

Vélemény, vita

AZ IDÉZETTSÉG SZAKTERÜLETI VÁLTOZÉKONYSÁGÁRÓL

Papp Zoltán

a fizikai tudomány kandidátusa,
Debreceni Egyetem, Környezetfizikai Tanszék
zpapp@tigris.unideb.hu

Manapság természettudós körökben a tudományos teljesítmények rangsorolását illetően a kapott (független) hivatkozások számának, illetve az idézettségnek (hivatkozások száma osztva a publikációk számával) van a legnagyobb tekintélye. Elterjedt vélekedés, hogy ezek a mutatók az elvégzett munka mennyisége mellett valamiképpen annak értékét, jelentőségét, minőségét is kifejezik, s ezért minden másnál alkalmasabbak a munka eredményességének megítélésére (lásd például Bencze et al., 1996). A kapott hivatkozások száma egyedi teljesítmények értékelésekor is az egyik legfontosabb adat. Így van ez egyszerű állás- és kutatási pályázatoknál csakúgy, mint akadémiai doktori vagy habilitációs eljárásokban. Esetenként e mutató valamilyen küszöbértékének eléréséhez kötik az eljárás, pályázat sikerét.

Vajon mennyire megérdemelt a fenti mutatók tekintélye, nimbusza? Igazságos-e ezek értékéhez kötni személyes karriereket meghatározó eljárások, pályázatok sikerét? Ha e mutatók értéke egy adott kutató esetében nagy, ez vajon mekkora részben köszönhető az ő személyes érdemeinek, és mekkora

részben rajta kívülálló, tőle független körülmények hatásának? A fentiek szempontjából nem mindegy, hogy e mutatók értékét személyes érdemek 90 %-ban határozzák meg, vagy éppen csak 10 %-ban befolyásolják!

Egy korábbi írásomban tágabb perspektívában már foglalkoztam ezzel a témával (Papp, 2004). Írtam arról, hogy a hivatkozás elsődleges motívuma nem a hivatkozott munka minősége, hanem annak a hivatkozó közleménnyel való tematikai rokonsága, ami egyáltalán nem látszik alátámasztani azt, hogy az idézettség a minőséggel lenne kapcsolatos. Úgy foglaltam állást, hogy a hivatkozásszámra már csak azért sem lenne szabad komoly döntéseket alapozni, mert értéke túl nagy mértékben függ a figyelembe vett hivatkozások fajtájától, a társszerzők számától (ha nem számolunk saját részesedést), azok hírnevétől és nemzetiségétől, a kutatási terület kutatói létszámától, a hivatkozott cikkek sajátos tartalmától, szubjektív tényezőktől, szerencsétől, becsületességtől.

Bár fenti írásomat többen bírálták (Bencze, 2004; Braun, 2004; Marton, 2004; Vinkler, 2004), előbbi véleményemmel kapcsolatos konkrét kritikát csak Vinkler Péter fogalma-

zott meg. Lényegi ellenvetése az volt, hogy az egy cikkre jutó idézetszám általában nem függ a kutatási területet művelő kutatók létszámától. Azt azonban elismerte, hogy az idézetség különböző területeken akár többszörösen is különbözhet.

Függetlenül az előbbi vitától, a szakértők körében már régóta ismert tény, hogy az egy cikkre jutó idézetek száma erős szakterületi függést mutat (lásd például Braun, 2001). Emiatt különböző területeken dolgozó kutatók összehasonlításakor a kapott hivatkozások számát nem, vagy csak igen óvatosan szabadna használni. Ezzel szemben gyakran tapasztalhatjuk, hogy a teljesítmények értékelésével hivatalból foglalkozók hajlamosak ezt a helyzetet figyelmen kívül hagyni.

Az egyedi teljesítmények értékelésére való használhatóság szempontjából tulajdonképpen mellékes, hogy mi okozza az idézetség szakterületi függését. Sokkal fontosabb, hogy ez a függés mekkora mértékű és milyen szerkezetű. Vagyis hányszoros eltérések lehetnek különböző területek között, és ezek a területek mekkorák, és milyen kapcsolatban vannak egymással? Ha egy tudomány különböző ágai vagy az ezeket alkotó még kisebb részterületek között is sokszoros (egy nagyságrendet elérő) eltérések lehetnek az idézetség átlagában, akkor az idézetszámot az adott tudomány keretein belül nem tekinthetjük minőségjelzőnek. Hiszen értékét a személyes érdemekenél sokkal erősebben befolyásolhatja, hogy a szerző melyik részterületen dolgozott.

Márpedig az idézetség egy tudomány különböző részterületei között is fluktuál. Ennek okai egyszerűek. Nyilvánvaló, hogy egy adott tudományterülettel, témacsoporttal, témával foglalkozó kutatóközösség tagjai saját cikkeikben csaknem kizárólag e közösség tagjainak korábbi cikkeire hivatkoznak. Ez azért van így, mert a hivatkozást a tematikai rokonság motiválja, és mert a különböző közösségek által kutatott dolgok, jelensé-

gek és az azok leírásához, vizsgálatához, megértéséhez szükséges fogalomrendszerek, módszerek és tudáshalmazok jelentősen különböznek. A kutatók általában csak saját kutatási területük tartalmát, fogalomrendszerét, módszereit, tudáshalmazát ismerik a kutatáshoz szükséges mélységben, így még akkor sem tudnának más területekre áthivatkozni, ha akarnának. Ezért az A tudomány művelői nem hivatkoznak a B tudomány művelőire, és megfordítva. De ugyanez a helyzet az A tudomány A_1, A_2, \dots, A_i ágai esetében is, és a sor folytatható az A_1 ág $A_{11}, A_{12}, \dots, A_{1i}$ részterületeivel, az A_{11} részterület $A_{111}, A_{112}, A_{113}, \dots, A_{11i}$ alterületeivel, és így tovább. Persze, minél mélyebbre hatolunk, annál kisebb lesz a különbség az egymáshoz közeli részterületek, alterületek stb. tartalma, fogalomrendszerei, módszerei és tudáshalmazai között, ezért lehetségessé és egyre gyakoribbá válik a közeli csoportok közötti áthivatkozás. Az áthivatkozást azonban még itt is gátolhatja az, hogy a kutatóknak csak a saját témáikhoz közvetlenül kapcsolódó eredmények figyelésére van idejük. Tematikus és hivatkozási szempontból is a kutatás legkisebb (nem individuális) egységei azok az ún. elemi témacsoportok, amelyeknél a csoporton belül maradó hivatkozások száma még épp meghaladja a csoportból kifelé menőkét, de amelyek már nem oszthatók tovább ugyanilyen tulajdonságú, még kisebb csoportokra.

Amint azt korábban már leírtam, eléggé kézenfekvő, hogy egy a kölcsönös hivatkozások szempontjából túlnyomóan zárt kutatói csoport hivatkozási mutatóit e csoport saját jellemzői határozzák meg, azokra külső körülmények nemigen hatnak (Papp, 2006). Esetleg a tematikailag leginkább rokon, szomszédos csoportok jellemzőinek lehet némi hatásuk, ha gyakori az onnan/oda áthivatkozás. Ezért közeli csoportok hivatkozási mutatóiban is lehetnek nagy eltérések, ha jelentősen különböznek azok a csoportjel-

lemzők, melyek e mutatókat meghatározzák. De vajon mekkorák az eltérések, és melyek az ezeket befolyásoló csoportjellemzők? Először foglalkozzunk az utóbbiakkal!

A kutatási területeknek, illetve kutatói csoportoknak azok a jellemzői, melyektől a hivatkozási mutatók közvetlenül függenek, a szakemberek előtt nem ismeretlenek. Segítségül hívhatjuk például a Vinkler Péter egyik legutóbbi cikkében leírtakat (Vinkler, 2003). Használjunk mi is az általa használt-hoz hasonló formalizmust, és vizsgáljunk egy a kölcsönös hivatkozás szempontjából nagyjából zárt kutatási területet! Jelöljük P_y -nal az y évben a vizsgált területen megjelent cikkek számát. A szerzők ezekben a cikkekben túlnyomóan azokra a régebbi cikkekre hivatkoztak, amelyek ugyanezen a területen korábban (a terület művelésének kezdetétől az $y-1$ évig) megjelentek. E régebbi cikkek összességét nevezhetnénk a terület meglévő cikkbázisának. Jelöljük az ebbe tartozó cikkek számát $P_{1 \rightarrow y-1}$ -gyel. Jelöljük továbbá r_y -nal az y évben megjelent cikkekben található hivatkozások átlagos (cikkenkénti) számát. Ezek után világos, hogy a területen az y évben összesen $r_y P_y$ számú hivatkozás történt, és ezek a hivatkozások $P_{1 \rightarrow y-1}$ számú korábbi cikk között oszlottak meg. Tehát annak az esélye, hogy egy, a cikkbázishoz tartozó cikk az y évben hivatkozást kapjon, átlagosan

$$CC_y = \frac{r_y P_y}{P_{1 \rightarrow y-1}}$$

ahol CC_y -t hivatkozási esélynek nevezhetnénk. A hivatkozási esély értéke évről évre változhat, miként a kifejezésében szereplő mennyiségek egyenként is, és az átlagon belül az y évhez közelebbi évek (frissebb) cikkeinek általában nagyobb az esélyük, mint a korábbiaknak. Az $y-n$ évben (az y évnél n évvel korábban) megjelent cikkek által összesen (az y évvel bezárólag) kapott hivatkozások száma átlagosan:

$$C_{y-n}^{\text{tot}} = \sum_{i=y-n+1}^y o_i CC_i$$

ahol o_i egy ún. elavulási tényező, amivel azt vesszük figyelembe, hogy a frissebb cikkek hivatkozási esélye nagyobb, a régieké pedig kisebb az átlagosnál. (Az $y-n+1$ -hez közeli i értékekre általában $o_i > 1$, az y -hoz közeli i -kre viszont $o_i < 1$ is lehet, főleg, ha n közelít y -hoz.) A C_{y-n}^{tot} mennyiség egyben az $y-n$ évben megjelent cikkek átlagos idézettsége (az y évvel bezárólag). Látható tehát, hogy a fenti modell szerint az átlagos idézettség az r_y , P_y , $P_{1 \rightarrow y-1}$ és o_i mennyiségektől függ, melyek a kutatási terület mennyiségi jellemzői, értékük területről területre, és egy területen belül időben is változhat.

A cikkenként adott hivatkozások r_y száma különböző területeken többszörösen is különbözhet. Ezt onnan lehet tudni, hogy nagyobb szakterületek folyóirathalmazainak IF (impakt faktor)-átlagait is ez a paraméter határozza meg: az IF-átlag (az IF definíciója miatt) általában kevéssel több, mint a $k_y r_y$ szorzat fele, ahol $k_y (< 1)$ az y évi hivatkozások közül azok részaránya, melyek az $y-1$ és $y-2$ évben megjelent cikkekre irányulnak (Vinkler, 2003). A nagy tudományok IF-átlagait pedig meghatározták, és közöttük többszörös eltéréseket találtak. Például 1998-ban az alapvető élettudományok (fundamental life sciences), a fizika, illetve a matematika és számítástudományok esetében az IF-átlag rendre 3,1, 1,5 és 0,5 volt (Amin – Mabe, 2000), aminek oka a fentiek szerint a megfelelő r_y (és k_y) értékek közti eltérés lehetett. A fentebb leírtak alapján és az IF-ek széles értéktartománya ismeretében nemigen várható, hogy r_y értékei az egy tudományon belüli tudományágak, illetve azok részterületei, alterületei stb. között kevésbé fluktuálnának. Inkább a még nagyobb, akár az egy nagyságrendet is elérő mértékű ingadozásra számíthatunk, hiszen a különböző témákkal foglalkozó, de ugyanazokban a folyóiratok-

ban közlő csoportok hivatkozási esélyeinek fluktuációit az IF elsimítja, kiátlagolja.

Hogy r_y milyen körülményektől függ, azt még a szakértők sem látják világosan, erről korábban már írtam (Papp, 2006). A fentiek szerint csábítónak tűnik azt feltételezni, hogy ha a terület cikkbázisa (P_{1y-1}) nagyobb, akkor r_y is nagyobb, hiszen nagyobb készletből több hivatkozásra alkalmas cikk akadhat. Ez az összefüggés azonban csak az elemi témacsoportok szintjén működhet, szemben a korábbiakkal. (A fenti egyenletek egy teljes tudományra vagy tudományágra is értelmezhetők.) Bármely szerző ugyanis túlnyomórészt saját elemi témacsoportjának cikkbázisából választ hivatkozásokat, nem valamely nagyobb körből. A cikkbázis terjedelme viszont nagyjából a terület múltbeli kutatói létszámával arányos, tehát úgy tűnik, hogy az elemi témacsoportok szintjén r_y is arányos a kutatói létszámmal. Ez a következtetés nem mond ellent annak, hogy az idézettség *általában* nem függ a kutatási terület kutatói létszámától (Vinkler, 2004), hiszen ez utóbbi megállapítás nyilván nagyobb területek adatain alapul. Nem arról van szó, hogy például egy fizikai témájú cikk idézettsége a fizika vagy valamely részterülete teljes népességétől függne, hanem arról, hogy függhet saját elemi témacsoportjának létszámától. Ez persze csak elmélkedés, de az adatbázisokon alapuló részletes vizsgálatokkal ki lehetne deríteni, hogy helyes-e.

A P_y/P_{1y-1} arány általában jóval kisebb egynél, azonban egy még fiatal, éppen felfutó, létszámban gyorsan gyarapodó területen akár többször nagyobb lehet, mint egy nagy múltú, de már visszafejlődő, elnéptelenedő területen. A hivatkozási esélyben és az idézettségben ez önmagában is többszörös szóródást okozhat.

A fenti modell hallgatólagosan azon a feltevésen alapszik, hogy a cikkek hivatkozási időperiódusa egy év, vagyis ugyanazon évben megjelent cikkekre még nem hivatkoz-

hatunk, de az eggyel korábbi évben megjelent cikkekre már igen. Ez az időperiódus azonban a valóságban területről területre, csoportról csoportra változhat, és egy évnél rövidebb és hosszabb is lehet. Gyanítható, hogy az idézettséget ez is befolyásolhatja. Ebből is látszik, hogy a fenti modellt tovább kellene finomítani ahhoz, hogy a valóságot jobban le tudja írni, és valószínű, hogy ehhez még más dolgokat is figyelembe kellene venni. Egész külön szakirodalma van például annak, hogy a hivatkozás aktusát (a tematikai rokonság mellett) milyen további, többnyire szubjektív tényezők befolyásolhatják.

Az itt bemutatott gondolatmenetek és adatok mindenesetre azt valószínűsítik, hogy a hivatkozási esély és az idézettség értékeiben szakterületenként, és különösen az elemi témacsoportok szintjén, esetenként sokszoros, akár tízszeresnél is nagyobb arányú eltérések lehetnek. További vélemények és adatok is ennek lehetőségére utalnak. Verő József a földtudományok felől vizsgálódva számolt be arról a benyomásáról, hogy különböző területek idézettségi átlaga között „szinte nagyságrendnyinek becsülhető” különbségek lehetnek (Verő, 1998). Csaba György a biológia oldaláról magától értetődőnek tartja, hogy a különböző területek kutatói létszámai közötti, nagyságrendekkel mérhető különbség az idézettségben is megmutatkozik (Csaba, 2004). A *Physical Review* című, a fizikában vezető folyóirat 110 évfolyamából (1893-2003) a száz legtöbbszór idézet cikk közül mindössze három volt kísérleti témájú (az 1948-1955 évekből), a többi mind elméleti, ami az elméleti fizika erős idézettségi előnyéről tanúskodik (Redner, 2005). A fizikán belül az idézettségen alapuló új sztár-mutató, a Hirsch-index is erősen az elméleti fizikusoknak kedvez (Bencze, 2006). A *Nature* 2005 nyarán szerkesztőségi cikkben (Editorial, 2005) jegyzi meg, hogy: „The citation rate of our papers

also varies sharply between disciplines. ... Clearly, these reflect differences in disciplinary dynamics, not in quality.”¹

Az idézetelemzés hívei gyakran érvelnek azzal, hogy a kapott hivatkozások száma képes megmutatni a tudományos cikk „hatását”, „felhasználhatóságát”, és ezért a cikk, illetve az azt megalapozó munka „minőségének” mérőszámaként használható. A fentiek alapján azonban úgy tűnik, hogy a munka minőségének hatása az idézettségi mutatókban nagyon nehezen érhető tetten. A cikkek ugyanis többnyire csak saját elemi témacsoportjukban vagy annak közvetlen közelében, az itt tevékenykedők körében tudnak hivatkozásokban kifejeződő hatást kiváltani. A közölt eredményeket csak az e körben dolgozók tudják felhasználni. Ezért a kiváltott hatást e csoport(ok) mennyiségi jellemzői általában jóval erősebben befolyásolják, mint a munka és a cikk minősége. A hatás a minőséggel többnyire csak az elemi témacsoportokon belül korrelálhat, feltéve, hogy a köztük lévő kapcsolatot nem torzítják el az itt csak megemlített, de részletesen nem tárgyalt szubjektív tényezők. Ritkán, de előfordul, hogy a kiugróan sok hivatkozás okát valóban a magának utat törő, az elemi témacsoportok határait áthágó, kiemelkedő minőségben valószínűsíthetjük. Hisz a legtöbb hivatkozást azok kapják, akik olyan ötletekkel állnak elő, melyekkel más csoportokból is magukhoz csábítják a kutatókat, új csoportokat képesek alapítani, melyek gyorsan növekedve akár nagyobb szakterületekké is fejlődnek. Ám ezek ritka esetek, és még ezekben is a kapható hivatkozásszám az átcsábítható kutatók számától, vagyis a szomszédos témacsoportok létszámától függ. Általában, az egyes tudományok vagy azok kisebb-nagyobb

részterületei keretében, a hatás nem korrelál a minőséggel, és amint láttuk, ennek oka az idézettség többszörös mértékű és mikroszintű fluktuációja, melyet a minőséggel össze nem függő körülmények okoznak.

Az idézetelemzés hívei újabban azzal próbálják menteni kedvenc mutatójukat a szakterületi változékonyság buktatójától, hogy olyan adatbázis-technikai módszereket, illetve korrigált, relatív mutatószámokat keresnek, melyekkel a hivatkozásszámokból kiindulva szakterület-független rangsorokat lehet készíteni. Ezek az eljárások azonban többnyire bonyolultak, vitathatók, tökéletlenek. A fentiek szerint ezeknek az elemi témacsoportok szintjéig kellene lehatolniuk ahhoz, hogy valóban a minőséggel kapcsolatba hozható mutatókat eredményezzenek. Ennek kivitelezhetősége és a mindennapi gyakorlatban való alkalmazhatósága azonban reménytelennek látszik.

Úgy tűnik tehát, hogy a hivatkozásszám, illetve idézettség kitüntetett szerepe az egyéni teljesítmények értékelésében indokolatlan, alaptalan, jogtalan. Használata igazságtalanságokhoz, feszültségekhez vezet a tudományos közösségekben azzal, hogy mások kárára kedvez egyes területek képviselőinek. Ha egy habilitációs vagy doktori eljárás sikeréhez egy egész tudományra érvényesen ugyanazt a küszöbhivatkozásszámot kell elémi, ezt – ugyanolyan „minőségű” munkát végezve – egyesek 5-10 év alatt könnyedén teljesíthetik, mások akár egész életükben sem. Az igazságosság nevében joggal szólalhatunk fel az ilyen gyakorlat ellen.

Nem vitás, hogy a tudományos teljesítményeket valahogyan rangsorolni kell. Az IF-ről és az idézettségről azonban kiderül, hogy szakterületi függésük mértéke és szerkezete alkalmatlanná teszi őket erre a célra. Ezért továbbra is meggyőződés, hogy az értékelést főként a közlemények számára kell alapozni. Itt legalábbis nehezen elképzelhető, hogy nagyságrendnyi hibát követünk el,

¹ Cikkeink idézettsége szakterületek szerint is élesen eltér. ... Világos, hogy ezek nem minőségi, hanem a szakterületek dinamikájában megmutatkozó különbségeket tükröznek.

feltéve persze, hogy a közleményszámból saját, egyéni részesedéseket számolunk. Az értékelésekben továbbra is fontos szempont maradhat, hogy a közlemények milyen folyóiratokban jelentek meg, illetve kaptak-e hivatkozásokat, de hogy az IF-ek és hivat-

kozásszámok mekkorák, annak nem szabad nagy jelentőséget tulajdonítani.

Kulcsszavak: *tudományos teljesítmény mérése, hivatkozásszám, idézettség, szakterületi változékonyság*

IRODALOM

- Amin, Mayur – Mabe, Michael A. (2000): Impact Factors: Use and Abuse. *Perspectives in Publishing*. 1, 1–6.
- Bencze Gyula (2006): H-index: egy új javaslat az egyéni tudományos teljesítmény értékelésére. *Magyar Tudomány*. 51, 1, 88–91.
- Bencze Gyula – Berényi D. – Tolnai M. (1996): Az egyéni tudományos teljesítmény értékelésének általános szempontjai. *Magyar Tudomány*. 41, 7, 862–869.
- Braun Tibor (2001): Pendlebury listája. *Magyar Tudomány*. 46, 5, 600–601.
- Braun Tibor (2004): A tudományos teljesítmény mérésének problémáiról. *Magyar Tudomány*. 49, 4, 530–531.
- Csaba György (2004): A biológia művelésének neuralgikus pontjai. *Természet Világa*. 135, 9, 394–397.
- Editorial (2005): Not-so-deep Impact. *Nature*. 435, 1003–1004.
- Marton János (2004): A tudomány és a metria. *Magyar*

- Tudomány*. 49, 6, 788.
- Papp Zoltán (2004a): A tudományos teljesítmény mérésének problémáiról. *Magyar Tudomány*. 49, 2, 232–240.
- Papp Zoltán (2004b): Válasz Braun Tibornak. *Magyar Tudomány*. 49, 4, 532–533.
- Papp Zoltán (2006): Mire és hogyan nem lenne szabad használni az impakt faktort? *Magyar Tudomány*. 51, 1, 99–103.
- Redner, Sidney (2005): Weighing the Value of Physical Review Citation Statistics – Redner Replies. *Physics Today*. 58, 12, 12–12.
- Verő József (1998): Összehasonlíthatók-e egyes tudományterületek? *Magyar Tudomány*. 43, 8, 994–1001.
- Vinkler Péter (2003): A Garfield-tényező. *Magyar Tudomány*. 48, 12, 1604–1610.
- Vinkler Péter (2004): Adalékok a tudománymetria néhány kérdésének megértéséhez. *Magyar Tudomány*. 49, 6, 789–793.

HOZZÁSZÓLÁS

Boda Miklós

elnök, Nemzeti Kutatási és Technológiai Hivatal

Bencze Gyula a *Magyar Tudomány* augusztusi számában megjelent *A tudomány szabadsága kinek a szabadsága?* című írásában hivatkozott cikkem a tudományos igazságok megítéléséről szólt. Arról, hogy az új, paradigmaváltó tételeket, „igazságokat” az adott kor tudósainak többsége számos esetben nem fogadta el. Ennek hazánkban legismertebb példája talán pont Semmelweis Ignác, hogy maradjunk a nőgyógyászatnál.

Bencze úr cikkéből nem derült ki, hogy szerinte mivel foglalkoznak a diákok, PhD-hallgatók, kutatók, oktatók, ha csak a tudomány fokozattal rendelkezőket tekintjük

a tudomány művelőinek. Szerintem nem attól lesz valaki a tudomány művelője, hogy tudományos fokozattal rendelkezik, hanem attól, hogy tudományos munkát végez. A tudomány művelése dinamikus folyamat. Úgy gondolom, nem kellene leragadnunk Zenon nyilánál. Az ország tudása pedig, csakúgy mint egy puzzle, egyéni tudásból, a puzzle darabjaiból tevődik össze a történelem folyamán.

Bencze úrnak, az Ericssonnál folyó kutatások irányításával és a mobiltelefont használókkal kapcsolatos példáját sajnós a legnagyobb jóindulattal sem tudom értelmezni.