

## A KÉPESSÉGEK FEJLŐDÉSI ÜTEMÉNEK EGYSÉGES KIFEJEZÉSE: A GAMMA KOEFFICIENS

**Csapó Benő**

*Szegedi Tudományegyetem, Neveléstudományi Tanszék  
MTA-SZTE Képességkutató Csoport*

A képességek fejlesztésének tervezése, a fejlődés követése, a lemaradások diagnosztizálása szükségessé teszi a fejlődési folyamatok részletes megismerését. Az iskolában számos olyan képesség fejlesztése folyik, amelynek a fejlődése több éven át tart, miközben az egyes tanulók között olyan mértékű különbségek alakulhatnak ki, amelyek több évnyi fejlődésnek felelnek meg. Az ilyen képességek vizsgálata során lehetőség van arra, hogy különböző életkorú tanulók fejlettségét ugyanazzal a tesztekkel mérjük fel, azaz a fejlődést *keresztmetszeti vizsgálatokkal* térképezzük fel.

Az iskolai gyakorlatban még nem terjedtek el a széles életkori intervallumban használható tesztek, a pedagógiai kutatásnak azonban régóta alapvető eszközei. A nemzetközi összehasonlító felmérések (például az IEA TIMSS és az OECD PISA; lásd: *Beaton, Martin, Mullis, Gonzalez, Smith és Kelly*, 1996; OECD, 2001) az utóbbi években eltávolodtak az iskolában közvetlenül elsajátított tananyag felmérésétől; így az e projektek keretében kidolgozott tesztek is fel lehet használni szélesebb életkori intervallumban.

A keresztmetszeti felmérések tesztjei általában különböző egységekben fejezik ki a tanulók teljesítményeit. (Például nyerspont, százalékpont, standard skálák, a modern tesztelmélet eszközeivel készített skálák stb.) Ezek a megoldások azonban – bár jól használhatók egy-egy képesség vizsgálata során – nem teszik lehetővé a különböző tesztekkel végzett fejlődési folyamatok összehasonlítását. Az empirikus eredmények felhalmozódásával a társadalomtudományokban is mind több területen jelentkezik az igény az eredmények szintetizálására, ezért az adatelemzési eljárások is mindinkább egységesülnek.

E tanulmányban egy olyan számítási eljárásra teszünk javaslatot, amely lehetőséget nyújt a keresztmetszeti vizsgálatok során mért fejlődés mértékének egységes kifejezésére, és így a különböző eszközökkel kapott adatok összehasonlítására. A bevezetett számítási eljárást értelmezzük és elhelyezzük a statisztikai eljárások kontextusában, majd különböző fejlődés-vizsgálatok adatainak összehasonlító elemzésével illusztráljuk alkalmazását. Végül értékeljük használhatóságát, összegezzük előnyeit és korlátjait, és megfogalmazzuk a további elemzések feladatait.

## A fejlődés egységes mértéke: a gamma koefficiens

### A hatás-méret közlésének elterjedése

A keresztmetszeti vizsgálatok során mért fejlődés egységes mértékének kifejezésére egy olyan mutató bevezetését javasoljuk, amely statisztikai természetét tekintve megfelel a társadalomtudományi kutatásban mind nagyobb szerepet játszó *hatás-méretnek* (az effektus nagysága, *effect size*). A hatás-méret előtérbe kerülésének, felértékelődésének számos közvetlen oka van. Szükségessége a kutatás különböző területein felmerült, de ha összefoglalóan jellemezni akarjuk felértékelődésének indokait, azt mondhatjuk, hogy a társadalomtudományok terén elért eredmények összehasonlíthatóságának, szintetizálhatóságának, integrálhatóságának igényei vezettek elterjedéséhez.

Egyrészt több statisztikai elemzést használó tudományág – közöttük a pszichológia számos területe – dolgozik viszonylag kis mintákkal. Ilyen esetekben különösen fontos annak megmutatása, hogy az adott effektus statisztikailag szignifikáns. Az eredmények szignifikanciáját alapvetően két tényező befolyásolhatja: az elemzésbe bevont minták mérete és a megfigyelt effektus mérete (például két minta átlaga közötti különbség mértéke). Ha viszonylag kis minták esetén megfelelő valószínűségi szinten a vizsgált effektus statisztikailag szignifikánsnak bizonyul, az általában egyben azzal is jár, hogy a megfigyelt effektus elér egy már gyakorlatilag is jelentős mértéket. Részben ennek tulajdonítható, hogy számos területen szokássá vált a kutatási eredmények olyan jellegű publikálása, amely csak a szignifikanciavizsgálat (hipotézisvizsgálat) eredményeit közli. Ez a megoldás azonban nem tette lehetővé az effektus nagyságának összehasonlítását más hasonló vizsgálatokban talált értékekkel. E bizonytalanság az érintett tudományágak módszertani kérdésekre érzékeny kutatóinak körében évtizedek óta vita tárgya volt. Az egyik álláspont képviselői még azt a javaslatot is felvetették, hogy a tudományos folyóiratok ne közöljenek többé szignifikancia-tesztet. Az *American Psychological Association* (APA) 1996-ban egy munkacsoportot bízott meg a kérdés vizsgálatával, amely végül 1999-ben fogalmazta meg ajánlásait. A munkacsoport nem javasolta a szignifikancia-tesztek elhagyását, viszont határozottan ajánlotta a hatás méretét jellemző mutatók közlését (lásd *Thompson*, 2002).

Másrészt az egyes kutatási területeken halmozódó eredmények nyomán mind gyakoribbá váltak az összegző („*review*”-típusú) tanulmányok, és felmerült az igény az eredmények kvantitatív integrálására. Egyes markáns elemzési technikákkal rendelkező területeken kezdett terjedni a meta-analízis, amely a különböző publikációkban közölt adatok másodelemzése révén az eredmények számszerű összegzésére alkalmas. Az ilyen elemzés célja lehet például annak megállapítása, hogy azonos jellegű kísérletek során átlagosan milyen fejlesztő hatást tapasztaltak. Mivel azonban a különböző vizsgálatok más-más eszközökkel mérik, és esetleg a publikációk különböző formában fejezik ki a fejlődés mértékét, meg kellett találni az egységes, standard kifejezési formát. Erre alkalmasnak bizonyult a hatás-méret, ezért elvárássá vált azoknak az adatoknak a közlése, amelyek a hatás-méret kiszámításához szükségesek, hogy ezáltal az adott publikációban közölt eredmények egy későbbi meta-analitikus szintézisben felhasználhatók legyenek.

Az 1990-es évek eleje óta pedig elvárás egyes kutatási területek – például a fejlesztő kísérletek – eredményeinek közvetlen hatás-méretben való megadása. Különösen fontossá vált a meta-elemzés a pedagógiában (*Kulik és Kulik, 1989*), ahol nagyon sok egymástól jelentősen különböző eszközzel végzett felmérés, kísérlet eredményeinek szintetizálására van szükség. Sok olyan meta-analitikus tanulmány jelent meg, amely a készségek, képességek fejlesztésével kapcsolatos kísérletek eredményeit szintetizálja (lásd például *Goossens, 1992; Klauer, 2001*).

A meta-analízis alkalmazásának különösen nagy lökést adott *Cohen* könyve, amely számos statisztikai elemzéssel kapcsolatban részletesen bemutatja a hatás-méret fogalmát, számítási módját (standard hivatkozássá a második kiadás vált, lásd *Cohen, 1988*). A hatás-méret *Cohen* eredeti értelmezésében általánosan megfogalmazva a null-hipotézistől való eltérés mértéke, konkrétan pedig az egyes statisztikai próbákhoz használt próbafüggvényből lehet kiemelni a hatás-méretet jellemző kifejezést. A hatás-méret egy dimenzió nélküli szám, lényegében egy arányszám. A számunkra a következőkben lényeges hatás-méret a két minta átlaga közötti különbség jellemzésére szolgáló *t*-próbaéhoz kapcsolódik. Ha mindkét mintát ugyanabból a normális eloszlású populációból vettük, akkor feltehetően szórásuk is megegyezik, így az átlagok különbségét jellemző hatás-méret (amit *Cohen* nyomán *d*-vel jelölnek) egyszerűen a két átlag különbségének és a (közös) szórásnak a hányadosa.

A hatás-méret tehát ebben az esetben a különbség standard mértéke: különbség szórásegységekben kifejezve. A gyakorlati számítási eljárások ettől az egyszerű alapelvtől eltérőek lehetnek, attól függően, hogy a konkrét esetben mennyiben különbözhetnek a körülmények az előzőekben vázolt feltevésektől, és ezeket az eltéréseket hogyan kívánjuk figyelembe venni, korrigálni. Például ha nem egyenlő a szórás a két mintában, akkor pontosabb becslést ad a szereplő szórások átlaga, mint ha csak egyik szórást használnánk. Ha a minták nagysága sem egyenlő, akkor súlyozott átlagolást kell végezni. Egy olyan bonyolult összehasonlításnál pedig, amikor egy kísérleti és egy kontroll-csoportban végbement fejlődést kívánunk elemezni – mivel mindkét csoportban egy elő- és egy utó-mérés eredményeinek a különbsége adja a fejlődést –, négy szórás figyelembe vételére lenne szükség. Néha az igényesebb elemzések használják a bonyolultabb képleteket, viszont ha a minták méretei kellően nagyok (pl. 30-nál nagyobbak) és a méretük nem különbözik lényegesen, akkor a különböző képletekkel kapott hatás-méretetek nem különböznek érdemben egymástól. Ezért terjed például az a (konzervatívabb) megoldás is, hogy a kísérleti hatás mértékét a kontroll csoport második (ezért várhatóan kissé nagyobb) mérésének szórásában fejezik ki.

### A gamma koefficiens

A keresztmetszeti vizsgálat során mért fejlődés jellemzésére egy olyan standard mértékegység bevezetését javasoljuk, amelyik alkalmas arra, hogy a különböző tesztekkel végzett fejlődés-vizsgálatok eredményeit egymással összehasonlítsuk. Ezt a mértéket *gamma koefficiensnek*, vagy egyszerűen *gammának* ( $\gamma$ ) nevezzük. A gamma egyfajta grádiens, a fejlődést leíró függvény meredekségét jellemzi.

A *gamma* definíciójaként a következő képletet javasoljuk:

$$\gamma = \frac{(\bar{x}_2 - \bar{x}_1)}{((s_1 + s_2) / 2) \cdot T}$$

ahol

$x_1$  és  $x_2$  az első és a második időpontban mért teljesítmények átlaga,  
 $s_1$  és  $s_2$  az első és a második időpontban mért teljesítmények szórása,  
 $T$  pedig a két mérési pont között eltelt idő években.

A *gamma* értékét tehát úgy számítjuk ki, hogy a két mérés átlagának különbségét elosztjuk a két mérés szórásának átlagával. Az így kapott érték a különbség standard mértéke (hatás-méret). Ezt a két mérés között eltelt évek számával elosztva kapjuk a *gamma* értéket, ami az *évenkénti fejlődés standard mértéke*.

A két szórás figyelembevételét azért javasoljuk, mert közöttük jelentős különbségek lehetnek. Általában a fejlődéssel arányosan a második szórás nagyobb az elsőnél. Viszont ha a teszt a második esetben már nem differenciál eléggé, akkor esetleg a szórás csökkenhet is. Nem látjuk viszont szükségesnek a minták méreteivel való súlyozást, mivel keresztmetszeti vizsgálatok mintáit nagyjából azonos méretűre szokás választani.

A *gamma* értéke a populáció, illetve a minta évenkénti változását fejezi ki. Értéke akkor 1, ha a minta elemei – a mintát alkotó tanulók – évente átlagosan egy szórásnyit fejlődnek. Ebből az értelmezésből következik, hogy a *gamma* a két egymást követő évfolyam standard  $z$  skálán kifejezett eredményei közötti különbséggel egyezik meg. A konkrét számításokat tekintve kisebb (a tapasztalat szerint elhanyagolható mértékű) különbségek adódhatnak abból, hogy a kétféle számítás a szórásokat különbözőképpen veszi figyelembe.

A *gamma* érték a keresztmetszeti és a longitudinális felmérések során mért fejlődés kifejezésére egyaránt alkalmazható. Keresztmetszeti mérések esetében mindazoknak a feltételeknek fenn kell állniuk, amelyek egyébként biztosítják, hogy az adatfelvétel megfelelően becsli egy hosszmetzeti vizsgálat eredményeit. Akkor kapunk megfelelő eredményt, ha a mérésben résztvevő összes korosztály pontosan ugyanazokat a teszteket oldja meg, vagy a teszteredmények egy egységes skálára átszámíthatóak. A *gamma* érték ideális eszköz lehet a fejlődés kifejezésére, ha a felméréseket közvetlenül az egymást követő évfolyamokon végezzük el, mégpedig a tanév végén. Ilyen esetben a *gamma* az adott tanév során bekövetkezett fejlődést tükrözi.

Mivel a *gamma* az időegységre eső fejlődés mértéke, így matematikai természetét tekintve analóg a fejlődésgörbe differenciálhányadosával. Ha a fejlettséget egészen pontosan fel tudnánk mérni, úgy, hogy adataink minden időpontban pontosan megfeleljenek a populáció fejlettségének, folytonos görbét kapnánk. Ennek a görbének a differenciálhányadosa egy másik görbét rajzolna fel, amelynek értéke megfelelne a fejlődésgörbe meredekségének. A *gamma* – természetesen nagyon durva felbontással, hiszen általában csak egymástól több éves távolságra levő pontokon mérjük fel a fejlődést – a fejlődésgörbe meredekségét tükrözi.

A *gamma* pszichológiai természetét tekintve a képességek fejleszthetőségéhez, módosíthatóságához (*modifyability*) áll legközelebb. Ha feltételezzük, hogy a képességek fejlődéséhez (fejlesztéséhez) szükséges környezeti ingerek, információk a fejlődés teljes szakaszán egyenletesen rendelkezésre állnak, akkor a képesség abban a szakaszban fejlődik a leggyorsabban, amelyben a legérzékenyebb a fejlesztő hatásokra, vagyis a fejleszthetőség mértéke a legmagasabb. *Piaget* elmélete feltételezi a fejlesztő információk egyenletes rendelkezésre állását, és így kisebb mértékben számol a környezeti hatás variabilitásával. A tapasztalat szerint azonban nem mindig áll rendelkezésre a fejlesztő stimulusok kellő gazdagsága, különösen azoknak a készségeknek az esetében, amelyek közelebb állnak az iskolában tanított tartalmakhoz. A fejlesztő hatás tehát nem egyenletes, hanem függ a tantervektől és egyéb külső tényezőktől. Helyesebb tehát, ha a gamma koefficiens a fejleszthetőség (*modifyability*) és a fejlesztő hatások függvényének tekintjük. A gamma tehát csak akkor tükrözi a módosíthatóságot, ha a fejlesztő hatások egyenlenségei nem torzítják el a fejlődési folyamatot.

### A gamma alkalmazása

Az előzőkben bevezetett gamma minden olyan felmérés esetén kiszámítható, ahol ismerjük a két mérési pontban kapott átlagot és szórást. Amikor a mérési pontok között több év különbség van, akkor természetesen csak az évi átlagos fejlődést tudjuk megadni, ami különbözhet az egyes években bekövetkező konkrét változásoktól. De lényegében ez a helyzet akkor is, amikor a fejlődési görbék megrajzolásánál a mérési pontokat egyenes szakaszokkal kötjük össze, hiszen ha a fejlődés nem lineáris, azok sem írják le pontosan az egyes mérési pontok közötti valódi átmenetet.

A gamma koefficienssel való számítások illusztrálására két különböző típusú keresztmetszeti felmérés adatainak elemzését mutatjuk be. Az első csoportba tartozó, több mérési pontban elvégzett felmérések kifejezetten a fejlődési folyamatok feltárására irányultak, így a gamma értékek néhány iskolai évfolyam közvetlen hatásáról adnak képet. A második csoportba tartozó vizsgálatok keretében csak két, egymástól négy tanévnyi távolságra levő korosztályt mértünk fel, ezeket azonban részletesebben, több teszttel. Ez utóbbi esetekben a gamma kiszámítása a felmért tudás változásának átlagos mértékéről adhat képet.

A gamma alkalmazását nagyrészt korábban már publikált adatok másodelemzése révén mutatjuk be. Ez a megoldás lényegében egyfajta meta-analízis, ami megfelel a gamma és a hatás-mérték szemléletének, és illusztrálja annak szélesebb körű alkalmazási lehetőségeit is. A bemutatott adatokat általában a tanévek végén gyűjtöttük össze; az adatgyűjtés időpontját csak abban a kivételes esetben jelezzük külön, amikor a felmérésekre a tanév elején került sor. A táblázatokban a mérésre használt tesztek neve (a megfelelő képességek megnevezése) mellett feltüntetjük az évet, amelyben a felméréseket elvégeztük.

### A képességek fejlődési ütemének részletes jellemzése

Először azokat az adatokat mutatjuk be, amelyeket kifejezetten a fejlődési folyamatok tanulmányozása érdekében gyűjtöttünk össze. Ezek az elemzések a képességek fejlődésének részletesebb jellemzésére alkalmasak.

#### *A műveleti képességek fejlődése*

A műveleti képességek felmérésére 1997 őszén, tehát a tanév elején került sor, országos reprezentatív mintákon, a harmadiktól a tizenegyedik évfolyamig terjedő életkori intervallumban, a páratlan évfolyamokon. A kombinatív képesség felmérésének részletes eredményeit korábban már publikáltuk (lásd Csapó, 2001b). A rendszerezési képesség felmérése Nagy József tesztjeinek (lásd Nagy, 1987) felhasználásával történt, a logikai képesség felmérését Vidákovich Tibor végezte (a tesztet illetően lásd pl. Vidákovich, 2002). A megfelelő gamma értékeket az 1. táblázatban foglaltuk össze.

1. táblázat. A kombinatív, rendszerezési és logikai képesség tanév eleji felmérésének adatai alapján számított gamma értékek

<i>Teszt (felmérés)</i>	<i>1. mérés (évfolyam)</i>	<i>2. mérés (évfolyam)</i>	<i>Fejlesztő évfolyamok</i>	<i>Fejlődés (gamma)</i>
Kombinatív 97	3.	5.	3–4.	0,19
Kombinatív 97	5.	7.	5–6.	0,39
Kombinatív 97	7.	9.	7–8.	0,10
Kombinatív 97	9.	11.	9–10.	0,20
Rendszerezés 97	3.	5.	3–4.	0,28
Rendszerezés 97	5.	7.	5–6.	0,31
Rendszerezés 97	7.	9.	7–8.	0,21
Rendszerezés 97	9.	11.	9–10.	0,18
Logikai 97	3.	5.	3–4.	0,17
Logikai 97	5.	7.	5–6.	0,18
Logikai 97	7.	9.	7–8.	0,02
Logikai 97	9.	11.	9–10.	0,11

A táblázatban feltüntettük az első és a második felmérés idejét, megjelölve az évfolyamot, amelyek kezdetén az adatgyűjtés történt, és azokat az évfolyamokat is, amelynek fejlesztő hatásával a megfelelő periódusban végbement fejlődést kapcsolatba hozhatjuk. Mivel ebben az esetben a tanév elején gyűjtöttük az adatokat, például a harmadik és az ötödik tanév elején mért adatok alapján számított gamma értékek a harmadik és a negyedik tanévben végbement fejlődést tükrözik.

A táblázat adatait – miként a további hasonló táblázatokat is – kellő körültekintéssel kell értelmeznünk, mindig szem előtt tartva, milyen úton számítottuk ki a gamma értékeket. Mindenképpen óvatosan kell kezelnünk azokat az adatokat, amelyek esetében az első mérés az általános, a második a középiskola idejére esik. Ekkor ugyanis a második mérés adataiból hiányoznak az időközben lemorzsolódó – többnyire a képességteszteken is gyengén teljesítő – tanulók. A lemorzsolódás egyébként általában is pozitív irányba torzítja a gamma értékeket, tehát a valódi értékek inkább kisebbek, mint amit az adatainkból számítottunk. A táblázatok értékeit áttekintve azonban megállapíthatjuk, hogy ez a hatás valószínűleg nem jelentős, nincsenek ugyanis váratlanul magas gamma értékek az iskolafokokon átnyúló mérések esetében.

Az 1. táblázatban feltüntetett gamma értékek 0,02 és 0,39 között változnak. Természetesen a 0 sem számít alsó határnak, hiszen bizonyos készségek esetében – és felméréseink között számos ilyen esettel találkoztunk, például a később bemutatandó korrelatív gondolkodás előfordulhat a negatív irányú változás is. Magasak, vagy alacsonyok ezek az értékek? Ha nincs más összehasonlítási alapunk, ezeket a gamma értékeket mindenek előtt egymással lehet összevetni.

A táblázatban szereplő tizenkét gamma érték átlaga 0,20. A képességek fejlesztésével foglalkozó irodalom a fejlesztő kísérletek esetében ezt az értéket alacsonynak tartja, egy magasnak számító hatás-méret érték ennek körülbelül a négyszerese.

#### *Az arányossággal kapcsolatos feladat alapján számított gamma értékek*

Annak érdekében, hogy a képességek iskolai fejlődéséről minél árnyaltabb képet kapjunk, néhány további képesség gamma értékeit is kiszámítjuk. Az 1993–94 években végeztünk egy felmérést, amelynek alapvető célja az induktív gondolkodás vizsgálata volt, azonban a feladatok között szerepelt egy olyan feladat, amely az egyenes arányosság megértését, az alkalmazás készségét vizsgálta. Azt tapasztaltuk, hogy a fejlődés egy szabályos, elnyúlt logisztikus görbével jellemezhető. A felmérés eredményei alapján számított gamma értékeket a 2. táblázatban tüntettük fel. Az előző táblázathoz hasonlóan itt is megadtuk a két mérési pontot, és azt, hogy mely évfolyamokra tehető a fejlődés. A gamma értékeket a korábban közölt adatok alapján (Csapó, 1994. 66. o.) a megfelelő átlag és szórás adatok felhasználásával számítottuk ki.

#### *2. táblázat. Az arány fogalmának fejlődését jellemző gamma értékek*

<i>Teszt (felmérés)</i>	<i>1. mérés (évfolyam)</i>	<i>2. mérés (évfolyam)</i>	<i>Fejlesztő évfolyamok</i>	<i>Fejlődés (gamma)</i>
Arány 94	3.	5.	4–5.	0,15
Arány 94	5.	7.	6–7.	0,19
Arány 94	7.	9.	8–9.	0,34
Arány 94	9.	11.	10–11.	0,11

Ezek az adatok már tanév végi felmérésekből származnak, ennek megfelelően a gamma által kifejezett fejlődés az első felmérést követő tanévben ment végbe. Például ha a két mérési pont az ötödik és a hetedik évfolyam végén volt, akkor a felmért fejlődés maga a hatodik és a hetedik évfolyamon ment végbe.

Természetesen a tanév eleji és az év végi felméréseket nem tekinthetjük egyenrangúnak, hiszen más lehet a tanulók motiváltsága, továbbá a tesztek megoldásában szerepet játszhat az is, hogy mennyire vannak „tréningben”. Ha azonban mindkét évfolyamot a tanévnek ugyanabban az időszakában mérjük fel, az említett jelenség mindkét évfolyamon nagyjából azonos módon jelentkezik. Ezért ha a két mérési pont különbségét tekintjük, azokban az említett jelenségnek már minimális szerepe lehet csak. Itt is érvényes továbbá, hogy az iskolai tantervekhez kevésbé kötődő általános képességek felméréseinek eredményeit kevésbé befolyásolja az időpont, mint a gyorsan tanulható, gyorsan fejlethető, tananyaghoz, tantárgyhoz kötődő készségeket.

A táblázatban szereplő négy gamma érték átlaga ebben az esetben is 0,2, ami azt jelzi, hogy az arány feladat megoldásának készsége, a lineáris összefüggések megértése is lassan fejlődik. Azt várnánk, hogy egy ilyen egyszerű matematikai művelet kialakulása nagyrészt az általános iskola alsó tagozatában lejátszódik, azonban ebben az időszakban még nagyon lassú a fejlődés. A 4–5. évfolyam fejlesztő hatása átlagosan csak 0,15 gamma értéket jelez, ami azért elgondolkodtató, mert ebben az időszakban a tanulók már a matematika tanítása keretében formálisan is sokat foglalkoznak az arányosságot tartalmazó feladatok megoldásával. A 6–7. évfolyam fejlesztő hatásának tulajdonítható, mindössze 0,19-es a gamma érték pedig azt jelzi, hogy az egyes tantárgyakban (fizika, kémia) előforduló, az arányosságra épülő tartalmak közvetítése sem eredményez átütő változást. A leggyorsabb fejlődés szakasza – érdekes módon – a 8–9. évfolyamra esik (0,34). A középiskola középső két évének a fejlesztő hatása ismét csekély mértékű (0,11), habár az adatok szerint (lásd Csapó, 1994) a tanulók számottevő csoportja még ekkor sem tudja a feladatot megoldani, tehát a fejlődés nem azért lassul le, mert a megfelelő készség már mindenkinél kifejlődött.

Ebben a tanulmányban nem célunk a felmerült problémák tartalmi elemzése, csak azt jelezzük, hogy a gamma használata milyen kérdéseket világít meg, hogyan helyezi az eredményeket egy újabb interpretációs keretbe, és segítségével milyen további kutatási feladatokat fogalmazhatunk meg. Az arány feladat esetében adódik a kérdés, meg lehet-e gyorsítani a fejlődést az iskola korábbi éveiben, illetve a 8–9. évfolyamokra eső magas értékben játszhat-e szerepet az iskolafokokozatok közötti átmenet vagy az eközben bekövetkező szelekció.

#### *Az induktív gondolkodás fejlődésének jellemzése*

Az induktív gondolkodás fejlődésének részletes feltérképezésére irányuló vizsgálat (Csapó, 1994) adatait felhasználva az induktív gondolkodás résztesztekre külön is kiszámítottuk a gamma értékeket, ezeket a 3. táblázat mutatja be.



A képességek fejlődési ütemének egységes kifejezése: a gamma koefficiens

3. táblázat. Az induktív gondolkodás komponenseinek fejlődését jellemző gamma értékek

<i>Teszt (felmérés)</i>	<i>1. mérés (évfolyam)</i>	<i>2. mérés (évfolyam)</i>	<i>Fejlesztő évfolyamok</i>	<i>Fejlődés (gamma)</i>
Számanalógiák 94	3.	5.	4–5.	0,21
Számanalógiák 94	5.	7.	6–7.	0,63
Számanalógiák 94	7.	9.	8–9.	0,28
Számanalógiák 94	9.	11.	10–11.	0,11
Szóanalógiák 94	3.	5.	4–5.	0,29
Szóanalógiák 94	5.	7.	6–7.	0,34
Szóanalógiák 94	7.	9.	8–9.	0,36
Szóanalógiák 94	9.	11.	10–11.	0,09
Számsorok 94	3.	5.	4–5.	0,19
Számsorok 94	5.	7.	6–7.	0,36
Számsorok 94	7.	9.	8–9.	0,14
Számsorok 94	9.	11.	10–11.	0,13
Betűsorok 94	3.	5.	4–5.	0,29
Betűsorok 94	5.	7.	6–7.	0,28
Betűsorok 94	7.	9.	8–9.	0,16
Betűsorok 94	9.	11.	10–11.	0,09
Átkódolás 94	3.	5.	4–5.	0,23
Átkódolás 94	5.	7.	6–7.	0,15
Átkódolás 94	7.	9.	8–9.	0,17
Átkódolás 94	9.	11.	10–11.	0,08
Kizárás 94	3.	5.	4–5.	0,37
Kizárás 94	5.	7.	6–7.	0,29
Kizárás 94	7.	9.	8–9.	0,28
Kizárás 94	9.	11.	10–11.	0,15

Amint a táblázatban látjuk, az egyes résztesztekkel felmért különböző képességek között is jelentős eltérések lehetnek. A legnagyobb gamma értéket (0,63) itt a számanalógiák esetében találtuk, mégpedig a 6–7. évfolyamokon. Úgy tűnik, a megfelelő matematikai háttér elsajátítása jelentősen felgyorsítja a fejlődést. A szóanalógiák teljesítményei a 6–9. évfolyamokon egyaránt gyorsan, 0,35 körüli gamma értékkel változnak. A legintenzívebb fejlődési szakaszban általában ilyen nagyságú értékeket találtunk. A számsorok a 6–7. évfolyamon 0,36; a kizárás a 4–5. évfolyamon 0,37 gamma értékkel fejlődik a leggyorsabban. Az előző táblázatban szereplő értékekkel való összehasonlítás-

hoz figyelembe kell vennünk, hogy ez utóbbi táblázatban szereplő tesztek sok könnyű feladatot tartalmaztak, némelyikük pedig közelebb állt az iskolában gyakorolt készségekhez. A gamma értékek itt már szélesebb skálán szóródnak. A kiszámított értékek átlaga 0,24, vagyis kissé magasabb az előzőeknél, de még mindig alacsony.

Az induktív gondolkodás fejlődését a már elemzett vizsgálaton túl egy országos reprezentatív mintán is felmértük 1999-ben (Csapó, 2001a). Mivel az induktív gondolkodás nagyon jól jellemzi a tanulók általános intellektuális fejlődését, a tesztet több felmérésben is felhasználtuk a tanulók gondolkodásának jellemzésére. Háttérváltozóként szerepelt például a 2000-ben elvégzett idegen nyelvi felmérésben (Csapó, 2001c), és egy olyan vizsgálatban is, amely az iskolákban folyó fejlesztő munkát elemezte, és egyes önkormányzatok által fenntartott összes iskolára kiterjedt. Ebben az esetben az ötödiktől a tizenegyedikig minden évfolyamot felmértük, így évenkénti felbontású adatokat kaptunk. Az e felmérések alapján számított gamma értékeket a 4. táblázatban foglaljuk össze. Itt mutatjuk be az 1994-es felmérés átfogó adatait is. Az 1994-es felmérésben részletesebb (és könnyebb) teszteket használtunk, míg a másik három vizsgálat pontosan ugyanarra a rövidebb, egy tanórán megoldható tesztre épült. (A 2000-ben végzett, önkormányzati felmérést a „00ö” jelzi.)

4. táblázat. Az induktív gondolkodás különböző felméréseinek összehasonlítása

Teszt (felmérés)	1. mérés (évfolyam)	2. mérés (évfolyam)	Fejlesztő évfolyamok	Fejlődés (gamma)
Induktív 94	3.	5.	4–5.	0,39
Induktív 94	5.	7.	6–7.	0,49
Induktív 94	7.	9.	8–9.	0,30
Induktív 94	9.	11.	10–11.	0,12
Induktív 99	5.	7.	6–7.	0,38
Induktív 99	7.	9.	8–9.	0,24
Induktív 99	9.	11.	10–11.	0,23
Induktív 00	6.	8.	7–8.	0,35
Induktív 00	8.	10.	9–10.	0,22
Induktív 00ö	5.	6.	6.	0,40
Induktív 00ö	6.	7.	7.	0,43
Induktív 00ö	7.	8.	8.	0,37
Induktív 00ö	8.	9.	9.	0,11
Induktív 00ö	9.	10.	10.	0,24
Induktív 00ö	10.	11.	11.	0,17

Az eddig elemzett gamma értékek többségét úgy kaptuk, hogy a mérési pontok legalább kétévnyi távolságra voltak egymástól. Ebben az esetben a két tanévre számítva egy átlagos gamma értéket kapunk. Ez az átlagos gamma – hacsak nem teljesen egyenletes a fejlődés a két évfolyamon – egy gyorsabb és egy lassúbb értékre bomlana, ha módunk lenne azokat a két tanévre külön-külön kiszámítani. Ebből a szempontból kedvező elemzési lehetőséget kínál a 2000-ben elvégzett önkormányzati felmérés adatbázisa, mivel ott az induktív gondolkodásról évenkénti felbontásban állnak rendelkezésünkre az adatok. Ez a minta nem reprezentatív, de a gamma értékek tekintetében a reprezentativitás hiánya csak az általános iskola–középiskola átmenetnél jelenthet problémát. Amint a táblázatból kitűnik, a kilencedik évfolyamon inkább a vártnál alacsonyabbak a gamma értékek, ami arra utal, hogy a két minta nem teljesen azonos összetételű. A többi esetben viszont ugyanazoknak az iskoláknak a tanulói vettek részt a gamma számításához tartozó két mérésben, így nincs okunk az adatok pontosságában kételkedni.

A táblázatban szereplő gamma értékek átlaga 0,3, ami azt jelzi, hogy az induktív gondolkodás fejlődése a felmért teljes életkori intervallumban viszonylag gyors. A leggyorsabb változás minden esetben a 6–7. évfolyamra esik. A páros és páratlan évfolyamokon végzett, valamint az évenkénti felmérés alapján pedig azt is valószínűsíthetjük, hogy a fejlődés a hatodik évfolyam során kissé gyorsabb, mint a hetedikben.

A gamma értékeket technikailag nagy pontosságú adatokból számítottuk ki. A bizonytalanság, az értékek pontatlansága illetve eltérése így nagyrészt abból fakad, hogy a fejlődési folyamatból ritkán, többévnyi távolságban vettünk mintát, így a fejlődés merekségét csak közelítő pontossággal jellemezhetjük.

Az 1994-ben végzett felmérés keretében felvettük a Raven intelligencia-tesztet is (Csapó, 1994). E teszt úgy kapcsolódott a felmérés tematikájához, hogy szintén induktív jellegű feladatokat tartalmaz, bár a bonyolult mártixos elrendezésben közölt információk feldolgozásában, a feladatok megoldásában más, formálisabb, műveletvégzés jellegű készségek is szerepet játszhatnak. A kiszámított gamma értékeket az 5. táblázatban tüntettük fel. Természetesen ebben az esetben a teszten elért nyerspontokat használtuk fel, azaz az adatokat nem standardizáltuk, nem számítottuk át más skálára.

5. táblázat. A Raven teszt segítségével mért intelligencia fejlődését jellemző gamma értékek

Teszt (felmérés)	1. mérés (évfolyam)	2. mérés (évfolyam)	Fejlesztő évfolyamok	Fejlődés (gamma)
Raven IQ 94	3.	5.	4–5.	0,43
Raven IQ 94	5.	7.	6–7.	0,21
Raven IQ 94	7.	9.	8–9.	0,47
Raven IQ 94	9.	11.	10–11.	0,17

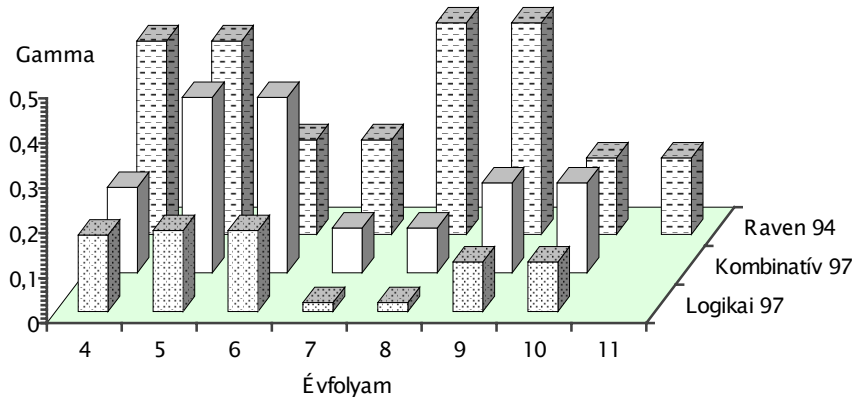
A Raven IQ esetében két gyorsan változó szakasz (0,43 és 0,47 gamma értékkel) fog közre egy lassúbbat. Ez a különös jelenség számos kérdést vet fel. Például felvethetjük, hogy az e tekintetben szabálytalan fejlődés vajon a feladatok megoldásában szerepet játszó gondolkodási mechanizmusok szerkezeti átrendeződésének tulajdonítható – ami átmeneti teljesítménycsökkenést eredményez –, a 6–7. évfolyamok fejlesztő hatása különösen alacsony e tekintetben, vagy esetleg a két magasabb gamma érték tulajdonítható valamilyen különlegesen erőteljes fejlesztő hatásnak. A négy kiszámított gamma érték átlaga egyébként 0,32, ami magasabb az eddig elemzett más képességek esetében tapasztaltaknál.

#### *A gamma értékek grafikus összehasonlítása*

Az egyes képességekre jellemző fejlődési tempót, illetve a különböző felmérésekből származó adatokat szemléletesebben össze tudjuk hasonlítani, ha a megfelelő gamma értékeket grafikusán ábrázoljuk. Az összehasonlító ábrázolással az a célunk, hogy bemutassuk, az egyes iskolai évek milyen hatást gyakorolnak a vizsgált képességek fejlődésére. Mivel a mérési pontok különbözhetnek, és nem minden esetben állnak rendelkezésünkre évenkénti adatok, az egyes évek fejlesztő hatását jellemző adatokat a több évre számított átlaggal becsüljük. Ez természetesen a konkrét esetekben kisebb-nagyobb pontatlanságokat eredményezhet, arra azonban ez a megoldás is alkalmas, hogy a képességek fejlődésének fő jellemzőit, illetve az iskola különböző hatású szakaszait összehasonlítsuk. Az ilyen jellegű bizonytalanságok egyébként általában is jellemzőek a kutatási eredmények szintetizálására, hiszen mindig többé-kevésbé eltérő feltételek között, más-más eszközökkel nyert eredményeket kell integrálni, és a fő tendenciák kiszűrése érdekében gyakran kell engedelményeket tenni a részletek tekintetében.

A elemzett képességek között három olyat találtunk, amelynek a fejlődési folyamata szabálytalan abban az értelemben, hogy a változás üteme megtörik, gyorsabb és lassabb szakaszok követik egymást. A kombinatív és a logikai képesség, valamint a Raven teszt esetében találtunk ilyen szabálytalanságokat, és ez éppen ellentétes azzal, amit az induktív gondolkodás esetében találtunk. Ezeknek a képességeknek a gamma értékeit az 1. ábrán ábrázoltuk. A három képesség esetében nem ugyanazokról az évfolyamokról áll rendelkezésünkre a fejlődés mértéke, ráadásul adataink mindegyik esetben csak kétéves felbontásúak. E különbözőség ellenére megmutatkozik az az alapvető hasonlóság, mely szerint két gyorsabb fejlődési szakasz közrefog egy lassúbbat. A mérési pontok különbözősége miatt nem lehet pontosan megállapítani, hogy ez a lelassulás pontosan egybeesik-e. A rendelkezésünkre álló adatok szerint a kombinatív és a logikai képesség esetében ez a lassúbb szakasz valószínűleg kicsit (kb. egy évvel) későbbre esik. A lelassulás lehetséges okaival itt nem foglalkozunk részletesen, a gondolkodási stratégiák minőségi átrendeződése tűnik a legvalószínűbb magyarázatnak.

A képességek fejlődési ütemének egységes kifejezése: a gamma koefficiens



1. ábra

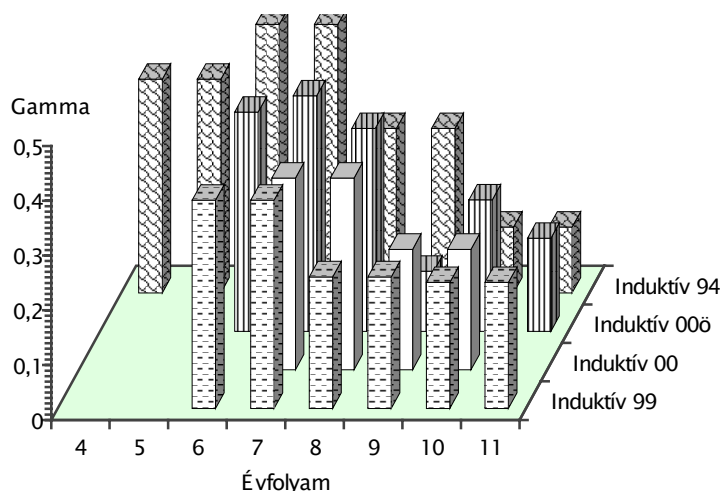
A kombinatív képesség, a logikai képesség és a Raven teszt megoldása alapján számított gamma értékeinek összehasonlítása

A 2. ábrán az induktív gondolkodásnak a különböző vizsgálatok alapján számított gamma értékeit hasonlíthatjuk össze. Az ábrán itt is évfolyamonként tüntettük fel a fejlődés mértékét. Itt egy felmérés valóban évenkénti pontosságú adatokat szolgáltat, a többi esetben, amikor csak két év átlaga áll rendelkezésünkre, a megfelelő évfolyamokon ezt a közös átlagot ábráztuk.

Teljes egyezést természetesen itt sem várhatunk, mivel mindegyik felmérésnek van valami sajátossága: az 1999-es felmérés (Induktív 99) a páratlan, a 2000-es (Induktív 00) a páros évfolyamokra terjedt ki, az önkormányzati felmérés (Induktív 00ö) mindegyik évfolyamot magában foglalta, a szegedi vizsgálat pedig más tesztrendszeren alapult. Az ábra azonban a felmérések együttes adataival még így is meglehetősen konzisztens képet mutat az induktív gondolkodás fejlődéséről. Az ábra alapján egy lassan induló, felgyorsuló, majd ismét lelassuló folyamat képe áll össze, ami megfelel egy szabályos logisztikus fejlődésmenetnek.

Az eddigi elemzéseink során csak olyan képességekkel foglalkoztunk, amelyek az áttekintett folyamat teljes szakaszán pozitív irányba változtak, fejlődtek. Ezekről összesítésként elmondhatjuk, hogy a képességek éves fejlődését jellemző gamma érték ezekben átlagosan 0,20–0,25 körül van, a leggyorsabb fejlődési szakaszokban pedig körülbelül 0,4-nek adódnak.

Csapó Benő



2. ábra

*Az induktív gondolkodás különböző vizsgálatok alapján számított gamma értékeinek összehasonlítása*

### **Az iskola fejlesztő hatása a tudás néhány minőségi jellemzője tekintetében**

A következő példák két hasonló témájú kutatási programból származnak, mindkét program az iskolában elsajátított tudás minőségi jellemzőit, a megértés mélységét, a megtanultak hasznosíthatóságát, a gondolkodás fejlődését vizsgálta. Az egyik kutatási program a matematikai és természettudományi tudáshoz (Csapó, 2002a), a másik a humán műveltséghez és a társadalomtudományi tudáshoz (Csapó, 2002b) kapcsolódott. Mindegyik kutatási program mintáját szegedi iskolákból választottuk, és mindkét esetben a hetedik és a tizenegyedik évfolyamokon végeztük el a felméréseket.

A természettudományi és matematikai tudás vizsgálatának eredményeit a 6. táblázatban foglaltuk össze. Ebben az esetben mindegyik képesség értékeléséhez csak egy felmérés-pár (két mérési pont) áll rendelkezésünkre. A táblázatban megadtuk a gamma kiszámításához felhasznált átlag és szórás adatokat. Az induktív gondolkodást és a logikai képességet (deduktív gondolkodás, Vidákovich, 2002) korábban már jellemeztük, a korrelatív gondolkodás (Bán, 2002), a természettudomány alkalmazása (B. Németh, 2002), a matematikai megértés (Dobi, 2002) és a természettudományi tévképzetek (Korom, 2002) teszteket a szerzőik által írott könyvfejezetek mutatják be.

A táblázat adatai csak a négy év során bekövetkezett átlagos fejlődésről adnak képet. Az induktív és a deduktív gondolkodás fejlődését jellemző gamma értékek összhangban vannak a korábbi, részletesebb felmérésekből számított adatokkal. A matematikai megértés és a természettudományos tévképzetek 0,2 körüli gamma értéke szintén összevethető más, lassúbb fejlődést jelző adatokkal. A természettudomány alkalmazása teszt jelzi, hogy a felmért iskolák e tekintetben gyakorolták a legnagyobb hatást a tanulókra, bár a

A képességek fejlődési ütemének egységes kifejezése: a gamma koefficiens

részletes elemzések (B. Németh, 2002) azt is megmutatták, hogy az alkalmazás képessége szoros kapcsolatban áll az induktív gondolkodás fejlettségével. A legsúlyosabb problémát a korrelatív gondolkodással kapcsolatos vizsgálat hozta a felszínre, amely a vizsgált időszakban egy kisebb mértékű negatív változást jelzett (Bán, 2002).

6. táblázat. A matematikai és természettudományos tudás minőségét jellemző mutatók változása

Teszt	7. évfolyam		11. évfolyam		Fejlődés (gamma)
	átlag	szórás	átlag	szórás	
Induktív gondolkodás	45,35	15,09	64,44	14,69	0,32
Deduktív (logikai) gondolkodás	61,41	26,89	76,07	15,28	0,17
Korrelatív gondolkodás	56,73	15,98	53,51	15,77	-0,05
Természettudomány alkalmazása	33,08	15,72	55,47	16,63	0,35
Matematikai megértés	33,51	14,03	46,98	17,10	0,22
Természettudományi tévképzetek	50,10	14,56	63,74	18,27	0,21

A humán műveltséggel kapcsolatos felmérés eredményei alapján számított gamma értékeket a 7. táblázatban összegeztük. A felmérésekről, valamint az angol tudásszint, az alkalmazható nyelvtudás (Bukta és Nikolov, 2002), a környezetkultúra (Kárpáti, 2002), az írásbeli szövegalkotás (Molnár E. K., 2002), a történelmi gondolkodás (Szebenyi és Vass, 2002) és a kritikai gondolkodás (Molnár L., 2002) felméréséről ebben az esetben is a vizsgálatokat bemutató fejezetek nyújtanak részletes tájékoztatást.

7. táblázat. A társadalomtudományi tudás és a humán műveltség minőségét jellemző mutatók változása

Teszt	7. évfolyam		11. évfolyam		Fejlődés (gamma)
	átlag	szórás	átlag	szórás	
Angol tudásszint	48,89	17,99	71,28	19,76	0,30
Alkalmazható angol nyelvtudás	36,27	17,73	51,24	25,75	0,17
Környezetkultúra	76,60	15,85	68,21	16,11	-0,13
Írásbeli szövegalkotás	6,06	1,53	6,47	1,55	0,07
Történelmi gondolkodás	57,09	16,41	67,46	14,84	0,17
Kritikai gondolkodás	32,63	11,61	39,50	13,50	0,14

A táblázatban bemutatott gamma értékek ebben az esetben is csak a négy év alatt bekövetkezett átlagos fejlődésről adnak áttekintést. Az előző elemzésekkel összhangban itt

is azt látjuk, hogy azon a területen jelentősebb a változás, amelyik közelebb áll az iskolában tanított tartalmakhoz. Ez esetünkben az angol tudásszint (0,30). Már ugyanannak a készségrendszernek az alkalmazásorientált, kommunikációközpontú vizsgálata sokkal kisebb fejlődésről ad képet (0,17). A többi gamma érték kifejezetten az alacsony sávba esik, mindegyik 0,2 alatt marad. Aggaszó jelenségre utal az anyanyelvi szövegalkotás képességének inkább a stagnálására, mint a fejlődésére utaló gamma értéke (0,07). Itt is találtunk egy negatív, jelentősebb romlást mutató gamma értéket, ami a környezetkultúrával összefüggő készségek iskolai fejlesztésének súlyos problémáira utal.

### **A gamma koefficiens használatának értékelése: lehetőségek és korlátok**

#### **A gamma használatának előnyei**

A gamma koefficiens használatának szükségességét alátámasztó megfontolásokat az elméleti bevezetőben már bemutattuk, az alkalmazás gyakorlati hasznát pedig az empirikus elemzések illusztrálták. Összegzésként elmondhatjuk, hogy a gamma alkalmas az adatok egységes kezelésére, meta-elemzések elvégzésére, a különböző forrásokból származó vizsgálati eredmények szintetizálására. Emellett a fejlődést több oldalról is bemutatja, a grafikus ábrázolások olyan összefüggéseket is megjelenítenek, amelyeket a fejlődési vonalakat ábrázoló grafikonok önmagukban nem tükröznek. A gamma koefficiens használata lehetővé teszi, hogy a spontán fejlődést, azaz az iskola természetes fejlesztő hatását és a fejlesztő kísérletek eredményeit közvetlenül összehasonlítsuk.

#### **A gamma használatának korlátjai**

A gamma koefficiens itt bemutatott értelmezése a pszichometriai hagyomány és a klasszikus tesztelmélet szemléletmódját követi. Ebből következően alapvetően normatív jellegű, a felmért minta (populáció) jellemző adataiból állítja elő a fejlődés mértékét. Ennek megfelelően – mint a klasszikus tesztelmélet minden eljárása – mintafüggő. Mint minden statisztikai szemléletű eljárás, a gamma számítása is feltételezi, hogy a felmért tanulók között különbségek vannak, azaz a vizsgált változóknak megfelelő szórása van. Ahogy a szórás, úgy a gamma sem független a vizsgált változó eloszlásának alakjától, szigorúan véve csak az azonos eloszlású (a legtöbb pedagógiai mérés esetében hallgatólagosan feltételezett normális eloszlású) változó esetén használható. Az eloszlás alakját tekintve azonban sok más statisztikai elemzés is hasonló előfeltevéseket fogalmaz meg, és így – bizonyos sajátos esetektől eltekintve – a gamma számítása jól illeszkedik a tesztekkel kapcsolatban szokásos egyéb (elsősorban a klasszikus tesztelmülethez kötődő) számítások kontextusába.

Nem használható a gamma olyan esetekben, ha a mintának nincs megfelelő szórása. Az alacsony szórás érték ugyanis megnövelné a fejlődés gamma statisztikájának értékét. Ez előfordulhat a fejlődés kezdő szakaszán, amikor még a tanulók többsége egyöntetűen



alacsony pontszámot ér el, és a telítődési szakaszban, amikor már mindenki egyöntetűen magas teljesítményt ér el. Ezek azonban meglehetősen triviális kizáró okok, ritkán fordulnak elő, és ilyen esetekben már magunknak a teszteknek a használata sem célszerű, hiszen azok nem rendelkeznek kellő differenciáló erővel. Kivételt képez a kritériumorientált mérés, amikor egy külső, a minta aktuális sajátosságaitól független kritériumhoz viszonyítjuk az eredményeket, ilyen esetekben annak megállapítása is fontos lehet, hogy mindenki tökéletes teljesítményt nyújtott (pl. a *mastery learning* programok esetében), ezekben az esetekben azonban a gamma kiszámítása nem értelmezhető.

Nem triviális kizáró ok lehet a minta szélsőségesen megnövekedett szórása, ami viszont alacsony gamma értéket eredményezhet. Ez például a polarizált fejlődés, a bimodális eloszlású teljesítmények esetében fordulhat elő. A természetes fejlődési folyamatok azonban – bár gyakran a különbségek, és ezáltal a szórás növekedésével járnak – általában ritkán idéznek elő ilyen szélsőséges helyzeteket. Mindamelllett az oktatási rendszer sajátosságai – többnyire diszfunkcionális működése – létrehozhatnak ilyen szélsőségesen differenciált fejlődési ütemet. Általában tehát a gamma koefficiens kiszámítása előtt érdemes a szórások illetve az eloszlás alakjának ellenőrzésére különös gondot fordítani.

### További kutatási feladatok

A valószínűségi tesztelméletek új skálázási lehetőséget kínálnak. Egyrészt a hídfeladatok használatával egy másik megoldást adnak annak a problémának a megoldására, hogy a különböző évfolyamok teljesítményeit ugyanazon a skálán fejezzük ki. Ez a megoldás azonban – bár számos előnnyel jár – önmagában nem teremt kapcsolatot a klasszikus statisztikai elemzésekkel, például a hatás-méret számításokkal. Szükség lenne tehát olyan jellegű elemzésekre, amelyek megteremtik az IRT modellek (valószínűségi tesztelméletek) és a gamma számítása közötti kapcsolatot. Erre egyébként a logit egységekben kifejezett személyparaméterek felhasználásával egyszerű lehetőség adódik, a kétféle eljárás összehasonlítása azonban további elemzéseket igényel. (Az induktív gondolkodás IRT skálázásával kapcsolatban lásd *Csapó, 1994.*)

Az előzőekben már említést tettünk az iskola „természetes” fejlesztő hatásáról, amelyet a gamma érték kiszámítása révén egyes pedagógiai kísérletek hatás-méretben kifejezett hozzáadott fejlesztő hatásával hasonlíthatunk össze. A kétfajta fejlesztő hatás azonban lényegi természetét tekintve nem különbözik egymástól, tehát azokat egységes fogalmi keretben lehet kezelni. Így az iskola „hagyományos” és a kísérletek során alkalmazott fejlesztő hatását hasonlíthatjuk össze. Ez lényegében csak annyit jelent, hogy a csoport-kontrollcsoport kísérleti elrendezés esetében a kísérleti csoportban mért fejlődésből nem vonjuk ki a kontroll csoport fejlődését, azaz nem a kísérlet hozzáadott fejlesztő hatására számítunk hatás-méretet, hanem a kísérleti csoportban mért teljes fejlődést vesszük figyelembe, arra számítjuk ki a gamma értéket, és ezt vetjük össze a kontroll csoportra – azaz a szokásos keretek között folyó iskolai oktatásra – számított gamma értékekkel.

Az előző gondolatmenetet tovább folytatva, minden egyes iskolát, az ott folyó munkát jellemezhetjük a megfelelő képességekre gyakorolt fejlesztő hatással. A pedagógiai

kísérletezés gondolatmenetét követve tehát az egyes iskolákat tekinthetjük úgy, mint a tantervek megvalósításának kísérleteit, amelyek eredményei az ottani feltételektől függetlenül különbözőek lehetnek. A megfelelő gamma értékek kiszámítása lehetővé teszi az iskolák fejlesztő hatásának összevetését, ami viszont a pedagógiai hozzáadott érték koncepciójának egy lehetséges megközelítése. Az ilyen jellegű elemzéseket viszont a pedagógiai hozzáadott érték számításának más lehetőségeivel, például a regressziós modellekkel (lásd Csapó, 2002c) kell összevetni.

## Irodalom

- Bán Sándor (2002): Gondolkodás a bizonytalanról: valószínűségi és korrelatív gondolkodás. In: Csapó Benő (szerk.): *Az iskolai tudás*. 2. kiadás. Osiris Kiadó, Budapest. 231–260.
- Beaton, A. E., Martin, M. O., Mullis I. V. S., Gonzalez, E. J., Smith, T. A. és Kelly, D. L. (1996): *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. TIMSS International Study Center, Boston College, Chestnut Hill, MA, USA.
- B. Németh Mária (2002): Iskolai és hasznosítható tudás: a természettudományos ismeretek alkalmazása. In: Csapó Benő (szerk.): *Az iskolai tudás*. 2. kiadás. Osiris Kiadó, Budapest. 123–148.
- Bukta Katalin és Nikolov Marianne (2002): Nyelvtanítás és hasznosítható nyelvtudás: az angol mint idegen nyelv. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 169–192.
- Cohen, J. (1988): *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates, New Jersey.
- Csapó Benő (1994): Az induktív gondolkodás fejlődése. *Magyar Pedagógia*, **94**. 1–2. sz. 53–80.
- Csapó Benő (2001a): Az induktív gondolkodás fejlődésének elemzése országos reprezentatív felmérés alapján. *Magyar Pedagógia*, **101**. 3. sz. 373–391.
- Csapó Benő (2001b): A kombinatív képesség fejlődésének elemzése országos reprezentatív felmérés alapján. *Magyar Pedagógia*, **101**. 4. sz. 511–530.
- Csapó Benő (2001c): A nyelvtudást és a nyelvtanulást befolyásoló tényezők. *Iskolakultúra*, **11**. 8. sz. 25–35.
- Csapó Benő (2002a, szerk.): *Az iskolai tudás*. 2. kiadás. Osiris Kiadó, Budapest.
- Csapó Benő (2002b, szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest.
- Csapó Benő (2002c): Az osztályok közötti különbségek és a pedagógiai hozzáadott érték. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 269–297.
- Dobi János (2002): Megtanult és megértett matematikatudás. In: Csapó Benő (szerk.): *Az iskolai tudás*. 2. kiadás. Osiris Kiadó, Budapest. 177–199.
- Goossens, L. (1992): Training scientific reasoning in children and adolescents: A critical commentary and quantitative integration. In: A. Demetriou, M. Shayer és A. Efklides (szerk.): *Neo-Piagetian Theories of Cognitive Development. Implications and Applications for Education*. Routledge and Kegan, London, 160–182.
- Kárpáti Andrea (2002): A vizuális műveltség. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 91–133.
- Klauer, K. J. (2001): Training des induktiven Denkens. In: Klauer, K. J. (szerk.): *Kognitives Training*. 165–209.
- Korom Erzsébet (2002b): Az iskolai tudás és a hétköznapi tapasztalat ellentmondásai: természettudományos tévképzetek. In: Csapó Benő (2002. szerk.): *Az iskolai tudás*. 2. kiadás. Osiris Kiadó, Budapest. 149–176.

A képességek fejlődési ütemének egységes kifejezése: a gamma koefficiens

- Kulik, J. A. és Kulik, C. L. C. (1989): The concept of meta-analysis. *International Journal of Educational Research*, **13**, 3. sz. 227–340.
- Molnár Edit Katalin (2002a): Az írásbeli szövegalkotás. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest, 193–216.
- Molnár László (2002): A kritikai gondolkodás. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 217–237.
- Nagy József (1987): *A rendszerezési képesség kialakulása*. Gondolkodási műveletek. Akadémiai Kiadó, Budapest.
- OECD (2001): *Knowledge and skills for life. First results from the OECD Program for International Students Assessment (PISA) 2000*. OECD, Paris.
- Szebenyi Péter és Vass Vilmos (2002): Történelmi tévképzetek, történelemszemlélet, nemzeti azonosságtudat. In: Csapó Benő (szerk.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 135–167.
- Thomson, B. (2002). What future quantitative social science research could look like: Confidence intervals for the effect sizes. *Educational Researcher*, **31**, 3. sz. 25–32.
- Vidákovich Tibor (2002): *Tudományos és hétköznapi logika: a tanulók deduktív gondolkodása*. In: Csapó Benő (szerk.): *Az iskolai tudás*. 2. kiadás. Osiris Kiadó, Budapest. 201–230.

Csapó Benő

## ABSTRACT

BENŐ CSAPÓ: COMPARATIVE MEASURE FOR THE DEVELOPMENT OF SKILLS AND COMPETENCIES IN CROSS-SECTIONAL SURVEYS: INTRODUCING THE GAMMA COEFFICIENT

Procedural components of knowledge (skills, abilities, competencies, expertise etc.) have been receiving growing attention both in the development of curricula and in the design of evaluation instruments. The development of these components of knowledge span several years, and in order to monitor their development practically the same (or parallel) tests may be administered to students of different ages. This way, the occasions when the achievements of students of different ages have to be compared become more frequent. This paper introduces a simple and standard way of expressing changes occurring in the course of one academic year, and illustrates how the proposed method has been put into practice by re-analyzing and comparing the results of some recent large-scale assessment projects. To express developmental data obtained from cross-sectional assessment in a standard form, a coefficient called *gamma* is proposed. The gamma coefficient is computed so that the difference of the means (measured in the same period of the academic year but in different grades) is divided by the mean of the two standard deviations, and then this ratio is divided by the number of years between the two grades.

$$\text{gamma} = \{(\text{Mean2}-\text{Mean1}) / [(\text{StdDev1}+\text{StdDev2})/2]\} / (\text{Grade2}-\text{Grade1})$$

Thus gamma is the standard measure of development of the assessed sample (population); it is the change in a skill that takes place during one academic year, expressed in standard deviation units. Gamma, defined in this way, is analogous with the measure of effect size (denoted by 'd') that is used for presenting the results of training experiments and as the basis of meta-analyses. Therefore, the natural development of a skill and the development as a result of specific training can be compared directly. The use of gamma is illustrated by data from previous surveys on combinative, inductive, logical and proportional reasoning, classification skills, word problem solving skills and the Raven Intelligence Test. Gamma values were computed from the means and standard deviations of the achievements in two neighboring measurement points. The gamma values presented in this paper typically range from ca. 0.1 to 0.4 with some interesting exceptions and anomalies. The results presented in this paper show that the use of gamma allows for several fruitful comparisons and poses questions for further investigations. (E.g. are low gamma values due to the low sensibility/modifiability of a given skill in a period, or to the weakness of stimuli that effect the development during that period?) Standard methods of processing data and presenting results make it easier to synthesize the research findings.

Magyar Pedagógia, 102. Number 3. 391–410. (2003)

Levelezési cím / Address for correspondence: Department of Education, University of Szeged, H-6722 Szeged, Petőfi sgt. 30–34. E-mail: csapo@edpsy.u-szeged.hu.