathoz kapcsolódó tanácsaival segítette munkámat. A tanári kérdőív szakmai lektorálását Nikolov Marianne, Petneki Katalin, Sominé Hrebik Olga és Szénich Alexandra végezte, köszönöm nekik a kérdőív fejlesztéséhez és javításához fűzött észrevételeiket. Köszönettel tartozom Pap-Szigeti Róbertnek az elektronikus kérdőív létrehozásáért, az adatfelvételben nyújtott segítségéért. Hasonlóan köszönöm a felmérésben résztvevő nyelvtanároknak a kérdőív kitöltését.

Köszönet illeti a Neveléstudományi Doktori Iskola tanárait és hallgatóit a munkámat segítő észrevételeikért, és külön köszönöm Fejes József Balázsnak, Kasik Lászlónak és Tóth Editnek a dolgozat egyes fejezeteihez fűzött javaslataikat, a témával kapcsolatban folytatott beszélgetésekben megfogalmazott ötleteiket.

Végül köszönöm édesanyámnak és testvéremnek, hogy támogatásukkal biztosították számomra az elmélyült munkához szükséges háttérrel.

IRODALOM

- 100/1997. (VI. 13.) Korm. Rendelet az érettségi vizsga vizsgaszabályzatának kiadásáról
(<http://net.jogtar.hu/jr/gen/getdoc.cgi?docid=99700100.kor>, megtekintés: 2010. 01. 25.)
- A 2010. január 1-től hatályos, Az érettségi vizsga részletes követelményeiről szóló 40/2002. (V. 24.) OM rendelet (2010): Oktatási és Kulturális Minisztérium, Budapest.
(<http://www.oh.gov.hu/kozoktatás/kapcsolodo-jogszabaly/2010-január-1-tol>, megtekintés: 2010.01.25.)
- Akkreditációs Kézikönyv (2007): Oktatási Hivatal Nyelvvizsgáztatási Akkreditációs Központ, Budapest
(<http://www.nyak.hu/nyat/doc/kezikonyv2007/AK20070920.pdf>, megtekintés: 2010.01.25.)
- Albers, H. G. és Bolton, S. (1995): *Testen und Prüfen in der Grundstufe. Einstufungstests und Sprachstandsprüfungen. Fernstudieneinheit 7*. Langenscheidt, Berlin.
- Alderson, J. C. (1991): Language testing in 1990s: How far have we come? How much further have we to go? In: Anivan, S. (szerk.): *Current developments in language testing*. SEAMEO Regional Language Centre, Singapore. 1–26.
- Alderson, J. C. (1999): *Testing is too important to be left to testers*. Paper presented at the Third Annual Conference on Current Trends in English Language Testing, United Arab Emirates University, Al Ain and Zayed University, Dubai Campus. May, 1999.
(<http://www.ling.lancs.ac.uk/groups/crile/docs/crile53alderson.pdf>, megtekintés: 2007. 06. 11.)
- Alderson, J. C. (2001): Szertefozsló mítoszok: Számít-e a heti óraszám? *Modern Nyelvoktatás*, 7. 2–3. sz. 19–28.
- Alderson, J. C. (2002): Testing proficiency and achievement. In: Coleman, J. A., Grotjahn, R. és Raatz, U. (szerk.): *University language testing and the C-Test. Fremdsprachen in Lehre und Forschung 31*. AKS-Verlag, Bochum. 15–30.
- Alderson, J. C. (2004): Foreword. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. ix–xii.
- Alderson, J. C. és Banerjee, J. (2001): Impact and washback research in language testing. In: Elder, C., Brown, A., Iwashita, N., Grove, E., Hill, K., Lumley, T., McNamara, T. és O'Loughlin, K. (szerk.): *Experimenting with uncertainty: essays in honour of Alan Davies*. Cambridge University Press, Cambridge. 150–161.
- Alderson, J. C., Clapham, C. és Wall, D. (1995): *Language test construction and evaluation*. Cambridge University Press, Cambridge.
- Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S. és Tardieu, C. (2004): *The development of specifications for item development and classification within the Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Reading and listening. Final report of the Dutch CEF Construct Project*. (http://eprints.lancs.ac.uk/44/1/final_report.pdf, megtekintés: 2007. 06. 11.)
- Alderson, J. C. és Hamp-Lyons, L. (1996): TOEFL preparation courses: A study of washback. *Language Testing*, 13. 3. sz. 280–297.
- Alderson, J. C. és Huhta, A. (2005): The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*, 22. 3. sz. 301–320.
- Alderson, J. C., Nagy, E. és Öveges, E. (2000, szerk.): *English language education in Hungary. Part II. Examining Hungarian learners' achievements in English*. The British Council Hungary, Budapest.
- Alderson, J. C. és Wall, D. (1993): Does washback exist? *Applied Linguistics*, 14. 2. sz. 115–129.
- Alderson, J. C. és Wall, D. (1996, szerk.): Special issue: Washback. *Language Testing*, 13. 3. sz.
- Al-Jamal, D. és Ghadi, N. (2008): English Language General Secondary Certificate Examination: Washback in Jordan. *The Asian EFL Journal Quarterly*, 10. 3. sz. 158–186.
- ALTE (1998): Handbook of language examinations and examination systems. Joint Committee on Testing Practices, Washington D.C.
- Altmayer, C. (2004): Sprachkultur und Mehrsprachigkeit: Neuerscheinungen zur europäischen Sprachenpolitik. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 9. 2. sz. 1–10.
- Andrews, S. (2004): Washback and curriculum innovation. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 37–50.
- Andrews, S., Fullilove, J. és Wong, Y. (2002): Targeting washback: a case-study. *System*, 30. 2. sz. 207–223.
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A. és van der Vleuten, C. P.M. (2006): The wheel of competency assessment. Presenting quality criteria for Competency Assessment Programs. *Studies in Educational Evaluation*, 32. 2. sz. 153–177.

- Bachman, L. F. (1990): *Fundamental considerations in language testing*. Oxford University Press, Oxford.
- Bachman, L. F. (2005): Building and supporting a case for test use. *Language Assessment Quarterly*, **2**. 1. sz. 1–34.
- Bachman, L. F. és Palmer, A. S. (1996): *Language testing in practice*. Oxford University Press, Oxford.
- Bachman, L. F. és Palmer, A. S. (2005): *Language assessment in practice*. Oxford University Press, Oxford.
- Bailey, K. M. (1996): Working for washback: A review of the washback concept in language testing. *Language Testing*, **13**. 3. sz. 257–279.
- Bailey, K. M. (1999): *Washback in language testing. TOEFL Monograph Series, Ms. 15*. Educational Testing Service, Princeton, NJ.
- Banerjee, J., Kaftandjieva, F., Takala, S. és Verhelst, N. (2004): *Szintillesztési módszertani segédlet. A nyelvvizsgák illesztése a Közös európai referenciakerethez című kézikönyv előzetes, kísérleti verziójához*. Nyelvvizsgát Akkreditáló Testület – PH Nyelvvizsgáztatási Akkreditációs Központ, Budapest.
- Bárdos Jenő (2001): Az idegen nyelvi mérés és értékelés fejlődése. In: Csapó Benő és Vidákovich Tibor (szerk.): *Neveléstudomány az ezredfordulón. Tanulmányok Nagy József tiszteletére*. Nemzeti Tankönyvkiadó, Budapest. 254–267.
- Bárdos Jenő (2002): *Az idegen nyelvi mérés és értékelés elmélete és gyakorlata*. Nemzeti Tankönyvkiadó, Budapest.
- Bárdos Jenő (2003): A nyelvtudás megítélésének korlátai. *Iskolakultúra*, **13**. 8. sz. 28–39.
- Bárdos Jenő (2004): *Nyelvpedagógiai tanulmányok*. Iskolakultúra könyvek 24, Pécs.
- Bárdos Jenő (2005): A magyarországi nyelvvizsgák és az európai mobilitás. *TEE Szemle*, 3–12.
- Bárdos Jenő (2006): A nyelvtudás-fogalom metamorfózisai – kritikai elemzés. In: Silye Magdolna, Kurtán Zsuzsa, Sturcz Zoltán és Wiwczarowski, T. B. (szerk.): *PORTA LINGUA – 2006 Utak és perspektívák a hazai szaknyelvtanításban és –kutatásban*. Debreceni Egyetem, Agrártudományi Centrum, Center Print Nyomda, Debrecen. 15–22.
- Barker, F. és Khalifa, H. (2009): Cambridge ESOL. *Research Notes*, **35**. 1–36.
- Barrett, S. (2001): The impact of training on rater variability. *International Education Journal*, **2**. 1. sz. 49–58.
- Bartels, N. (2006): *The construct of cognition in language teacher education and development*. Inaugural-Dissertation zur Erlangung des Doktorgrades der Philosophie des Fachbereiches 05: Sprache, Literatur, Kultur der Justus-Liebig-Universität Gießen.
- Báthory Zoltán (1985): *Tanítás és tanulás*. Tankönyvkiadó, Budapest.
- Biggs, J. B. (1995): Assumptions underlying new approaches to educational assessment. *Curriculum Forum*, **4**. 2. sz. 1–22.
- Bolton, S. (1996): *Probleme der Leistungsmessung. Lernfortschrittstests in der Grundstufe. Fernstudieneinheit 10*. Langenscheidt, Berlin.
- Bond, L. A. (1996): Norm- and Criterion-referenced testing. *Practical Assessment, Research and Evaluation*, **5**. 2. sz. (<http://pareonline.net/getvn.asp?v=5&n=2>, megtekintés: 2006. május 8.)
- Boócz-Barna Katalin (2004): Az Európa Tanács dokumentumai a nyelvtanításban In: Einhorn Ágnes (szerk.): *Kézikönyv a német érettségi vizsga reformjához*. Országos Közoktatási Intézet, Budapest. 283–297.
- Boócz-Barna, K. és Majorosi, A. (2002): Ein Dokument für die Schublade oder doch etwas mehr? Gemeinsamer Europäischer Referenzrahmen für Sprachen: lehren, lernen, beurteilen. *Deutschunterricht für Ungarn*, **17**. 1–2. sz. 21–32.
- Brualdi, A. (1999): *Traditional and modern concepts of validity*. ERIC Document, ED435714. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.
- Bukta Katalin és Nikolov Marianne (2002): Nyelvtanítás és hasznos nyelvtudás: az angol mint idegen nyelv. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 169–192.
- Burrows, C. (2001): Searching for washback: The impact of assessment in the Certificate in Spoken and Written English. In: Brindley, G. és Burrows, C. (szerk.): *Studies in immigrant English language assessment. Volume II*. National Centre for English Language Teaching and Research Macquarie University, Sydney. 95–187.
- Burrows, C. (2004): Washback in classroom-based assessment: A study of the washback effect in the Australian adult migrant English program. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 113–128.

- Caine, N. A. (2005): *EFL examination washback in Japan: Investigating the effects of oral assessment on teaching and learning. Doctoral Dissertation.* Language and Literacy Studies in Education, The University of Manchester, Manchester.
- Canale, M. (1983): On some dimensions of language proficiency. In: Oller, J. W. Jr.: *Issues in Language Testing Research.* Newbury House, Rowley, Mass. 333–342.
- Canale, M. és Swain, M. (1980): Theoretical bases of communicative approaches to language learning and testing. *Applied Linguistics*, 1. sz. 1–47.
- Carroll, J. B. (1963): A model of school learning. *Teachers College Record*, 64. 723–733.
- Celce-Murcia, M. (2008): Rethinking the role of communicative competence in language teaching. In: Solder, E. A. és Jordá, M. P. S. (szerk.): *Intercultural language use and language learning.* Springer Schiencie and Business Media, B.V. 41–57.
- Celce-Murcia, M., Dörnyei, Z. és Thurrell, S. (1995): Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6. 2. sz. 5–35.
- Chen, L. (2002): *Taiwanese junior high school English teachers' perceptions of the washback effect of the basic competence test in English. Doctoral Dissertation.* The Ohio State University, Ohio.
- Cheng, L. (1997): How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11. 1. sz. 38–54.
- Cheng, L. (1998): *The washback effect of public examination change on classroom teaching: an impact study of the 1996 Hong Kong Certificate of Education in English on the classroom teaching of English in Hong Kong secondary schools. Doctoral Dissertation.* University of Hong Kong, Hong Kong.
- Cheng, L. (1999): Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education*, 15. 3. sz. 253–271.
- Cheng, L. (2003): Looking closely at the impact of a public examination change on classroom teaching. *Journal of Classroom Interaction*, 38. 1. sz. 1–10.
- Cheng, L. (2004): The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods.* Lawrence Erlbaum Associates, Mahwah, NJ. 147–170.
- Cheng, L. (2005): *Changing language teaching through language testing: A washback study. Studies in language testing 21.* Cambridge University Press, Cambridge.
- Cheng, L. és Curtis, A. (2004): Washback or backwash: A review of the impact of testing on teaching and learning. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods.* Lawrence Erlbaum Associates, Mahwah, NJ. 3–17.
- Cheng, L., Watanabe, Y. és Curtis, A. (2004, szerk.): *Washback in language testing: Research contexts and methods.* Lawrence Erlbaum Associates, Mahwah, NJ.
- Chomsky, N. (1965): *Aspects of the theory of syntax.* Massachusetts Institute of Technology (MIT) Press, Cambridge.
- CITO (2008): Building a european bank of anchor items for foreign language skills. EBAFLS. General report. Cito B.V., Arnheim.
- Cumming, A., Kantor, R. és Powers, D. E. (2001): *Scoring TOEFL essays and TOEFL 2000 prototype writing tasks: An investigation into raters' decision making and development of a preliminary analytic framework.* Educational Testing Service, Princeton, NJ.
- Csapó Benő (1978): A „mastery learning” elmélete és gyakorlata. *Magyar Pedagógia*, 74. 1. sz. 60–73.
- Csapó Benő (1987): A kritériumorientált értékelés. *Magyar Pedagógia*, 87. 3. sz. 247–266.
- Csapó Benő (1996): Standard érettségi: lehetőségek és dilemmák. *Új Pedagógiai Szemle*, 46. 9. sz. 17–26.
- Csapó Benő (2001): A nyelvtanulást és a nyelvtudást befolyásoló tényezők. *Iskolakultúra*, 11. 8. sz. 25–35.
- Csapó, B. (2003): *Cognitive factors of the development of foreign language skills.* Paper presented at the 10th Biennial Conference of the European Association for Research on Learning and Instruction. Padova, Italy, August 26–30, 2003.
- Csapó Benő (2007): Hosszmetzeti felmérések iskolai kontextusban: az első átfogó magyar iskolai longitudinális kutatási program elméleti és módszertani keretei. *Magyar Pedagógia*, 109. 4. sz. 321–355.
- Csapó Benő (2008): A közoktatás második szakasza és az érettségi vizsga. In: Fazekas Károly, Köllő János és Varga Júlia (szerk.): *Zöld könyv a magyar közoktatás megújításáért.* Ecostat, Budapest. 71–93.
- Csapó Benő, Molnár Gyöngyvér és Kinyó László (2009): A magyar oktatási rendszer szelektivitása a nemzetközi összehasonlító vizsgálatok eredményeinek tükrében. *Iskolakultúra*, 19. 3–4. sz. 3–13.

- Csapó, B. és Nikolov, M. (2001): *Hungarian students' performances on english and german tests: Results of a large-scale assessment project*. Paper presented at ALTE European Year of Languages Conference: European Language Testing Issues in a Global Context. Barcelona, Spain, 5–7 July 2001.
- Csapó, B. és Nikolov, M. (2002): *The relationship between students' foreign language achievement and general thinking skills*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April 1–5, 2002.
- Csapó Benő és Nikolov Marianne (2008): A tanulók idegen-nyelv tudása az általános iskola és a középiskola végén. In: Csíkos Csaba (szerk.): *PÉK 2008 – VI. Pedagógiai Értékelési Konferencia: Program – Tartalmi összefoglalók*. SZTE, Neveléstudományi Doktori Iskola, Szeged. 26.
- Csapó, B. és Nikolov, M. (2009): The cognitive contribution to the development of proficiency in a foreign language. *Learning and Individual Differences*, **19**. 2. sz. 209–218.
- Csapó Benő, Székely László és Tóth Edit (2009): Az iskolai különbségek változása a longitudinális vizsgálat adatainak tükrében. In: Molnár Gyöngyvér és Kinyó László (szerk.): *PÉK 2009 – VII. Pedagógiai Értékelési Konferencia: Program – Tartalmi összefoglalók*. SZTE Neveléstudományi Intézet, Szeged. 44. o.
- Dávid Gergely (1992): Objektivitás, értékelés, standardizáció. *Iskolakultúra*, **2**. 1. sz. 26–29.
- Dávid Gergely (2002): A nyelvvizsgák akkreditációjának tanulságai. In: Kárpáti Eszter és Szűcs Tibor (szerk.): *Nyelvpedagógia*. Iskolakultúra könyvek 12, Pécs. 159–165.
- Dávid Gergely (2005): Nyelvvizsgaszintek validálása: lehetőségek és korlátok. In: Sillye Magdolna, Kurtán Zsuzsa, Sturcz Zoltán és Wiwczarowski, T. B. (szerk.): *PORTA LINGUA – 2005 Szakmai nyelvtudás – szaknyelvi kommunikáció cikkek, tanulmányok a hazai szaknyelvtanításról és -kutatásról*. Debreceni Egyetem, Agrártudományi Centrum, Center Print Nyomda, Debrecen. 279–295.
- Dávid Gergely (2006): *Az emelt szintű érettségi és az államilag elismert nyelvvizsgák a vizsgázói teljesítmények tükrében. Összegző tanulmány*. Nyelvvizsgát Akkreditáló Testület – PH Nyelvvizsgáztatási Akkreditációs Központ.
(http://www.euroexam.org/content/EuroExam/david_gergo/erettsegi_tanulmany.pdf, megtekintés: 2007. 06. 22.)
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. és McNamara, T. (1999): *Dictionary of language testing. Studies in Language Testing 7*. Cambridge University Press, Cambridge.
- DeAyala, R. J., Dodd, B. G. és Koch, W. R. (1991): Partial credit analysis of writing ability. *Educational and Psychological Measurement*, **51**. 1. sz. 103–114.
- Dévény Ágnes (2003): Az idegen nyelvi tesztelés etikai problémái. *Új Pedagógiai Szemle*, **53**. 7–8. sz. 134–145.
- Dévény Ágnes (2008): *Az idegen nyelvi közvetítés feladat helye, szerepe a kritériumfügő nyelvvizsgán. Doktori disszertáció*. Eötvös Lóránd Tudományegyetem, Pedagógiai és Pszichológiai Kar, Neveléstudományi Doktori Iskola, Budapest.
- Djigunovič, M. J., Nikolov Marianne és Ottó István (2006): Horvát és magyar nyelvű tanulók angol nyelvtudása nyolcadik osztályban. *Magyar Pedagógia*, **106**. 107–186.
- Eckes, T. (2003): Qualitätssicherung beim TestDaF: Konzepte, Methoden, Ergebnisse. *Fremdsprachen und Hochschule*, **69**. sz. 43–68.
- Eckes, T. (2004): Facetten des Sprachtestens: Strenge und Konsistenz in der Beurteilung sprachlicher Leistungen. In: Wolff, A., Ostermann, T. és Chlosta, C. (szerk.): *Integration durch Sprache. Materialien Deutsch als Fremdsprache*, **73**. FaDaF, Regensburg. 485–518.
- Eckes, T. (2005): Examining rater effects in TestDaF writing and speaking performance assessments: A many-facet Rasch analysis. *Language Assessment Quarterly*, **2**. 3. sz. 197–221.
- Eckes, T. (2008): Rater types in writing performance assessments: A classification approach to rater variability. *Language Testing*, **25**. 2. sz. 155–185.
- Eckes, T., Ellis, M., Kalnberzina, V., Pizorn, K., Springer, C., Szollás, K. és Tsagari, C. (2005): Progress and problems in reforming public language examinations in Europe: cameos from the Baltic States, Greece, Hungary, Poland, Slovenia, France and Germany. *Language Testing*, **22**. 3. sz. 355–377.
- Einhorn Ágnes (1998, szerk.): *Német nyelv. Mérés – értékelés – vizsga 5*. Országos Közoktatási Intézet, Budapest.
- Einhorn Ágnes (1999): Nyelvtudás és vizsga. *Educatio*, **8**. 3. sz. 543–556.
- Einhorn Ágnes (2004, szerk.): *Az érettségiről tanároknak 2005. Német nyelv*. Országos Közoktatási Intézet, Budapest.
(http://www.okm.gov.hu/letolt/kozokt/erettsegi2005/tanaroknak/nemet/nemet_nyelv.htm, megtekintés: 2006. 03. 17.)

- Einhorn Ágnes (2006a): A német nyelvi érettségi 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*. Országos Közoktatási Intézet, Budapest. 153–172.
- Einhorn Ágnes (2006b): A vizsgafeladat fejlesztésének folyamata és kritériumai. *Új Pedagógiai Szemle*, **56**. 1. sz. 67–74.
- Einhorn Ágnes (2007a): *A 2005. évi érettségi vizsga eredményeinek elemzése. Német nyelv*. Oktatókutató és Fejlesztő Intézet, Budapest. (<http://tamop311.ofi.hu/egyeb/2005-os-erettsegi/2005-evi-erettsegi-090617-2>, megtekintés: 2010. 01. 25.)
- Einhorn Ágnes (2007b): Az idegen nyelvi érettségi vizsga reformja. In: Vágó Irén (szerk.): *Fókuszban a nyelvtanulás*. Oktatókutató és Fejlesztő Intézet, Budapest. 73–105.
- Einhorn Ágnes (2008a): *A 2006. évi érettségi vizsga eredményeinek elemzése. Német nyelv*. Oktatókutató és Fejlesztő Intézet, Budapest. (<http://tamop311.ofi.hu/egyeb/2006-os-erettsegi/2006-evi-erettsegi-090617-2>, megtekintés: 2010. 01. 25.)
- Einhorn Ágnes (2008b): Fejlesztés vagy vizsgatréning? A fejlesztő- és mérőfeladatok különbségei. In: Bánkuti Zsuzsa és Lukács Judit (szerk.): *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés*. Oktatókutató és Fejlesztő Intézet, Budapest. 147–168.
- Einhorn Ágnes (2008c): Hogy vizsgázott az új érettségi németből? In: Bánkuti Zsuzsa és Lukács Judit (szerk.): *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés*. Oktatókutató és Fejlesztő Intézet, Budapest. 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*. Országos Közoktatási Intézet, Budapest. 127–137.
- Emelt szintű írásbeli érettségi vizsga. Német nyelv*. (2006): Oktatási Minisztérium, Budapest. (http://www.okm.gov.hu/letolt/okev/doc/erettsegi_2006/e_nemet_06maj_fl.pdf, megtekintés: 2010. 01. 25.)
- Engelhard, G. (1992): The measurement of writing ability with a many-facet Rasch model. *Applied Measurement in Education*, **5**. 3. sz. 171–191.
- Engelhard, G. és Myford, C. M. (2003): *Monitoring faculty consultant performance in the Advanced Placement English Literature and Composition Program with a many-faceted Rasch model*. College Entrance Examination Board, New York.
- English, F. W. és Steffy, B. E. (2001): *Deep curriculum alignment*. Scarecrow Press, Maryland.
- Érettségi 2009. május-június. Tapasztalatok* (2009): Oktatási Hivatal, Budapest. (http://www.oh.gov.hu/letolt/okev/doc/gyorsadatok_2009_tavaszi.ppt, megtekintés: 2010. 01. 25.)
- Fazekas Márta (2006): Standardok a nyelvi érettségis: régi-új érettségik és nyelvvizsgák. In: Kósa Barbara és Simon Mária (szerk.): *Új vizsga – új tudás? Az új érettségi hatása az iskolakezdéstől a záróvizsgáig*. Országos Közoktatási Intézet, Budapest. (http://www.oki.hu/oldal.php?tipus=cikk&kod=uj_vizsga_uj_tudas-6vita-fazekas, megtekintés: 2007. 04. 14.)
- Felmérés a 2008-ban Magyarországon érettségizettek idegennyelv-tudásáról* (2009): Oktatási és Kulturális Minisztérium, Budapest. (http://www.okm.gov.hu/letolt/kozokt/nyelvkutatas_090622.pdf, megtekintés: 2009. 09. 30.)
- Ferman, I. (2004): The washback of an EFL national oral matriculation. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 191–210.
- Figueras, N. (2005): Testing, testing, everywhere, and not a while to think. *ELT Journal*, **59**. 1. sz. 47–54.
- Fischer, G. H. (1981): On the existence and uniqueness of Maximum-likelihood estimates in the Rasch model. *Psychometrika*, **46**. 1. sz. 59–77.
- Frederiksen, J. R. és Collins, A. (1989): A systems approach to educational testing. *Educational Researcher*, **18**. 9. sz. 27–32
- Gates, S. (1995): Exploiting washback from standardized tests. In: Brown, J. D. és Yamashita, S. O. (szerk.): *Language testing in Japan*. Japan Association for Language Teaching, Tokyo. 101–106.
- Glaboniat, M. és Müller, M. (2006): Note „Sehr gut!“ – Aber in Bezug worauf? Referenzrahmen und Profile Deutsch in ihren Auswirkungen auf Prüfungen und Tests. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*, **34**. sz. 14–21.
- Glaboniat, M., Müller, M., Rusch, P., Schmitz, H. és Wertenschlag, L. (2002): *Profile Deutsch A1–B1. Version 1.0*. Langenscheidt, Berlin, München, Wien, Zürich és New York.
- Glaboniat, M., Müller, M., Rusch, P., Schmitz, H. és Wertenschlag, L. (2005): *Profile Deutsch A1–C2. Version 2.0*. Langenscheidt, Berlin, München, Wien, Zürich és New York.
- Gronlund, N. E. (1988): *How to construct achievement tests*. Englewood Cliffs, NJ: Prentice Hall.

- Grotjahn, R. (2000a): Testtheorie: Grundzüge und Anwendungen in der Praxis. In: Wolff A. és Tanzer H. (szerk.): *Sprache – Kultur – Politik. Beiträge der 27. Jahrestagung DaF 1999. Materialien Deutsch als Fremdsprache*. FaDaF, Regensburg. 304–341.
- Grotjahn, R. (2000b): Determinanten der Schwierigkeit von Leseverstehensaufgaben: Theoretische Grundlagen und Konsequenzen für die Entwicklung des TestDaF. In: Bolton, S. (szerk.): *TestDaF: Grundlagen für die Entwicklung eines neuen Sprachtests Beiträge aus einem Expertenseminar*. Goethe-Institut – Gilde Verlag, Köln – München. 7–55.
- Grotjahn, R. (2002): *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*. AKS-Verlag, Bochum.
- Hamp-Lyons, L. (1997): Washback, impact and validity: Ethical concerns. *Language Testing*, **14**. 3. sz. 295–303.
- Hamp-Lyons, L. (1998): Ethical test preparation practice: the case of the TOEFL. *TESOL Quarterly*, **32**. 2. sz. 329–337.
- Hayes, B. és Read, J. (2004): IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 97–111.
- Horváth Zsuzsanna és Lukács Judit (2006, szerk.): *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Országos Közoktatási Intézet, Budapest.
- Hughes, A. (1989): *Testing for language teachers*. Cambridge University Press, Cambridge.
- Hughes, A. (1993): *Backwash and TOEFL 2000*. Unpublished manuscript, University of Reading.
- Hymes, D. H. (1972): On communicative competence. In: Pride, J. B. és Holmes, J. (szerk.): *Sociolinguistics: Selected readings*. Penguin, Harmondsworth. 269–293.
- Imre Anna (2007): Nyelvoktatás, nyelvtanulás, nyelvtudás a középfokú oktatásban. In: Vágó Irén (szerk.): *Fókuszban a nyelvtanulás*. Oktatókutató és Fejlesztő Intézet, Budapest. 107–136.
- Józsa Krisztián és Nikolov Marianne (2005): Az angol és német nyelvi készségek fejlettségét befolyásoló tényezők. *Magyar Pedagógia*, **105**. 3. sz. 307–337.
- Kapitánffy Johanna (2001): Az idegen nyelvek oktatásának fejlesztése a közoktatásban. *Iskolakultúra*, **11**. 8. sz. 71–75.
- Katona Lucia (1995): A nyelvtudás fogalmának új értelmezése. *Iskolakultúra*, **5**. 1–2. sz. 67–75.
- Katona Lucia és Dörnyei Zoltán (1991): Az idegen nyelvtudás mérésének új módja a C-teszt. *Iskolakultúra*, **1**. 4. sz. 42–51.
- KER (2002): Közös Európai Referenciakeret: nyelvtanulás, nyelvtanítás, értékelés. OM – PTMIK, Budapest – Pilisborosjenő
- Kerentanterv (2000): Tantárgyi füzetek 2. Idegen nyelv. Oktatási Minisztérium, Budapest
- Kiss-Gulyás, J. (2001): Experiencing the exam content and material. In: Glover, G. (szerk.) *English language teaching in Hungary. Part III*. The British Council, Budapest. 40–58.
- Kontra József (2009): A parciális kredit modell egy alkalmazása. In: Pšenáková Ildikó, Mező Ferenc és Viczayová Ildikó (szerk.): *Képzés és gyakorlat II*. Konstantin Filozófus Egyetem, Nyitra. 99–108.
- Kormos, J. (1999): Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams. *Language Testing*, **16**. 2. sz. 163–188.
- Kósa Barbara és Simon Mária (2006, szerk.): *Új vizsga – új tudás? Az új érettségi hatása az iskolakezdéstől a záróvizsgáig*. Országos Közoktatási Intézet, Budapest.
- Középszintű írásbeli érettségi vizsga. Német nyelv. (2006): Oktatási Minisztérium, Budapest. (<http://www.okm.gov.hu/main.php?folderID=266&articleID=7293&ctag=articlelist&iid=1>, megtekintés: 2010. 01. 25.)
- Közoktatási Statisztikai Kiadvány (2008): Oktatási és Kulturális Minisztérium, Budapest. (http://db.okm.gov.hu/statisztika/ks07_fm/index.html, megtekintés: 2009. 06. 10.)
- Krekeler, C. (2002): Die Grammatik fehlt! Fehlt die Grammatik? Rückwirkungsmechanismen von TestDaF und DSH. *InfoDaF*, **29**. 5. sz. 441–458.
- Krekeler, C. (2005): *Grammatik und Fachbezug in Sprachtests für den Hochschulzugang*. Dissertation beim Fachbereich für Geisteswissenschaften der Universität Duisburg-Essen.
- Krumm, J-H. (2006): Müssen jetzt alle dasselbe können? Vor- und Nachteile der Globalisierungsprozesse im Sprachunterricht. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*, **34**. sz. 30–33.
- Kunnan, A. J. (2004): Test fairness. In: Milanovic, M. és Weir, C. (szerk.): *European language testing in a global context*. Cambridge University Press, New York. 27–48.
- Lado, R. (1961): *Language testing. The construction and use of foreign language tests*. Longman, London.
- Leitzel, T. C. és Vogler, D. E. (1994): *Curriculum alignment: Theory to practice*. ERIC Document ED371812. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.

- Lienert, G. A. és Raatz, U. (1994): *Testaufbau und Testanalyse*. Beltz, Psychologie Verlags Union, Weinheim.
- Linacre, J. M. (1994): *Many-faced Rasch Measurement*. MESA press, Chicago.
- Linacre, J. M. (2002): What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, **16**. 2. sz. 878.
- Lumley, T. (2002): Assessment criteria in a large-scale writing test: What do they really mean to the raters? *Language Testing*, **19**. 246–276.
- Lynch, B. K. (2001): Rethinking assessment from a critical perspective. *Language Testing*, **18**. 4. sz. 351–372.
- Madaus, G. F. (1988): The influence of testing on the curriculum. In: Tanner, L. N. (szerk.): *Critical issues in curriculum: Eighty-seventh yearbook of the National Society for the Study of Education*. University of Chicago Press, Chicago. 83–121.
- Major Éva (2007): *A 2005-ös érettségi vizsga eredményeinek elemzése – Angol nyelv*. Oktatókutatató és Fejlesztő Intézet, Budapest (<http://tamop311.ofi.hu/egyeb/2005-os-erettsegi/2005-evi-erettsegi-090617-8>, megtekintés: 2010. 01. 25.)
- Major Éva (2008): *A 2006. évi érettségi vizsga eredményeinek elemzése – Angol nyelv*. Oktatókutatató és Fejlesztő Intézet, Budapest. <http://tamop311.ofi.hu/egyeb/2006-os-erettsegi/2006-evi-erettsegi-090617-9>, megtekintés: 2010. 01. 25.)
- Malone, M. (2000): *Simulated oral proficiency interviews: Recent developments*. ERIC Document ED447729. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.
- Markee, N. (1993): The diffusion of innovation in language teaching. *Annual Review of Applied Linguistics*, **13**. 229–243.
- Masters, G. N. (1982): A Rasch model for partial credit scoring. *Psychometrika*, **47**. 2. sz. 149–174.
- McEwen, N. (1995): Educational accountability in Alberta. *Canadian Journal of Education*, **20**. 1. sz. 27–44.
- McNamara, T. (1996): *Measuring second language performance*. Longman, London.
- McNamara, T. (2000): *Language Testing*. Oxford University Press, Oxford.
- Messick, S. (1994): The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, **23**. 2. sz. 13–23.
- Messick, S. (1995): Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, **50**. 9. sz. 741–749.
- Messick, S. (1996): Validity and washback in language testing. *Language Testing*, **13**. 3. sz. 241–256.
- Milanovic, M. (1997, szerk.): *User's guide for examiners*. A Common European Framework of Reference for Language Learning and Teaching, Council for Cultural Cooperation, Council of Europe Press, Strasbourg.
- Miyata-Boddy, N. és Langham, C. S. (2000): *Communicative language testing – an attainable goal?* The British Council, Tokyo. 75–82.
- Molnár Gyöngyvér (2003): Az ismeretek alkalmazásának vizsgálata modern tesztelméleti eszközökkel. *Magyar Pedagógia*, **103**. 4. sz. 423–446.
- Molnár Gyöngyvér (2005): Az objektív mérés megvalósításának lehetősége: a Rasch-modell. *Iskolakultúra*, **15**. 3. sz. 71–80.
- Molnár Gyöngyvér (2006): A Rasch-modell alkalmazása a társadalomtudományi kutatásokban. *Iskolakultúra*, **16**. 12. sz. 99–113.
- Molnár Gyöngyvér (2008): A Rasch-modell kiterjesztése nem dichotóm adatok elemzése: a rangskálás és parciális kredit modell. *Iskolakultúra*, **18**. 1–2. sz. 66–77.
- Molnár Gyöngyvér és Józsa Krisztinán (2006): Az olvasási képesség értékelésének tesztelméleti megközelítései. In: Józsa Krisztinán (szerk.): *Az olvasási képesség fejlődése és fejlesztése*. Dinasztia Tankönyvkiadó, Budapest. 155–174.
- Morrow, K. (1981): Communicative language testing: Revolution or evolution. In: Alderson, J. C. és Huges A. *111-Issues in Language Testing* (ELT documents 111). The British Council, London. 9–25.
- Morrow, K. (1986): The evaluation of tests of communicative performance. In: Portal, M. (szerk.): *Innovations in Language Testing: Proceedings of the IUS/NFER conference*. NFER/Nelson, London. 1–13.
- Nagy József (1975): *A témazáró tesztek reliabilitása és validitása*. Acta Universitatis Szegediensis de A.J. Nominatae, Sectio Paedagogica et Psychologica, Series Specifica, Szeged.
- Nagy József (1979): *Köznevelés és rendszerszemlélet*. OOK, Veszprém.
- Nagy József (2002): *XXI. század és nevelés*. Osiris Kiadó, Budapest.
- Nemzeti alaptanterv (1995): Korona Kiadó, Budapest.

- Nemzeti alaptanterv (2003): Oktatási Minisztérium, Budapest.
- Nemzeti alaptanterv (2007): Oktatási Minisztérium, Budapest. (<http://www.okm.gov.hu/main.php?folderID=391>, megtekintés: 2009. 09. 30.)
- Neuner, G. és Hunfeld, H. (1993): *Methoden des Fremdsprachlichen Deutschunterrichts. Eine Einführung. Fernstudieneinheit 4*. Langenscheidt, Berlin, München, Wien, Zürich és New York.
- Nikolov Marianne (1999): Osztálytermi megfigyelés átlagos és hátrányos helyzetű középiskolai angol csoportokban. *Modern Nyelvoktatás*, 5. 4. sz. 9–31.
- Nikolov Marianne (2003): Angolul és németül tanuló diákok nyelvtanulási attitűdje és motivációja. *Iskolakultúra*, 13. 8. sz. 61–73.
- Nikolov Marianne (2007): A magyarországi nyelvoktatás-fejlesztési politika és annak gyakorlati megvalósulása a nemzetközi trendek tükrében. In: Vágó Irén (szerk.): *Fókuszban a nyelvtanulás*. Oktatókutatás és Fejlesztő Intézet, Budapest. 43–72.
- Nikolov, M. és Csapó, B. (2002): *Twelve-year-olds' attitudes towards classroom activities and their performances on tests of English and German as a foreign language*. Paper presented at the American Association of Applied Linguists Annual Conference, Salt Lake City, USA.
- Nikolov Marianne és Józsa Krisztián (2003): *Az idegen nyelvi készségek fejlettsége angol és német nyelvből a 6. és 10. évfolyamon a 2002/2003-as tanévben*. OKÉV, Budapest. (http://www.om.hu/letolt/okev/doc/orszmer2003/idegen_nyelv_beliv.pdf, megtekintés: 2010. 01. 23.)
- Nikolov, M. és Józsa, K. (2006): Relationships between language achievements in English and German and classroomrelated variables. In: Nikolov, M. és Horváth, J. (szerk.): *UPRT 2006: Empirical studies in English applied linguistics*. Lingua Franca Csoport, Pécs. 197–224.
- Nikolov Marianne, Ottó István és Öveges Enikő (2008): *Az idegennyelv-tanulás és tanítás helyzete és fejlesztésének lehetőségei a szakképző intézményekben*. Oktatási és Kulturális Minisztérium – Oktatásért Közalapítvány, Budapest. http://www.oktatasert.hu/files/idegennyelvi_kutatas_jelentes.pdf, megtekintés: 2010. 01. 23.)
- Nikolov Marianne, Pércsich Richárd és Szabó Gábor (2000): A puding próbája: Alapszintű angol feladatsorok bemérésének tapasztalatai. *Modern Nyelvoktatás*, 6. 4. sz. 3–28.
- Nikolov, M., Tóth E. és Víg, T. (2009): *The impact of background factors on Hungarian students' performances on English and German reading comprehension tests*. Paper presented at the 13th Biennial Conference of the European Association for Research on Learning and Instruction. Amsterdam, The Netherlands.
- Niska, H. (1997): *Testing community interpreters: a theory, a model and a plea for research*. Stockholm University, Sweden. (<http://lisa.tolk.su.se/00TEST.HTM>, megtekintés: 2006. május 8.)
- Noble, A. J. és Smith, M. L. (1994): *Measurement-driven reform: Research on policy, practice, repercussion*. CSE Technical Report 381, Tempe, Arizona State University, Arizona.
- North, B., Avermaet, P. V., Figueras, N., Takala, S. és Verhelst, N. (2003): *Nyelvvizsgák szintillesztése a Közös európai referenciakerethez. Kézikönyv. Előkészítő, kísérleti változat*. Nyelvvizsgát Akkreditáló Testület – PH Nyelvvizsgázati Akkreditációs Központ, Budapest.
- OECD (2005): The Rasch Model. In: *OECD: PISA 2003 Data Analysis Manual*. OECD, Paris. 53–70.
- Oller, J. W. (1979): *Language tests at school*. Longman, London.
- Park, T. (2004): An investigation of an ESL placement test of writing using many-facet Rasch measurement. *Working Papers in TESOL & Applied Linguistics*, 4. 1. sz. 1–21.
- Perlmann-Balme, M. (2001): Formen und Funktionen von Leistungsmessung und -kontrolle. In: Helbig, G., Götze, L., Henrici, G. és Krumm, H.-J. (szerk.): *Deutsch als Fremdsprache: ein internationales Handbuch*. de Gruyter, Berlin és New York. 994–1006.
- Perlmann-Balme, M. (2006): „Das kann ich schon!” Kompetenzen testen, prüfen, zertifizieren. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*, 34. sz. 5–13.
- Petneki Katalin (1998): A tantervek és az érettségi vizsga. In: Einhorn Ágnes (szerk.): *Német nyelv. Mérés – értékelés – vizsga 5*. Országos Közoktatási Intézet, Budapest. 57–85.
- Petneki Katalin (2007): *Az idegen nyelvek oktatása Magyarországon az ezredfordulón*. JATE-Pressz, Szeged.
- Petneki, K. (2009): Problematik der Niveaustufen im neuen Abitur für Deutsch als Fremdsprache in Ungarn. In: Wolff, A., Hunstiger, A. és Koreik, U. (szerk.): *Chance Deutsch: Schule – Studium – Arbeitswelt. Beiträge der 34. Jahrestagung DaF 2006*. FaDaF, Regensburg, 167–178.
- Popham, W. J. (1987): The merits of measurement-driven instruction. *Phi Delta Kappa*, 68. 679–682.
- Prodromou, L. (1995): The backwash effect: from testing to teaching. *ELT Journal*, 49. 1. sz. 13–25.
- Qi, L. (2004): Has a high-stakes test produced the intended changes? In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 171–190.

- Qi, L. (2005): Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, **22**. 2. sz. 142–173.
- Qi, L. (2007): Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education*, **14**. 1. sz. 51–74.
- Quetz, J. (2001): Der Gemeinsame Europäische Referenzrahmen. *Info DaF*, **28**. 6. sz. 553–563.
- Quetz, J. (2003): A1–A2–B1–B2–C1–C2. Der Gemeinsame Europäische Referenzrahmen. *Deutsch als Fremdsprache. Zeitschrift zur Theorie und Praxis des Deutschunterrichts für Ausländer*, 1. sz. 42–48.
- Rasch, G. (1960): *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research, Copenhagen.
- Rea-Dickins, P. és Scott, C. (2007, szerk.): Special issue: Investigating washback in language testing and assessment. *Assessment in Education*, **14**. 1. sz.
- Richardson, V. (1996): The role of attitudes and beliefs in learning to teach. In: Sikula, J. (szerk.): *Handbook of Research on Teacher Education*. Second Edition, MacMillan, New York, 102–119.
- Robinson, P. (1996): Task based testing, performance-referencing and ESL program development. In: Robinson P. (szerk.): *Task complexity and second language syllabus design: Data-based studies and speculations*. University of Queensland Working Papers in Language and Linguistics (Special Issue). CLTR, Brisbane. 95–116.
- Saif, S. (1999): *Theoretical and empirical considerations in investigating washback. A study of ESL/EFL learners. Doctoral Dissertation*. University of Victoria.
- Saif, S. (2006): Aiming for positive washback: A case study of international teaching assistants. *Language Testing*, **23**. 1. sz. 1–34.
- Saville, N. és Hawkey, R. (2004): The IELTS impact study: Investigating washback on teaching materials. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 73–96.
- Schifko, M. (2001): Prüfungen, Zertifikate, Abschlüsse als Planungskategorien für den Unterricht. In: Helbig, G. (szerk.): *Deutsch als Fremdsprache: ein internationales Handbuch. Handbücher zur Sprach- und Kommunikationswissenschaft*, 19. de Gruyter, Berlin és New York. 827–834.
- Schoonen, R. (2005): Generalizability of writing scores: An application of structural equation modeling. *Language Testing*, **22**. 1. sz. 1–30.
- Seiffert, C. (2003): Was ist und was will „Profile deutsch?“ Deutsch als Fremdsprache. *Zeitschrift zur Theorie und Praxis des Deutschunterrichts für Ausländer*, 4. sz. 238–240.
- Shohamy, E. (1999): Language testing: impact. In: Spolsky, B. (szerk.): *Concise encyclopaedia of educational linguistics*. Elsevier, Amsterdam. 711–714.
- Shohamy, E. (2001): *The power of tests: a critical perspective of the uses of language tests*. Longman, Harlow.
- Shohamy, E., Donitsa-Schmit, S. és Ferman, I. (1996): Test impact revisited: Washback effect over time. *Language Testing*, **13**. 3. sz. 298–317.
- Sigott, G. (1996): *Quantifying language ability*. ERIC Document ED394323. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.
- Skehan, P. (1998): *A cognitive approach to language learning*. Oxford University Press, Oxford.
- Smith, M. L. (1991): Put to the test: the effects of external testing on teachers. *Educational Researcher*, **20**. 5. sz. 8–11.
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C. és Cherland, M. (1990): *The role of testing in elementary schools. CSE ethnical Report 321*. CRESST. University of California.
- Spratt, M. (2005): Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, **9**. 1. sz. 5–29.
- Stecher, B., Chun, T. és Barron, S. (2004): The effects of assessment-driven reform on the teaching of writing in Washington state. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 53–71.
- Stevenson, D. K. (1985): Pop validity and performance testing. In: Lee, Y. P., Fok, A. C. Y. Y., Lord, R. és Low, G.: *New directions in language testing. Papers presented at the International Symposium on Language Testing, Hong Kong*. Pergamon Press, Oxford. 111–118.
- Szabó, G. (2006): Anchors aweigh! An analysis of the impact of anchor item's number and difficulty range on item difficulty calibrations. In: Nikolov, M. és Horváth, J. (szerk.): *UPRT 2006: Empirical studies in English applied linguistics*. Lingua Franca Csoport, Pécs. 249–262.
- Szabó, G. (2008): *Applying item response theory in language test item bank building*. Peter Lang, München.
- Szabó Gábor (2009): Ponthatár vagy képességszint? Az IRT-alapú teljesítményszint-minimumok megállapításának gyakorlati kérdései a nyelvtudásmérésében. *Modern Nyelvoktatás*, **15**. 4. sz. 3–13.

- Szénich Alexandra (2007): A felsőfokú továbbtanuláshoz szükséges németnyelv-tudás. *Iskolakultúra*, **17**. 8–10. sz. 49–59.
- Takala, S. (1998): *Language testing: Recent developments and persistent dilemmas*. ERIC Document ED460636. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.
- Tannenbaum, R. J. és Wylie, E. C. (2005): *Mapping English language proficiency test scores onto the Common European Framework*. Educational Testing Services, Princeton, NJ. (<http://ets.org/Media/Research/pdf/RR-05-18.pdf>, megtekintés: 2010. 01. 23.)
- Tannenbaum, R. J. és Wylie, E. C. (2008): *Linking English-language test scores onto the Common European Framework of Reference: An application of standard-setting methodology*. Educational Testing Services, Princeton, NJ. (http://www.ea.etsglobal.org/fileadmin/free_resources/ETS_Global_master/CEFRL_complete_study.pdf, megtekintés: 2010. 01. 23.)
- Taylor, L. (2005): Washback and impact. *ELT Journal*, **59**. 2. sz. 154–155.
- Tsagari, D. (2007): *Review of washback in language testing: How has been done? What more needs doing?* Washington DC: ERIC Clearinghouse on Assessment and Evaluation, ED497709
- Tsagari, D. (2009): Revisiting the concept of washback: investigating FCE in Greek language schools. *Cambridge ESOL: Research Notes*, **35**. 5–10.
- Tschirner, E. (2001): Die Evaluation fremdsprachlicher mündlicher Handlungskompetenz: Ein Problem-aufriß. In: Grotjahn, R. (szerk.): *Fremdsprachen Lehren und Lernen*, 30. Themenschwerpunkt Leistungsmessung und Leistungsevaluation. Narr, Tübingen. 87–115.
- Turner, C. E. (2001): The need for impact studies of L2 performance testing and rating: Identifying areas of potential consequences at all levels of testing cycle. In: Elder, C., Brown, A., Iwashita, N., Grove, E., Hill, K., Lumley, T., McNamara, T. és O'Loughlin, K. (szerk.): *Experimenting with uncertainty: essays in honour of Alan Davies*. Cambridge University Press, Cambridge. 138–149.
- University of Cambridge ESOL Exams. (<http://www.cambridgeesol.org/exams/index.htm>, megtekintés: 2006. 08. 20.)
- Vágó Irén (2007): Nyelvtanulási utak Magyarországon. In: Vágó Irén (szerk.): *Fókuszban a nyelvtanulás*. Oktatókutató és Fejlesztő Intézet, Budapest. 137–174.
- Vámos Ágnes (2007): Kétszintű érettségi vizsga a két tanítási nyelvű középiskolákban. *Új Pedagógiai Szemle*, **57**. 3–4. sz. 104–113.
- Verhelst, N. (2004): Az item-válasz elmélet (IRT). In: Banerjee, J., Kaftandjieva, F., Takala, S. és Verhelst, N.: *Szintillesztési módszertani segédlet. A nyelvvizsgák illesztése a Közös európai referenciakerethez című kézikönyv előzetes, kísérleti verziójához*. Nyelvvizsgát Akkreditáló Testület – PH Nyelvvizsgáztatási Akkreditációs Központ, Budapest. 151–201.
- Vidakovic, I. és Galaczi, E. (2009): ILEC Speaking: revising assessment criteria and scales. *Cambridge ESOL: Research Notes*, **35**. 29–34.
- Vidakovich Tibor (2008): *Kompetenciaértékelési rendszerek – értékelésméleti szempontból*. Plenáris előadás a VI. Pedagógiai Értékelési Konferencián, Szeged.
- Vidakovich Tibor és Vígh Tibor (2008): A kompetenciaértékelés minőségi mutatói. Tematikus előadás. In: Perjés István és Ollé János (szerk.): *VIII. Országos Neveléstudományi Konferencia: Hatékony Tudomány, Pedagógiai Kultúra, Sikeres Iskola. Program és Összefoglalók*. Magyar Tudományos Akadémia, Pedagógiai Bizottság, Budapest. 291.
- Vígh, T. (2007): Bewertung von Schreibfertigkeit im Abitur für Deutsch als Fremdsprache. *Deutschunterricht für Ungarn*, **22**. 1–2. sz. 81–95.
- Vígh Tibor (2008): Egy IRT-alapú nyelvi feladatbank létrehozásának módszertani kérdései: A német érettségi vizsgafeladatok elemzésének eredményei. *Magyar Pedagógia*, **108**. 1. sz. 29–51.
- Vígh Tibor és Tóth Edit (2010, megj. alatt): A különböző háttérváltozók hatása a tanulók angolnyelvtudásának szintjére az általános iskola és a középiskola végén. In: Vajda Zoltán (szerk.): *Bölcsész-műhely 2009*. JATEPress, Szeged.
- Wainer, H. és Messick, S. (1983, szerk.): *Principles of modern psychological measurement*. Lawrence Erlbaum Associates, Hillsdale, N.J.
- Wall, D. (1996): Introducing new tests into traditional systems: Insights from general education and from innovation theory. *Language Testing*, **13**. 3. sz. 334–354.
- Wall, D. (1997): Impact and washback in language testing. In: Clapham, C. és Corson, D. (szerk.): *Testing and assessment: the Kluwer encyclopaedia of language in education. Volume 7*. Kluwer Academic, Netherlands. 291–302.
- Wall, D. (2000): The impact of high-stakes testing on teaching and learning: can this be predicted or controlled? *System*, **28**. 4. sz. 499–509.
- Wall, D. (2005): *The impact of high-stakes examinations on classroom teaching. Studies in Language Testing 22*. Cambridge University Press, Cambridge.

- Wall, D. és Alderson, J. C. (1993): Examining washback: The Sri Lankan impact study. *Language Testing*, **10**. 1. sz. 41–69.
- Watanabe, Y. (1996a): Investigating problems of washback in Japanese EFL classrooms: problems of methodology. In: Wigglesworth, G. és Elder, C. (szerk.): *The language testing cycle: from inception to washback. Australian review of applied linguistics series, 13*. ANU Printing Services, Canberra. 208–239.
- Watanabe, Y. (1996b): Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, **13**. 3. sz. 318–333.
- Watanabe, Y. (2004a): Methodology in washback studies. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 19–36.
- Watanabe, Y. (2004b): Teacher factors mediating washback. In: Cheng, L., Watanabe, Y. és Curtis, A. (szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ. 129–146.
- Weigle, S. C. (1998): Using FACETS to model rater training effects. *Language Testing*, **15**. 2. sz. 263–287.
- Wiseman, S. (1961): The efficiency of examinations. In: Wiseman, S. (szerk.): *Examinations and English education*. Manchester University Press, Manchester.
- Wolfe, E. W. (1997): The relationship between essay reading style and scoring proficiency in a psychometric scoring system. *Assessing Writing*, **4**. 1. sz. 83–106.
- Woods, D. (1996): *Teacher cognition in language teaching: Beliefs, decisionmaking and classroom practice*. Cambridge University Press, Cambridge.
- Wu, M. (2006a): *PISA Training Workshop: Application of Item Response Theory (IRT) to PISA (ConQuest)*. Hong Kong PISA Centre, Hong Kong.
- Wu, M. (2006b): *How Well Do the Data Fit the Model?* Kézirat.
- Wu, M., Adams, R. J. és Wilson, M. R. (1998): *ACER ConQuest. Generalised Item Response Modelling Software*. ACER Press, Australia.
- Young, R. F. (2000): *Interactional competence: Challenges for validity*. ERIC Documents ED444361. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.
- Zeidler, B. (2009): Niveaustufen vergleichbar machen. In: Wolff, A., Hunstiger, A. és Koreik, U. (szerk.): *Chance Deutsch: Schule – Studium – Arbeitswelt. Beiträge der 34. Jahrestagung DaF 2006*. FaDaF, Regensburg. 197–216.
- Zertifikate und Prüfungen für Deutsch als Fremdsprache (2006): *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*, 34. sz. 45.

ÁBRÁK JEGYZÉKE

1. ábra	A kommunikatív kompetencia modellje (<i>Canale</i> , 1983 alapján: <i>Bárdos</i> , 2004. 150. o.)	10
2. ábra	A nyelvi kompetencia szerkezete <i>Bachman</i> (1990. 87. o.) modelljében.....	11
3. ábra	A nyelvtudás funkcionális-dinamikus modellje (<i>Celce-Murcia</i> , <i>Dörnyei</i> és <i>Thurrell</i> , 1995. 10. o., idézi <i>Celce-Murcia</i> , 2008. 44. o.)	13
4. ábra	A kommunikatív kompetencia funkcionális-dinamikus modellje (<i>Celce-Murcia</i> , 2008. 45. o.).....	14
5. ábra	A visszahatás alapmodellje (<i>Bailey</i> , 1996. 264. o.).....	37
6. ábra	A visszahatás konceptuális modellje (<i>Saif</i> , 1999. 69. o.)	38
7. a ábra	A tradicionális visszahatás-elmélet: az inger válasz modell.....	40
7. b ábra	Az 1990-es évek visszahatás felfogása: a „fekete doboz” modell.....	40
7. c ábra	A visszahatás <i>Burrows</i> (2004. 126. o.) által javasolt értelmezése: az innovációs modell	40
8. ábra	A teszthatást szintetizáló koncepcionális modell (<i>Tsagari</i> , 2009. 8. o.)	42
9. ábra	Az idegennyelv-tudás vizsgálat mintáinak longitudinális kapcsolata (<i>Csapó</i> , 2007. 341. o.)...	49
10. ábra	Az idegennyelv-tudás szintjét meghatározó tényezők rendszere	50
11. ábra	A visszahatás működését szintetizáló modell	67
12. ábra	A hallott szöveg értése feladatsor 15. itemének jelleggörbéje	108
13. ábra	Az olvasott és hallott szöveg értése, és a nyelvhelyesség feladatlap minta- és itemtérképe ...	111
14. ábra	A nyelvhelyesség feladatlap 3. és 15. itemének jelleggörbéje	112
15. ábra	A nyelvhelyesség feladatlap 18. és 20. itemének jelleggörbéje	114
16. a ábra	A 2005. évi középszintű német érettségi nyelvhelyesség feladatlap minta- és itemtérképe....	115
16. b ábra	A 2006. évi középszintű német érettségi nyelvhelyesség feladatlap minta- és itemtérképe ...	115
16. c ábra	A 2005. évi emelt szintű német érettségi nyelvhelyesség feladatlap minta- és itemtérképe ...	115
17. ábra	Az értékelési szempontok összefüggése	118
18. ábra	Az első feladat szókincs, kifejezőmód thurstoni küszöbértéke és modellilleszkedése	120
19. ábra	Az emelt szintű íráskészség feladatlap minta- és skálapont-térképe	123
20. ábra	A második feladat nyelvtan, helyesírás értékelési szempont karakterisztikus görbéje	125
21. ábra	A második feladat tartalom értékelési szempont karakterisztikus görbéje.....	126
22. ábra	A középszintű íráskészség feladatlap minta- és skálapont-térképe	130

TÁBLÁZATOK JEGYZÉKE

1. táblázat. Az autentikus nyelvhasználat jellemzői és az abból levezethető célok (<i>Morrow</i> , 1981 alapján).....	9
2. táblázat. <i>Bachman</i> és <i>Palmer</i> (1996) kommunikatív nyelvi képességek modellje	12
3. táblázat. A kommunikatív tesztelés kritériumai a kompetenciaértékelés rendszerében.....	18
4. táblázat. A kommunikatív tesztelés kritériumai és jellemzői.....	19
5. táblázat. A kompetencia és a performancia tesztek összehasonlítása (<i>McNamara</i> , 2000; <i>Grotjahn</i> , 2000a; <i>Krekeler</i> , 2005 alapján).....	24
6. táblázat. Közös Referenciaszintek: általános szintrendszer (<i>KER</i> , 2002. 33. o.)	27
7. táblázat. Különböző vizsgarendszerek szintjei (<i>Akkreditációs Kézikönyv</i> , 2007; <i>Bárdos</i> , 2005. 10. o.; <i>University of Cambridge ESOL Exams; Zertifikate und Prüfungen für DaF</i> , 2006. 45. o.)..	29
8. táblázat. A megelőző és az utólagos tantervi összehangolás jellemzői (<i>English</i> és <i>Steffy</i> , 2001. idézi <i>Chen</i> , 2002. 31. o. alapján).....	35
9. táblázat. A teszthatás mértékét és típusát befolyásoló tényezők (<i>Spratt</i> , 2005 alapján)	46
10. táblázat. A nyelvi készségek fejlettségét meghatározó néhány háttérváltozó hatása, regresszióanalízis, r ² (%).....	52
11. táblázat. A szintek közötti különbségek az íráskészség feladatokban a létrehozandó szövegfajta, szövegössz és az értékelési kritériumok szerint.....	56
12. táblázat. A 2005. és 2006. május-júniusi érettségi adatai angolból és németből	61
13. táblázat. A 2007., 2008. és 2009. május-júniusi érettségi vizsga adatai angol és német nyelvből.....	62
14. táblázat. A közép- és emelt szintű vizsga eredményei 2005-ben és 2006-ban készségenként angolból és németből (%)	63
15. táblázat. A 2008. évi május-júniusi vizsgaidőszak szintillesztési eljárásának eredménye angolból	64
16. táblázat. A 2008. évi május-júniusi vizsgaidőszak szintillesztési eljárásának eredménye németből ...	65
17. táblázat. A feladatlapok átlaga, szórása és Cronbach- α -ja (N=1092)	79
18. táblázat. A minta eloszlása a teljesítmény szerint (N=736)	79
19. táblázat. A feladatok és a feladatlap átlaga, szórása, Cronbach- α -ja a nyerspontok alapján.....	80
20. táblázat. A mintába került nyelvtanárok iskolatípus szerinti eloszlása (N=160)	81
21. táblázat. A csökkentett elemszámú minta jellemzői (N=149).....	81
22. táblázat. A minta eloszlása településtípus szerint (N=149).....	82
23. táblázat. A minta eloszlása a tanári pályán eltöltött évek száma szerint (N=149)	82
24. táblázat. A továbbképzések megoszlása időtartam, gyakoriság és hely szerint (%)	83
25. táblázat. A kérdőív fejlesztési folyamata	83
26. táblázat. A tanári kérdőív változórendszere	85
27. táblázat. „A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról” kérdőív-rész létrehozásának szempontjai (<i>Petneki</i> , 2007. 121-122. o. alapján)	86
28. táblázat. „A tanárok nézetei az idegen nyelvi tesztelés kritériumairól” kérdőív-rész létrehozásának szempontjai a kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek vizsgálatához	87
29. táblázat. A tanárok vizsgával kapcsolatos nézeteivel és attitűdjeivel végzett faktoranalízis eredménye	89
30. táblázat. „A tanárok nézetei az oktatási folyamatot befolyásoló tényezőkről” kérdőív-részben lévő változókkal végzett faktoranalízis eredménye	90
31. táblázat. „A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról” kérdőív-rész változóival végzett faktoranalízis eredménye.....	91
32. táblázat. „A tanárok nézetei az idegen nyelvi tesztelés kritériumairól” kérdőív-részben lévő változókkal végzett faktoranalízis eredménye	92
33. táblázat. „A tanárok nézetei a vizsgafeladatok oktatási folyamatban való felhasználásának céljairól” kérdőív-rész változóival végzett faktoranalízis eredménye	93
34. táblázat. A teljes kérdőív-re végzett faktoranalízis eredménye	94
35. táblázat. A kérdőív egyes részeihez tartozó itemek száma és megbízhatósága.....	95
36. táblázat. A középszinten vizsgázók száma és aránya a vizsgafajták szerint angolból és németből	97

37. táblázat. Az emelt szinten vizsgázók száma és aránya a vizsgafajták szerint angolból és németből ...	97
38. táblázat. A közép- és emelt szinten vizsgázók százalékos eloszlása a képzéstípusok szerint angolból és németből	98
39. táblázat. A közép- és emelt szinten vizsgázók százalékos eloszlása az iskolatípusok szerint angolból és németből.....	99
40. táblázat. A közép- és emelt szinten vizsgázók százalékos eloszlása régiók szerint angolból és németből	99
41. táblázat. A középszintű vizsga eredményei készségenként és a teljes vizsgán angolból (%)	100
42. táblázat. A középszintű vizsga eredményei készségenként és a teljes vizsgán németből (%)	101
43. táblázat. Az iskolatípusok közötti különbségek hatása a középszintű vizsgán elért teljesítményekre angolból és németből az eta-négyzet (%) alapján.....	102
44. táblázat. A középszintű vizsga készségenkénti különbségek mértéke angolból és németből a felnőttoktatásban tanulók szerint a t-értékek alapján.....	103
45. táblázat. A nemek közötti különbségek hatása a középszintű vizsgán elért teljesítményekre angolból és németből az eta-négyzet (%) alapján.....	104
46. táblázat. Az emelt szintű vizsga eredményei készségenként és a teljes vizsgán angolból (%)	105
47. táblázat. Az emelt szintű vizsga eredményei készségenként és a teljes vizsgán németből (%)	105
48. táblázat. Az iskolatípusok és nemek közötti különbségek hatása az emelt szintű vizsgán elért teljesítményekre angolból és németből az eta-négyzet (%) alapján.....	106
49. táblázat. A feladatlapok nehézségi indexének jellemzői.....	109
50. táblázat. A nem illeszkedő itemek paraméterének jellemzői	113
51. táblázat. Az első feladat összpontszámra mint függő változóra számított regresszióanalízis eredménye	118
52. táblázat. A második feladat összpontszámára mint függő változóra számított regresszióanalízis eredménye	118
53. táblázat. Az értékelési szempontok nehézségi indexe és modellilleszkedése	119
54. táblázat. Az értékelési szempontok thurstoni küszöbértéke (γ) és modellilleszkedése (outfit) az első feladatban	121
55. táblázat. Az értékelési szempontok thurstoni küszöbértéke (γ) és modellilleszkedése (outfit) a második feladatban	121
56. táblázat. Az értékelési szempontokhoz tartozó skálák minőségi mutatói	125
57. táblázat. Az értékelési szempontok nehézségi indexe és modellilleszkedése	128
58. táblázat. A tanárok nézetei a vizsga jelentőségéről.....	133
59. táblázat. A vizsgáztatási tapasztalat szerinti különbségek a tanárok vizsgával kapcsolatos nézeteiben.....	134
60. táblázat. A vizsga következményeivel kapcsolatos attitűdök	135
61. táblázat. A tankönyvváltás szerinti különbségek a tanórára való felkészüléssel kapcsolatos attitűdökben	136
62. táblázat. A nyelvtanítási gyakorlatot befolyásoló tényezők jellemzői	138
63. táblázat. A nyelvoktatási gyakorlat jellemzői és az érettségi hatása	139
64. táblázat. A környezetre és a tanárokra vonatkozó tényezők szerinti különbségek a nyelvoktatási gyakorlatot jellemző faktorok megítélésében	141
65. táblázat. Az érettségire felkészítő csoport szerinti különbségek a vizsga nyelvoktatási folyamatra gyakorolt hatásában a tanárok nézetei alapján.....	142
66. táblázat. A tanított évek száma szerinti különbségek az érettségi nyelvoktatási folyamatra gyakorolt hatásában a tanárok nézetei alapján.....	143
67. táblázat. A vizsgáztatói tapasztalat szerinti különbségek az érettségi nyelvoktatási folyamatra gyakorolt hatásában a tanárok nézetei alapján.....	144
68. táblázat. A tanárok nézetei az idegen nyelvi tesztelés kritériumairól	146
69. táblázat. A kommunikatív tesztelés jóságmutatóinak fontosságáról kialakított nézetek mint függő változóra számított regresszióanalízis eredménye	147
70. táblázat. A tanárok nézetei a vizsgafeladatok oktatási folyamatban való alkalmazásának céljairól ..	147
71. táblázat. Regressziós modellek a vizsga nyelvoktatási folyamatra gyakorolt hatásával kapcsolatos nézetekre mint függő változókra	149

MELLÉKLETEK JEGYZÉKE

1. melléklet. A korábbi és a jelenlegi nyelvi érettségi vizsga összehasonlítása	173
2. melléklet. A szövegalkotás értékelési szemponthoz tartozó kétféle skálaszerkezet (<i>Emelt szint. Javítási-és értékelési útmutató, 2006. 17. és 20. o.</i>).....	174
3. melléklet. A kérdőívhez adaptált források és eredmények	175
4. melléklet. Az elektronikus változatú tanári kérdőív.....	177
5. melléklet. A felkérő levél szövege	186
6. melléklet. A tájékoztató levél szövege.....	187
7. melléklet. A megköszönő levél szövege	188
8. melléklet. „A tanárok nézetei az érettségi és a nyelvoktatási gyakorlat közti viszonyról” kérdőív-rész módosított változóival végzett faktoranalízis eredménye	189
9. melléklet. A teljes kérdőív-re végzett faktoranalízis eredménye.....	190
10. melléklet. A teljesítmények átlaga, szórása iskolatípusok szerinti bontásban a középszintű vizsgán angolból és németből.....	193
11. melléklet. A 2008. évi május-júniusi vizsgaidőszak szintillesztési eljárásának eredménye közép- és emelt szinten iskola- és képzéstípusok szerint	194
12. melléklet. A teljesítmények átlaga, szórása képzéstípusok szerinti bontásban a középszintű vizsgán angolból és németből.....	195
13. melléklet. A teljesítmények átlaga, szórása nemek szerinti bontásban a középszintű vizsgán angolból és németből	196
14. melléklet. A teljesítmények átlaga, szórása iskolatípusok szerinti bontásban az emelt szintű vizsgán angolból és németből.....	197
15. melléklet. A teljesítmények átlaga, szórása nemek szerinti bontásban az emelt szintű vizsgán angolból és németből	198
16. melléklet. 2006. évi középszintű német érettségi olvasott szöveg értése feladatlap itemeinek jellemzői.....	199
17. melléklet. 2006. évi középszintű német érettségi hallott szöveg értése feladatlap itemeinek jellemzői.....	200
18. melléklet. 2006. évi középszintű német érettségi nyelvhelyesség feladatlap itemeinek jellemzői	201
19. melléklet. A nyelvi komponensekhez tartozó értékelési szempontok karakterisztikus görbéi	202
20. melléklet. A szövegalkotás értékelési szempontok karakterisztikus görbéi.....	203
21. melléklet. A kommunikatív komponensekhez tartozó értékelési szempontok karakterisztikus görbéi	204
22. melléklet. A teljes kérdőív faktoranalízise során létrejött faktorok korrelációs mátrixa.....	205
23. melléklet. Az iskolai gyakorlatban alkalmazott tankönyvek és feladatgyűjtemények gyakorisága...	206

MELLÉKLETEK

1. melléklet. A korábbi és a jelenlegi nyelvi érettségi vizsga összehasonlítása

	<i>A korábbi idegen nyelvi érettségi vizsga</i>	<i>Az új típusú idegen nyelvi érettségi vizsga</i>
<i>Általános jellemzők</i>	<ul style="list-style-type: none"> – A vizsga egy szintből állt. – A vizsga eltért a gimnáziumok és szakközépiskolában tanulók számára.* – A vizsga kétnyelvű (a fordításnak központi szerepe volt). – A vizsga nem standardizált. – A vizsga elsősorban a grammatikai és lexikai ismereteket mérte. 	<ul style="list-style-type: none"> – A vizsga két szintből áll. – A vizsga egységes a gimnáziumok és szakközépiskolák számára. – A vizsga egynyelvű. – A vizsga törekszik a standardizációra. – A vizsga a kommunikatív nyelvhasználatot méri.
<i>A mért készségek jellemzői</i>	<ul style="list-style-type: none"> – A receptív készségek mérése háttérbe szorult: a vizsga az olvasott szöveg értését a nyelvtani feladatokhoz képest aránytalanul, a hallott szöveg értését egyáltalán nem mérte. – Az íráskészség mérése különböző súlyllyal jelent meg a vizsgákon (többnyire tisztázatlan célú fogalmazás és fordítás formájában). A beszéd-készség mérése iskolánként eltérő módon történt. 	<ul style="list-style-type: none"> – A receptív készségek mérése egyenlő arányban történik a vizsgán. A hallott szöveg értését külön vizsgarészben mérik. – A produktív készségek mérése a kommunikatív keret és irányító szempontok megadásával történik. A beszéd-készség mérése és értékelése útmutatók segítségével történik.
<i>A teszt-feladatok jellemzői</i>	<ul style="list-style-type: none"> – A feladatok nem kontextualizáltak: az írásbeli részben a mondatokat a teszt-feladat írója készíti el. – A feladatok sematikusak és többnyire zártak. – A feladatokat a szóbeli vizsgarészben a tanár szabadon készítette útmutató nélkül. 	<ul style="list-style-type: none"> – A feladatok autentikus szövegekre és szituációra épülnek. – A feladatok változatosak, nem csak zártak, összetett tevékenységet várnak. – A feladatokat a középszintű vizsga szóbeli részén a tanárok készítik a vizsgaleírás alapján; emelt szinten központi feladatok készülnek.
<i>Az értékelési rendszer jellemzői</i>	<ul style="list-style-type: none"> – Az írásbeli vizsgarész értékelése nem részletező és szabályozott, a tanár szabadon mérlegelte, hogy milyen megoldásokat fogadott el. – A produktív készségek értékelése globális útmutatóval, értékelési kritériumok és skálák nélkül történt. – Az értékelés eredménye nem összehasonlítható, nem megbízható. 	<ul style="list-style-type: none"> – Az írásbeli vizsgarész értékelése központi útmutató alapján: csak a megadott megoldás fogadható el, a nyelvi hibákat funkció szerint kell megítélni. – A produktív készségek értékelése részletes útmutatóval, értékelési kritériumok és analitikus skálákkal történik. – Az értékelés eredménye összehasonlítható, megbízhatóbb.

* *Megjegyzés:* Az érettségi vizsgafeladatokat három intézmény bocsátotta ki: az Országos Közoktatási Szolgáltató Iroda, a Nemzeti Szakképzési Intézet és az Országos Felsőoktatási Felvételi Iroda. Német nyelvből például egy vizsgaévben 13 feladatsort adtak ki (*Einhorn*, 1998).

2. melléklet. A szövegalkotás értékelési szempontához tartozó kétféle skálaszerkezet
(Emelt szint. Javítási- és értékelési útmutató, 2006. 17. és 20. o.)

<i>Emelt szint, 1. feladat</i>	<i>Emelt szint, 2. feladat</i>
3 pont	5 pont
A szöveg felépítése, az irányító szempontok elrendezése logikus. A gondolati tagolás megfelelő: van bevezetés, tárgyalás és befejezés. A vizsgázó megfelelően használja a szövegösszefüggés tartalmi és nyelvi elemeit: létrejön a szövegkohézió. A formai tagolás megfelelő.	A szöveg felépítése, az irányító szempontok és az ezekhez kapcsolódó gondolatok elrendezése logikus. A gondolati tagolás megfelelő. A vizsgázó megfelelően használja a szövegösszefüggés tartalmi és nyelvi elemeit: létrejön a szövegkohézió. A szöveg formai tagolása megfelelő.
2 pont	4–3 pont
A szöveg felépítése, az irányító szempontok elrendezése nagyrészt logikus. A vizsgázó törekszik a gondolati tagolásra: van bevezetés vagy befejezés. A vizsgázó nagyrészt megfelelően használja a szövegösszefüggés tartalmi és nyelvi elemeit. A mondatok nagyrészt szervesen kapcsolódnak egymáshoz.	A szöveg felépítése, az irányító szempontok és az ezekhez kapcsolódó gondolatok elrendezése nagyrészt logikus. A vizsgázó törekszik a gondolati tagolásra. A vizsgázó nagyrészt megfelelően használja a szövegösszefüggés tartalmi és nyelvi elemeit. A mondatok nagyrészt szervesen kapcsolódnak egymáshoz. Amennyiben a szöveg formailag nem megfelelően tagolt, az alacsonyabb pontszámot kell adni.
1 pont	2–1 pont
A mondanivaló nem mindenütt logikus elrendezésű. A vizsgázó nem törekszik a gondolati tagolásra: hiányzik a bevezetés és a befejezés. Az írásmű még szövegnek tekinthető, bár a mondatok több helyen nem kapcsolódnak egymáshoz.	A mondanivaló nem mindenütt logikus elrendezésű. Sok a gondolati ismétlés. A vizsgázó nem törekszik a gondolati tagolásra. Az írásmű szövegnek tekinthető, bár a mondatok több helyen nem kapcsolódnak egymáshoz. Amennyiben a szöveg formailag nem megfelelően tagolt az alacsonyabb pontszámot kell adni.
0 pont	0 pont
Nem jött létre szöveg. A mondanivaló nem logikus elrendezésű, a gondolatok esetlegesen követik egymást. Az írásmű tagolatlan és áttekinthetetlen, túlnyomórészt izolált mondatok halmazából áll.	Nem jött létre szöveg. A mondanivaló nem logikus elrendezésű, a gondolatok esetlegesen követik egymást. Az írásmű tagolatlan és áttekinthetetlen, túlnyomórészt izolált mondatok halmazából áll.

3. melléklet. A kérdőívhez adaptált források és eredmények

<i>Az I–II. kérdőív rész megírásához felhasznált források és eredmények</i>		
<i>Cheng (1998. 348. o.) kérdőív részlet</i>	<i>Átlag</i>	<i>Szám</i>
What do you see as the major reasons for the HKEA (Hong Kong Examination Authority) to change the present HKCEE in English? 1=strongly disagree, 2=disagree, 3=undecided, 4=agree, 5=strongly agree		
1. To meet the demands of tertiary education.	3,77	3.
2. To prepare students for their future career.	3,85	1.
3. To refine testing methods.	3,30	7.
4. To narrow the gap between HKCEE and UE.	4,01	5.
5. To cope with the present decline in English standards.	2,91	4.
<i>Chen (2002. 175–180) kérdőív részlet</i>	<i>Átlag</i>	<i>Szám</i>
Please read each statement below and write down the number that mostly describes your current teaching situation regarding the level of agreement on the factors associated with your perceptions of impact of the BCT [...] 1=strongly disagree, 2=disagree; 3=slightly disagree, 4=slightly agree, 5=agree, 6=strongly agree	-	
44. I will feel embarrassed if my students perform less well on the BCT than other students taught by my colleagues.	-	12.
45. I will feel guilty if my students do not succeed on the BCT.	-	15.
46. Students' test results influence how people judge me as a good teacher.	-	13.
47. The BCT gives me important feedback about how I teach	-	14.
56. The BCT is able to test my students' language ability.	-	2.
63. The BCT impacts nationwide curriculum innovation	-	7.
<i>Al-Jamel és Ghadi (2008. 185–186. o.) kérdőív részlet</i>	<i>Átlag</i>	<i>Szám</i>
What are the effects of the General Secondary Certificate Examination on my teaching ? 1=strongly disagree, 2=disagree, 3=no opinion, 4=agree, 5=strongly agree		
25. I feel embarrassed if my students perform less well on the GSCE than other students taught by my colleagues.	2,73	12.
26. I get feedback on my teaching from the GSCE.	3,61	14.
32. I regard the GSCE as a test, which will influence students' future career.	3,60	1.
35. The GSCE is able to test my students' language ability.	2,40	2.
<i>Az III. kérdőív rész megírásához felhasznált források és eredmények</i>		
<i>Cheng (1998) kérdőív részlet</i>	<i>Átlag</i>	<i>Szám</i>
The factors that most influence your teaching are? 1=strongly disagree, 2=disagree, 3=undecided, 4=agree, 5=strongly agree		
1. Professional training	3,94	20.
2. Academic seminars or workshops	3,51	21.
3. Teaching experience and belief	4,40	22.
7. Textbooks	3,24	23.
8. Public examinations	3,92	25.
9. Learners' expectations	3,91	26.

A 3. melléklet folytatása

<i>Caine (2005. 109–115.) kérdőív-részlet</i>	<i>Átlag</i>	<i>Szám</i>
To what extent do the following factors influence your teaching? Circle your answers. Influence: 1=a little – 4=a lot		
a) professional training	2,86 (3,58)	20.
b) teaching experience	3,43 (4,29)	22.
e) textbooks	2,29 (2,87)	23.
g) external examinations	2,57 (3,22)	25.
h) learners' expectations	3,00 (3,75)	26.

Az V. kérdőív-rész megírásához felhasznált források és eredmények

<i>Cheng (1998) kérdőív-részlet</i>	<i>Átlag</i>	<i>Szám</i>
What do you think are the basic functions of mock tests in school? 1=strongly disagree, 2=disagree, 3=undecided, 4=agree, 5=strongly agree		
1. To give feedback to teachers.	4,03	51.
2. To assess students' learning difficulties.	3,83	54.
3. To motivate students	3,40	52.
4. To direct students' learning	3,56	55.
5. To prepare students for public examination	4,38	53.

Megjegyzés: Cheng (1998) két alkalommal, a vizsga bevezetése előtt (1994) és után (1995) kérdezte meg a tanárokat ugyanazzal a kérdőívvel. A táblázat csak a második felmérés adatait tartalmazza.

Caine (2005) négyfokú skálán kérdezte a tanárok véleményét. Az eredményeket az összehasonlíthatóság miatt ötfokú skálára transzformáltuk és zárójelben adtuk meg.

4. melléklet. Az elektronikus változatú tanári kérdőív

Tanári kérdőív

Kedves Kolléga!

Köszönjük, hogy válaszaival hozzájárul ahhoz, hogy a 2005-ben bevezetett kétszintű idegen nyelvi érettségi nyelvoktatásra gyakorolt hatásait megismerjük, amelynek alapján a döntéshozók számára javaslatokat fogalmazhatunk meg.

A kérdőív névtelen, titkos és kitöltése önkéntes. A gyűjtött egyéni adatokat kizárólag kutatási célokra használjuk és bizalmasan kezeljük.

A kérdőív használhatósága teljes mértékben az Ön őszinteségétől függ. Kérjük, hogy válaszaival járuljon hozzá, hogy a munkánk eredményeként kialakuló kép minél valóságosabb legyen.

A kérdőív hét oldalból áll. Egy oldal kitöltése a kipróbálás alapján átlagosan 2-3 percet vesz igénybe. Az oldal adatait elmentheti és az adatfelvétel időszakában bármikor folytathatja a kérdőív kitöltését.

Köszönjük a felmérésben való részvételét és együttműködését! A kutatás eredményeiről az adatfeldolgozást és -elemzést követően egy rövid összefoglalást küldünk.

Bejelentkezés

Kérdőív kódszáma:

© 2009 SZTE Neveléstudományi Intézet

Tanári kérdőív

Kérem, adja meg, hogy melyik idegen nyelvet tanítja!

Amennyiben Ön több idegen nyelvet is tanít, kérem, *válasszon ki egyet, amelyiken érettségire is felkészít.* Kérem, hogy a további válaszait a megjelölt idegen nyelvre vonatkoztassa!

A választott idegen nyelv:

Angol Német Francia Egyéb:

© 2009 SZTE Neveléstudományi Intézet

A 4. melléklet folytatása

Tanári kérdőív

Kedves Kolléga!

Az alábbiakban kiválaszthatja, hogy a kérdőív melyik oldalát kívánja kitölteni. Az egyes oldalak mentésekor automatikusan a következő oldal kitöltésével folytathatja a válaszadást. A korábban már mentett (de nem feltétlenül teljesen kitöltött) oldalak neve mellett ✓ jel, a még nem mentett oldalak mellett pedig ✗ jel látható.

Az egyes oldalakon a korábban már értékelt kijelentések szövegét sötétkék betűkkel írtuk.

A kérdőív oldalainak kitöltése

[1. oldal ✗](#) | [2. oldal ✗](#) | [3. oldal ✗](#) | [4. oldal ✗](#) | [5. oldal ✗](#) | [6. oldal ✗](#) | [7. oldal ✗](#)

Kérjük, az egyes oldalak kitöltése után minden esetben nyomja meg a **"Mentés, majd tovább..."** gombot! Amennyiben ezt nem teszi meg (pl. a Vissza gombbal visszatér a menüoldalhoz), úgy bevitt adatait a rendszer nem rögzíti.

[Kilépés](#)

Köszönjük közreműködését!

© 2009 SZTE Neveléstudományi Intézet

Tanári kérdőív

1. oldal

Kérem, olvassa el az alábbi kijelentéseket a 2005-ben bevezetett kétszintű idegen nyelvi érettségi vizsgáról, majd pontozza 1-től 5-ig az állításokat aszerint, hogy mennyire ért velük egyet! Válaszát a megfelelő szám bejelölésével adhatja meg. (A korábban már értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja.)

- 1: határozottan nem értek egyet
2: nem igazán értek egyet
3: részben egyetértek, részben nem
4: nagyjából egyetértek
5: határozottan egyetértek

A kétszintű idegen nyelvi érettségi vizsga...	1	2	3	4	5
1. ... a diákoknak a korábbi vizsgánál nagyobb jogosítványokat nyújt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. ... a diákok nyelvtudását a korábbi vizsgánál jobban méri.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. ... a korábbi vizsgánál jobban szolgálja a felsőoktatás igényeit.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. ... illeszkedik a Közös Európai Referenciakeret szintrendszeréhez.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. ... a Nemzeti alaptantervben deklarált célokat a korábbi vizsgánál jobban méri.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. ... elősegíti a nyelvi készségek fejlesztését a nyelvoktatásban.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. ... a nyelvoktatásban tartalmi és módszertani modernizációt indított el.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. ... a korábbi vizsgánál jobban segíti a javítói és értékelői munkámat.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Mentés, majd tovább a 2. oldalra](#)

[Vissza az oldalak választásához \(mentés nélkül\)](#)

© 2009 SZTE Neveléstudományi Intézet

A 4. melléklet folytatása

Tanári kérdőív

2. oldal

Kérem, olvassa el az alábbi kijelentéseket a 2005-ben bevezetett kétszintű idegen nyelvi érettségi vizsgáról, majd pontozza 1-től 5-ig az állításokat aszerint, hogy mennyire ért velük egyet! Válaszát a megfelelő szám bejelölésével adhatja meg. (A korábban már értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja.)

- 1: határozottan nem értek egyet
 2: nem igazán értek egyet
 3: részben egyetértek, részben nem
 4: nagyjából egyetértek
 5: határozottan egyetértek

		1	2	3	4	5
9.	Jónak tartom, hogy az érettségien van külön nyelvhelyességet mérő vizsgarész.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10.	Feleslegesnek tartom a középszintű vizsgán az A2 szintű feladatokat, a B1 szint elérhető az érettségizőknek.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11.	Mióta bevezették a kétszintű érettségit, több időre van szükségem az óráimra való felkészüléshez.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12.	Kínosan érzem magam, ha az érettségien a diákjaim rosszabbul teljesítenek, mint a kollégám diákjai.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13.	Az érettségi vizsgaeredmények befolyásolják, hogy a tanulók és a kollégák engem jó tanárnak tartanak-e.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14.	Úgy érzem, hogy az érettségi megbízható visszajelzést ad a tanításom minőségéről.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15.	Felelősnek érzem magam, ha a tanulók nem sikeresek az érettségi vizsgán.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16.	Szükségesnek tartom, hogy a közvetítési készség mérése is megjelenjen az emelt szintű érettségien.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17.	Az óráimon használt tankönyv alkalmas az érettségien mért készségek fejlesztésére.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18.	Szükségesnek érzem a tankönyvi feladatok kiegészítését az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatokkal.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19.	Az érettségire felkészítő feladatgyűjteményekben szereplő mérőfeladatokkal megfelelően tudom szimulálni a vizsgaszituációt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mentés, majd tovább a 3. oldalra

[Vissza az oldalak választásához \(mentés nélkül\)](#)

Tanári kérdőív

3. oldal

Melyek azok a tényezők, amelyek befolyásolják az Ön nyelvtanítási gyakorlatát? Kérem, 1-től 5-ig pontozza a tényezőket aszerint, hogy mennyire ért velük egyet! Válaszát a megfelelő szám bejelölésével adhatja meg. (A korábban már értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja.)

- 1: határozottan nem értek egyet
2: nem igazán értek egyet
3: részben egyetértek, részben nem
4: nagyjából egyetértek
5: határozottan egyetértek

A nyelvtanítási gyakorlatomat befolyásolják...	1	2	3	4	5
20. ... a felsőoktatásban, a nyelvtanárképzésben tanultak.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. ... a módszertani továbbképzésen tanultak.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. ... a korábbi nyelvoktatási tapasztalataim.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. ... az aktuálisan alkalmazott nyelvkönyv.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. ... a kétszintű nyelvi érettségi vizsgakövetelményei.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. ... a kétszintű nyelvi érettségi korábbi vizsgafeladatai.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26. ... a nyelvtanulók elvárásai.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mentés, majd tovább a 4. oldalra

[Vissza az oldalak választásához \(mentés nélkül\)](#)

A 4. melléklet folytatása

Tanári kérdőív

4. oldal

A továbbiakban állításokat olvashat a nyelvtanítási gyakorlatról. Kérem, a baloldali oszlopban azt jelölje, hogy mennyire jellemző Önre a felsorolt állítások! A jobb oldali oszlopba 1-5-ig terjedő skálán írja be, hogy véleménye szerint mennyire köszönhető a kétszintű idegen nyelvi érettségi hatásának, hogy az adott nyelvtanítási gyakorlat jellemző Önre. Kérem, hogy válaszainak megadásakor egy nyelvi érettségire készülő csoportra gondoljon. (A korábban már mindkét szempont szerint értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja. A kitöltendő négyzetek között tabulátorral is mozoghat.)

- 1: egyáltalán nem jellemző
2: kevéssé jellemző
3: jellemző is meg nem is
4: többnyire jellemző
5: teljes mértékben jellemző

Egy példa:

Ha az Ön nyelvoktatási gyakorlatára többnyire jellemző az olvasott szöveg értésének fejlesztése, amelyben az érettségi hatása kevéssé jellemző, akkor az első sorban látható módon jelölje a választ!

	Mennyire jellemző?	Érettségi hatása
<i>Az olvasott szöveg értésének fejlesztése nagy szerephez jut az óráimon.</i>	4	2
27. A hallott szöveg értésének fejlesztése nagy szerephez jut az óráimon.	<input type="checkbox"/>	<input type="checkbox"/>
28. Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.	<input type="checkbox"/>	<input type="checkbox"/>
29. A nyelvórákon a célnyelvet használom.	<input type="checkbox"/>	<input type="checkbox"/>
30. A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítetem fel.	<input type="checkbox"/>	<input type="checkbox"/>
31. A nyelvórákon autentikus szövegekkel dolgozom.	<input type="checkbox"/>	<input type="checkbox"/>
32. A diákoknak fordítási feladatokat adok.	<input type="checkbox"/>	<input type="checkbox"/>
33. A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.	<input type="checkbox"/>	<input type="checkbox"/>
34. A nyelvórákat a nyelvi szituációk gyakorlására építem.	<input type="checkbox"/>	<input type="checkbox"/>
35. Változatos munkaformákat alkalmazok.	<input type="checkbox"/>	<input type="checkbox"/>
36. Változatos, szövegközpontú feladattípusokat használok.	<input type="checkbox"/>	<input type="checkbox"/>
37. A produktív készségek mérésénél értékelési szempontokat és skálákat használok.	<input type="checkbox"/>	<input type="checkbox"/>

Mentés, majd tovább az 5. oldalra.

[Vissza az oldalak választásához \(mentés nélkül\)](#)

A 4. melléklet folytatása

Tanári kérdőív

5. oldal

Mennyire fontosak az Ön számára az alábbi szempontok, ha saját maga készít a tanulóknak mérőfeladatokat dolgozathoz? Kérem, 1-től 5-ig pontozza az állításokat aszerint, hogy mennyire tartja fontosnak! Válaszát a megfelelő szám bejelölésével adhatja meg. (A korábban már értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja.)

- 1: egyáltalán nem fontos
2: nem igazán fontos
3: fontos is meg nem is
4: elég fontos
5: nagyon fontos

A saját magam által készített mérőfeladat...	1	2	3	4	5
38. ... mérjen megbízhatóan, pontosan.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. ... épüljön autentikus szövegekre.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. ... eredménye segítse elő a nyelvtanuló előrehaladását.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41. ... legyen objektíven értékelhető.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42. ... mérje külön az egyes nyelvi készségeket.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43. ... tükrözzön reális, a valóságban is várható nyelvi szituációt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. ... a tanuló szókincsét célzottan mérje.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45. ... ne csak a nyelvtudást mérje, hanem az általános műveltséget, háttérismeretet is.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46. ... mérje a tananyag elsajátításának mértékét.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47. ... a tanuló formális nyelvtani ismereteit mérje.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48. ... épüljön a Közös Európai Referenciakeret szintrendszerére.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
49. ... igazodjon a kétszintű idegen nyelvi érettségi szemléletéhez.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
50. ... épüljön olyan szövegekre, szituációkra, amelyeket a diákok szívesen oldanak meg.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mentés, majd tovább a 6. oldalra

[Vissza az oldalak választásához \(mentés nélkül\)](#)

Tanári kérdőív

6. oldal

A továbbiakban a központilag összeállított érettségi vizsgafeladatok tanítási-tanulási folyamatban való alkalmazásának céljait olvashatja. Kérem, 1-től 5-ig pontozza az állításokat aszerint, hogy mennyire ért velük egyet! Válaszát a megfelelő szám bejelölésével adhatja meg. (A korábban már értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja.)

- 1: határozottan nem értek egyet
2: nem igazán értek egyet
3: részben egyetértek, részben nem
4: nagyjából egyetértek
5: határozottan egyetértek

A központilag összeállított vizsgafeladatok tanítási-tanulási folyamatban való alkalmazásának célja, hogy...	1	2	3	4	5
51. ... a diáknak a vizsgán várható teljesítményéről nyújtson információkat.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
52. ... a diákokat tanulásra motiválja.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
53. ... a diákokat felkészítse az érettségi vizsgán a konkrét teendőkre.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
54. ... megállapítsa, az adott készséggel milyen minőségben rendelkezik a diák.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
55. ... a diákok nyelvi készségeit fejlessze.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mentés, majd tovább a 7. oldalra

[Vissza az oldalak választásához \(mentés nélkül\)](#)

Tanári kérdőív

7. oldal

Kérem, válaszoljon a következő kérdésekre! (A korábban már értékelt kijelentések sötétkék betűvel látszanak, de az ezekre adott értékelést is megváltoztathatja.)

A további válaszait az első oldalon megadott idegen nyelvre vonatkoztassa!

56. Milyen iskolatípusban tanít? Többet is bejelölhet!

- Gimnáziumban Szakközépiskolában Szakiskolában

57. Milyen településtípuson található az intézménye? Jelölje be a megfelelőt!

- Nem város Kisváros Megyei jogú város Főváros, kerület:

58. Mi az Ön neme? Jelölje be a megfelelőt!

- Nő Férfi

59. Mi az Ön végzettsége a tanított idegen nyelvből? Jelölje be a megfelelőt!

- Tanár szakos főiskolai végzettség
 Három éves nyelvtanári végzettség
 Tanár szakos egyetemi végzettség

60. Melyik évben szerezte diplomáját idegen nyelvből?

61. Hány évet tanított már?

62. Jellemzően milyen szintre készíti fel a diákjait? Jelölje be a megfelelőt!

- Középszintre Emelt szintre Közép- és emelt szintre

63. Jellemzően milyen típusú nyelvi csoportban készíti fel a diákjait az érettségire? Többet is bejelölhet!

- Normál óraszámú csoportban Emelt óraszámú csoportban Kéttannyelvű csoportban

64. Vett-e részt az érettségire vonatkozó vizsgáztatói képzésen?

- Igen Nem

A 4. melléklet folytatása

65. Melyik szinten érettségiztetett már? *Többet is bejelölhet!*

- Még nem vizsgáztattam egyik szinten sem.
- Középszinten.
- Emelt szinten írásbeli javítóként vizsgáztattam.
- Emelt szinten szóbeli vizsgán vizsgáztattam.

66. Vett-e részt nyelvvizsgáztatói képzésen?

- Igen
- Nem

67. Hány módszertani továbbképzésen vett részt 2005. szeptember 1. óta?

- 0-30 órás hazai konferencia/tanfolyam/továbbképzés: db
- 31-60 órás hazai tanfolyam/továbbképzés: db
- 61-90 órás hazai tanfolyam/továbbképzés: db
- 91-120 órás hazai tanfolyam: db
- 120 óránál hosszabb hazai képzés: db
- Külföldi tanulmányút/tanfolyam/továbbképzés: db

68. Váltott-e tankönyvet az új érettségi hatására?

- Igen
- Nem

69. Használ-e az idegen nyelvi érettségire felkészítő feladatgyűjteményt?

- Igen
- Nem

Ha igen, melyiket használja leggyakrabban? *Kérem, hogy a feladatgyűjtemény címét és szerzőjét adja meg!*

70. Melyik tankönyvet használja leggyakrabban az érettségire való felkészülés során? *Kérem, hogy a tankönyv címét és szerzőjét adja meg!*

Mentés, majd tovább az összegző oldalra

[Vissza az oldalak választásához \(mentés nélkül\)](#)

© 2009 SZTE Neveléstudományi Intézet

Megjegyzés: A kérdőív háttere sötétkék színű volt, a válaszadó rész világoskék háttérű. Az elektronikus kérdőív színét módosítottuk, hogy a dolgozatban jobban olvasható legyen.

5. melléklet. A felkérő levél szövege

Kedves Kolléga!

A Szegedi Tudományegyetem Neveléstudományi Intézetének kutatócsoportja felmérést készít a 2005-ben bevezetett kétszintű idegen nyelvi érettségi jelenlegi működéséről. Az érettségi bevezetésének egyik célja a nyelvoktatás tartalmi és módszertani modernizációja. Jelenleg azonban nincsenek adataink arról, hogy a nyelvtanárok hogyan látják belülről a vizsga működését és a napi tanítási-tanulási folyamatban az érettségi hatását. Ennek megismeréséhez egy kérdőív kitöltésére szeretnénk felkérni.

Tisztelettel kérjük, vegyen részt a kérdőív kitöltésében, amellyel hozzájárul ahhoz, hogy az érettségi hatásmechanizmusait megismerhessük. A felmérés eredménye alapján célunk, hogy a döntéshozók számára javaslatokat fogalmazzunk meg a vizsgafejlesztés szempontjából, így nagyon fontos az Ön véleménye. A kutatás eredményeiről az adatfeldolgozást és -elemzést követően egy rövid összefoglalót küldünk Önnek.

A kérdőív névtelen és teljesen titkos, kitöltése önkéntes és a kipróbálás alapján átlagosan 25 percet vesz igénybe. A kérdőívet azonban elmentheti és az adatfelvétel időszakában, 2009. május 4–17. között bármikor folytathatja a kérdőív kitöltését.

A kérdőívek névtelenségéről úgy gondoskodunk, hogy az SZTE Neveléstudományi Intézet honlapjára lépnek a válaszadók, ahol a kérdőív hozzáféréséhez egy kódszámot adnak meg. A kódszám azt szolgálja, hogy vissza tudjon térni a kérdőívhez, és garantáljuk, hogy ezt az adatfeldolgozás és -elemzés során semmilyen formában nem hozzuk kapcsolatba Önnel.

Kérjük, kövesse az alábbi linket és használja a kódot a kérdőív eléréséhez:

Kérdőív kitöltésének helye: <http://www.edu.u-szeged.hu/nt/tanarikerdoiv/index.php>

Kérdőív kódszáma: s0649cdfgjk

Köszönjük a részvételét és együttműködését a kérdőív kitöltésében! A kutatással kapcsolatban felmerülő kérdéseit az alábbi e-mail címre küldje.

Üdvözlettel:

Vígh Tibor
Középiskolai német nyelvtanár
SZTE, Neveléstudományi Intézet munkatársa
Postacím: 6722, Szeged, Petőfi Sándor sgt. 30–34.
E-Mail cím: vightibor@edu.u-szeged.hu

6. melléklet. A tájékoztató levél szövege

Kedves Kolléga!

A Szegedi Tudományegyetem Neveléstudományi Intézetének kutatócsoportja nevében engedjék meg, hogy felhívjam szíves figyelmüket a kérdőívünkre, amellyel a 2005-ben bevezetett kétszintű idegen nyelvi érettségi jelenlegi működését, a tanítási-tanulási folyamatra gyakorolt hatását vizsgáljuk.

A kérdőív adatfelvételi időszakát meghosszabbítottuk: a felmérés május 20-án zárul. Tisztelettel kérem, hogy akik szívesen részt vennének a vizsgálatunkban, töltsék ki a kérdőívet.

A kérdőív elérhetősége: <http://www.edu.u-szeged.hu/nt/tanarikerdoiv/index.php>

A kérdőív hozzáférésehez szükséges automatikusan generált kódszámot a május 3-án kiküldött felkérő levél tartalmazza.

Az adatfelvétellel és a kutatással kapcsolatban felmerülő kérdéseiket az alábbi e-mail címre küldjék. Köszönöm a részvételüket és együttműködésüket a kérdőív kitöltésében!

Üdvözlettel:

Vígh Tibor
Középiskolai német nyelvtanár
SZTE, Neveléstudományi Intézet munkatársa
Postacím: 6722, Szeged, Petőfi Sándor sgt. 30–34.
E-Mail cím: vightibor@edu.u-szeged.hu

7. melléklet. A megköszönő levél szövege

Kedves Kolléga!

A Szegedi Tudományegyetem Neveléstudományi Intézetének kutatócsoportja nevében szeretném megköszönni mindazoknak, akik kitöltötték a kérdőívünket a 2005-ben bevezetett kétszintű idegen nyelvi érettségi jelenlegi működéséről, a tanítási-tanulási folyamatra gyakorolt hatásáról.

Az adatfelvételi időszakot lezártuk és megkezdjük az adatok elemzését, amelynek legfontosabb eredményeiről egy rövid összefoglalást küldünk Önnek.

Köszönöm a részvételüket és együttműködésüket a kérdőív kitöltésében!

Üdvözlettel:

Vígh Tibor

Középiskolai német nyelvtanár

SZTE, Neveléstudományi Intézet munkatársa

Postacím: 6722, Szeged, Petőfi Sándor sgt. 30–34.

E-Mail cím: vightibor@edu.u-szeged.hu

8. melléklet. „A tanárok nézetei az érettségi és a nyelvtanítási gyakorlat közti viszonyról” kérdőív rész módosított változóival végzett faktoranalízis eredménye

Nr.	Kérdőívétel	Faktor				
		1	2	3	4	5
35b	Változatos munkaformákat alkalmazok.	0,84				
30b	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.	0,83				
36b	Változatos, szövegközpontú feladattípusokat használok.	0,79				
34b	A nyelvórákat a nyelvi szituációk gyakorlására építem.	0,78				
29b	A nyelvórákon a célnyelvet használom.	0,77				
31b	A nyelvórákon autentikus szövegekkel dolgozom.	0,75				
33b	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.	0,70				
27b	A hallott szöveg értésének fejlesztése nagy szerephez jut az óráimon.	0,63				
28b	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.	0,61				
28a	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.		0,81			
35a	Változatos munkaformákat alkalmazok.		0,64			
36a	Változatos, szövegközpontú feladattípusokat használok.		0,57			
31a	A nyelvórákon autentikus szövegekkel dolgozom.			0,83		
33a	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.			0,76		
29a	A nyelvórákon a célnyelvet használom.				0,80	
30a	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.		0,43		0,67	
34a	A nyelvórákat a nyelvi szituációk gyakorlására építem.				0,62	
37b	A produktív készségek mérésénél értékelési szempontokat és skálákat használok.					0,86
37a	A produktív készségek mérésénél értékelési szempontokat és skálákat használok.					0,81
<i>Megmagyarázott variancia (%)</i>		27,56	10,36	9,64	9,12	8,63

Megjegyzés: A táblázatban a 0,40-nél nagyobb factorsúlyokat tüntettük föl. A KMO-index: 0,817; a Bartlett-teszt szignifikanciája: $p < 0,001$; a teljes megmagyarázott variancia: 65,33%.

9. melléklet. A teljes kérdőívre végzett faktoranalízis eredménye

Nr.	Kérdőívtétel	Faktor												
		1	2	3	4	5	6	7	8	9	10	11	12	13
35b	Változatos munkaformákat alkalmazok.	0,86												
30b	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.	0,84												
36b	Változatos, szövegközpontú feladattípusokat használlok.	0,80												
31b	A nyelvórákon autentikus szövegekkel dolgozom.	0,75												
34b	A nyelvórákat a nyelvi szituációk gyakorlására építem.	0,74												
29b	A nyelvórákon a célnyelvet használom.	0,72												
33b	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.	0,69												
27b	A hallott szöveg értésének fejlesztése nagy szerephez jut az óráimon.	0,56												
28b	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.	0,54												
5.	a Nemzeti alaptantervben deklarált célokat a korábbi vizsgánál jobban méri.	0,83												
6.	elősegíti a nyelvi készségek fejlesztését a nyelvtanításban.	0,83												
2.	a diákok nyelvtudását a korábbi vizsgánál jobban méri.	0,79												
3.	a korábbi vizsgánál jobban szolgálja a felsőoktatás igényeit.	0,76												
8.	a korábbi vizsgánál jobban segíti a javítói és értékelői munkámat.	0,74												
7.	a nyelvtanításban tartalmi és módszertani modernizációt indított el.	0,73												
4.	illeszkedik a Közös Európai Referenciakeret szintrendszeréhez.	0,70												
14.	Úgy érzem, hogy az érettségi megbízható visszajelzést ad a tanításom minőségéről.	0,60												
1.	a diákoknak a korábbi vizsgánál nagyobb jogosítványokat nyújt.	0,47												
52.	a diákokat tanulásra motiválja.											0,76		
54.	megállapítsa, az adott készséggel milyen minőségben rendelkezik a diák.											0,76		
55.	a diákok nyelvi készségeit fejlessze.											0,70		
53.	a diákokat felkészítse az érettségi vizsgán a konkrét teendőkre.											0,68		
51.	a diákoknak a vizsgán várható teljesítményéről nyújtson információkat.											0,61		

A 9. melléklet folytatása

Nr.	Kérdőív-tétel	Faktor												
		1	2	3	4	5	6	7	8	9	10	11	12	13
31a	A nyelvórákon autentikus szövegekkel dolgozom.				0,72									
33a	A nyelvtant autentikus szövegek és szituációk segítségével is gyakoroltatom.				0,71									
36a	Változatos, szövegek központú feladattípusokat használlok.				0,41									
48.	épüljön a Közös Európai Referenciakeret szintrendszerére.					0,74								
39.	épüljön autentikus szövegekre.				0,48	0,65								
45.	csak a nyelvtudást, az általános műveltséget, háttérismeretet ne mérje.					0,52		0,42						
50.	épüljön olyan szövegekre, szituációkra, amelyeket a diákok szívesen oldanak meg.					0,52								
49.	igazodjon a kétszintű idegen nyelvi érettségi szemléletéhez.					0,46								
13.	Az érettségi vizsgaeredmények befolyásolják, hogy a tanulók és a kollégák engem jó tanárnak tartanak-e.						0,85							
12.	Kínosan érzem magam, ha az érettségin a diákjaim rosszabbul teljesítenek, mint a kollégám diákjai.						0,69							
15.	Felelősnek érzem magam, ha a tanulók nem sikeresek az érettségi vizsgán.						0,61							
44.	a tanuló szókincsét célzottan mérje.							0,74						
47.	a tanuló formális nyelvtani ismereteit mérje.							0,69						
46.	mérje a tananyag elsajátításának mértékét.							0,59						
28a	Az egyéni különbségeket figyelembe véve, differenciáltan tanítom a diákokat.								0,78					
35a	Változatos munkaformákat alkalmazok.								0,76					
37b	A produktív készségek mérésénél értékelési szempontokat és skálákat használlok.									0,81				
37a	A produktív készségek mérésénél értékelési szempontokat és skálákat használlok.									0,77				
29a	A nyelvórákon a célnyelvet használom.										0,66			
30a	A tanulókat folyamatosan, a nyelvi készségek kommunikatív használatára készítem fel.								0,48		0,58			
34a	A nyelvórákat a nyelvi szituációk gyakorlására építem.											0,62		

A 9. melléklet folytatása

Nr.	Kérdőív-tétel	Faktor												
		1	2	3	4	5	6	7	8	9	10	11	12	13
43.	tükrözzön reális, a valóságban is várható nyelvi szituációt.												0,73	
42.	mérje külön az egyes nyelvi készségeket.					0,41							0,57	
41.	legyen objektíven értékelhető.												0,74	
40.	eredménye segítse elő a nyelvtanuló előrehaladását.											0,47	0,50	
38.	mérjen megbízhatóan, pontosan.												0,49	
18.	Szükségesnek érzem a tankönyvi feladatok kiegészítését az érettségire felkészítő feladatgyűjteményekből származó mérőfeladatokkal.													0,79
11.	Mióta bevezették a kétszintű érettségit, több időre van szükségem az óráimra való felkészüléshez.													0,63
<i>Megmagyarázott variancia (%)</i>		10,88	10,28	5,64	5,03	4,69	4,56	4,34	4,25	3,76	3,75	3,47	3,36	3,15

Megjegyzés: A táblázatban a 0,40-nél nagyobb faktorsúlyokat tüntettük föl. A KMO-index: 0,729; a Bartlett-teszt szignifikanciája: $p < 0,001$; a teljes megmagyarázott variancia: 67,16%.

10. melléklet. A teljesítmények átlaga, szórása iskolatípusok szerinti bontásban a középszintű vizsgán angolból és németből

Vizsgarész	Iskolatípus	Angol									Német								
		2007			2008			2009			2007			2008			2008		
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értése	gimnázium	77,13	14,73	0,10	65,88	22,87	0,15	82,15	19,52	0,12	82,80	17,47	0,16	72,96	19,65	0,19	72,22	22,84	0,21
	szakközépiskola	65,80	16,55	0,11	45,82	20,57	0,14	62,52	22,22	0,15	66,57	19,83	0,16	54,11	17,88	0,15	50,18	18,62	0,16
Nyelvhelyesség	gimnázium	55,53	26,13	0,18	66,22	21,97	0,15	63,63	25,28	0,16	55,26	18,00	0,17	62,91	20,13	0,19	66,68	21,43	0,20
	szakközépiskola	35,27	22,61	0,15	45,63	20,25	0,14	40,85	23,58	0,15	41,96	14,78	0,12	44,69	16,41	0,13	46,99	16,64	0,14
Hallott szöveg értése	gimnázium	69,20	19,87	0,14	72,62	18,30	0,12	69,00	18,36	0,11	78,49	17,69	0,16	65,85	22,12	0,21	58,48	18,59	0,17
	szakközépiskola	55,89	18,65	0,12	57,65	17,80	0,12	54,00	16,92	0,11	64,58	19,34	0,15	45,32	19,93	0,16	43,54	14,71	0,12
Íráskészség	gimnázium	72,03	26,01	0,18	74,01	27,04	0,18	77,16	25,38	0,16	63,48	30,27	0,28	67,90	29,73	0,28	70,01	29,25	0,27
	szakközépiskola	49,16	31,18	0,21	47,59	32,64	0,22	51,08	32,40	0,21	36,68	30,64	0,24	39,73	31,15	0,26	41,13	31,26	0,26
Beszédkészség	gimnázium	73,12	23,59	0,16	76,49	23,61	0,16	77,69	23,14	0,14	72,72	24,20	0,22	75,94	24,32	0,23	77,02	24,12	0,22
	szakközépiskola	57,23	26,05	0,17	58,45	26,44	0,18	58,58	26,87	0,18	55,73	26,08	0,20	56,70	26,32	0,22	57,76	26,44	0,22
Teljes vizsga	gimnázium	70,47	18,34	0,13	71,19	19,79	0,13	74,63	19,16	0,12	71,75	18,83	0,17	69,41	20,59	0,19	68,78	20,38	0,19
	szakközépiskola	54,08	18,77	0,12	51,23	19,72	0,13	54,33	20,71	0,14	53,87	18,58	0,15	48,12	18,72	0,15	47,68	17,70	0,15

Megjegyzés: A gimnáziumban és szakközépiskolában tanuló diákok közötti különbségek szignifikánsak $p < 0,001$ szinten.

11. melléklet. A 2008. évi május-júniusi vizsgaidőszak szintillesztési eljárásának eredménye közép- és emelt szinten iskola- és képzéstípusok szerint

Középszinten a B1 szintet határoló pontértéket elérők aránya (%)

Változó	Csoport	Olvasás		Hallás		Írás	
		A	N	A	N	A	N
Iskolatípus	Gimnázium	68	71	64	62	46	47
	Szakközépiskola	32	32	30	22	16	14
Képzéstípus	Nappali	52	50	48	41	32	30
	Felnőttoktatás nappali	18	24	18	13	4	5
	Felnőttoktatás esti/levelező	37	45	32	34	10	14
	Nincs jogviszonya	59	47	52	36	38	23

Jelölések: A: Angol; N: Német.

Forrás: Felmérés a 2008-ban Magyarországon érettségizettek idegennyelv-tudásáról, 2009.

Emelt szinten a B2 szintet határoló pontértéket elérők aránya (%)

Változó	Csoport	Olvasás		Hallás		Írás	
		A	N	A	N	A	N
Iskolatípus	Gimnázium	84	88	90	74	73	55
	Szakközépiskola	73	81	83	67	58	49
Képzéstípus	Nappali	80	85	87	70	68	52
	Nincs jogviszonya	79	80	82	69	68	52

Jelölések: A: Angol; N: Német.

Forrás: Felmérés a 2008-ban Magyarországon érettségizettek idegennyelv-tudásáról, 2009.

Megjegyzés: A felnőttoktatásban nagyon kevesen tesznek érettségi vizsgát (mindkét nyelvből kb. 20–30-an), így adataikat nem tartalmazza a táblázat.

12. melléklet. A teljesítmények átlaga, szórása képzéstípusok szerinti bontásban a középszintű vizsgán angolból és németből

Vizsgarész	Képzéstípus	Angol									Német								
		2007			2008			2009			2007			2008			2009		
		Átlag	Szórás	Std. Hiba	Átlag	Szórás	Std. Hiba	Átlag	Szórás	Std. Hiba	Átlag	Szórás	Std. Hiba	Átlag	Szórás	Std. Hiba	Átlag	Szórás	Std. Hiba
Olvasott szöveg értése	nappali	71,80	16,51	0,08	56,75	23,98	0,12	73,79	22,77	0,11	74,29	20,28	0,13	63,04	20,82	0,13	60,91	23,42	0,15
	Fo. levelező/esti	67,53	16,44	0,41	47,48	20,33	0,52	64,33	21,52	0,56	68,49	21,22	0,72	59,56	20,34	0,74	56,39	21,91	0,88
	Fo. nappali	57,19	16,23	0,49	37,75	16,71	0,48	50,14	20,00	0,61	60,38	18,97	0,52	50,44	17,83	0,48	45,62	17,72	0,54
Nyelvhelyesség	nappali	45,85	26,35	0,13	56,93	23,41	0,11	53,82	26,89	0,12	48,09	17,51	0,11	53,28	20,26	0,13	56,76	21,36	0,14
	Fo. levelező/esti	38,10	25,87	0,64	45,76	20,61	0,52	43,63	25,11	0,65	45,05	18,45	0,62	47,98	18,15	0,66	48,22	18,29	0,74
	Fo. nappali	25,57	18,41	0,55	36,50	17,73	0,51	29,97	20,39	0,62	39,78	14,53	0,40	42,45	17,34	0,47	41,44	16,01	0,49
Hallott szöveg értése	nappali	63,02	20,23	0,10	65,86	19,42	0,09	62,60	19,10	0,09	71,26	19,58	0,12	55,14	23,21	0,15	50,84	18,20	0,12
	Fo. levelező/esti	54,92	20,41	0,50	58,07	18,89	0,48	54,61	18,16	0,47	64,80	22,39	0,76	51,31	22,64	0,82	48,47	18,31	0,74
	Fo. nappali	47,03	16,43	0,49	50,97	16,76	0,48	45,79	15,49	0,47	59,28	19,75	0,55	39,40	18,55	0,50	39,89	14,35	0,44
Íráskészség	nappali	61,71	30,38	0,15	62,52	32,20	0,16	66,51	30,91	0,14	49,80	33,05	0,21	53,98	33,20	0,22	55,95	33,29	0,21
	Fo. levelező/esti	43,85	31,50	0,78	43,28	31,54	0,80	45,02	32,40	0,84	34,03	29,75	1,00	38,64	31,64	1,15	38,05	30,75	1,24
	Fo. nappali	28,22	27,98	0,84	27,67	27,00	0,78	27,79	29,14	0,88	23,54	26,56	0,73	23,70	25,59	0,70	25,29	25,42	0,78
Beszédkészség	nappali	65,52	26,01	0,13	68,34	26,44	0,13	69,61	26,47	0,12	63,79	26,50	0,16	66,03	27,03	0,18	67,36	26,95	0,17
	Fo. levelező/esti	60,22	26,03	0,64	60,01	25,96	0,66	59,20	25,82	0,67	59,75	26,41	0,89	59,46	27,24	0,99	62,82	26,62	1,07
	Fo. nappali	49,08	24,98	0,75	49,67	25,50	0,74	47,04	25,18	0,76	48,79	25,57	0,71	50,65	25,30	0,69	50,12	26,31	0,81
Teljes vizsga	nappali	62,83	20,12	0,10	62,26	21,99	0,11	66,08	22,04	0,10	62,44	20,56	0,13	58,46	22,17	0,14	58,19	21,65	0,14
	Fo. levelező/esti	54,09	18,89	0,47	51,12	18,85	0,48	54,02	19,97	0,52	55,02	19,73	0,67	51,42	20,55	0,74	50,73	19,45	0,78
	Fo. nappali	42,67	15,54	0,47	40,59	15,85	0,46	40,82	17,68	0,54	46,68	16,39	0,45	40,87	15,39	0,42	40,06	14,70	0,45

Jelölések: Fo.: Felnőttoktatás. A nappali, valamint a felnőttoktatásban tanuló nappali és levelező/esti tagozatos diákok közötti különbségek szignifikánsak $p < 0,001$ szinten.

13. melléklet. A teljesítmények átlaga, szórása nemek szerinti bontásban a középszintű vizsgán angolból és németből

Vizsgarész	Nem	Angol									Német								
		2007			2008			2009			2007			2008			2009		
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értéke	fiú	72,00	16,48	0,11	56,75	24,17	0,16	74,39	22,49	0,15	71,32	20,95	0,19	60,49	20,68	0,20	57,47	22,83	0,22
	lány	70,55	16,88	0,11	55,19	23,74	0,16	71,74	23,40	0,15	74,98	20,05	0,16	63,56	20,92	0,17	62,06	23,59	0,19
		F>L			F>L			F>L			F<L			F<L			F<L		
Nyelvhelyesség	fiú	45,29	26,50	0,18	56,05	23,83	0,16	55,06	27,06	0,17	46,21	17,34	0,16	50,39	19,70	0,19	52,64	20,56	0,20
	lány	44,79	26,31	0,17	56,00	23,21	0,15	51,15	26,81	0,17	48,60	17,58	0,14	54,15	20,47	0,17	58,20	21,60	0,17
		n.s.			n.s.			F>L			F<L			F<L			F<L		
Hallott szöveg értéke	fiú	63,74	20,31	0,14	66,70	19,17	0,13	64,02	19,66	0,13	69,36	19,92	0,18	52,92	23,55	0,22	49,64	17,88	0,17
	lány	60,95	20,34	0,13	63,79	19,79	0,13	60,16	18,61	0,12	71,29	19,83	0,16	55,14	22,99	0,19	50,81	18,41	0,15
		F>L			F>L			F>L			F<L			F<L			F<L		
Íráskészség	fiú	57,50	31,56	0,22	58,74	33,15	0,22	62,67	32,33	0,21	41,86	32,54	0,30	45,48	33,36	0,32	47,33	33,60	0,32
	lány	62,65	30,28	0,20	63,02	32,18	0,21	67,23	30,86	0,19	52,51	33,05	0,26	56,60	32,98	0,27	59,05	32,78	0,26
		F<L			F<L			F<L			F<L			F<L			F<L		
Beszédkészség	fiú	63,46	26,15	0,18	66,10	26,63	0,18	67,18	27,03	0,17	58,65	26,53	0,24	60,78	27,26	0,26	62,09	27,24	0,26
	lány	66,21	26,09	0,17	68,97	26,54	0,17	70,33	26,33	0,16	66,03	26,34	0,21	68,11	26,75	0,22	69,59	26,71	0,21
		F<L			F<L			F<L			F<L			F<L			F<L		
Teljes vizsga	fiú	62,33	20,21	0,13	61,01	22,31	0,15	65,30	22,67	0,15	58,26	20,43	0,19	54,04	22,04	0,21	53,62	21,27	0,21
	lány	61,58	20,41	0,14	61,61	21,96	0,14	65,09	22,01	0,14	63,75	20,56	0,16	59,72	22,05	0,18	59,79	21,63	0,17
		F>L			F<L			F>L			F<L			F<L			F<L		

Jelölések: F: fiú, L: lány; >/< szignifikánsan nagyobb/kisebb érték (p<0,001); n. s. nem szignifikáns.

14. melléklet. A teljesítmények átlaga, szórása iskolatípusok szerinti bontásban az emelt szintű vizsgán angolból és németből

Vizsgarész	Iskolatípus	Angol									Német								
		2007			2008			2009			2007			2008			2008		
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értéke	gimnázium	69,61	21,13	0,45	71,64	16,04	0,34	70,66	20,08	0,39	80,37	20,08	0,63	83,35	12,15	0,39	90,50	9,31	0,27
	szakközépiskola	61,02	21,87	0,73	65,10	15,84	0,52	62,31	21,43	0,69	72,55	22,16	0,90	79,90	14,93	0,65	86,33	10,19	0,40
Nyelvhelyesség	gimnázium	69,57	18,88	0,40	70,44	17,06	0,36	64,17	16,97	0,33	71,23	16,89	0,53	67,68	17,86	0,57	67,59	16,83	0,50
	szakközépiskola	61,28	19,56	0,65	63,09	17,26	0,57	56,23	16,64	0,54	64,13	17,29	0,70	61,67	18,23	0,79	59,37	15,36	0,61
Hallott szöveg értéke	gimnázium	79,48	12,39	0,26	78,84	14,93	0,31	74,66	18,09	0,35	70,38	17,19	0,54	76,55	17,79	0,56	81,32	13,70	0,40
	szakközépiskola	75,22	13,37	0,45	73,82	15,99	0,53	66,88	18,42	0,59	64,21	16,69	0,67	71,92	19,76	0,86	76,18	13,99	0,55
Íráskészség	gimnázium	76,86	17,12	0,36	76,52	17,57	0,37	78,41	16,07	0,32	69,01	24,23	0,75	76,38	20,10	0,64	73,89	21,23	0,63
	szakközépiskola	69,66	19,13	0,64	69,75	19,21	0,63	72,03	17,47	0,56	58,32	26,53	1,07	72,57	21,87	0,95	67,51	21,92	0,86
Beszédkészség	gimnázium	71,08	23,35	0,49	73,80	22,69	0,48	74,29	21,61	0,42	73,17	23,63	0,74	76,38	21,09	0,67	78,52	20,66	0,61
	szakközépiskola	60,55	24,58	0,82	63,80	24,78	0,82	64,59	24,56	0,79	61,75	28,20	1,14	70,30	22,57	0,98	67,01	23,89	0,94
Teljes vizsga	gimnázium	73,00	15,28	0,32	73,92	14,61	0,31	72,10	15,55	0,31	72,51	16,89	0,53	75,76	14,42	0,46	78,03	13,43	0,40
	szakközépiskola	65,21	16,17	0,54	66,78	14,95	0,49	64,08	16,28	0,52	63,85	18,27	0,74	70,96	16,12	0,70	70,96	13,71	0,54

Megjegyzés: A gimnáziumban és szakközépiskolában tanuló diákok közötti különbségek szignifikánsak $p < 0,001$ szinten.

15. melléklet. A teljesítmények átlaga, szórása nemek szerinti bontásban az emelt szintű vizsgán angolból és németből

Vizsgarész	Nem	Angol									Német								
		2007			2008			2009			2007			2008			2009		
		Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba	Átlag	Szórás	Std. hiba
Olvasott szöveg értése	fiú	70,27	20,34	0,54	70,39	15,90	0,40	71,40	20,26	0,50	78,89	20,88	0,91	82,21	13,68	0,62	89,67	9,66	0,41
	lány	65,49	22,64	0,48	68,51	16,97	0,37	65,69	21,36	0,44	77,18	21,58	0,60	81,59	13,69	0,40	88,43	10,16	0,27
		F>L*			F>L*			F>L*			n.s.			n.s.			F>L**		
Nyelvhelyesség	fiú	69,09	18,63	0,49	68,74	17,23	0,44	62,78	17,99	0,44	70,39	17,43	0,76	65,59	18,97	0,86	63,58	18,09	0,77
	lány	66,10	20,18	0,43	66,98	18,00	0,39	60,89	17,11	0,35	68,29	17,56	0,49	64,92	18,61	0,54	64,52	16,88	0,45
		F>L*			F>L**			F>L*			F>L**			n.s.			n.s.		
Hallott szöveg értése	fiú	80,38	11,60	0,31	79,02	15,33	0,39	74,96	18,00	0,44	69,35	16,99	0,74	76,52	18,35	0,83	79,45	14,54	0,62
	lány	76,92	13,68	0,29	75,13	15,83	0,34	69,97	19,01	0,39	67,92	17,53	0,49	73,41	19,35	0,56	79,10	14,16	0,38
		F>L*			F>L*			F>L*			n.s.			F>L**			n.s.		
Íráskészség	fiú	72,78	18,57	0,49	71,11	19,26	0,49	75,24	17,16	0,42	60,92	26,67	1,16	70,64	21,80	0,99	66,91	23,89	1,02
	lány	76,16	17,71	0,37	76,21	17,70	0,38	77,32	16,61	0,34	66,65	25,20	0,70	75,90	20,72	0,60	72,20	21,83	0,59
		F<L*			F<L*			F<L*			F<L*			F<L*			F<L*		
Beszédkészség	fiú	68,04	24,04	0,64	69,45	24,29	0,62	71,57	23,57	0,58	67,17	26,21	1,14	71,63	23,53	1,07	71,69	24,62	1,05
	lány	69,25	24,39	0,52	70,39	24,22	0,52	70,64	23,18	0,47	69,90	26,15	0,73	74,06	22,30	0,65	74,35	22,79	0,61
		n.s.			n.s.			n.s.			F<L**			F<L**			F<L**		
Teljes vizsga	fiú	71,79	15,28	0,40	71,43	15,11	0,38	70,85	16,22	0,40	69,01	17,89	0,780	72,98	15,68	0,71	73,92	15,05	0,64
	lány	70,45	16,64	0,35	71,10	15,59	0,34	68,58	16,55	0,34	69,66	18,18	0,508	73,67	15,80	0,46	75,39	14,20	0,38
		F>L**			n.s.			F>L*			n.s.			n.s.			n.s.		

Jelölések: F: fiú, L: lány; >/< szignifikánsan nagyobb/kisebb érték (*p<0,001; **p<0,05); n. s. nem szignifikáns.

16. melléklet. 2006. évi középszintű német érettségi olvasott szöveg értése feladatlap
 itemeinek jellemzői

<i>Feladat</i>	<i>Item</i>	<i>Nehézségi index</i>	<i>Standard hiba</i>	<i>Infít paraméter</i>	<i>Diszkrimináció</i>
1.	1	-2,78	0,065	1,04	0,08
	2	-1,58	0,060	1,02	0,21
	3	-0,76	0,055	0,95	0,40
	4	-0,34	0,053	0,93	0,42
	5	-1,59	0,060	0,99	0,24
	6	-1,58	0,060	0,99	0,25
	7	-1,17	0,058	1,02	0,25
	8	0,41	0,050	0,96	0,46
2.	9	0,49	0,050	0,94	0,49
	10	0,74	0,049	1,06	0,38
	11	0,86	0,049	0,99	0,47
	12	2,91	0,052	0,94	0,42
	13	2,33	0,050	1,02	0,37
	14	-0,14	0,052	1,06	0,33
	15	0,40	0,050	0,87	0,57
3.	16	1,41	0,048	1,08	0,36
	17	-0,52	0,054	0,97	0,36
	18	-0,45	0,054	0,99	0,34
	19	-1,60	0,060	0,94	0,33
	20	0,59	0,049	1,08	0,35
	21	0,54	0,050	1,32	0,13
	22	0,55	0,049	1,01	0,41
	23	-0,87	0,056	0,91	0,46
	24	1,34	0,048	0,95	0,51
	25	0,84	0,262	1,02	0,41

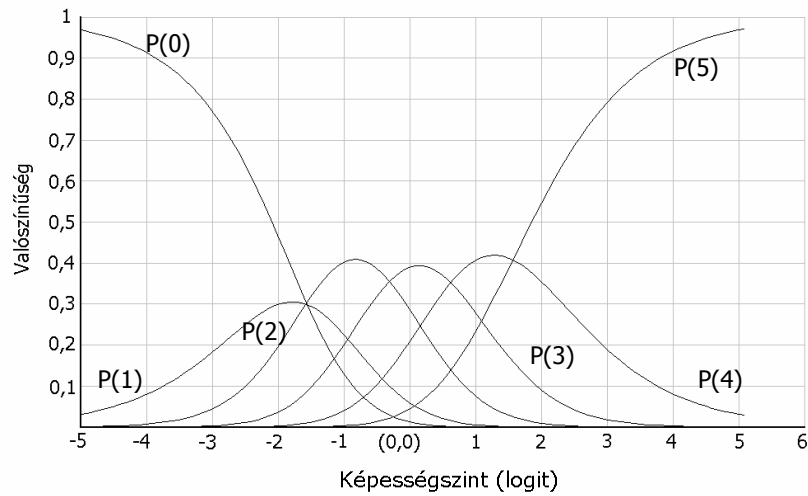
17. melléklet. 2006. évi középszintű német érettségi hallott szöveg értése feladatlap
 itemeinek jellemzői

<i>Feladat</i>	<i>Item</i>	<i>Nehézségi index</i>	<i>Standard hiba</i>	<i>Infit paraméter</i>	<i>Diszkrimináció</i>
1.	1	-1,27	0,069	1,04	0,25
	2	-2,54	0,081	1,07	0,18
	3	-0,95	0,066	1,09	0,28
	4	0,79	0,056	0,99	0,45
	5	-0,77	0,065	1,00	0,36
	6	-0,45	0,062	1,18	0,19
	7	-1,61	0,073	1,07	0,21
2.	8	2,52	0,059	1,08	0,34
	9	1,28	0,056	0,97	0,49
	10	-1,02	0,067	0,99	0,39
	11	-0,80	0,065	0,98	0,41
	12	2,47	0,059	1,06	0,36
	13	-2,48	0,081	1,02	0,20
	14	-1,77	0,074	1,01	0,28
3.	15	1,95	0,056	1,01	0,41
	16	-1,01	0,067	0,89	0,44
	17	0,44	0,057	0,97	0,45
	18	1,58	0,056	1,03	0,40
	19	1,91	0,056	1,02	0,45
	20	0,87	0,056	0,93	0,46
	21	1,75	0,056	0,88	0,51
	22	1,61	0,056	0,95	0,48
	23	-1,59	0,073	0,99	0,31
	24	-0,91	0,308	1,02	0,31

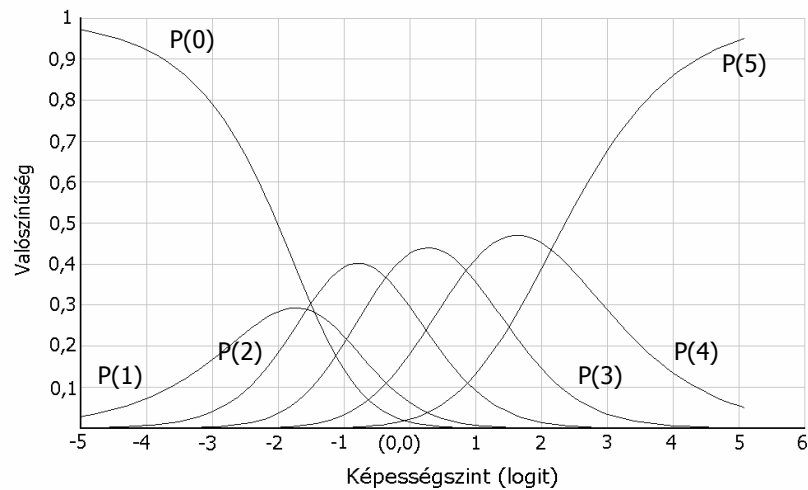
18. melléklet. 2006. évi középszintű német érettségi nyelvhelyesség feladatlap itemeinek jellemzői

<i>Feladat</i>	<i>Item</i>	<i>Nehézségi index</i>	<i>Standard hiba</i>	<i>Infít paraméter</i>	<i>Diszkrimináció</i>
1.	1	-0,17	0,049	1,01	0,35
	2	-1,51	0,056	1,11	0,11
	3	1,18	0,048	1,17	0,14
	4	-0,58	0,051	0,91	0,43
	5	-0,38	0,050	1,00	0,37
	6	-2,20	0,060	0,95	0,32
	7	-0,55	0,051	1,00	0,33
2.	8	-0,22	0,050	1,01	0,33
	9	-0,42	0,050	0,99	0,37
	10	0,93	0,047	0,96	0,40
	11	-0,45	0,051	0,94	0,44
	12	0,05	0,049	0,97	0,44
	13	-1,49	0,056	1,04	0,21
	14	0,25	0,048	0,97	0,37
	15	-1,40	0,056	1,07	0,14
3.	16	-1,02	0,053	0,88	0,47
	17	-1,11	0,054	1,08	0,25
	18	0,64	0,048	0,88	0,54
	19	0,55	0,048	1,00	0,37
	20	-0,60	0,051	1,09	0,25
	21	0,34	0,048	1,09	0,31
	22	1,07	0,048	1,04	0,30
	23	0,57	0,048	0,92	0,47
	24	-0,77	0,052	0,90	0,43
4.	25	2,34	0,051	1,06	0,23
	26	0,81	0,047	1,11	0,23
	27	1,48	0,048	1,00	0,35
	28	1,10	0,048	0,98	0,39
	29	0,04	0,049	0,95	0,43
	30	1,53	0,273	0,92	0,45

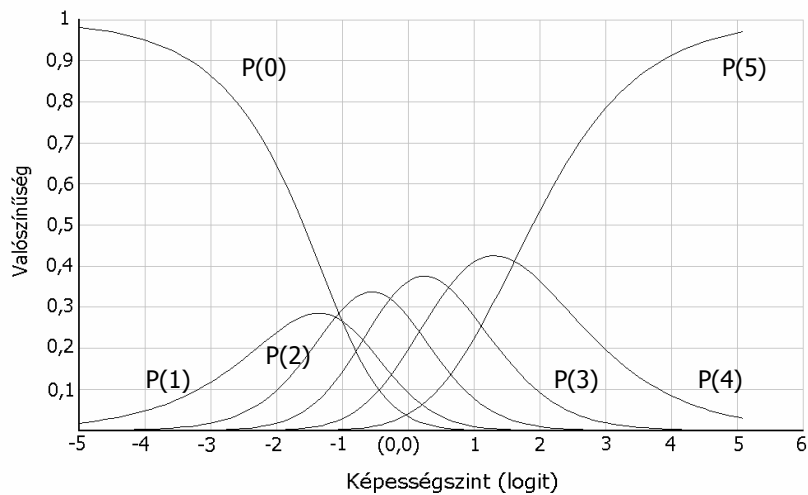
19. melléklet. A nyelvi komponensekhez tartozó értékelési szempontok karakterisztikus görbéi



Az első feladat szókincs, kifejezésmód értékelési szempont karakterisztikus görbéje

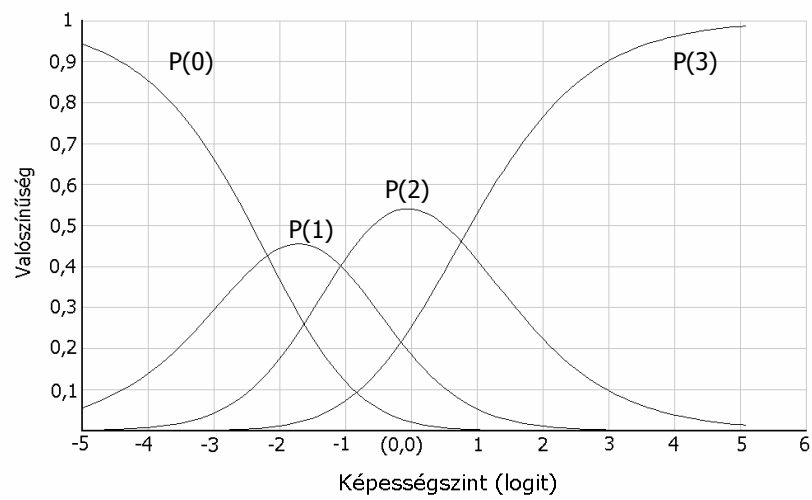


Az első feladat nyelvtan, helyesírás értékelési szempont karakterisztikus görbéje

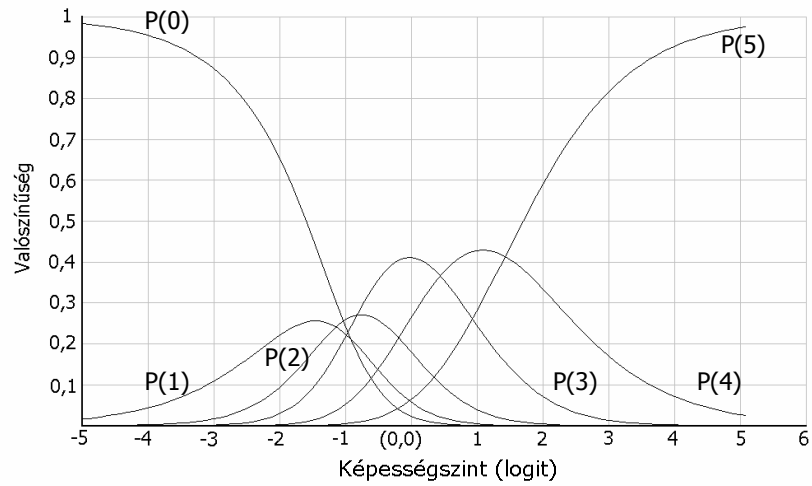


A második feladat szókincs, kifejezésmód értékelési szempont karakterisztikus görbéje

20. melléklet. A szövegalkotás értékelési szempontok karakterisztikus görbéi

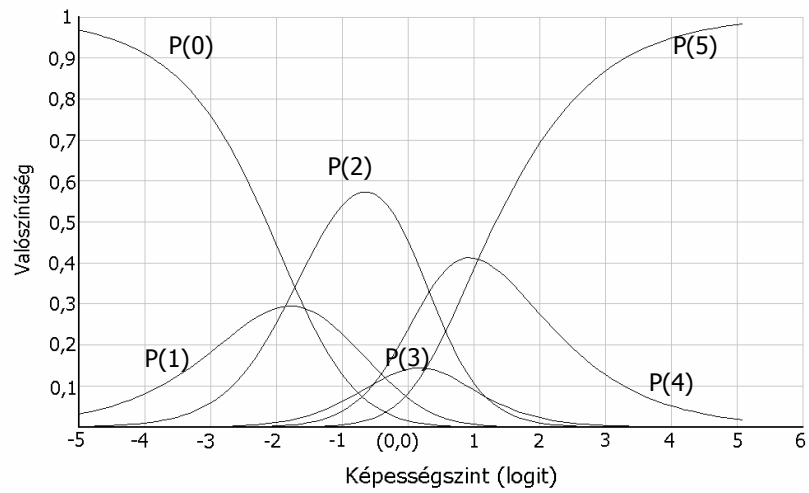


Az első feladat szövegalkotás értékelési szempont karakterisztikus görbéje

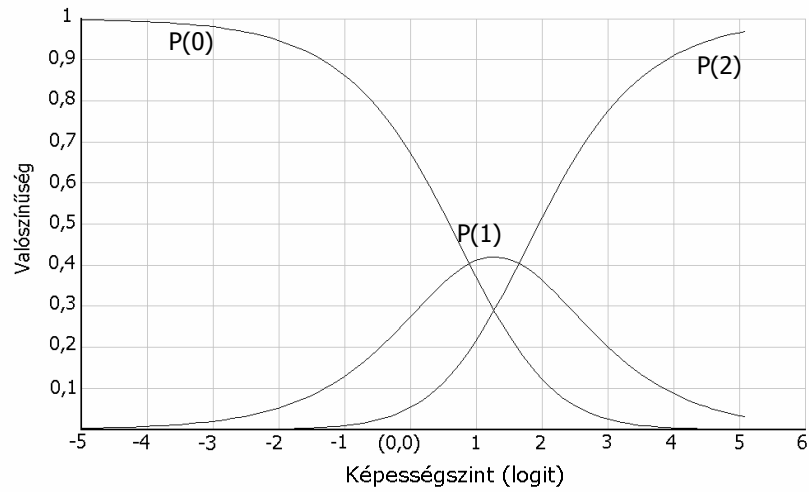


A második feladat szövegalkotás értékelési szempont karakterisztikus görbéje

21. melléklet. A kommunikatív komponensekhez tartozó értékelési szempontok karakterisztikus görbéi



Az első feladat tartalom értékelési szempont karakterisztikus görbéje



Az első feladat formai jegyek és hangnem értékelési szempont karakterisztikus görbéje

22. melléklet. A teljes kérdőív faktoranalízise során létrejött faktorok korrelációs mátrixa

Faktor	Faktor												
	1	2	3	4	5	6	7	8	9	10	11	12	
1. A tanárok nézetei a vizsga jelentőségéről	1												
2. A vizsga következményeivel kapcsolatos attitűdök	n.s.	1											
3. A tanórára való felkészüléssel kapcsolatos attitűdök	n.s.	0,38**	1										
4. A vizsga nyelvoktatási gyakorlatra kifejtett hatásával kapcsolatos nézetek	n.s.	0,17*	0,18*	1									
5. Az eredményes tanítási módszerekről kialakított nézetek	n.s.	n.s.	n.s.	n.s.	1								
6. A vizsga és a tanítási gyakorlat közti kapcsolatról kialakított nézetek	n.s.	n.s.	n.s.	n.s.	0,61**	1							
7. A vizsga és az oktatás tartalma közti kapcsolatról kialakított nézetek	n.s.	n.s.	n.s.	n.s.	0,43**	0,33**	1						
8. A vizsga és az osztálytermi értékelési eljárások közti kapcsolatról kialakított nézetek	0,23**	n.s.	n.s.	n.s.	n.s.	0,16*	0,22**	1					
9. A klasszikus tesztelmélet jószágmutatóinak fontosságáról kialakított nézetek	n.s.	n.s.	n.s.	n.s.	0,17*	n.s.	0,27**	0,21*	1				
10. A kommunikatív tesztelés jószágmutatóinak fontosságáról kialakított nézetek	0,23**	0,18*	0,26**	0,33**	0,23**	0,18*	0,43**	0,28**	0,36**	1			
11. A validitás fontosságáról kialakított nézetek	n.s.	0,19*	0,17*	n.s.	n.s.	n.s.	n.s.	n.s.	0,25**	0,25**	1		
12. A tanárok nézetei a vizsgafeladatok oktatási folyamatban való felhasználásának céljairól	0,18*	0,19*	n.s.	0,21*	n.s.	n.s.	n.s.	n.s.	0,20*	0,25**	0,24**	1	

Jelölések: ** p<0,01; *p<0,05; n.s. nem szignifikáns.

23. melléklet. Az iskolai gyakorlatban alkalmazott tankönyvek és feladatgyűjtemények gyakorisága

Angol		Német	
Tankönyvcím	Használók száma	Tankönyvcím	Használók száma
Challenge	2	Abiturtraining	13
Channel	3	Blick Mittelstufe	2
Channel Your English	6	Delfin	7
Enterprise	6	Deutsch mit Grips	3
Exam Excellence	9	em Brückenkurs, Hauptkurs	3
Finals	1	Geni@1	2
First Certificate Star	3	Grünes Licht	1
Headway-sorozat	3	Pinpong Neu	2
Matrix könyvek	8	Schrittweise	2
Matura	5	Sowieso	1
New English File sorozat	17	Start Neu	10
New Headway-sorozat	3	Tangram	12
New Opportunities-Pre-Intermediate	3	Tangram Aktuell	1
Opportunities	3	Unterwegs	15
Snapshot Elementary	1	Wortwörterlich	2
Twenty but one	1		
Upstream	1		
<i>Összesen</i>	78	<i>Összesen</i>	76

Angol nyelvi érettségire felkészítő feladatgyűjtemény szerzője és címe	Használók száma
Alderson, J. C. és Cseresznyés, M. (2003): <i>Into Europe: Prepare for modern english exams. Reading and Use of English.</i> Teleki László Alapítvány, Budapest	4
Bajnóczi Beatrix és Kirsi, Haavisto (2006): <i>888 kérdés és válasz angol nyelvből szóbeli nyelvvizsgára készülőknek.</i> Maxim Könyvkiadó, Szeged.	1
Batár Levente (2005): <i>Készüljünk az új érettségire! Angol Írásbeli gyakorlatok.</i> Nemzeti Tankönyvkiadó, Budapest.	6
Bukta Katalin, Gróf Szilvia és Sulyok Andrea (2006): <i>7 próbaérettségi angol nyelvből.</i> Maxim Könyvkiadó, Budapest	11
Harrison, Mark; Whitehead, Russel (2007): <i>Oxford Érettségi Felkészítő Tesztek Közép- és Emelt (+Cd) Oxford University Press, Oxford</i>	14
Hevesi Katalin, Pusztai Balázs: <i>Készüljünk az új érettségire! Angol szóbeli gyakorlatok – emelt szint.</i> Nemzeti Tankönyvkiadó	1
Magyarics Péter (2004): <i>Készüljünk az új érettségire! Angol feladatsorok.</i> Nemzeti Tankönyvkiadó, Budapest.	30
Mitchell, H.Q. (2007): <i>Matura Plus 2 Student's Book.</i> MM Publications – ELT Hungary, Budapest.	4
<i>Összesen</i>	71

A 23. melléklet folytatása

<i>Német nyelvi érettségire felkészítő feladatgyűjtemény szerzője és címe</i>	<i>Használók száma</i>
Husztiné Varga Klára (2007): <i>Plusz 7 próbaérettségi német nyelvből középszint</i> . Maxim Kiadó, Szeged.	9
Lázár Györgyné (2004): <i>Készüljünk az új érettségire! Német feladatsorok. Középszint</i> . Nemzeti Tankönyvkiadó, Budapest.	20
Maros Judit (2006): <i>Abitraining Oberstufe</i> . Nemzeti Tankönyvkiadó, Budapest.	5
Péntek Bernadett (2002/2004): <i>Középfokú kurzuskönyv a német szóbeli vizsgára</i> . Nordwest, Győr.	4
Péntek Bernadett és Husz Anikó (2002/2004): <i>Német középszint érettségi komplett feladatsorok</i> . Nordwest, Győr.	3
Rixer Márta és Sominé Hrebik Olga (2006): <i>7 próbaérettségi német nyelvből – középszint</i> . Maxim Kiadó, Szeged.	16
Somló Katalin (2007): <i>Német írásbeli gyakorlatok</i> . Nemzeti Tankönyvkiadó, Budapest.	5
Somló Katalin (2007): <i>Készüljünk az új érettségire. Abiturtraining Mittelstufe</i> . Nemzeti Tankönyvkiadó, Budapest.	3
<i>Összesen</i>	65