

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 tézisei



Témavezető: Dr. Vidákovich Tibor

Pedagógiai értékelés
doktori program

Szeged, 2010

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 minden esetben visszavezethetők a visszacsatolás jelentőségének felismerésére. Ezekben az innovációs folyamatokban fontos szerepe van a közoktatásban megfelelően működtetett vizsgáknak, amelyek 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 változást előidéző reformok mellett lényeges szerepe van a vizsgán elért eredmények vizsgálatából és a tanítási-tanulási folyamatban résztvevők tapasztalatainak feltárásából kiinduló visszacsatolásoknak, melyek hozzájárulnak a vizsga továbbfejlesztéséhez.

A vizsgaeredményekkel számszerűsíthető a tanulók tudása, amelyek befolyásolhatják a tanulók továbbhaladását, iskolafokozat-váltását. A vizsgaeredmények optimalizálhatják az oktatási folyamatot, és alapot szolgáltatnak a kitűzött célok megvalósulásának értékeléséhez. Mindezek mellett a pedagógiai munkában érdekelt személyek tapasztalatainak felméréséből és értelmezéséből értékes visszajelzést adhatunk a vizsgafejlesztés számára. Ezekkel az információkkal a vizsgafejlesztők megismerhetik, milyen a vizsga hatása a tanítási-tanulási folyamatra, és ez összhangban van-e az eredeti szándékaikkal.

A vizsgák befolyásolják 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. Ez a több évtizede ismert visszacsatolási mechanizmus és az ezzel kapcsolatos kérdések gyakran merülnek fel újra új mérési-értékelési rendszerek, érettségi, felvételi vizsgák bevezetésével vagy módosításával. Így részletesen kell ismerni a vizsga alapján elinduló hatások folyamatát, működését és befolyásoló tényezőit, és ezek alapján modellezni a jelenséget.

A KUTATÁS ELŐZMÉNYEI ÉS CÉLKITŰZÉSEI

Mindezen kérdések vizsgálata hazánkban a 2005-ben bevezetett kétszintű érettségi vizsga esetében is relevánsak (a vizsga jellemzőiről lásd: 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 kevés információval rendelkezünk az érettségi tényleges hatásáról, és kevésbé tisztázott a két vizsgaszint egymáshoz való viszonya, szükséges a tartalmi standardok továbbfejlesztése, a vizsgaszintek összekapcsolása nemzetközi mérési rendszerek teljesítményszintjeivel (Csapó, 2008). Az eredmények monitorozása, a vizsgafunkciók működésének ellenőrzése, és az ezek alapján történő vizsgafejlesztés a jelenlegi gyakorlatra kevésbé jellemző. Fontos vizsgálandó kérdés, hogyan kell ezt a vizsgát úgy továbbfejlesztetni, hogy kimutatható hatása a fejlesztők szándékával összhangban legyen.

Az értekezésben az érettségi kötelező vizsgatárgyai közül az idegen nyelvi érettségi példáján keresztül mutatjuk be a mérés-értékelés vizsgafejlesztésben betöltött szerepét, a vizsga közoktatásra gyakorolt hatását a tanárok nézetei alapján. A nyelvi érettségi vizsgák működésének és hatásának vizsgálatát négy körülmény teszi indokolttá. Egyrészt az idegen nyelvek tudása iránti igény hazánkban az 1990-es években jelentősen megnövekedett, majd az Európai Unió tagságunkkal a használható nyelvtudás presztízse tovább nőtt. Másrészt a nyelvi érettségik fejlesztési munkálatainak eredményeként (lásd Alderson, Nagy és Öveges, 2000; Einhorn, 2004) a vizsga igazodik a Közös Európai Referenciakeret (KER, 2002) nyelvtudásmodelljéhez és szintjeihez. Harmadrészt a vizsga elhelyezhető az idegennyelv-tudást mérő

vizsgák rendszerében, mivel az emelt szintű vizsgán elért eredmény alapján nyelvvizgabizonyítvány 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 2000-es évektől elinduló nagymintás nyelvtudást mérő kutatásokban és az idegen nyelvi érettségi vizsgán is megvalósulnak.

A magyar közoktatási rendszerben végzett idegen nyelvi felmérések és az érettségi vizsga a nyelvi tudásszint-mérés nemzetközi vizsgálatának tendenciáihoz három szempontból illeszkedik. Egyrészt a kommunikatív tesztelés alapelveinek alkalmazásával kapcsolódik a nyelvtudás mérésének és értékelésének aktuális kérdéseire. Másrészt az európai nyelvtudásmérési projektekben gyakran vizsgált terület a nyelvtudást mérő vizsgák szintjeinek validálása a Referenciakeret szintrendszeréhez és ezek alapján nyelvi feladatbank létrehozása. Harmadrészt a kommunikatív tesztelés nyelvoktatásra gyakorolt hatását, a tesztelés folyamatának modellezését számos vizsgálatban elemzik. A nemzetközi szakirodalom feltárása számos, a hazai kontextusban empirikusan vizsgálható kérdéseket határoz meg. Ugyanakkor ezeket az elmúlt évtizedben a hazai szakirodalomban nem, vagy nem eléggé hangsúlyozzák és vizsgálják.

Az értekezés célja, hogy feltárja a nyelvi mérés-értékelés nemzetközi tendenciáit és bemutassa hazai jellemzőit, amely lehetővé teszi e két terület legfontosabb tanulságainak szintetizálását egy elméleti modellben. A nyelvi mérés-értékelés mechanizmusainak modellezésével célunk az érettségi komplex és többszempontú elemzése, a modell empirikus validálása. Mindezek alapján fogalmazzuk meg következtetéseinket az érettségi továbbfejlesztésére, a további kutatási feladatokra és a vizsgára felkészítő gyakorlatra vonatkozóan.

A KUTATÁS KÉRDÉSEI ÉS MÓDSZEREI

I. A szakirodalmi elemzés és szintézis eszközével célunk a nyelvi mérés-értékelés nemzetközi tendenciáinak feltárása. Ennek keretében azt vizsgáltuk,

(1) miként integrálódik a kommunikatív tesztelés a tesztelmélet egészébe, és hogyan jelennek meg a kompetenciaértékelés minőségi mutatói a kommunikatív kompetencia mérése és értékelése során;

(2) milyen tesztelési módszereket alkalmaznak a kommunikatív tesztelés gyakorlatában.

A kommunikatív tesztelés nyelvoktatásra gyakorolt hatásával kapcsolatban szakirodalmi elemzéssel feltártuk,

(3) hogyan értelmezhető és jellemezhető a tesztelés folyamata;

(4) melyek a tesztelés folyamatát befolyásoló tényezők.

A kétszintű idegen nyelvi érettségi vizsgakövetelményeinek alapján azt elemeztük,

(5) hogyan illeszkedik a vizsga a meghatározott nyelvtudás-modellhez és a kommunikatív tesztelés minőségi mutatóihoz;

(6) milyen hipotézisek fogalmazhatók meg a nemzetközi vizsgálatok tanulságai alapján a vizsga hatására vonatkozóan.

A nemzetközi tendenciák feltárása és a hazai jellemzők szintetizálása alapján azt vizsgáltuk,

(7) hogyan modellezhetők a nyelvi mérés és értékelés mechanizmusai.

A nemzetközi szakirodalom alapján vizsgáltuk az egyes visszahatás-modellek fejlődését, a tesztelést meghatározó tényezőket, majd ezek alapján hoztuk létre a szintetizáló modellt, melyet a nyelvi érettségi példáján értelmeztünk. A modell empirikus validálásának céljából választottuk ki és indokoltuk az elemzendő célterületeket. Empirikus kutatásainkban a vizsgaeredmények alapján vizsgáltuk az érettségi jelenlegi működését, belső és külső validitását, az íráskészség mérési-értékelési eljárásainak minőségét. A vizsga tanítási-tanulási folyamatra gyakorolt hatását a tanárok visszajelzései alapján elemeztük.

II. A jelenlegi idegen nyelvi érettségi vizsgákon elért teljesítmények alapján a vizsgafejlesztésre és az igények és célokra való visszacsatolás érdekében azt vizsgáltuk,

- (8) hogyan változnak az idegen nyelvi érettségin elért teljesítmények az egyes vizsgarészekben;
- (9) milyen különbségek találhatók az érettségizők nyelvi készségeinek fejlettségében iskolatípusok, képzési típusok és nemek szerint az angol és a német nyelvi érettségin elért teljesítmények összehasonlítása alapján.

E kutatási kérdések megválaszolásához az Oktatási Hivatal (OH) szabadon elérhető adatbázisait használtuk. Az adatelemzés első fázisában a teljesítményeket vizsgáltuk, majd a nyelvi készségek fejlettségében megjelenő különbségek feltárását végeztük el, és ezek teljesítményre gyakorolt hatását vizsgáltuk mindkét nyelvnél.

III. A receptív készségeket mérő vizsgarészeknek a tanulók nyelvtudás-szintjéhez és a KER-hez (2002) való illeszkedésének vizsgálatát elsősorban a vizsgafejlesztési folyamatra irányuló visszacsatolás miatt a vizsgaminőség javítása érdekében tartottuk indokoltnak. Vizsgálatunkban feltártuk,

- (10) mennyire fedi le az érettségi vizsgafeladatok empirikus nehézsége a vizsgázók képességtartományát;
- (11) hogyan építhető ki egy, a Referenciakeret szintrendszerére épülő nyelvi feladatbank.

Az elemzés első fázisában a 2006. évi középszintű német érettségi olvasott és hallott szöveg értése és a nyelvhelyesség vizsgarészek itemszintű adatait vizsgáltuk a Rasch-moddal (*Rasch*, 1960). A három feladatlapot közös skálára kalibráltuk és a feladatok nehézségi indexét vizsgáltuk, amelyet három dimenzióban a minta- és itemtérképen ábrázoltunk. A második fázisban a 2005. és 2006. évi közép- és emelt szintű feladatlapok minta- és itemtérképét vizsgáltuk egyenként, és európai projektek eredményeit is tanulmányoztuk (például *CITO*, 2008).

IV. Az íráskészség vizsgarész mérési-értékelési eljárásainak vizsgálatával célunk olyan minőségi mutatók és paraméterek definiálása, amelyek mentén a különböző évekből származó eredmények összehasonlíthatók és változásai nyomon követhetők. E cél elérése érdekében azt vizsgáltuk,

- (12) milyen paraméterekkel jellemezhetők az értékelési szempontok és skálapontok;
- (13) mennyire jól működnek az egyes értékelési szempontok analitikus skálái.

E kutatási kérdések megválaszolására a 2006. évi emelt szintű német érettségi adatbázisát használtuk. Először az adatokat közös skálára kalibráltuk a parciális kredit modellel (*Masters*, 1982) és az értékelési szempontok nehézségét vizsgáltuk, amelyet a vizsgázók íráskészség-szintjével kapcsolunk össze. Ezt követően a skálák működését elemeztük a karakterisztikus görbéik segítségével az értékelők megítéléseinek vizsgálata céljából. A feltárt paraméterek használhatóságát a középszintű német érettségi alapján is elemeztük.

V. Az angolt és a németet tanító nyelvtanárok visszajelzései alapján célunk feltárni, hogy a vizsgával közvetlenül kapcsolatban álló tanárok hogyan látják belülről a vizsga működését és a tanítási-tanulási folyamatban az érettségi hatását. A kutatási kérdések a következők:

- (14) Milyen nézetekkel és attitűdökkel rendelkeznek a nyelvtanárok a vizsgával kapcsolatban és ezek milyen vizsgával kapcsolatos tapasztalatokkal vannak összefüggésben?
- (15) Milyen tényezők határozzák meg a tanárok nyelvoktatási gyakorlatát és ebben milyen szerepe van a vizsgakövetelményeknek és a vizsgafeladatoknak?
- (16) Meghatározzák-e a nyelvoktatás tartalmát, módszereit, mérési-értékelési gyakorlatát a tanárok vizsgával kapcsolatos nézetei és attitűdjei?
- (17) Mennyiben segíti elő az idegen nyelvi érettségi a kommunikatív tesztelés mérési és értékelési eljárásainak elterjedését a tanítási-tanulási folyamatban?

A kutatási kérdések megválaszolására kérdőíves felmérést végeztük, amelyben 149 nyelvtanár vett részt. A minta reprezentatív a tanított idegen nyelv és az iskolatípus szerint. A kérdőív

szerkezetét, a vizsgálandó tényezőket és a kérdéseket a nemzetközi szakirodalom alapján készítettünk (például *Spratt*, 2005; *Cheng*, 1998; *Chen*, 2002). A kérdőív reliabilitása (Cronbach- $\alpha=0,89$) megfelelő, a validitását faktoranalízis elvégzésével ellenőriztük. Először a tanárok vizsgával kapcsolatos nézeteit és attitűdjeit, majd az idegen nyelvi érettségi hatását vizsgáltuk a tanárok nézetei alapján. Ezt követően feltártuk, hogy a nézetek és attitűdök mennyiben határozzák meg a nyelvoktatás folyamatát és mérési-értékelési eljárásait.

A KUTATÁS EREDMÉNYEI

A kommunikatív tesztelés az idegen nyelvi mérés és értékelés rendszerében

(1) A kommunikatív tesztelés a pedagógiai értékelés sajátos jellegű önálló szakterülete, amely a klasszikus tesztelméletre építve határozza meg a kommunikatív tesztek minőségi mutatóit, funkcióit és típusait (*Bachman*, 1990). 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 (*Messick*, 1996). Ezek közé tartozik a tesztfejlesztési folyamatra vonatkozó autentikusság, interaktivitás és hatékonyság (*Bachman és Palmer*, 1996), a teszt lebonyolítására és következményeire vonatkozó igazságosság, összehasonlíthatóság, döntések reprodukálhatósága (*Kunnan*, 2004) és teszthatások (*Alderson és Wall*, 1993). E kritériumok összhangban vannak a kompetenciaértékelés minőségi mutatóival (*Baartman és mtsai*, 2006).

(2) A kommunikatív tesztek típusai közül kiemelendők a kompetencia és performancia tesztek (*McNamara*, 2000), és az ezekkel szorosan összefüggő közvetett és közvetlen tesztek, amelyek egy-egy skála két végpontját jelenítik meg. A tesztfejlesztők feladata, hogy e tesztípusok valamilyen egyensúlyát megtalálják (*Alderson, Clapham és Wall*, 1995). A nyelvtudást mérő vizsgák alkalmazzák ezeket a tesztelési módszereket, és ezt a folyamatot a *KER* (2002) is elősegíti, amely alapján 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 hatása a tanítási-tanulási folyamatra

(3) A nyelvtudást mérő vizsgák hatásának fogalmi szintű megnevezése nem egységes, ugyanazt a jelenséget többféleképpen, más-más jelentéstartalommal definiálják a kutatók (*Bailey*, 1999). Ezek szintetizálása alapján a teszthatás komplex jelenség, mely közvetlen és közvetett hatást fejt ki a vizsgával kapcsolatban álló személyekre, a tanítási-tanulási folyamatra, a tanulás tartalmára és minőségére (*Alderson és Wall*, 1993); tágabb értelemben véve a vizsgaeredmények alapján hozott döntésekkel a közoktatásra és a társadalomra (*Hamp-Lyons*, 1997).

(4) Az elméleti kutatások során készült modellek nem határozzák meg pontosan a folyamatényező jellemzőinek leírását és kölcsönhatását a vizsgaeredményekkel (*Wall*, 2005). E modellekben a kutatók kiemelik az igényeknek és céloknak megfelelő vizsgafejlesztési folyamatot (*Saif*, 2006), a létrejövő vizsga presztízsének fontosságát a hatások elindulásában (*Watanabe*, 2004), a tanárok nézeteinek, feltételezéseinek, ismereteinek és attitűdjeinek (*Burrows*, 2004) meghatározó szerepét. Az empirikus vizsgálatok eredményei (lásd például *Cheng, Watanabe és Curtis*, 2004) csak az adott kontextusban értelmezhetők és nehezen általánosíthatók, de megállapítható, hogy a vizsgának a tanítás tartalmára általában nagyobb hatása van, mint az oktatási módszerekre (*Alderson*, 2004), az osztálytermi értékelésre gyakorolt hatásokat azonban kevés kutatásban vizsgálták.

A kétszintű nyelvi érettségi a kommunikatív tesztelés rendszerében

(5) A kétszintű nyelvi érettségi a *KER* (2002) alapján definiálja mérési modelljét a vizsgakövetelményekben, melynek illeszkedését vizsgáltuk a nyelvtudás-modellhez, valamint a vizsga jellemzőit hasonlítottuk össze a kommunikatív tesztelés kritériumaival, funkcióival és típusaival. 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. A vizsgán elsősorban szűkebb értelemben vett performancia tesztekkel használnak.

(6) A kétszintű idegen nyelvi érettségi a kommunikatív nyelvhasználat mérésével igazodik a nyelvtanulók igényeihez, magas presztízzsel rendelkezik, mert záróvizsgaként, felvételi vizsgaként és nyelvvizsgaként is működik. A teszthatásokat vizsgáló nemzetközi kutatások alapján a vizsga hatásmechanizmusainak elindulása szempontjából minden tényező megjelenik, és ezért az érettségi meghatározhatja a tanárok vizsgával kapcsolatos nézeteit és attitűdjeit, amelyek befolyásolhatják a nyelvoktatás folyamatát és a tanárok mérési-értékelési eljárásokkal kapcsolatos nézeteit.

A teszthatás működését szintetizáló modell

(7) A szakirodalmi feltárás eredményeként létrehozott és az empirikus vizsgálataink alapján szolgáló modell magában foglalja egy nyelvtudást mérő vizsga tanítási-tanulási folyamatra gyakorolt hatásának mechanizmusát. Eszerint a vizsga először az oktatási rendszerben résztvevők nézeteit és attitűdjeit befolyásolja, amelyek aztán meghatározzák a tanítási folyamatot, a tanulás tartalmát és minőségét. A modell továbbá kétféle típusú 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 az idegen nyelvi érettségi vizsgákat a vizsgaeredmények és a nyelvtanárok visszajelentései alapján vizsgáltuk.

A vizsga jelenlegi működése

(8) A vizsgaeredmények alapján megállapítható, hogy 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 teljesíthetősége 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 nyelvszakra a felvételi funkcióját.

(9) Az iskolatípusok szerinti különbségek középszinten jelentősek, és minden évben folyamatosan emelkednek. A két nyelv esetében hasonló tendenciák figyelhetők meg: a gimnáziumban érettségizők szignifikánsan jobban teljesítenek a szakközépiskolában tanulóknál mindkét vizsgaszinten, a nappali tagozaton tanulók szignifikánsan magasabb teljesítményt mutatnak a felnőttoktatásban tanulóknál. A produktív készségeknél a lányok teljesítenek jobban a fiúknál, de a nemek közötti különbségek teljesítményre gyakorolt hatása a teljes vizsgán nem jelentős.

A vizsga belső és külső validitása

(10) A receptív készségeket mérő vizsgarészek feladat- és itemszintű vizsgálatai a 2006. évi középszintű német érettségien azt mutatták, hogy a feladatkészítők által megállapított nehézségi sorrend többnyire megegyezett az empirikus nehézség növekedésével. A vizsgafeladatok itemnehézségi indexei nem fedik le teljesen a vizsgázók nyelvi készségszintjei 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 közül az egyik az átlagos képességszint alatt, a másik az e felett lévő vizsgázók nyelvtudását méri. Azonosítottunk egy magas képességszintet is, amit a feladatsorok nem mérnek.

(11) Az egyes készségek minta- és itemtérképeinek önálló 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 felhasználhatók a referenciafeladatok kiválasztásában, az érettségi feladataihoz tartozó szövegek és itemek jellemezésében. A létrejövő feladatbank viszonyítási pontként funkcionálhat a további vizsgafeladatok fejlesztésekor és hozzájárulhat ahhoz, hogy az idegen nyelvi érettségi szintjeinek a KER-hez való illeszkedését empirikusan igazoljuk.

Az íráskészség mérési-értékelési eljárásainak minősége

(12) A 2006. évi német érettségi íráskészség-vizsgálatának eredményei szerint 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 mindkét vizsgaszinten). A skálapontok működése akkor megfelelő, ha küszöbértékei az íráskészség-szint növekedésével arányosan emelkednek. Ez a vizsgán megvalósult, de mindkét szinten a szövegalkotásnál és a szókincs, kifejezőmódnál aránytalanul. A skálapontokhoz tartozó modellilleszkedés-paraméter magas értéke azt jelezheti, hogy különböző képességszintű vizsgázóknál azonos pontszámot adnak az értékelők. Ezt emelt szinten elsősorban a nyelvtan, helyesírásnál a magasabb képességtartományban állapítottuk meg. A minta-és itemtérkép alapján azonosítottunk egy magasabb íráskészség-szintet, amelyet az egyes skálapontok küszöbértékei nem fednek le.

(13) Vizsgálatunkban kimutattuk, hogy az egyes értékelési skálák működése eltér szempontként. 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.

A vizsga nyelvoktatási folyamatra gyakorolt hatása a tanárok visszajelzései alapján

(14) Eredményeink szerint a nyelvi érettségit a tanárok elfogadhatónak és a korábbinál megbízhatóbbnak és érvényesebbnek tartják. A vizsgával kapcsolatos nézetek formálódásában meghatározók a tanárok vizsgáztatási tapasztalatai.

(15) A tanárok nézetei szerint a nyelvoktatási gyakorlatot jelentős mértékben befolyásolják a korábbi nyelvtanítási tapasztalatok és az érettségi vizsgákövetelményei. A vizsgadokumentumokban lévő alapelvek megismerése hatással van arra, hogy a tanárok a mérőfeladataik elkészítésében fontosnak tartják a kommunikatív tesztelés kritériumait.

(16) A teszthatást vizsgáló nemzetközi kutatások eredményeivel összhangban a tanárok szerint a vizsga a nyelvoktatás tartalmára nagyobb hatást fejt ki, mint a módszerekre.

(17) Feltételeztük, hogy a vizsga értékelési alapelvei befolyásolják a tanárok mérési-értékelési eljárásokkal kapcsolatos nézeteit. Ezt megerősítve vizsgálatunkban kimutattuk, hogy a tanárok vizsgával kapcsolatos nézetei és attitűdjei az osztálytermi mérés-értékelés több aspektusára fejtik ki hatásukat. Felértékeli a tanárok számára a kommunikatív tesztelés kritériumainak fontosságát és a vizsgafeladatok funkció szerinti felhasználását.

KÖVETKEZTETÉSEK ÉS JAVASLATOK

Végül az empirikus vizsgálataink eredményeiből levonható következtetéseket fogalmazzuk meg. Ezt követően teszünk javaslatokat a további vizsgálandó területekre.

A vizsga jelenlegi működésének elemzése középszinten jelentős különbséget mutatott a vizsgázók nyelvi készségének fejlettségében iskolatípusok és képzési típusok szerint. Ez a különbség az évek múlásával növekszik. Emelt szinten a követelményszintek, a feladatsorok minősége és az érettségi felvételi vizsga funkciójának teljesülése jelent problémát. Ezek további vizsgafejlesztési igényeket és újabb fejlesztési célokat jelölhetnek ki. Szükséges az érettségizők teljesítményében megjelenő különbségek okainak feltárása, diagnosztizálása, a mérési-értékelési eljárások minőségének ellenőrzése, felülvizsgálata. Erre azonban az OH adatbázisai nem alkalmasak, így szükségesek a feladat- és itemszintű elemzések.

A vizsga belső és külső validitásának vizsgálatában a valószínűségi tesztelmélet modelljét alkalmaztuk, így az eredményeink módszertanilag általánosíthatók. A követelményszintek empirikus bemérése más vizsgatárgyban is lényeges kutatási feladat. A mérés színvonalának és a döntések reprodukálhatóságának növelése céljából fontos lenne például a matematikai, természettudományi és az olvasás-szövegértési feladatsorok nehézségének összekapcsolása nemzetközi mérések (például a PISA-vizsgálatok) kompetenciaszintjeivel, a vizsgaszintek közötti tartalmi és mennyiségbeli különbségek meghatározása és feladatbank kiépítése.

Az íráskészség vizsgarész elemzésével az értékelési szempontok és skálák empirikus validálásának folyamatára mutattunk példát. E vizsgarész alapján feltárt paraméterek segítségével a vizsgaeredmények összehasonlíthatók és változásai nyomon követhetők. Ezek az elemzések az idegen nyelvi érettségin az alacsony megbízhatósággal rendelkező beszédkészség esetében is kellő alapot szolgáltatnak. Mindezek mellett a vizsgálatunk módszerei, a bemutatott minőségi mutatók olyan vizsgatárgyakban és vizsgálatokban is értelmezhetők, alkalmazhatók és elemezhetők, ahol az értékelés analitikus skálákkal történik.

A tanári kérdőíves felmérés a vizsgafejlesztés számára felhívja a figyelmet a vizsgával kapcsolatos képzések jelentőségére, melyek a vizsga filozófiájának megismertetését szolgálják. Ugyanakkor e vizsgálatunk korlátai sokrétűek: a nyelvi érettségi hatásait 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. Szükségesek a tantárgy-pedagógiai vizsgálatok, amelyek egy-egy területre kiterjedően részletesen elemzik a tanárok tanítási gyakorlatát és a tanulók nyelvtanulási folyamatát.

Vizsgálataink egyértelműen jelzik a vizsgafejlesztés és a további kutatások szükségességét. Ezek 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 megállapítás csak korlátozásokkal érvényes: eredményeink elsősorban az idegen nyelvi érettségre 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.

FELHASZNÁLT IRODALOM

- 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., Clapham, C. és Wall, D. (1995): *Language test construction and evaluation*. Cambridge University Press, Cambridge.
- 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.
- 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. és Palmer, A. S. (1996): *Language testing in practice*. Oxford University Press, Oxford.
- Bailey, K. M. (1999): *Washback in language testing. TOEFL Monograph Series, Ms. 15*. Educational Testing Service, Princeton, NJ.
- Báthory Zoltán (1985): *Tanítás és tanulás*. Tankönyvkiadó, Budapest.
- 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.
- 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. (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., Watanabe, Y. és Curtis, A. (2004, szerk.): *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates, Mahwah, NJ.
- CITO (2008): Building a European bank of anchor items for foreign language skills. EBAFLS. General report. Cito B.V., Arnhem.
- 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.
- 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.)
- Hamp-Lyons, L. (1997): Washback, impact and validity: Ethical concerns. *Language Testing*, **14**. 3. sz. 295–303.
- Horváth Zsuzsanna és Lukács Judit (2006, szerk.): *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Országos Közoktatási Intézet, Budapest.
- KER (2002): Közös Európai Referenciakeret: nyelvtanulás, nyelvtanítás, értékelés. OM – PTMIK, Budapest – Pilisborosjenő
- Kósa Barbara és Simon Mária (2006, szerk.): *Új vizsga – új tudás? Az új érettségi hatása az iskolakezdéstől a záróvizsgáig*. Országos Közoktatási Intézet, Budapest.
- 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.
- Masters, G. N. (1982): A Rasch model for partial credit scoring. *Psychometrika*, **47**. 2. sz. 149–174.
- McNamara, T. (2000): *Language testing*. Oxford University Press, Oxford.
- Messick, S. (1996): Validity and washback in language testing. *Language Testing*, **13**. 3. sz. 241–256.
- Nagy József (1979): *Köznevelés és rendszerszemlélet*. OOK, Veszprém.
- 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.
- Rasch, G. (1960): *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research, Copenhagen.
- Saif, S. (2006): Aiming for positive washback: A case study of international teaching assistants. *Language Testing*, **23**. 1. sz. 1–34.
- 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.

- Wall, D. (2005): *The impact of high-stakes examinations on classroom teaching. Studies in Language Testing* 22. Cambridge University Press, Cambridge.
- Watanabe, Y. (2004): 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.

AZ ÉRTEKEZÉS TÉMAKÖRÉHEZ KAPCSOLÓDÓ PUBLIKÁCIÓK

- Nikolov M., Tóth E. és Vígh T. (2009): The impact of background factors on Hungarian students' performances on English and German reading comprehension tests. Tematikus előadás (absztrakt). In: Riljaarsdam, G. (szerk.): *Forstoring Communities of Learners. 13th Biennial Conference of the European Association for Research on Learning and Instruction*. Graduate School of Teaching and Learning, UvA, Amsterdam. 180.
- Vidákovich Tibor és Vígh Tibor (2009): Az idegen nyelvi értékelés és oktatás kapcsolata. In: Kozma Tamás és Perjés István (szerk.): *Új kutatások a neveléstudományban 2008. Hatékony tudomány, pedagógiai kultúra, sikeres iskola*. Magyar Tudományos Akadémia Pedagógiai Bizottsága, Budapest. 159–168.
- Vígh Tibor (2005): A kommunikatív tesztelés elméleti alapjai. *Magyar Pedagógia*, **105**. 4. sz. 381–407.
- Vígh, T. (2006): Der Prüfungsteil Schreiben im neuen Abitur für DaF. In: Orosz, M. és Albrecht, T. (szerk.): *Jahrbuch der ungarischen Germanistik 2005*. Gesellschaft Ungarischer Germanisten – Deutscher Akademischer Austauschdienst, Budapest – Bonn. 335–351.
- Vígh Tibor (2007a): A kommunikatív tesztelés az idegen nyelvi mérés és értékelés rendszerében. Tematikus előadás (absztrakt). In: Korom Erzsébet (szerk.): *PÉK 2007 – V. Pedagógiai Értékelési Konferencia: Program – Tartalmi összefoglalók*. SZTE, Neveléstudományi Doktori Iskola, Szeged. 78.
- Vígh Tibor (2007b): A vizsgák tanítási-tanulási folyamatra gyakorolt hatásának elméleti és empirikus kutatása. *Magyar Pedagógia*, **107**. 2. sz. 141–161.
- Vígh Tibor (2008a): A nyelvi szintek kialakításának és nyelvi feladatbank kiépítésének vizsgálata az idegen nyelvi érettségi vizsgán. Tematikus előadás (absztrakt). 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. 149.
- Vígh Tibor (2008b): A vizsgák tanítási-tanulási folyamatra gyakorolt hatásának modellezése. Tematikus előadás (absztrakt). In: Csíkos Csaba (szerk.): *PÉK 2008 – VI. Pedagógiai Értékelési Konferencia: Program – Tartalmi összefoglalók*. SZTE, Neveléstudományi Doktori Iskola, Szeged. 97.
- Vígh Tibor (2008c): Egy IRT-alapú nyelvi feladatbank létrehozásának módszertani kérdései. A német nyelvi érettségi vizsgafeladatok eredményei. *Magyar Pedagógia*, **108**. 1. sz. 29–51.
- Vígh, T. (2008d): Theoretische und empirische Erforschung von Rückwirkungsmechanismen des Abiturs für DaF. In: Tichy, E. és Masát, A. (szerk.): *Jahrbuch der ungarischen Germanistik 2007*. Gesellschaft Ungarischer Germanisten – Deutscher Akademischer Austauschdienst, Budapest – Bonn. 302–317.
- Vígh Tibor (2009a): Az íráskészség mérési-értékelési eljárásainak minőségi mutatói az emelt szintű német érettségiben. Tematikus előadás (absztrakt). In: Bárdos Jenő és Sebestyén József (szerk.): *IX. Országos Neveléstudományi Konferencia: Neveléstudomány - Integritás és integrálhatóság. Program és tartalmi összefoglalók*. Magyar Tudományos Akadémia, Pedagógiai Bizottság, Veszprém. 210–211.
- Vígh, T. (2009b): Der Prüfungsteil Schreiben im neuen ungarischen Abitur für Deutsch als Fremdsprache. In: Hunstiger, A. és Koreik, U. (szerk.): *Chance Deutsch: Schule – Studium – Arbeitswelt. 34. Jahrestagung des Fachverbandes Deutsch als Fremdsprache 2006 an der Leibniz Universität Hannover. Materialien Deutsch als Fremdsprache Band 78*. Universitätsverlag Göttingen, Göttingen. 149–165.
- Vígh Tibor és Vidákovich Tibor (2008): Az idegen nyelvi tesztelés és oktatás kapcsolata: a nyelvtudást mérő vizsgák visszahatásának elméleti és empirikus kutatása. Referátum (absztrakt). 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. 148.