

Randomizált pedagógiai kísérletek a 21. században: in tempore opportuno

Csíkos Csaba*

Tanulmányunkban a randomizált pedagógiai kísérletek 21. századi amerikai elterjedésének és a fogalom recepciójának felvázolására vállalkozunk. Elsőként a randomizált kísérletek fogalmának részletes elemzését adjuk, majd a randomizált kísérletek történetiségével, emberi és szervezeti kihívásaival foglalkozunk. Az Amerikai Pedagógiai Kutatók Szervezete (AERA) által kiadott hat impakt faktoros folyóirat utóbbi 13 évfolyamának cikkein, a címet, az absztraktot és a kulcsszavakat bevonva, tartalomelemzést végeztünk, amely jól jelzi a randomizált kísérletek megjelenését, jelenlegi és várható előretörését a nívós folyóiratokban. Korábbi munkánkban (Csíkos, 2011) már kifejtettük az Amerikai Egyesült Államokban 2001-ben elfogadott 'No Child Left Behind' elnevezésű törvény és a rá egy évre, 2002-ben elfogadott Education Sciences Reform törvény lehetséges hatását a neveléstudományi kutatásokra. Mostani elemzésünk célja kettős. A randomizált kísérletek fogalmának még pontosabb körüljárásával az elmélet és a gyakorlat számára használható leírást szeretnénk adni erről a kísérleti típusról, továbbá a fogalom 21. századi recepcióját kísérjük figyelemmel a tartalomelemzés módszerével.

Kulcsszavak: randomizált kísérletek (randomized experiments), pedagógiai kísérletek (educational experiments), No Child Left Behind, tartalomelemzés (content analysis)

Randomizált pedagógiai kísérletek

A randomizált szó – bármennyire is idegen hangzású – pontos, mással nem helyettesíthető szakkifejezése a pedagógiai kísérletek egy típusának. A köznyelvben, főleg az ifjúság körében gyorsan terjedő random jelzőből képezve egyrészt a véletlenszerűséget fejezi ki, másrészt pedig befejezett melléknévi igenévként utal egy tevékenységre, amelyet a kutatónak szükséges elvégeznie. Oximoronként hat önmagában ez az egyetlen szó, hiszen a véletlenszerűség és a szándékos cselekvés összekapcsolása ölt benne testet.

A véletlenszerűség a kutatásban nem valamiféle baleset vagy kellemetlen mellékhatás, hanem szükségszerűség. A tudományos kutatás egy vagy több pontján olyan jelenségekkel szükséges foglalkoznunk, amelyek véletlenszerűen fordulnak elő. A mérési hiba véletlenszerűségétől a statisztikai következtetéseinkben meglévő bizonytalanságon át elfogadjuk a matematikai értelemben vett, számszerűsíthető bizonytalanságot, mint a kutatási tevékenység velejáróját. Ugyanakkor a pedagógiai kísérletek világában körültekintő tervezés és kivitelezés adja annak lehetőségét, hogy a kísérlet kizárólag és pontosan a vizsgálni kívánt beavatkozás megfigyelésére szolgáljon. Meddig terjed egy kísérlet kivitelezése során a tudatos kutatói döntések köre, és mely pontokon szükséges a véletlenszerűségre (legalábbis az egyforma valószínűségekre) hagyatkozni?

Könnyen belátható, hogy a tudatos tervezésnek és beavatkozásnak határt szab az a kíváncsi, hogy amennyiben a megfigyelt hatás létezik, azt el lehessen különíteni számtalan további, a kísérleti eredményeket leíró tulajdonságokra ugyancsak ható tényezőtől. Például bármely iskolai fejlesztő program esetén joggal feltételezhető, hogy a tanulókat számtalan egyéb fejlesztő hatás is éri. Ha egy kísérletben egy bizonyos konkrét hatást szeretnénk kimutatni, akkor az összes többi tényező hatását ki kell szűrünk. Ennek a kiszűrésnek matematikailag

* Csíkos Csaba egyetemi docens, Szegedi Tudományegyetem Bölcsészettudományi Kar, Pedagógiai Értékelés és Tervezés Tanszék, e-mail cím: csikoscs@edpsy.u-szeged.hu

és a kutatói döntésben egyaránt adekvát módja, ha a véletlen mintaválasztás segítségével megadjuk az esélyt, hogy valamennyi szóba jöhető hatótényező egyformán nyilvánuljon meg.

Kikerülve a véletlen filozófiai fogalmának mélységeit és nehézségeit, a matematika nyelvén megfogalmazott tudományos kijelentésekben a valószínűség fogalmát használjuk. A mintavétel (Csikos, 2009) során a reprezentativitás záloga lehet, ha a vizsgálni kívánt, elvileg végtelenül tág jelenségkör elemei és egyedei egyforma valószínűséggel kerülnek bele a vizsgált mintába. Az azonos valószínűséget a véletlen számok táblázata vagy más, az emberi tudatos beavatkozástól független eszköz biztosítja. Akkor is a valószínűségek matematikai nyelvén fogalmazzuk meg kijelentéseinket, amikor statisztikai összehasonlítást vagy összefüggés-vizsgálatot végzünk. Látszólagos paradoxon, hogy a lehető legbiztosabb, legmegalapozottabb következtetést megfogalmazni óhajtó kutató akkor jár el helyesen, ha a kutatás során egyszer vagy többször a véletlenszerűségre hagyatkozik.

A pedagógiai kísérletekben tudatos kutatói beavatkozással állítunk elő olyan helyzeteket, amelyekre hosszasan kellene várnunk, ha a megfigyelés módszerét alkalmaznánk (Mill, 1872). A tudatos kutatói beavatkozás során ugyanakkor biztosítanunk kell, hogy a megfigyelni kívánt, mesterségesen előállított hatást elkülöníthessük a többi, óhatatlanul jelen lévő hatástól. A vizsgálni nem kívánt, óhatatlanul jelen lévő hatásokat egyrészt akár eliminálni is lehet (mint ahogyan az a laboratóriumokban történik), másrészt pedig azokat lehet hagyni továbbra is szabadon érvényesülni, de azzal a megkötéssel, hogy a kísérleti és a kontroll körülmények között egyaránt és egyformán érvényesüljenek.

A kísérlet eredményei értelmezésének, azaz a kísérleti eredmények kiterjesztésének kardinális kérdése: Bekövetkezett volna-e a megfigyelt változás a kísérlet során, ha nem történik meg a kutató által tervezett és implementált fejlesztő hatás? Ez a „mi lett volna, ha” típusú megközelítés fogja adni a randomizált kísérletek fogalmi jelentőségét. A „mi lett volna” típusú kérdésekre – még gondolatkísérletben is – úgy adhatunk választ, ha a kiinduló állapot még egyszer előállítható. Ha a kiinduló állapot nem állítható elő újra (például amikor a heroinnak a memóriateljesítményre gyakorolt hatását figyelik meg), akkor ex post facto, másképpen felidézett kísérletről van szó. Az ex post facto kísérletekben nem történik kutatói beavatkozás, tehát a klasszikus definíció szerint nem valódi kísérletről van szó.

Hol és hogyan húzható meg a határ a randomizált és a kvázikísérletek között? (A kvázikísérlet kifejezés helyett a nem randomizált kifejezést pontosabbnak és kevésbé degradálónak érezzük, ezért törekszünk majd ez utóbbi használatára.) A különbségtétel lényegét abban találjuk meg, hogy az imént említett „mi lett volna” típusú kérdésnél elképzelt kiinduló állapot esetén a kísérleti beavatkozást véletlenszerűen kiválasztott mintaelemeken végezzük-e el. Ennek pontosabb megvilágítására egy nagyon egyszerű gondolatkísérletet végezzünk el először!

A papírhajókkal játszó kisgyermek észreveszi, hogy ha kavicsot helyez a papírhajóba, akkor az gyorsan elsüllyed. Tapasztalatát megerősítve látja több egymás utáni próbálkozásban, és szeretné anyukájának is bemutatni felfedezését: „Nézd, ha kavicsot teszek bele, pillanatok alatt elsüllyed!” Érti, hogy nem elegendő egyetlen hajót elsüllyesztenie a demonstráció során, így egymás után három kishajóba is kavicsot helyez. Azonban, tudatosan vagy tudattalanul, éppen a legkevésbé szépen hajtogatott papírhajók esnek áldozatul a kísérletnek. A kísérlet nem volt randomizált, mert a kiinduló állapotban a véletlenszerű kiválasztás nem valósult meg. Egy olyan tulajdonság is meghatározta, mely hajók kerülnek a „kísérleti csoportba”, amelyek a vizsgált tulajdonsághoz semmi köze, hiszen a szépen és a csúnyán hajtogatott papírhajók is gyorsabban elsüllyedtek volna a beléjük rakott kavicsból.

A kutató – ha valóban a világ összefüggéseire kíváncsi, és nem prekonceptiót hajszol – nem tudhatja előre, mely tulajdonságoknak lesz köze a kísérlet kimeneteléhez, így azokat a tulajdonságokat, amelyeket nem kíván

tudatosan befolyásolni, a véletlenszerűségekre kell bízni. Ezt szolgálja a randomizált mintavétel. A kisgyermek iménti kísérlete úgy lett volna még inkább érvényes, ha számos papírhajó közül véletlenszerűen választja ki a kísérleti hatásnak kitetteket. Még nyilvánvalóbb a randomizálás jelentősége, ha egy másik kisgyermek is sok papírhajót hajtogatott volna, és együtt vizsgálják meg, elsüllyednek-e a hajók a kavics hatására, ám csak egyikük papírhajóiba tesznek kavicsot. Ebben az esetben a kísérleti hatáshoz már talán közelebb álló tulajdonságban különbözik a kísérleti és kontroll minta (nem a hajtogatás szépségében, hanem esetleg technikájában lehet különbség köztük). Ez utóbbi esetben még akkor is fenntartásaink lennének a kísérlet eredményeivel kapcsolatban, ha néhány további tulajdonságról megbizonyosodna, hogy azok a kísérleti és kontroll papírhajók mezőnyében egyformák: magasság, súly.

Az iménti papírhajós, nem randomizált kísérletekben a kisgyermek igyekezett megteremteni a kísérleti és a kontroll csoportok ekvivalenciáját azáltal, hogy néhány kikutatott tulajdonság egyezőségét biztosította. Mindig lehetnek ugyanakkor olyan további tulajdonságok, amelyek egyformasága nincs biztosítva. És mivel számtalan sok mérhető tulajdonság és tényező azonosságát kellene biztosítani, az objektivitást és a kísérlet érvényességét az segíti elő, ha mindezt a véletlenszerűsége, tehát az azonos valószínűséggel kísérleti csoportba kerülésre bízjuk.

Feltűnő, hogy gyakran éppen a legizgalmasabb, a pedagógiai fejlesztő hatások törvényszerűségeinek kutatásában a laikusok által is gyakran feltett kérdésekre a legnehezebb valódi, azaz randomizált kísérletet tervezni. Néhány ilyen példa: a kétnyelvű környezet hatása a tanulói személyiségjegyekre; a hat és a hét éves kori iskola-kezdés összehasonlítása az iskolai előmenetel szempontjából; a tankönyvbevételek kérdései; a bukattatás vagy a gyorsítás következményei a személyiség fejlődésében. Közös ezekben a pedagógiai kérdésekben, hogy hosszú távú hatásokat feltételezünk, amelyek számos további tényezőnek köszönhetően alakulnak ki. De nem ez okozza a randomizálás nehézségét! Sokkal inkább két olyan további tényezőt sejtünk, amelyek csaknem kizárják a randomizált kísérlet lehetőségét: a kísérlet iránti elkötelezettség kérdése és a jogszabályokból következő választási szabadság (vagy annak hiánya).

A nem randomizált pedagógiai kísérletekben általában időrendi elsőbbsége van a kísérleti mintának, azaz jellemző, hogy a kísérletben résztvevő iskolák vagy pedagógusok egyetértésének megszerzését követően kerül sor a kontroll illesztésére és az egyforma kiinduló helyzetnek az igazolására. A kísérleti és kontroll csoportoknak akár egyidejű, akár időben egymást követő kiválasztása esetén felmerül a kérdés, hogy miért kerül egyik vagy másik tanuló, osztály, iskola a kísérleti és mások miért a kontroll csoportba. Úgy tűnik, sok esetben technikai jellegű a válasz: a kísérleti csoport résztvevőinek egyetértésének megszerzése történik meg (akár egyidejűleg a kontroll csoport résztvevőinek hozzájárulásával), vagyis mintha a randomizálás és a nem randomizálás közötti választásvonal egyszerűen a kísérletben történő részvétel fogalmának kiszélesítésével áthidalható lenne. A randomizált kísérletekben a kiinduló mintánk részvételét két lépésben biztosítjuk: elsőként a sorsolási kalapba kerüléshez szerezzük meg a hozzájárulást, majd a sorsolási kalapba kerültek egy része kísérleti, más része kontroll körülmények közé kerül. Hogy ez a valóságban miért nem ilyen egyszerű, arról a „Miért olyan ritka...” című részben szólnunk majd.

Az NCBL és az Education Sciences Reform (ESR) törvények a pedagógiai kutatásról

Korábbi munkánkban (Csíkos, 2011) már elemeztük a No Child Left Behind (NCBL) és az Education Sciences Reform törvények fogalomhasználatát, amely a pedagógiai kísérleteket piedesztálra helyezte a kutatási módsze-

rek között, és ezen belül az úgynevezett randomizált kísérletek fontosságát tette nyilvánvalóvá. Egészen pontosan két logikai lépés az, amit a törvények szövege alapján meg tudunk tenni, és amelyek végén a randomizált pedagógiai kísérletek preferenciája áll. Az NCBL törvényben az értelmező definíciók között szerepel a „tudományosan megalapozott kutatás” kifejezés. Megjegyezzük, maga a kutatás fogalom például a hazai jogrendben tudományos kutatást jelent, tehát mintha redundáns lenne a kutatás szóhoz külön jelzőt fűzni. Nyilvánvaló a szándék, hogy ezzel a szókapcsolattal a kutatások egy részének előnyben részesítése a cél, és valóban a törvény szóhasználatában is preference (előnyben részesítés) szerepel, amikor a „tudományosan megalapozott kutatás” típusain belül a kísérletek között a randomizált kísérletek preferálásáról van szó. A törvény szövege ugyanakkor a tudományosan megalapozott kutatás körében említi a kvázikísérleteket és más kísérleti elrendezéseket is, ahol a változók kontrollja megvalósul.

Az NCBL törvényben tehát a tudományos kutatás egyik jellegzetes módszertani típusát, a kísérleteket kitüntetett figyelem övezi, és ezt tovább fokozza, elaborálja az ESR törvény, melyben a kísérletek mellett az értékelés mint kutatási módszer érvényességének biztosítékaként is a véletlenszerű elrendezést nevesíti. Érdemes megnéznünk a randomizált kísérletek törvényi említését megelőző és az azt követő időszak tudományos és oktatáspolitikai törekvéseit, hogy ezáltal világossá váljanak a törvénysszöveget illető kritikák, melyek jobbnak tartották volna, ha a tudományos közösség önszabályozó folyamatainak termékeként válik meghatározóvá egy-egy kutatási módszer (ld. *Berliner, 2002; Moss, 2014*).

A randomizált kísérletek elsőként az agrártudományban terjedtek el (*Shadish, Cook és Campbell, 2002*) annak köszönhetően, hogy megfelelő kivitelezés esetén a vizsgált tulajdonságokra ható külső tényezőkben mutatózó variancia kiegyenlítődik anélkül, hogy ezeket a külső tényezőket természetellenes laboratóriumi körülmények között kellene izolálni. A randomizált kísérletek előnyei mellett sem állítható azonban, hogy csupán ez az egyetlen érvényes kutatási módszer van oksági kapcsolatok tudományos feltárására.

A 20. században, 1925 után lendült föl a randomizált kísérletek alkalmazása a pedagógiai és pszichológiai szakirodalomban, de „valójában nem sok tudományos kutatásban használatos manapság” (*Rubin, 1974. 688.*). A randomizált kísérletek részarányára vonatkozóan *Cook (2002)* több becslést is megemlít. Amikor külön kiemeli az oktatáspolitikai relevanciájú kísérleteket, mindössze hármat tud felsorolni. Tisztázni szükséges, mit ért oktatáspolitikai relevancia alatt: indirekt értelmezésében azok a kísérletek tartoznak ide, amelyeknek közük van az iskolához, tehát amelyek az oktatási-nevelési rendszer döntéshozói számára releváns információt hordoznak. Nem tekinti ebbe a körbe tartozónak azokat a laboratóriumi vagy földrajzi helyhez kötött vizsgálatokat, amelyeknél számos körülmény (melyek egyébként pedagógiai szempontból fontosak lennének) nem reprodukálható máshol. Az iskola mint intézmény lehet az a legkisebb egység, amely döntéshozói szinten releváns pedagógiai kísérletek alapegységeként szolgálhat. Akár egyetértünk *Cook* gondolatmenetével, akár nem, figyelemre méltó, hogy az általa említett három releváns kísérletnek is csupán egyike zajlott iskolai közegben.

A pedagógiai területen született disszertációk kevesebb, mint 1%-ában volt beszámoló randomizált kísérletről (1999-es amerikai adat), a jelentősebb iskolai fejlesztésről szóló folyóiratcikkekben szintén 1% körüli értéket mutat a randomizált kísérletek aránya. A National Reading Panel (ld. bővebben: *Józsa és Steklács, 2009*) vizsgálatban áttekintett publikációkban kicsivel 1% fölötti volt a randomizált kísérletek aránya. *Cook* cinikusnak ható megjegyzése szerint: nyilvánvaló, hogy lehet pedagógiai témákban is kísérletezni, csak éppen nem megszokott. Az USA kormánya 2002 óta több dokumentumban kifejezte szándékát az olyan oktatási reformok folytatása mellett, amelyek tudományosan megalapozott kutatási eredményekre épülnek. A gyengén teljesítő iskolák fejlesztéséről szóló, blueprintnek nevezett stratégiai tervben az ilyen iskolák fejlesztésének négy lépcsőfokát írják le. Legenyhébb esetben is igazgató váltás következik be, leg súlyosabb esetben az iskola bezárása, és tanulóinak

szétosztása az iskolakörzet jobban teljesítő iskoláiban. Az iskolakörzetek mindehhez 3+2 éves pénzügyi támogatást kapnak, mely elnyerésének prioritásai között szerepel, hogy az iskolakörzet vállalja „az elérhető legerősebb bizonyítékokra alapozott fejlesztő programok implementációját” (U.S. Department of Education, 2010. 27.). Az pedig, hogy mi számít legerősebb bizonyítéknak és ezáltal a finanszírozás egyik feltételének, a 2001-es és 2002-es törvények szövege alapján kikövetkeztethető.

Miért olyan ritka a randomizált kísérlet a pedagógiai kutatásban?

Cook (2002) hasonló című tanulmányában tág teret kap a lehetséges okok elemzése. Gondolatmenetét az alábbiakban kritikai elemzés alá vetjük. Kiindulásképpen bemutatja, hogy egy viszonylag közeli területen, tehát emberekkel foglalkozó kísérletező tudományágban, például az orvostudományban elterjedtek és megszokottak a randomizált kísérletek. Ennek fő okait abban látja, hogy ott a finanszírozó ügynökségek, a publikációs színterek ezt pártfogolják, és az egyetemi képzés és a klinikai tesztelés hagyományai is erősítik ezt a szemléletet. Iskolai keretek között hasonló kultúrát képvisel néhány rövid időtartamú prevenció program. Ugyanakkor a kormányzati dokumentumokban nem látja nyomát olyan jelentéseknek, amelyek a pedagógiában a randomizált kísérletek szerepét erősítenék. (Mindezt az NCBL törvény kiadásának „előestéjén” írja.)

A múltbéli tapasztalatok *Cook* szerint segítik annak megértését, miért nem terjedtek el az iskolai randomizált kísérletek. Egy 24 iskolára kiterjedő kísérletben, amelyben tehát a randomizáció szellemében a 24 az összes iskolát jelenti, melynek fele vált kísérleti, fele pedig kontroll iskolává, az egyik kontroll iskola igazgatója azonnal visszalépett a részvételtől, mihelyt kiderült, hogy nem kísérleti, hanem „csupán” kontroll iskolaként számítanak rájuk. Ráadásul a kísérleti iskolák közül is három visszalépett, mert a kísérlet három éves időtartama alatt lejárt a kísérletet felvállaló igazgatók mandátuma. Úgy tűnik, a randomizált kísérleti elrendezés kialakításának egyik fő akadály a lélektani természetű: a kísérletbe bevonás első lépcsőjében olyan szereplők válogatása történik, akik remélhetőleg vállalják akár a kísérleti, akár a kontroll csoport szerepét, ám a második lépésben, a véletlenszerű besorolást követően megváltozhat mindkét csoportba sorolt szereplőknek a hozzáállása a kísérlethez, annak lefolytatásához. A kísérletet vezető kutatónak döntő szerepe van abban, hogy a nem randomizált vagy kényelmi mintát elkerülve ragaszkodjon a két lépésben történő kiválasztáshoz. Előnyös lehet, kompenzálva a kontroll csoportba kerülők vélhető csalódottságát, ha a kontroll csoport számára elsőként ajánlja föl a kísérleti program kipróbálását, mihelyt lefut a tudományos vizsgálat. Szintén a kompenzáció kategóriába illik a részvétel olyan módon történő díjazása, amely a kontroll csoport státus erkölcsi és anyagi elismerését egyszerre megvalósítja: a kutatáshoz kapcsolódó elő- és utómérések lebonyolítása nyilvánvalóan szakmai feladat, amelynek megvalósítása szerves része a kísérletnek. A vezető kutató határozott kísérletvezetését tovább erősíti, ha a kutatást finanszírozó külső forrás eleve randomizált kísérletet favorizál.

További tényezők, amelyek előmozdíthatják a randomizált kísérleteket: rövid idejű fejlesztő program (ám ez ellentmond annak, hogy a döntéshozói relevanciájú programok általában hosszabb, tanévnyi időtartamúak); ha nem kell plusz továbbképzés a résztvevők számára; ha sikerül olyan célkitűzést megfogalmazni, ami fontosabb, mint a részvételért járó ellenszolgáltatás; ha különböző módszerek összehasonlítása történik; és ha a résztvevő csoportok nem érintkezhetnek egymással.

Ez utóbbi tényező olyan megfontolást érdemel, amelyet az internet széleskörű elterjedése előtti korszakban másképp volt célszerű megfogalmazni, mint mostanság. Az egymást átszövő virtuális hálózatok korában aligha életszerű azt feltételezni, hogy egy valóban izgalmas és releváns kísérlet zárt ajtók között maradhat. Az egymással információt cserélő pedagógusok, de még inkább az egymással minden lényeges-lényegtelen dolgot meg-

osztó diákok kommunikációjából pillanatok alatt kiderül, ha valahol lényeges innováció zajlik. A felhasznált tananyagok, oktatási segédeszközök, módszerek titkossága sem szavatolható, hiszen ami egyszer akár csak a képernyőn megjelenik, az eséllyel felkerülhet a világhálóra. Emiatt a kísérleti és kontroll csoportok közötti, illetve még általánosabban a pedagógiai kísérletet érintő kommunikáció aligha akadályozható meg. Elképzelhető, hogy napjainkban bármely kísérletnek ab ovo része az a lehetőség, hogy a benne felhasznált szellemi termékek részben vagy egészen nyilvánosságra kerülnek.

A kutatás tervezésekor mindezek után fontos alapelv, hogy kísérleti és kontroll osztályok ne egy intézmény falain belülről szerepeljenek. Ebből az következik, hogy ha az osztály a randomizáció alapegysége, akkor iskolánként egy-egy osztály szerepelhet, így végül földrajzilag és szervezetileg igen kiterjedt mintára lesz szükségünk, hogy statisztikai értelemben megbízhatóan tudjuk leírni a kísérlet eredményét. Ha iskolák jelentik a randomizáció alapegységét, akkor a változatosságot, a kísérletben a randomizáció segítségével kordában tartott heterogenitást számos iskola, összességében pedig sok ezer gyerek bevonása biztosíthatja. Ugyanakkor a nagyon nagy mintákon, hasonló kvantitatív mutatók mellett, a kísérleti hatás nagysága lecsökkenhet (Csíkos, 2009).

Shadish, Cook és Campbell (2002) beszámolnak arról a tapasztalatról, mely szerint az 1960-as években az USA-ban, *Johnson* elnök idején lebonyolított Great Society (Nagy Társadalom) mozgalom projektjei részben annak köszönhetőek sikertelenségüket, hogy szövetségi szintről rótták ki a résztvevőkre. Nincs arra vonatkozó tudományos bizonyíték, hogy általában ez volna a helyzet (ti., hogy helyi igényeket kielégítő kísérletek sikeresebbek, mint a központilag vezényelték), azonban egy érdekes és feltétlenül tudatosítandó jelenség napvilágra került. Bármiféle kísérletben részt venni annak a lehetőségnek a megengedését jelenti, hogy valamit jobban is lehet csinálni, mint ahogyan az jelenleg folyik. Másképpen megfogalmazva, a pedagógiai kísérleteket tervező kutató jó eséllyel találhat olyan kísérleti vagy kontroll osztályt vagy iskolát, amelynek akkor éppen nincs igénye vagy szüksége a fejlesztő kísérlet eredményeire. (A randomizáció arra is alkalmas, hogy ilyen osztályok és iskolák egyforma arányban szerepeljenek a kísérleti és a kontroll csoportban is.)

A remélt kísérleti eredmények iránti attitűd meghatározza, mely iskolák kerülnek bele a randomizáció előtt kialakítandó sorsolási kalapba. Mindez oda vezet, hogy a randomizált kísérletek eredményei arra az alapsokaságra lesznek általánosíthatók, amelynek leírásában szerepet kap valamiféle részvételi hajlandóság a pedagógiai innovációban. Ez a leírás már nem sokban különbözik a nem randomizált kísérletekbe bevont kísérleti osztályok és iskolák azon tulajdonságától, hogy ők azok, akik szeretnének egy pedagógiai kísérlet részesei lenni. Így vagy úgy, akár randomizált, akár nem randomizált kísérletek esetén megjelenik és hangsúlyos a részvételi hajlandóság emberi tényezője, mely elsősorban az iskolavezetéshez, másodsorban a tantestülethez mint szervezethez kapcsolható.

Cook (2002) végül súlyos következtetésre jut: elképzelhetőnek tartja, hogy az oktatási rendszer még nem érett meg a randomizált kísérletekre! Hiszen a randomizált kísérletek implementációja, melynek része a hajlandóság megfogalmazása arra vonatkozóan, hogy akár kísérleti, akár kontroll csoportként részesei legyenek az osztályok vagy iskolák egy kutatásnak, pozitív menedzseri szemléletet feltételez. A részvételi hajlandósághoz szükséges menedzseri szemlélet megragadása ugyanakkor a kvalitatív eszközökkel megragadható szervezeti kultúra ismeretét feltételezi. *Cook* végletes következtetésre jut ezek után: nem kellene jelenleg a randomizált kísérletek, mert vannak helyettük jobb és oktatáspolitikai szempontból releváns kutatási módszerek: intenzív esettanulmányok, és kvázi-kísérletek!

Tartalomelemzés az AERA folyóirat-cikkei alapján

Az Amerikai Egyesült Államok szövetségi jogalkotásában 2001-ben és 2002-ben jelentek meg azok a dokumentumok, amelyek nyíltan megfogalmazzák álláspontjukat a pedagógiai kutatás módszereiről, előtérbe helyezve a randomizált kísérleteket. Vajon milyen hatása lett ezeknek a jogszabályoknak az azóta eltelt időszak kutatásaira? Megnőtt-e a randomizált kísérleteket bemutató tudományos közlések aránya? E két kérdés egymás mellé helyezése kifejezi azt a meggyőződésünket, hogy a tudományos kutatás fázisai között alapvető a publikálás, és a szakmai párbeszédközösség elsősorban a publikált kutatásokra hagyatkozik, amikor a tudományos eredmények gyarapodását és a megjelenő új tendenciákat követi nyomon. Az Amerikai Neveléstudományi Kutatók Szervezete (AERA, American Educational Research Association) összesen hat folyóiratot ad ki, amelyek közül kettő a Thomson Reuters impakt faktor szerinti listáján (ld. bővebben Csíkos, 2013) az első tíz között van.

Top 10 folyóiratok a Web of Science rendszerében az Education and Educational Research területen:

1. Review of Educational Research (AERA lapja),
2. Educational Psychology,
3. Journal of Research on Educational Effectiveness,
4. Educational Research Review,
5. Learning and Instruction,
6. Journal of Research in Science Teaching,
7. Educational Researcher (AERA lapja),
8. Science Education,
9. Journal of the Learning Sciences,
10. Journal of Engineering Education.

További AERA lapok:

15. American Educational Research Journal
21. Educational Evaluation and Policy Analysis
63. Journal of Educational and Behavioral Statistics
80. Review of Research in Education

A vizsgálatunkban lehatárolt időszak a 2002 és 2014 közötti 13 évfolyam. Ismert jelenség, hogy a vezető folyóiratoknál, különösen a nem angol anyanyelvű szerzők esetén, átlagosan nagyjából egy éves átfutási idővel kell számolni a kézirat beadásától az oldalszámmal és doi azonosítóval ellátott megjelenésig. Emiatt a 2002-es évfolyam akár túl közelinek is számít a 2001-es megszövegezésű, ám 2002. január 8-án életbe léptetett NCBL-hez törvény szempontjából. Ugyanakkor ismerjük azt a jelenséget is, miszerint egy már elvégzett kutatás eredményei a folyóirat (ezen belül a főszerkesztő) által követett paradigmának megfelelően többféle címmel és hangsúllyal publikálhatók. Emiatt előfordulhat, hogy egy átdolgozás alatt lévő kéziratban, annak tartalmán és adatain nem változtatva, a szerkesztők kérésére megváltozik a cím, akár a randomizált szót is magába foglalóan.

A folyóirat-címek tartalomelemzésének bőséges szakirodalmi előzménye és tudományos háttere van. *Timonen* és *Paloheimo* (2008) fölfogását követjük abban, hogy címeket, kulcsszavakat és absztraktokat vizsgáltunk, amikor egy konkrét fogalom folyóirat-cikkekben való megjelenését kutatjuk. Szerintük így ugyanis biztosítható, hogy olyan cikkek kerüljenek kiválasztásra, amelyek szándékoltan, és nem csupán érintőlegesen „not just use it in passing” (u. o. 179.) tárgyalják az adott fogalmat. *Timonen* és *Paloheimo* a Social Science Citation Index korpuszát vették alapul, és az 1974 és 1992 közötti időszakban 273 cikk elemzésével dokumentálták a „knowledge

work" (tudásmunka) fogalom megjelenését és elterjedését. Ugyanezt a tartomelemzési mélységet választotta *Seale* (2008), aki kilenc, az egészségügy szociológiája területéhez kötődő amerikai és brit kiadású lapot hasonlított össze számos szempontból. A teljes absztraktszövegek bevonásával közel 800 ezer szavas korpuszt elemeztek, mely alapján az egyébként angliai szerző a brit folyóiratok számára fogalmaz meg javaslatot a nyilvánosság és a szakpolitika felé nyitás követelményével.

Szűkebb tartalmi körre terjedt ki *Mallen, Stevens és Adams* (2011) tartomelemzése, akik sporthoz kötődő folyóiratokban vizsgálták a környezeti fenntarthatóság fogalmát. Ők a címet és a kulcsszavakat figyelték, és az 1987 és 2008 közötti időszakban 21 folyóiratban összesen 17 vonatkozó cikket találtak (0,4%-a volt ez az összes 4639 folyóiratcikknél).

Bővebb tartalmi elemzés, azaz teljes cikkszövegek figyelembe vétele valósult meg *Blumer, Green, Knowles és Williams* (2012) transzneműekkel kapcsolatos kutatásában. A téma megjelenését keresték 17 folyóirat 10 839 cikkében az 1997 és 2009 közötti időszakból. Kijelölték előzetesen azoknak a kulcsszavaknak a listáját, amelyek előfordulása a keresett téma megjelenését jelzi, és teljes folyóirat-szövegeket vizsgáltak. Összesen így is mindössze 9 témájukhoz kötődő cikket találtak.

Látható, hogy egy-egy szakszó tudományos folyóiratokban megragadható jelenlétét különböző mélységű tartomelemzési technikákkal lehet nyomozni. Mostani témánk és erőforrásaink azt a lehetőséget preferálják, hogy amerikai vezető folyóiratok cikkeinek címét, absztraktját és kulcsszavait emeljük be az elemzésbe. Ezek alapján az amerikai jogalkotásnak a pedagógiai kutatásra, ezen belül kiemelten a publikációs stratégiára tett hatásának elemzése során az AERA hat folyóiratának 13 évfolyamából a címek, absztraktok és kulcsszavak között vizsgáltuk meg a random vagy randomized szavakat tartalmazókat. A randomized helyett gyakran a random assignment kifejezés használatos. Ugyanakkor a random szó előfordulhat olyan publikációk címeiben és absztraktjában, melyek nem kísérleti eredményekről vagy kutatás-módszertani elemzésről számolnak be.

1. *Review of Educational Research*, az AERA zászlóshajója, kerekén 5-ös impakt faktoral: A lapban számos metaanalízis jelenik (a metaanalízisekről magyar nyelven ld. *Csapó, 2007*), melyek egy-egy tartalmi terület korábbi empirikus kutatásait szintetizálják. Kiemelkedik témánk szempontjából ezek közül *McEwan* cikke, amely randomizált kísérletek metaanalízisét közli a fejlődő országok kisiskolásai körében végzett fejlesztő programok köréből. Figyelemre méltó tehát, hogy egy modern metaanalízisben a benne áttekintett kísérletek szűrőjeként jelent meg a randomizálás 2014-ben. Néhány további metaanalízis mellett a lapban még egy 2005-ös cikk számol be randomizált kísérletről, és feltűnően megritkulnak a randomizálást említő cikkek az ezt megelőző időszakban.

2. *Educational Researcher*: A szintén top 10-es lapban gyakran jelennek meg kutatás-módszertani hangsúlyú cikkek. Itt jelent meg 2002-ben *Slavin* – azóta sok száz hivatkozással rendelkező – cikke a bizonyítékokra alapozott oktatáspolitikáról (evidence-based policy), majd nagyjából évente egy-egy olyan cikket közölt a lap, amely a randomizált kísérleti megközelítésmódot tárgyalta. 2012-ben megugrott a randomizált kísérletekkel foglalkozó cikkek száma, annak köszönhetően, hogy a lapban megjelentetett két cikkre (egyik ráadásul ugyanabban a lapszámban jelent meg) három kommentár jellegű cikk született, és mind kulcsszóként használta a randomizált kísérletet.

3. *American Educational Research Journal*: 2005-től indulóan rendszeresen megjelennek randomizált kísérletek. 2009-ben egy ilyen sem volt, 2012-ben pedig öt, összesen a szóban forgó időszakban 22. Tekintetbe véve, hogy a 13 évfolyam lapszámaiban összesen 451 cikket találunk, ez magas aránynak tekinthető, különösen, ha figyelembe vesszük, hogy 1964-től 2002-ig számítva 29 cikk címében, absztraktjában vagy kulcsszavai között fordult elő a random vagy randomized szó.

4. *Educational Evaluation and Policy Analysis*: Profilja alapján ennél a lapnál magasabb előfordulási gyakoriságra számíthatunk. A vizsgált 13 évfolyamban 39 cikk vonatkozott randomizált kísérletekre, és ezek darabszáma az egyes évfolyamokon érdekes utat jár be, amelyet akár jellemzőnek is tarthatnánk, ha nagyobb számú megfigyelésre támaszkodhatna az elemzés. A kezdeti, évenkénti 2-3 előfordulás 2011-ben 4-re nő, majd 2012-ben 6, 2013-ban 8 cikket találunk, míg 2014-ben csak hármat.

5. *Journal of Educational and Behavioral Statistics*: A lap témakörei között szereplő matematikai statisztikai elemzések nagyban emelik a random szó előfordulását olyan esetekben is, amikor nem kísérletekről van szó. A 2012-ben a randomizált kontrollált kísérletek hatásméretének növekedéséről szóló tanulmány azonban beleillik elemzésünkbe, és hasonló módon végigtekintve a megjelent cikkeken 2012-ben még egye, majd 2011-ben szintén kettő, 2010-ben egy, a korábbi években viszont elvétve egy-egy cikk foglalkozott a randomizált kísérletekkel.

Végül, a 6. *Review of Research in Education*, amely vállaltan egy-egy témakör esszéisztikus feldolgozására törekszik, még nem szentelt lapszámot a randomizált kísérletek – alapvetően kutatás-módszertani és oktatáspolitikai színezetű, és tartalmilag kevésbé jól fókuszálható – témájának. A lap évente egyszer jelenik meg, így egy lapszám egy teljes évfolyamot jelent. Az utóbbi években a következő vezértémákat elemezték: ifjúság, kultúra, nyelv és műveltség; nevelés, nevelés és közjó; különleges bánásmód pedagógiája; legutóbb pedig nyelvpolitika és diverzitás volt a középponti téma.

A vezető amerikai pedagógiai tudományos folyóiratok utóbbi tizenhárom évfolyamában, tehát a 2001–2002-es jogalkotási mérföldkőtől napjainkig jól láthatóan jelen vannak a randomizált kísérletek: egyrészt kutatás-módszertani hangsúlyú elemzésekben, másrészt tényleges empirikus tanulmányokban megjelenő kísérleti formaként. A cikkeket összességében nagyobb számban megjelentető lapok (*Educational Researcher* és *American Educational Research Journal*) esetén nemcsak a téma jelenléte jól látható, hanem az is világos, hogy a 2002 előtti időszakhoz képest látványossá vált a randomizált kísérletek jelenléte. A közeljövő titka, hogy az *Educational Evaluation and Policy Analysis* lapnál 2014-ben megjelenni látszó, az azt megelőző két évhez képest szembetűnő visszaesés a téma „kifulladását” jelzi vagy visszatérést a 2002-től lényegében egyenletesnek tekinthető évi átlag 3-4 randomizált kísérleti témájú cikkmennyiséghez. Tekintettel arra, amit a randomizált kísérletekről – lehántva bármiféle oktatáspolitikai vagy jogalkotói hátszelet – elmondhattunk, nem számítunk a téma iránti érdeklődés és a téma publikációs jelenlétének lanygulására, és úgy sejtjük, a következő években is a megjelent cikkek néhány %-a a randomizált kísérletekhez fog kötődni a vezető amerikai lapokban. Ebből következően az európai és világszervezetek által kiadott lapokban is hasonlóra számítunk.

Tanulságok a hazai neveléstudomány számára

Falus (1993) szerint a pedagógiai kutatás során a személyiség fejlesztésében érvényesülő törvényszerűségeket kutatjuk. A fejlesztés mint kulcsszó kiemelése lehetővé teszi, hogy az önmagában vett fejlődésanalízist (amely során nincs lehetőség vagy igény a fejlesztés törvényszerűségeinek vizsgálatára) a pszichológiai tudomány tárgykörébe tartozónak tekintsük. A fejlesztés a tudományos kutatás módszertana szempontjából olyan ráhatás, amely tervezett, tudatos, emellett pedig kontrollált. A pedagógiai kísérletek kutatómódszertanának mindezekből következően a pedagógiai kutatómódszertan középpontjában a helye. A pedagógiai kísérletek kutatómódszertana pedig a randomizált kísérleteket tekinti a kísérletek családjában a tudományosság követelményét legmagasabb szinten megvalósító formának. Akkor is, ha más kísérleti típusok (vagy tágabban: más kutatómód-

szertani eszközök) könnyebben, reálisabban kivitelezhető, a randomizált kísérletek egyfajta szakmai etalonként állnak előttünk.

Bár a kutatás-módszertani precizitás és megalapozottság szempontjából világos a randomizált kísérletek szerepe és erényei, a tanítás-tanulás tudományos törvényszerűségeinek kutatásában egyrészt más módszerek is megkerülhetetlenek, másrészt pedig előfordulhat, hogy egy randomizált kísérlet a benne vizsgált téma vagy az alkalmazott mérőeszközök miatt mégsem vezet jelentős tudományos eredményekhez. Sokat árthatna a randomizált kísérletek fontosságába vetett hitünknek, ha központilag előírta, kötelezően, finanszírozást mellérendelve írta elő a randomizált kísérletek végzését valamilyen autoritás. Látva a tengerentúli fejleményeket, a hazai neveléstudomány számára a következő javaslatot fogalmazzuk meg: a meglévő projektek és infrastruktúra felhasználásával minél több randomizált kísérletet végezzünk el; alakuljon ki ezeknek egy hazai, terebélyesedő publikációs korpusza, és juttassuk el a kísérleti eredményeket a legilletékesebb ítések, a leendő és gyakorló pedagógusok képzési programjaiba. Mindezek után később a központi, hatósági jellegű hátszél is elfogadható lesz a randomizált kísérletek finanszírozását és elterjesztését illetően.

Szakirodalom

1. Blumer, M. L. C., Green, M. S., Knowles, S. J. and Williams, A. (2012): Shedding light on thirteen years of darkness: Content analysis of articles pertaining to transgender issues in marriage/couple and family therapy journals. *Journal of Marital and Family Therapy*, 38. 1. 244–256.
2. Cook, T. D. (2002): Randomized experiments in education: Why are they so rare? *Educational Evaluation and Policy Analysis*, 24. 3. 175–199.
3. Csapó Benő (2007): Hosszmetszeti felmérések iskolai kontextusban – az első átfogó magyar iskolai longitudinális kutatási program elméleti és módszertani keretei. *Magyar Pedagógia*, 107. 321–355.
4. Csíkos Csaba (2009): *Mintavétel a kvantitatív pedagógiai kutatásban*. Gondolat Kiadó, Budapest.
5. Csíkos Csaba (2011): *Pedagógiai kísérletek kutatómódszertana*. Gondolat Kiadó, Budapest.
6. Csíkos Csaba (2013): Az impakt faktor a neveléstudományi kutatásokban. *Iskolakultúra*, 23. 9. 3–16.
7. Falus Iván (1993): A pedagógiai kutatás metodológiai kérdései. In: Falus Iván (szerk.): *Bevezetés a pedagógiai kutatás módszereibe*. Keraban Kiadó, Budapest. 9–33.
8. Józsa Krisztián és Steklács János (2009): Az olvasástanítás kutatásának aktuális kérdései. *Magyar Pedagógia*, 109. 365–397.
9. Mallen, C., Stevens, J. and Adams, L. J. (2011): A content analysis of environmental sustainability research in a sport-related journal sample. *Journal of Sport Management*, 25. 240–256.
10. Mill, J. S. (1872): *A system of logic*. 8th edition. Longmans, Green, Reader, and Dyer, London.
11. Moss, P. A. (2014): Critiques of Educational Research. In: D. C. Phillips (ed.): *Encyclopedia of Educational Theory and Philosophy*. SAGE Publications, 263–270.