

A VIZSGÁK TANÍTÁSI-TANULÁSI FOLYAMATRA GYAKOROLT HATÁSÁNAK ELMÉLETI ÉS EMPIRIKUS KUTATÁSA

Vígh Tibor

Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola

A visszahatás fogalmát és folyamatát a pedagógiában a kibernetika segítségével megalkotott rendszerszemléletű tanítás-tanulás modellek pontosan meghatározzák (Báthory, 2000; Nagy, 1979). A pedagógiai törekvéseknek megfelelő innovációk és reformok, amelyek 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 a vizsgák visszahatásának, hiszen a közoktatásban megfelelően működtetett vizsgák az iskolarendszer kontrollját is jelentik, valamint igen régi reformstratégia az oktatásban lezajló változások elindításához a mérés és értékelés megváltoztatása (Noble és Smith, 1994).

Magyarországon a megreformált és 2005-ben bevezetett érettségi vizsga hatásától a tanítási-tanulási folyamat tartalmi és módszertani modernizációját várják (Kósa és Simon, 2006; Einhorn, 2007; Petneki, 2007). Így joggal merül fel a kérdés, hogy az új érettséginek milyen a hatása a közoktatásra, illetve mennyiben szolgálja az oktatási folyamat modernizációját (Eckes, mtsai és Tsagari, 2005). Megfogalmazódott az igény az érettségi hatásvizsgálatára (Horváth és Lukács, 2006), fontos e jelenség pontos ismerete.

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. Ennek feltehetően az az oka, hogy ezen a területen fontos kérdés, hogy a különböző nyelvtudást mérő vizsgákra felkészítő kurzusok a vizsga sikere mellett a tanulónak használható nyelvtudást is biztosítsanak. Ugyanakkor a visszahatás mechanizmusainak bemutatását egyetlen hazai publikáció sem tűzte ki célul. Elsősorban a nyelvi mérés és értékelés területén jelenik meg kritériumként (Bárdos, 2002), illetve az idegen nyelvi tesztelés etikai kérdéseinél került bemutatásra (Dévény, 2003).

Tanulmányunk szakirodalmi elemzésen alapul, célja a pedagógiai értékelés és az 1990-es években elinduló nyelvi mérés és értékelés területén publikált elméleti és empirikus kutatások eredményeinek áttekintése. Az első részében értelmezzük a vizsgák visszahatásának fogalmát, a különböző koncepciókban betöltött szerepét. Ezt követően olyan elméleti modellek fejlődését mutatjuk be, amelyek a visszahatás folyamatát és

működését értelmezik. A harmadik részben 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, valamint arra is, hogy a kutatási eredményekből milyen következtetések vonhatók le. A tanulmány összegezésésként a vizsgák tanítási-tanulási folyamatra gyakorolt hatásmechanizmusainak jellemzőit fogalmazzuk meg.

A visszahatás fogalma

Bailey (1999) a vizsgák hatását leíró fogalmak értelmezése során megállapítja, hogy kutatónként eltérő az értelmezés, nincs egységes definíció a tesztek/vizsgák¹ hatásának meghatározására. Az angol nyelvű szakirodalomban (1) a *washback* és *backwash*, valamint (2) az *impact* fogalmakat használják². A vizsgák visszahatásának 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 gyakran a tesztek validitásával is összekapcsolják, és a kettő viszonyát (5) a visszahatás-validitás (*washback validity*), (6) a szisztematikus validitás (*systematic validity*), valamint (7) a következményes validitás (*consequential validity*)³ fogalmával határozzák meg. A továbbiakban a visszahatás egy-egy aspektusát megjelenítő kifejezéseket, koncepciókat tekintjük át.

A *backwash* az általános pedagógiai kutatásokban a teszt negatív visszahatását jelenti (Cheng és Curtis, 2004), a nyelvpedagógiai kutatásokban pedig semleges jelentéstartalommal, 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, 1997a, 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.)

¹ A tanulmányban a teszt fogalmát tágabb értelemben használjuk, amelyben a vizsga kifejezés szinonimája.

² A *backwash*, a *washback* és az *impact* kifejezéseket magyarul a visszahatás fogalmával fordítjuk, és az angol kifejezést alkalmazzuk, mivel az angol nyelvben jobban elkülöníthető a kifejezések jelentéstartalma.

³ A fogalom magyar fordítását Dévény (2003. 138. o.) alapján adtuk meg.

Bachman és Palmer (1996) csak az *impact* fogalmát használja, amelynek értelmében 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 visszahatás ezen értelmezéséhez nagyon hasonló *Shohamy* (1999) definíciója: „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.

Andrews (2004), valamint *Andrews, Fullilove és Wong* (2002) a washback fogalmát tág, az *impact* kifejezést pedig technikai értelemben az effect szinonimájaként használják: „a washback a tesztek tanításra és tanulásra, az oktatási rendszerre és az oktatási folyamatban résztvevőkre vonatkozik” (*Andrews*, 2004. 37. o.).

A visszahatás fogalmának különböző értelmezései egyrészt azt mutatják, hogy a terminológiai vita nem zárult le, másrészt azt is, hogy a kutatók a visszahatás fogalmát a korábbihoz képest tágabban értelmezik (*Bárdos*, 2002. 74. o.).

A vizsgák visszahatásának értelmezésében az értékelés által vezérelt oktatás (*measurement driven instruction*) koncepciója jelentős szerepet tölt be (*Popham*, 1987, 1993, 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 értékévé válik (*McEwen*, 1995). A koncepció elsősorban olyan standardizált tesztekre vonatkozik (például érettségi, felvételi vizsga), amelyeknek jelentős tétje, következménye van a tanulóknak, valamint a tanároknak és az iskolára (például a közölt teszt-eredményekből vonnak le következtetéseket az oktatás színvonalára), tehát a tesztek jelentős mértékben meghatározzák a tanítási-tanulási folyamatot.

A vizsgák visszahatásának 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 teszt-eredmények alapján is módosíthatják. *Chen* (2002a) az 1. táblázatban közölt módon összegzi a tantervi összehangolás két formáját: a megelőző (*alignment by frontloading*) és az utólagos (*alignment by backloading*) összehangolás jellemzőit.

1. táblázat. A megelőző és az utólagos tantervi összehangolás jellemzői (*English és Steffy*, 2001, idézi *Chen*, 2002a. 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.

A terminológiai viták és a visszahatás különböző értelmezése mellett fontos kérdéssé vált az is, milyen a teszt visszahatásának és validitásának a 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 (pl. *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 visszahatását rendeljék hozzá a validitáshoz, mivel a tanítási-tanulási folyamatban így érhető el pozitív irányú teszthatás. 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 tanítás és tanulás közötti viszony minőségét mutatja. A kritérium bevezetésének egyrészt az a célja, hogy a teszt-fejlesztőket 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 fejlődését 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. További kritikaként *Wall* (2005) megjegyzi, hogy a szisztematikus validitás elsősorban elméleti megközelítés, mivel e validitástípus nem mutatható ki empirikusan (*Alderson és Wall*, 1993).

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 visszahatásnak ezért magának a tesztnek a tanulás és a tanítás különböző aspektusaira gyakorolt hatására kell vonatkoznia (*Cheng és Curtis*, 2004), így a visszahatás pozitív vagy negatív értéke nem befolyásolhatja a teszt validitását.

Messick (1996) validitáskoncepcióban a következményes validitás (*consequential validity*) a teszt validitása és visszahatása közti összefüggéseket írja le. *Messick* (1996)

elképzelése szerint a következményes validitás egy aspektusa a szerkezeti validitásnak⁴, így sem a tesztelés következményei általában, sem a visszahatás önmagában 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) értelmezésében is a teszthatás (*impact*) a használhatóságnak (*usefulness*) az egyik összetevője⁵. *Messick* (1996), valamint *Bachman és Palmer* (1996) kritériumai tehát megegyeznek abban, hogy a teszthatás a szerkezeti validitás csak egyik aspektusát jeleníti meg.

A vizsgák hatásának fogalmi szintű értelmezéséről és a különböző koncepciókban betöltött szerepéről nagyon eltérő, és néhány esetben pedig egymásnak ellentmondó elképzelések alakultak ki. A jelenség pontos megismeréséhez szükséges a következő kérdések tisztázása: Pontosan 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? A hatások minden tanulóra, minden tanárra egyaránt érvényesek és azonos mértékűek (*Krekeker*, 2002)? A továbbiakban ezen kérdésekre keresünk válaszokat a visszahatás elméleti és empirikus kutatásainak bemutatásával.

A visszahatás 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 visszahatás 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.

Alderson és Wall (1993) a visszahatás jelenségének kutatásához 15 hipotézist fogalmazott meg, amelyek igazolják vagy elvetik a visszahatás létezését. Fontos megjegyezni, hogy 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. Azoknak a teszteknek lehet hatása, amelyeknek fontos következményei vannak, és így azoknak, amelyeknek nincsenek következményei, nem lehet hatása. A tesztek hatással lehetnek minden tanulóra és tanárra, vagy csak néhányra.

⁴ A szerkezeti validitás további aspektusai: (1) tartalmi, (2) lényegi, (3) szerkezeti, (4) általánosíthatóság, (5) külső és (6) a hatás aspektusa (magyar nyelven *Vígh* (2005) összegzi).

⁵ További összetevők a reliabilitás, a szerkezeti validitás, az autentikusság, az interaktivitás és a praktikusság.

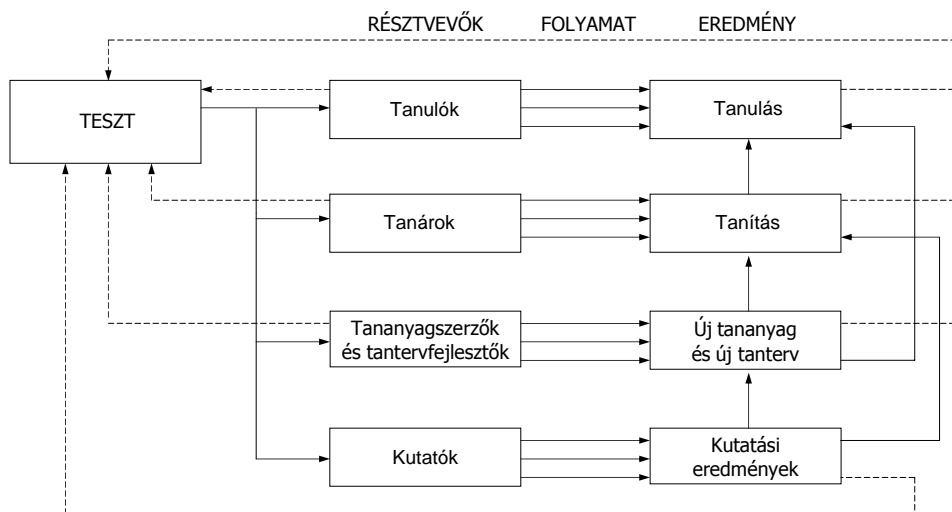
Az 1990-es évek második felében több kutató modellezte a visszahatás folyamatát. A továbbiakban az elméleti modellek fejlődését, változásait, illetve az egyes modellek közötti összefüggéseket mutatjuk be.

Hughes (1993, idézi *Bailey*, 1996. 262. o.) nem publikált munkájában bemutatott trichotóm modell (2. táblázat) a visszahatás alapmechanizmusát ábrázolja. A modell szerint a teszt először a résztvevők nézeteire és attitűdjeire van hatással, amelyek befolyásolják a tanítási-tanulási folyamatot, ezáltal meghatározzák a folyamat eredményét, a tanulás tartalmát és minőségét (*Bailey*, 1996).

2. táblázat. A visszahatás trichotóm modellje (*Hughes*, 1993, idézi *Bailey*, 1996. 262. o.)

- | |
|--|
| 1. résztvevők – tanulók, tanárok, adminisztrátorok, tananyag fejlesztők és készítők, akiknek a nézeteit és attitűdjeit a teszt befolyásolja. |
| 2. folyamat – a résztvevők minden olyan tevékenysége, amely hozzájárul a tanulás folyamatához. |
| 3. eredmény – a tanulás tartalma és minősége. |

Bailey (1996) visszahatás alapmodellje (1. ábra) kombinálja *Alderson* és *Wall* (1993) hipotéziseit és *Hughes* (1993) trichotóm modelljét (2. táblázat).



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

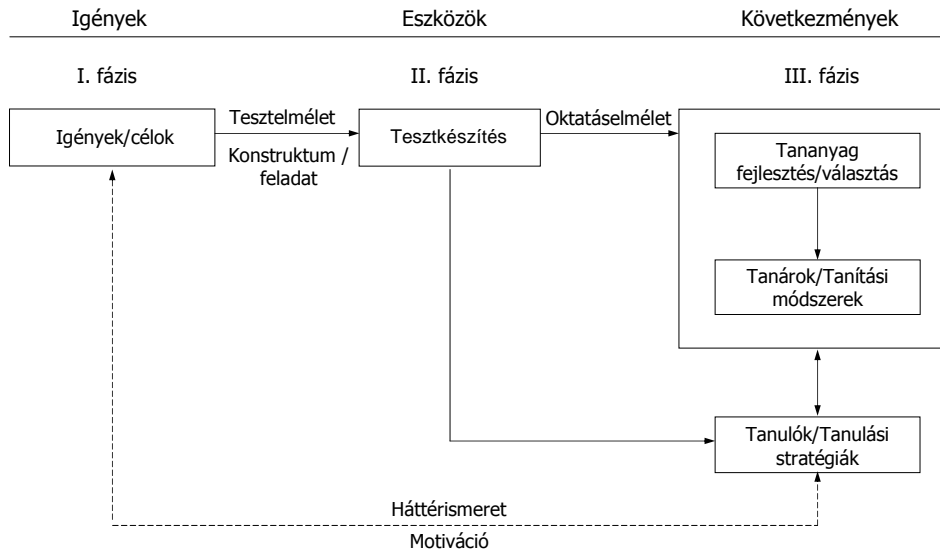
A modell – *Hughes* (1993) alapján – egyrészt három tényezőt nevez meg (résztvevők, folyamat és eredmény), másrészt *Alderson* és *Wall* (1993) hipotéziseit két csoportba sorolja. *Bailey* (1996) kétféle visszahatást különböztet meg. A tanulókra irányuló vissza-

hatás (*learner washback*) a teszteredmény által a tesztmegoldók felé közvetített információkra utal. A tanítási-tanulási folyamatra vonatkozó visszahatás (*program washback*) az oktatási rendszerrel közvetlen kapcsolatban állókra irányul, és a teszteredmény nekik szolgáltat információkat. 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 folyamaténező jellemzőinek leírását (*Tsagari, 2007*), valamint annak kölcsönhatását az eredményekkel (*Wall, 2005*), így valójában *Hughes (1993)* trichotóm modellje (2. táblázat) 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 elősegíti a visszahatás folyamatát. A modell jelentőségét a problémás pontok mellett mégis az adja, hogy a visszahatás folyamatát és működését ábrázolja.

Saif (1999) modelljében (2. ábra) egy újabb szempontot nevez meg, kiemeli az igényeknek és a céloknak megfelelő tesztfejlesztést a visszahatás folyamatában. A modell a visszahatás két tulajdonságát mutatja. Egyrészt azt ábrázolja, hogy figyelembe kell venni a társadalmi és az oktatással kapcsolatban állók igényeit, amely közvetlenül vagy közvetetten határozza meg a teszt típusát, célját és tartalmát, 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.

Saif (1999) modelljét a továbbiakban elsősorban a nyelvoktatásra vonatkoztatva értelmezzük. A modell alapján pozitív visszahatás 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 tartalma nem egyezik meg a tanítási és tanulási célokkal, a tanulók csak olyan képességeket alakítanak ki, amelyekkel a tesztet meg tudják oldani, de valós nyelvi szituációban nem tudnak hatékonyan cselekedni (*Krekeler, 2002, 2005*). Ennek magyarázata, hogy a nyelvoktatás középpontjában a vizsgán előforduló feladattípusok és feladatmegoldó stratégiák tanítása áll (*Krumm, 2006; Perlmann-Balme, 2006; Prodromou, 1995*), tehát a vizsgára való felkészítés „rejtett” tantervvé válik (*Glaboniat és Müller, 2006*). *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. Tehát az ebben az értelemben vett jó teszteknek valóban pozitív a hatásuk, a rossz teszteknek pedig negatív, vagyis a jó szerkezeti validitású teszteknek pozitív a visszahatása (*Messick, 1996*)? Illetve, ami a hazai kontextusban fontos kérdés: a kommunikatív tesztelés alkalmazása kommunikatív nyelvoktatást eredményez (*Alderson, 1998*)?



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

E kérdés megválaszolására *Burrows* (2004) koncepciója alapján bemutatjuk a teszt és annak hatása közötti összefüggésekről alkotott elképzelések változásait. A tradicionális visszahatás-elmélet (3.a ábra) alapján egy új teszt bevezetése szükségszerűen vezet visszahatáshoz, a tanítási-tanulási folyamat irányítóiból, a tanárokból egyszerű és egyforma reakciót vált ki. A visszahatás lehet pozitív vagy negatív irányú, és ez csak a teszt minőségétől, nem a tanárok felfogásától, attitűdjeitől függ. Ennek a tradicionális felfogásnak 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 fejlesztendő készségek mérése; (2) a mérendő készségek széles körű kiválasztása; (3) közvetlen és (4) kritériumorientált tesztelés alkalmazása; (5) a tesztelés alapelveinek megismertetése a tanárokkal.

A *Burrows* (2004) által „fekete doboznak” nevezett modell (3.b ábra) már abból indul ki, hogy összetettebb a visszahatás folyamata. 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 nézetei, feltételezései és ismeretei határozzák meg. A tanárok ugyanis egy új teszt bevezetésére *egyéni* tanítási reakciókat adnak. *Alderson és Wall* (1993) feltételezését kutatási eredmények is alátámasztották. *Wall és Alderson* (1993) azt találták, hogy a tanárok azonos típusú tesztre különböző, míg különböző tesztre nagyon

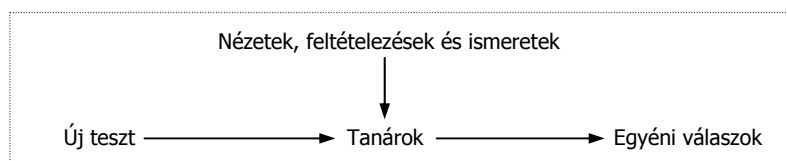
hasonló módon készítették fel diákjaikat. Ennek feltehetően az az oka, hogy a tesztelés alapelve, „filozófiája” többnyire rejtve marad, ezért szükséges a tanárok bevonása a tesztfejlesztés és a lebonyolítás folyamatába (Alderson, 1999). Ez különösen fontos, hiszen a tesztfejlesztők által szándékolt visszahatás jelentősen függ a tanároknak a teszt bevezetésére adott egyéni reakcióitól, amely meghatározza a tanárok vizsgával kapcsolatos attitűdjeit (Turner, 2001), a tanítási-tanulási folyamatot, és ebből adódóan a teszt visszahatását. A modell előnye tehát, hogy a korábbiakhoz képest felhívja erre az egyéni tényezőre a figyelmet, amely azonban a visszahatás kutatását jelentősen megnehezíti.

Burrows (2004) empirikus kutatása során csoportosította a tanárok reakcióit. A vizsgálat eredményeként alkotott modell (3.c ábra) azt ábrázolja, 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 tanítási-tanulási folyamatot leíró innováció modellezésére.



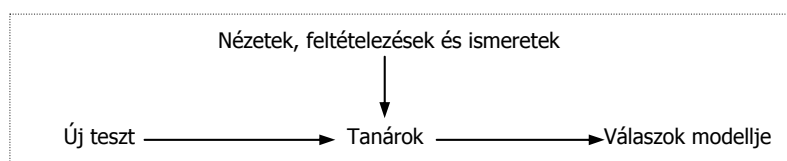
3.a ábra

Tradicionális visszahatás-elmélet: az inger válasz modell



3.b ábra

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



3.c ábra

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

A vizsgák hatásmechanizmusai a közoktatásban elinduló innovációval is jellemezhetőek (Markee, 1993; Wall, 2005). Markee (1993) innovációs modellje 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).

Markee (1993) arra is utal, hogy az innováció nem csak a tanítási-tanulási folyamatra vonatkozik, valamint minőségét jelentősen meghatározza az oktatáspolitikai. *Andrews* (2004) az alábbi javaslatokat teszi a döntéshozóknak a vizsga helyes megváltoztatásának és a pozitív visszahatás elérése érdekében:

- kerüljük a kényszerítő körülményeket,
- vegyük figyelembe az innováció komplexitását és előre nem megjósolható hatásait, 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 értelmeik a visszahatás folyamatának működését, valamint azt is, hogy hogyan érhető el visszahatás. Az elméleti megközelítések mindegyikében közös elem, hogy a hatást mint a tanítási-tanulási folyamatban és az oktatási rendszerben bekövetkező változásokat (*Cheng*, 1997a, 2005) ábrázolják. *Wall* (1996, 2005) ugyanakkor rámutat arra, hogy ezek a modellek nem tartalmazzák azokat a dimenziókat és faktorokat, amelyek a visszahatás folyamatának minőségét alapvetően befolyásolják.

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

- *sajátosság*: általános (olyan tesztelés, 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 magyarázza, a tanárok különbözőképpen tanítanak azonos teszt esetén);
- *hosszúság*: rövid (a visszahatás csak addig figyelhető meg, amíg a tanulók felkészülnek a tesztre), hosszú (a hatás a teszt letétele után is tapasztalható);
- *szándék*: a tesztfejlesztők által szándékolt és nem szándékolt hatás (ezt a dimenziót a tesztfejlesztés időszakában pontosan meg kell jelölni, hogy legyen visszahatás);
- *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:

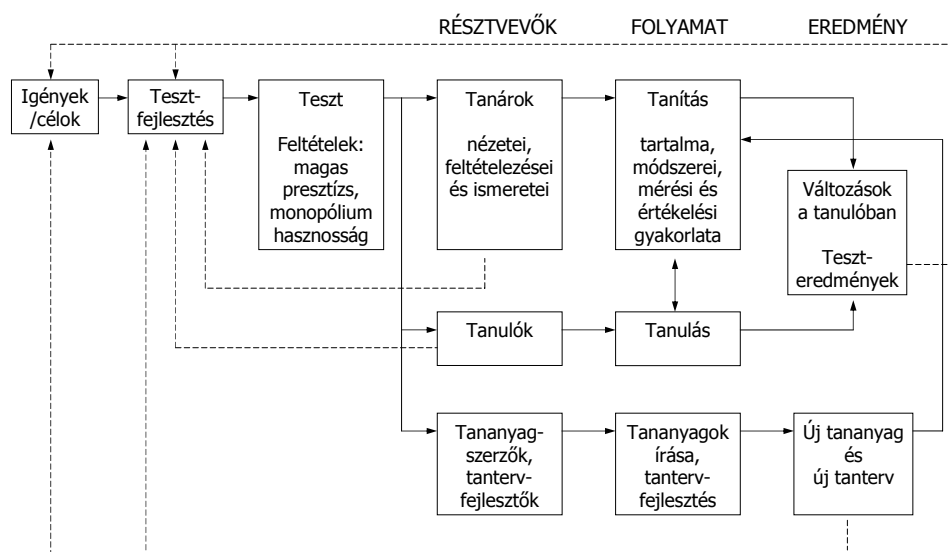
- *átláthatóság*: minél közelebb áll a tesztkonstrukció a valós igényekhez, annál erősebb a hatása (lásd *Saif* (1999) modelljét, 2. ábra);
- *hasznosság*: minél több lehetőséget biztosít a letett vizsga, annál erősebb a hatása;
- *monopólium*: minél kevesebb versenytársa van az adott vizsgának, annál erősebb a hatása;
- *félelem*: minél fontosabb a tanuló számára, hogy letegye a vizsgát, annál erősebb a hatása.

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

- *tesztfaktorok*: tesztelési módszerek, tartalmak, a mért készségek, a teszt céljai, azok a döntések, amelyeket a teszteredmények alapján hoznak;

- *presztízs*faktor: a teszt fontossága, státusza a közoktatási rendszeren belül;
- *személyes* faktorok: a tanárok képzettsége, nézeteik a tanítási-tanulási folyamatról;
- *mikrokörnyezeti* faktorok: az iskola felszereltsége, amelyben az oktatás, felkészítés zajlik;
- *makrokörnyezeti* faktorok: az a társadalom, amelyben a tesztet alkalmazzák.

A továbbiakban javaslatot teszünk egy szintetizáló modellre (4. ábra), amely a tesztnek a tanítási-tanulási folyamatra gyakorolt hatását ábrázolja. A modell figyelembe veszi a korábban ismertetett elméleti modelleket és a visszahatás jellegét leíró dimenziókat és faktorokat is, így szemléltetve a tesztek visszahatásának működését.



4. ábra
A visszahatás működését szintetizáló modell

A modell *Gates* (1995) és *Watanabe* (2004a) alapján tartalmazza azokat a fontosabb feltételeket (például a teszt magas presztízse, hasznossága, monopóliuma), amelyek elengedhetetlenek a visszahatás folyamatának elindulásához. *Saif* (1999, 2006, 2. ábra) modellje alapján magában foglalja az igényeknek és a céloknak megfelelő tesztfejlesztési folyamatot. A szintetizáló modellben a folytonos vonal *Hughes* (1993, 2. táblázat) és *Bailey* (1996, 1. ábra) modellje alapján a teszt közvetlen és közvetett hatását mutatja a teszttel kapcsolatban álló személyekre (a felkészítő tanárookra, a tanulókra, valamint a tankönyvszerzőkre és a tantervfejlesztőkre), a tanítási-tanulási folyamatra (a tanítás tartalmára, a módszerekre, az értékelési kultúrára), valamint annak eredményére (a tanulás tartalmára és minőségére), amit a tesztben mérünk. A szintetizáló modell tartalmazza a *Burrows* (2004, 3. c ábra) által megnevezett tanári tényezőt is, hiszen a tanárok nézetei, feltételezései és ismeretei alapvetően befolyásolják a visszahatás folyamatát és irányát. Mindezek mellett tartalmazza a tantervi összehangolás koncepcióját is (1. táblázat), tehát

a teszt, a tanítás és a tanterv közötti összefüggéseket. A szaggatott vonal a tesztfejlesztési folyamatra és a célokra irányuló visszahatást illusztrálja, amit elsődlegesen a tesztteredmények határoznak meg. Ezen kívül különböző felmérésekkel a tanítási-tanulási folyamatban résztvevők tesztel kapcsolatos meggyőződéseiről, attitűdjeiről is kaphatunk információkat, amelyek befolyásolhatják a tesztfejlesztési folyamatot, illetve elősegíthetik a tesztteredmények megfelelő értelmezését is. A különböző visszajelentési körökből szerzett pedagógiai információk így elősegítik az oktatás eredményességének növelését, az innovációs folyamatok elindítását.

A visszahatás olyan komplex jelenség, amelyet sok esetben egyéni tényezők is befolyásolnak. Szükséges tehát olyan empirikus kutatások elvégzése, amelyek segítségével megismerhetjük a visszahatás természetét és mechanizmusait.

A visszahatás folyamatát leíró empirikus kutatások

Alderson és Wall (1993) visszahatás-hipotézisei az empirikus kutatások elindításában is irányadóak voltak, amelyeket először *Wall és Alderson* (1993) vizsgáltak. Egy Sri Lankán végzett kutatásban a módosított alapszintű angol vizsga hatását elemezték, amely az olvasás és írás mérésére és értékelésére vonatkozott. 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, hat megfigyelési szakaszt tartalmazó kutatás eredményei szerint a vizsga hatást gyakorolt a tananyagra, a tanítás tartalmára, mérési és értékelési gyakorlatára. A vizsgának azonban nem, vagy csak nagyon kevés hatása volt a tanárok tanítási módszereire.

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*, 1997a, 1997b, 1999, 2004, 2005; *Shohamy, Donitsa-Schmit és Ferman*, 1996; *Shohamy*, 2001; *Stecher, Chun és Barron*, 2004; *Andrews és mtsai*, 2002; *Wall és Alderson*, 1993), egyetemi felvételi (*Watanabe*, 1996a, 1996b), valamint bevándorlók számára indított nyelvi kurzusok eredményét igazoló vizsgák (*Burrows*, 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 meg-

figyelésen 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*, 2002a, 2002b; *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*, 1997a) szerint a vizsgák jelentősen befolyásolták azt, mely nyelvi készségek tanítására fektettek nagyobb hangsúlyt a tanárok. *Shohamy és mtsai* (1996) tanulmánya azonban ettől eltérő képet mutat, mivel az eredményekből kitűnik, hogy az alacsony presztízsű vizsgának a tanítás tartalmára kevés hatása volt. *Watanabe* (1996b) ugyanakkor azt találta, hogy a tanárok nem feltétlenül tanították nagyobb hangsúllyal a vizsgán mért készségeket (például hallás és olvasás). *Nikolov* (1999) osztálytermi megfigyelésre és kérdőíves vizsgálatra 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 behelyettesítés nyelvtani gyakorlatok szerepeltek.” (*Nikolov*, 1999. 28. o.) A fordítás és a kitöltéses feladatok tipikus hagyományos szemléletű vizsgafeladatok, ezért joggal feltételezték a vizsgák visszahatását ezen a területen.

A tananyagokra gyakorolt hatás. A tanulmányok többsége szerint a vizsgáknak jelentős hatása volt arra, hogy a tanárok milyen tananyagot használnak mind a tanítás, mind a vizsgára való felkészítés során. Az is kiderül, hogy a tananyagok megváltoztak az új vizsga bevezetésének hatására (*Cheng*, 1997a; *Andrews és mtsai*, 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*, 1997a, 2003; *Shohamy és mtsai*, 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 nagyon eltérő. *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 visszahatásnak ezt a területét kevés kutatásban vizsgálták. *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 tanulók érzelmeire, attitűdjeire gyakorolt hatás. Az eredmények alapján megállapítható, hogy a tanárok sokszor aggódtak a vizsga eredménye miatt. *Shohamy* és *mtsai* (1996) kutatása szerint néhány tanár úgy érezte, hogy a tanulók sikere vagy kudarca a tanításuk minőségét tükrözi. A tanároknak sokszor voltak negatív attitűdjeik az új, vagy módosított vizsgával kapcsolatban (például *Alderson* és *Hamp-Lyons*, 1996; *Cheng*, 1997a). 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 teszt-eredményeket vizsgáló kutatások hiányoznak, amelyek bemutatják, hogy a tanulók többet vagy jobban tanultak-e egy bizonyos tesztre” (502. o.). A kevés kutatási eredmény 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 annál is inkább fontos, mivel 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) visszahatás-hipotéziseire 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 készítenek 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 3. táblázatban közölt, *Spratt* (2005) által összeállított tényezők – *Burrows* (2004) modelljéhez hasonlóan – felhívják a figyelmet a tanárok vizsgával kapcsolatos (1) meggyőződéseinek, (2) attitűdjeinek, valamint (3) képzettségének és (4) személyiségének fontosságára a visszahatás folyamatában.

3. táblázat. A visszahatás mértékét és típusát befolyásoló tényezők (Spratt, 2005. alapján)

Tanárokra vonatkozó faktorok	Források, az iskola, a vizsga
<p><u>A tanárok meggyőződése</u></p> <ul style="list-style-type: none"> – a vizsga megbízhatóságáról és tökéletességéről – a vizsga presztízsről 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. Mivel több publikáció (például *Chen*, 2002a; *Cheng*, 1997a; *Saville* és *Hawkey*, 2004) közli a kutatás eszközeit, valamint a kutatások pontos lebonyolítását (*Cheng*, 1997a; *Wall* és *Alderson*, 1993), ezek új kutatási kontextusban alkalmazhatók, így az eredmények összehasonlíthatók. A visszahatás jelenségének komplexitása miatt nagyon fontos egyrészt a kvalitatív és kvantitatív eljárások kombinálása, másrészt a mérőeszközök validálása. A vizsgák tanítási-tanulási folyamatára gyakorolt hatásának vizsgálatához *Alderson* és *Banerjee* (2001) tanácsokat fogalmaz meg, ezáltal elősegítik, hogy a jelenség folyamatának, működésének és mechanizmusainak feltárása során megbízható adatokat kapjunk.

Összegzés

A tanulmányban szakirodalmi áttekintés alapján értelmeztük a visszahatás fogalmát, valamint elméleti és empirikus kutatások eredményeinek felhasználásával bemutattuk a vizsgák hatásmechanizmusainak jellegét és működését, illetve vizsgálatának lehetőségeit.

A vizsgák hatásának elméleti és empirikus kutatását az 1990-es évek kezdete óta a nyelvi tesztelés területén *Alderson és Wall* (1993), valamint *Wall és Alderson* (1993) tanulmányainak hatására számos kutató tűzte ki célul, így igen gazdag angol nyelvű szakirodalma van. Ezt az is mutatja, hogy például a *Language Testing* (1996) és az *Assessment in Education* (2007) folyóirat különszámában közölt tanulmányokat a témáról. Ezen kívül számos tanulmánykötet (*Bailey*, 1999; *Cheng*, 2005; *Wall*, 2005), doktori disszertáció (például *Chen*, 2002a; *Cheng*, 1997a; *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. Ezzel szemben a hazai szakirodalomban a vizsgák hatásmechanizmusainak meghatározására jelentőségéhez képest jóval kevesebb figyelem irányult.

A szakirodalomban a vizsgák hatásának fogalmi szintű megnevezése nem egységes, ugyanazt a jelenséget többféleképpen, más-más jelentéstartalommal definiálják a kutatók. A fogalommeghatározások alapján a visszahatás két típusa különböztethető meg: a vizsgák hatásmechanizmusai *szűkebb értelemben* a tanárookra és a tanulókra, a tanítási-tanulási folyamatra vonatkozik, *tágabb értelemben* pedig a közoktatási rendszerre, illetve a társadalomra gyakorolt hatásról van szó.

A szűkebb és a tágabb értelemben vett visszahatás folyamatát leíró elméleti modellek és a vizsgák tanítási-tanulási folyamatra gyakorolt hatását vizsgáló empirikus kutatások eredményei alapján a vizsgák hatásmechanizmusa komplex jelenség, amely közvetlen és közvetett hatást fejt ki a teszttel kapcsolatban álló személyekre, a tanítási-tanulási folyamatra, valamint a tanulás tartalmára és minőségére. Ez a hatás előre meghatározhatatlan, mivel számos tényezőtől függ, melyek közül meghatározó jelentősége van a tanárok egyéni nézeteinek, feltételezéseinek és ismereteinek. A hatás iránya pozitív vagy negatív lehet, amelynek értéke nincs hatással a teszt validitására. A tanulmányban javasolt szintetizáló modell (4. ábra) elsősorban a folyamattényező pontosabb leírását tartalmazza a visszahatás működésében. Az empirikus kutatások eredményei ezen a területen alátámasztják a modell relevanciáját.

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.

A szerző a tanulmány kéziratának elkészítésekor a *Deutscher Akademischer Austauschdienst* támogatásával az Universität Kassel és a Szegei Tudományegyetem közötti *Germanistische Institutspartnerschaft* projekt keretében tanulmányi és kutatói ösztöndíjban részesült.

Irodalom

- 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. (1998): Testing and teaching: the dream and the reality. *Novelty*, 5. 4. sz. 23–37.
- 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>, 2007. június 11.
- 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. és mtsai. (szerk.): *Experimenting with uncertainty: essays in honour of Alan Davies*. Cambridge University Press, Cambridge. 150–161.
- 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 Wall, D. (1993): Does washback exist? *Applied Linguistics*, 14. 2. sz. 115–129.
- 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.
- Bachman, L. F. és Palmer, A. S. (1996): *Language testing 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.
- Bárdos Jenő (2002): *Az idegen nyelvi mérés és értékelés elmélete és gyakorlata*. Nemzeti Tankönyvkiadó, Budapest.
- Báthory Zoltán (2000): *Tanulók, iskolák – különbségek. Egy differenciális tanításmélet vázlatja*. OKKER Kiadó, 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. (2002a): *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.
- Chen, L. (2002b): *Washback of a public Exam on English Teaching*. ERIC Document ED472167. ERIC Clearinghouse on Assessment and Evaluation, Washington DC.

- Cheng, L. (1997a): *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. (1997b): How does washback influence teaching? Implications for Hong Kong. *Language and Education*, **11**. 1. sz. 38–54.
- 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.
- Dévény Ágnes (2003): Az idegen nyelvi tesztelés etikai problémái. *Új Pedagógiai Szemle*, 7–8. sz. 134–145.
- Einhorn Ágnes (2007): *Az idegen nyelvi érettségi vizsga reformja*. In: Vágó Irén (szerk.): *Fókuszban a nyelvtanulás*. Oktatókutató és Fejlesztő Intézet, Budapest. 73–105.
- 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.
- 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.
- 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.
- 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.
- Kósa Barbara és Simon Mária (2006, szerk.): *Új vizsga – új tudás? Az új érettségi hatása az iskolakezdéstől a záróvizsgáig*. Országos Közoktatási Intézet, Budapest.

A vizsgák tanítási-tanulási folyamatra gyakorolt hatásának elméleti és empirikus kutatása

- 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.
- Leitzel, T. C. és Vogler, D. E. (1994): *Curriculum Alignment: Theory to Practice*. ERIC Document ED371812. 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.
- McEwen, N. (1995): Educational Accountability in Alberta. *Canadian Journal of Education*, **20**. 1. sz. 27–44.
- Messick, S. (1996): Validity and washback in language testing. *Language Testing*, **13**. 3. sz. 241–256.
- 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 (1979): *Köznevelés és rendszerszemlélet*. Országos Oktatástechnikai Központ, Veszprém.
- Nikolov Marianne (1999): Osztálytermi megfigyelés átlagos és hátrányos helyzetű középiskolai angol csoportokban. *Modern Nyelvoktatás*, **5**. 4. sz. 9–31.
- 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.
- 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 (2007): *Az idegen nyelvek oktatása Magyarországon az ezredfordulón*. JATE-Pressz, Szeged.
- 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.
- Saif, S. (1999): *Theoretical and Empirical Considerations in Investigating Washback*. A study of ESL/EFL learners. Unpublished 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.
- 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.

- 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.
- 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
- Tschirner, E. (2001): Die Evaluation fremdsprachlicher mündlicher Handlungskompetenz: Ein Problemaufriss. 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. és mtsai (szerk.): *Experimenting with uncertainty: essays in honour of Alan Davies*. Cambridge University Press, Cambridge. 138–149.
- Vígh Tibor (2005): A kommunikativ tesztelés elméleti alapjai. *Magyar Pedagógia*, **105**. 4. sz. 381–407.
- 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.

ABSTRACT

TIBOR VÍGH: THEORETICAL AND EMPIRICAL RESEARCH ON THE WASHBACK-EFFECT OF EXAMS ON TEACHING AND LEARNING

Washback, also known as backwash in the educational literature, is a common term in applied linguistics referring to the influence of testing on teaching and learning and, in a broader sense, to its impact on the educational system or society. This paper gives an overview of the discussion of this phenomenon from different perspectives, both in general education and in language education. Part 1 discusses the definition of washback and its role in measurement-driven instruction, in curriculum alignment and in different interpretations of validity. Part 2 describes the function and mechanisms of washback in the development of theoretical models. As a conclusion, a synthesising model is offered, aiming to demonstrate the functioning and mechanism of washback. Part 3 reviews results of empirical studies from different cultures in order to show the methodology of research on washback and the areas affected most by a test. Also discussed are the problems that emerge when findings are generalised and when results are transformed into conclusions. Part 4 highlights the lessons from the theoretical and empirical literature reviewed. It is obvious that washback has often been discussed and that there are different points of views on what the construct comprises. Washback is a complex phenomenon that has a direct or indirect effect on the perceptions and attitudes of the participants (teachers and students, developers of instructional materials and curriculum designers) on the teaching and learning process (on the content of teaching, methods, and assessment practices) and on the quality of learning. The effect is elusive, as it depends on a number of aspects, the most important of which are teacher-related factors: their beliefs, attitudes, assumptions and knowledge. Washback can be positive or negative and its strength is independent of test validity. The study of washback requires mixed methodology, involving both qualitative and quantitative research techniques.

Magyar Pedagógia, **107**. Number 2. 141–161. (2007)

Levelezési cím / Address of correspondence: Vígh Tibor, Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola, H–6722 Szeged, Petőfi S. sgt. 30–34.