



Slavko SPLICHAL, Franc MALI*

IZVIRNI
ZNAŠTVENI
ČLANEK

OBJEKTIVNOST ALI ARBITRARNOST?

O možnostih in težavah standardizacije
kvantitativnih kazalcev v slovenskem družboslovju

Povzetek. Članek obravnava problem kvantificiranja kot socialne tehnologije, ki ima v znanosti za cilj standardizacijo (skupin) ljudi, organizacijo (znanstvenega) diskurza in enotnost znanstvene skupnosti. Značilen primer je raziskovanje znanstvene ustvarjalnosti in uspešnosti, še zlasti bibliometrično ocenjevanje kakovosti, pomembnosti in vpliva objavljenih znanstvenih del, ki po eni strani zadeva raziskovalce same, po drugi strani pa financerje oz. odločevalce. V preštevanju publicističnih enot v znanosti ni mogoče nedvoumno ločevati znanstvene aktivnosti, kakovosti, produktivnosti in napredka. Pri uporabi bibliometričnih metod se pogosto zanemarjajo razlike med posameznimi znanstvenimi področji ali pa sploh ni jasno določeno, kaj z njimi želimo meriti. Primeri sistemov bibliometričnega ocenjevanja, ki se uveljavljajo v Sloveniji kot tipičnem primeru "majhne znanstvene skupnosti" (zaradi jezikovne ekskluzivnosti) pa vendarle kažejo, da je kvantificiranje kljub vsem pomislekom lahko neprimerno manj arbitrarno in pristransko kot subjektivno izvedensko ocenjevanje.

Ključni pojmi: kvantifikacija, objektivnost, kazalci v znanosti, družbene vede, bibliometrija.

893

Mehanična objektivnost kvantifikacij

Kvantificiranje (štetje, merjenje, računanje) je socialna tehnologija, katere jedro je učinkovita strategija objektiviranja narave in družbe. Ideal objektivnosti je hkrati *znanstven* in *političen*. V skrajni konsekvenci pomeni vladavino (statističnih ali drugačnih) zakonov, ne človeka; zahteva podreditve (p)osebni interesov javnim standardom. V znanosti ima kvantifikacija kot socialna tehnologija za cilj standardizacijo ljudi, organizacijo (znanstvenega) diskurza in uveljavljanje vrednote *enotnosti* znanstvene skupnosti, neredko tudi za ceno globljega razumevanja pojavov (Porter 1995, 228). Ta proces je videti zlasti pomemben v "šibkih" znanstvenih skupnostih oz. na področjih, kjer metodološka rigoroznost "nadomešča"

* Dr. Slavko Splanhal, redni profesor na Fakulteti za družbene vede.

Dr. Franc Mali, docent na Fakulteti za družbene vede.

manjkajoči temeljni (npr. teoretski ali strateški) konsenz v skupini ali skupnosti. Idejo o nadzorovanju in "popravljanju" družbe najdemo na samem začetku razvoja socialne statistike v 19. stoletju. Ker je statistično razumevanje verjetnosti postavilo na glavo deterministično razumevanje stvarnosti, je s tem odpiralo pot socialnemu "inženirstvu". Sistematično zbiranje podatkov in odkrivanje statističnih zakonitosti se je začelo na področju deviantnih družbenih pojavov – samomorov, zločinov, norosti, bolezni sploh, prostitucije itd., odkrivanje teh zakonitosti pa je izhajalo iz ideje, da je s štetjem in klasificiranjem mogoče nadzorovati, usmerjati in izboljšati deviantne dele populacije glede na vladajoče standarde in vrednote (Hacking 1995). V sodobnosti se je informacijsko-usmerjevalna vloga statistike v proučevanju družbe "nevtralizirala", saj statistično sklepanje sega na številna in različna področja, od običnih teorij odločanja, operacijskih raziskav in analiz tveganja do tako specifičnega področja kot je analiza znanstvene ustvarjalnosti in uspešnosti. Seveda pa so ideje o nadzorovanju in usmerjanju ostale temelj kvantifikacijskih postopkov tudi na "nevtralnih" področjih, čeprav se včasih oz. v posameznih primerih zdi, da so kvantifikacije same sebi namen.

Ali lahko "mehanična objektivnost", ki jo zagotavlja kvantifikacija, v znanosti nadomesti subjektivno izvedensko presojo nikdar ni bilo (samo) strogo znanstveno, ampak vedno tudi kulturno in *politično* vprašanje. Kot npr. ugotavlja Porter, je primat kvantifikacije nastajal zaradi dveh vrst pritiskov. Uvajanje kvantitativnih kriterijev za odločanje na najrazličnejših področjih je bilo v veliki meri posledica zmagoslavja kvantifikacije v naravoslovnih, medicinskih in dužbenih vedah v 20. stoletju. Po drugi strani je kvantifikacija omogočala bolj demokratično odločanje kot pa individualna izvedenska mnenja, katerih "uporaba" v odločanju je po definiciji omejena na tehnokratske elite, ki so sposobne globljega razumevanja problematike. Vendar je politična določenost uporabe statistike in kvantifikacije hkrati protislovna: če na eni strani demokratizira procese odločanja, ker omogoča sodelovanje tudi manj kvalificiranih posameznikov, po drugi strani s tem zmanjšuje odgovornost izvedencev za sprejete odločitve, kar je klasični problem demokratičnih teorij. Analogija s parlamentarnimi volitvami in – še bolj – z javnomnenskimi glasovanji je povsem na mestu: če gre na eni strani za temelj demokratičnega konstituiranja politične družbe, pa po drugi strani obe instituciji hkrati tudi zmanjšujeta neposredno odgovornost odločevalcev, ki le sledijo natančno, "objektivno" določenemu razmerju politične moči in izmerjenemu javnemu mnenju. *Mehanična objektivnost zmanjšuje subjektivno odgovornost.*

Temeljno načelo kvantifikacije je pridobiti rezultate s postopkom, ki omogoča (zahteva) skoraj mehanično ocenjevanje: presoje ne smejo biti osebne, da bi s tem na minimum zmanjšali tudi možnost, da bi drugi dvomili v rezultate analize. Vztrajanje na mehanični objektivnosti in brezosebnosti ocenjevanja je torej delno tudi reakcija na zunanje pritiske na znanost. Mehanična objektivnost je, kot pravi Porter, zlasti cenjena na področjih, kjer ni stroge meje med tistimi, ki so "znotraj" (insiderji), in tistimi, ki so "zunaj" (outsiderji). Značilen primer so aplikativne, še zlasti ti. administrativne raziskave, ki so vedno izpostavljene kritikam tistih (naročnikov, uporabnikov), katerih interese zadevajo.

Med tovrstne raziskave nedvomno sodi tudi raziskovanje ustvarjalnosti in uspešnosti raziskovanja, ki zadeva raziskovalce same, hkrati pa tudi interese financerjev oz. odločevalcev. Ni naključje, da se je mehanična objektivnost intenzivno razvijala prav na takih mejnih prodočjih, ki so bila sistematično deležna majhnega zaupanja. V času, ko država financira velik del znanosti, zahteva po širši družbeni odgovornosti znanstvenikov spodbuja nastajanje ekstremnih oblik standardizacije in mehanične objektivnosti. Birokracija, ki odloča o financiranju raziskovanja, z uporabo strogih formalnih pravil in statističnih "dejstev" minimizira svojo odgovornost. Pri (samo)ocenjevanju raziskovalnega dela se tudi znanstvene skupnosti pogosto vedejo podobno kot birokracija.

Tudi uporabo in razvojno pot kvantitativnih kazalcev znanosti pogojujejo spremembe v pojmovanjih znanstvene politike oziroma odnosa med državnimi institucijami in znanstveno skupnostjo. Van Der Meulen (1997), Gibbons s soavtorji (1994) in vrsta drugih piscev ugotavljajo, da so šle te spremembe v smeri od prvotne skrbniške države do znanosti, za katero je značilen klientizem, preko odnosa, ki je temeljil na večjem medsebojnem zaupanju in se je izražal v ustanavljanju agencij in drugih posredovalnih institucij, do danes vedno bolj poudarjene vloge znanstvene politike v funkciji državnega nadzora znanosti. Do srede šestdesetih let, ko so na Zahodu državni skladi povečevali delež v znanosti tudi za 5 do 10% letno, državne administracije niso imele večjih pomislekov glede učinkovitosti do takrat prevladujočih kvalitativnih izvedenskih ocenjevanj v znanosti. V tem času je prevladovalo prepričanje, naj nadzor kakovosti znanstvenega dela temelji predvsem na načelu zaupanja: v znanstveno kompetenco, v motivacijo raziskovalcev, da skrbijo za kvaliteto svojih rezultatov, nenazadnje v institucijo akademske znanosti, ki je v okviru svoje dolge tradicije vzpostavila celo vrsto implicitnih in eksplicitnih nadzornih mehanizmov kakovosti raziskovalnih rezultatov. Na področju ocenjevanja znanosti se je uveljavilo prepričanje, da sistem deluje gladko in brez ovir. To je bil čas prepričanja o nezadržni rasti znanosti. Obsežno poročilo o stanju in perspektivah razvoja znanosti v svetu *Science – The Endless Frontier* Vannevarja Busha, najbolj vplivne osebnosti na področju vodenja znanstvene politike takoj po drugi svetovni vojni v ZDA, je že v naslovu izražalo neizmerni optimizem. Nekaj desetletij kasneje je opazen upad nekdanjega optimizma. V sredini devetdesetih let je vplivni ameriški lobist znanosti in predsednik ameriške družbe za napredek raziskovanja Leon Ledermann objavil poročilo s povsem drugačnim naslovom: *Science – The End of the Frontiers*.

Nastajanje novih znanstvenih področij in njihovo financiranje v razvitih državah Zahoda in tudi drugod v petdesetih in šestdesetih letih nista bili vprašljivi. Po prepričanju ključnih akterjev znanstvene politike v tistem času, je bilo izvedensko ocenjevanje v znanosti dovolj učinkovito. Stanje se je spremenilo, ko je financiranje znanosti zašlo v stagnacijo. Denarja za znanost je bilo vedno manj, s tem pa so naraščali dvomi državne administracije glede moči oziroma zadostnosti kvalitativnih izvedenskih ocen. Na te probleme je že sredi šestdesetih let opozarjal Derek de Solla Price (1963), znan širši strokovni javnosti zlasti po svoji logistični krivulji razvoja znanosti. Zanj je značilno, da po obdobju eksponentne rasti preide v linearno rast in potem v fazo saturacije. Ne da bi se na tem mestu soočili z

vsemi argumenti, ki govorijo za veljavnost takega občega razvojnega modela znanosti ali proti njej, se zdi, da so gibanja na področju državnega financiranja znanosti v zadnjih treh desetletjih sledila prav tej razvojni logiki.¹

Čeprav so povečane zahteve po zunanjem nadzoru kakovosti v znanosti izhajale predvsem iz spremenjene politike državnega financiranja znanosti, ni mogoče zanemariti tudi drugih dejavnikov, ki so spodbujali bolj formalizirane pristope k vrednotenju znanstvenih rezultatov. Širša javnost je bila vedno bolj kritična do razvoja znanosti in tehnike, hkrati pa je naraščala tudi kompetitivnost med znanstvenimi disciplinami in znotraj njih. Danes se nasploh uveljavlja prepričanje, da se struktura delovanja znanstvenega sistema temeljito spreminja. Med drugim naj bi naraščal pomen implicitnih komponent znanstvenega vedenja, neposredna posledica naj bi bili novi pristopi k ocenjevanju znanosti, predvsem zahteva po vključevanju širšega kroga izvedencev, ki niso več vezani na ozka disciplinarna področja. Takšna širitev izvedenske "baze" prav gotovo ne škoduje nadaljnjemu razvoju znanosti, kar dokazujejo tudi primeri z drugih področij intelektualnega ustvarjanja (npr. umetnosti), jo je pa seveda treba nadgrajevati z ustreznim razvojem ocenjevalnih metod in tehnik. Le tako se je mogoče izogniti zahtevam, da je treba v znanosti zmanjšati vlogo kvalitativnih izvedenskih ocen, še zlasti zato, ker je v sedanjem obdobju restriktivnega financiranja znanosti vedno težje dobiti nepristranske izvedence. Kvalitativne izvedenske ocene v znanosti se – še zlasti kot mehanizem za izbiro predlogov raziskovalno-razvojnih projektov – uporabljajo mnogo bolj pri "ex ante" kot pri "ex post" vrednotenjih znanosti. V zadnjem primeru gre največkrat za ocenjevanja velikih raziskovalnih projektov v obliki "top-down" pristopov, ki sledijo zelo različnim družbenim ciljem. Za potrebe teh vrst ocenjevanj raziskovalnih projektov so se predvsem na Zahodu oblikovali posebni znanstveni centri, ki so – ne smemo pozabiti, da gre za čas restriktivnega državnega financiranja raziskovanja in razvoja – opravili vrsto bibliometričnih analiz. Posamezni pristopi k vrednotenju so v veliki meri odvisni od institucionalne organizacije in modelov financiranja raziskovanja in razvoja v posameznih deželah.

Kvantitativni kazalci znanosti in njihova uporaba v znanstveni politiki

Kvantitativni kazalci znanosti dobijo svoj pravi pomen šele z jasno teoretsko definicijo pojma ali konstrukta, ki ga želimo meriti. Na to se pri uporabi kazalcev v znanstveni politiki vse prevečkrat pozablja. Ker predstavljajo samo kvantitativno mero izbranih teoretskih pojmov znanosti ali njihovih razsežnosti, ne izražajo znan(stve)osti v vseh njenih razsežnostih in se vedno nahajajo v polju dvojne inter-

¹ Po Priceovih predvidevanjih iz srede šestdesetih let bi se v nasprotnem primeru prej kot v eni človeški generaciji znašli v absurdnem položaju, da bi "že dva znanstvenika prišla na moškega, žensko, otroka oziroma psa, hkrati pa bi trošili za znanost dvakrat toliko denarja" (Price 1963, 30). Zdi se, da Priceova logistična krivulja rasti izraža največjo veljavnost v zvezi z vladnimi izdatki v RR. Vzemimo primer ZDA: rast vladnih sredstev v RR je do leta 1965 vsako leto za več kot 11% preseгла rast vseh drugih izdatkov v okviru državnega proračuna. Če se stopnja rasti izdatkov za RR po letu 1965 ne bi razpolovila, potem bi v ZDA že leta 1980, v celotnem proračunu države, 72% sredstev šlo za RR.

pretacije: v zvezi z njimi se ni potrebno samo vprašati, kako potekajo postopki operacionalizacije različnih definicij znanosti, temveč tudi, iz katerih teoretskih predpostavk te definicije sploh izhajajo. Ker takšna teoretska izhodišča praviloma ne vodijo k enoznačnim zaključkom, je tudi za kvantitativne mere znanosti mogoče reči, da se še vedno nahajajo bolj v polju splošne hevristike kot natančno operacionaliziranih definicij pojava, ki naj bi ga merile.

Za pojmi kot so kvaliteta, dosežek, pomen, vpliv, vidnost, učinkovitost (produktivnost) v znanosti, če naštejemo le nekaj najbolj pogostih lasnosti ali pojavov, ki naj bi jih merili, se praviloma skrivajo različni teoretski pristopi in razumevanja znanosti. Ni jih mogoče zvesti le na številčne vrednosti. Da bi pokazali temeljno razliko med kazalci in podatki, si pomagajmo z naslednjo analogijo: izmerjeni človekov pritisk je podatek brez pomena, dokler nekdo, ki v zadostni meri pozna fiziologijo, ne spozna tega podatka kot pokazatelja človekovega zdravja.

Problem merjenja znanstvenih rezultatov je bolj kompleksen, kot se morda zdi na prvi pogled, kar je mogoče pokazati tudi s primerjavo značaja znanstvenih in ekonomskih kazalcev – ne nazadnje kazalci znanstvene učinkovitosti radi uporabljajo ekonomsko metaforiko. Obstajajo določene vzporednice med kvantitativnimi kazalci znanosti in ekonomskimi kazalci. Kvantitativne kazalce znanosti si lahko predstavljamo kot informacijsko podporo specifičnemu “trgu” znanosti, znotraj katerega vlada visoka stopnja kompetitivnosti. Vendar ne smemo prezreti, da ekonomski kazalci operirajo pretežno z vrednostmi, ki so na trgu že dosegle denarno ovrednotenje. To ne velja za kazalce, ki se nanašajo na rezultate znanstvenega dela. Tu nimamo opravka z jasnimi razmerji med proizvajalci in odjemalci znanstvenih rezultatov. Prej je mogoče govoriti o neke vrste “kvazi” trgu, ki sicer lahko temelji na visoki stopnji notranje kompetitivnosti, vendar igra v fazi “inputa” ključno vlogo dejavnik državne podpore, v fazi “outputa” pa ideja kolektivnega prisvajanja rezultatov znanstvenega dela. Razen tega v primeru kvantitativnih kazalcev znanosti ne razpolagamo s tako standardiziranimi merskimi enotami kot je denarna enota v ekonomiji.

Primerjave znanstvenih in ekonomskih kazalcev nas seveda ne smejo voditi k morebitnemu napačnemu sklepu, da se s pojavom različnih oziroma nasprotujočih interpretacij istih (monetarnih) podatkov ne srečujemo tudi na področju ocenjevanja družbenoekonomskega stanja, na kar največkrat vplivajo politični predsodki. Očiten primer za to je ocena Dunajskega inštituta za ekonomska raziskovanja o nepripravljenosti Slovenije za vstop v Evropsko unijo spomladi 1999. Tudi ekonomski kazalci so – tako kot kazalci znanosti – “samo” teoretski konstrukti. Tudi v tem primeru podatki sami po sebi ne povedo ničesar, ampak šele, ko jih vstavimo v ustrezen teoretski okvir. Res pa je, da so za interpretacije ekonomskih podatkov na voljo bolj konsistentno zgrajene ekonomske teorije. Teorije o znanosti so manj konsistentne. Zato tudi potencialna prepričevalna moč kvantitativnih kazalcev, naj se obračajo na znanstveno ali zunajznanstveno javnost, ni enaka moči ekonomskih kazalcev. Kot ugotavljajo posamezni avtorji, ki poglobljeno proučujejo vlogo in značaj znanstvenih kazalcev, pri njih v primerjavi z ekonomskimi kazalci ne gre za absolutno, ampak kvečjemu za relativno zaostajanje (Weingart in Winterhager 1984; Hornbostel 1997).

Da ekonomski kazalci na današnji stopnji razvoja dajejo bolj izčrpne informacije, ni vzrok samo v višji stopnji socio-kognitivne zrelosti ekonomskih znanosti, temveč tudi v večjem izkustvu pri uporabi teh kazalcev. Čeprav se bodo razlike med obema vrstama kazalcev v prihodnosti nedvomno zmanjševale, si vsaj v bližnji prihodnosti ni mogoče zamisliti, da bi scientometriki razvili neki "supra" teoretski okvir na področju kvantitativnih analiz znanosti.

V zvezi z razlikami in podobnostmi med kvantitativnimi kazalci v znanosti in ekonomiji je vredno opozoriti še na eno okoliščino. Oboji so imeli podoben razvoj, čeprav so nastali v različnih obdobjih: najprej so se razvili kot analitični instrument v akademskih tipih raziskovanj in šele kasneje se je njihova uporaba prenesla na področje ekonomskih in znanstvenih politik.² S tem sta se spremenila njihova funkcija in značaj, saj nadaljnji razvoj vedno bolj določajo (neposredne) potrebe posameznih politik, ki – če uporabimo Van Raanovo formulacijo – "vedno bolj sledijo specifičnim manipulacijam podatkov" (Van Raan 1988, 1). V okviru akademskih proučevanj znanosti o znanosti, je bil razvoj kvantitativnih kazalcev vezan predvsem na vprašanja, kako izboljšati njihovo zanesljivost in veljavnost. Pri njihovi uporabi v politiki so stvari drugačne: tako kot drugi tipi racionalizacij tudi kvantitativni kazalci znanosti nasploh kot način "objektiviranega" znanstvenega vedenja utrjujejo družbeno legitimnost politike. Poleg tega so tudi koristen instrument za praktično politično delovanje. To je eden od razlogov, da so kvantitativni kazalci v znanosti postali predmet različnih interesnih konfliktov in s tem povezanih interpretacij.

V nekritičnem zaupanju v moč kvantitativnih kazalcev znanosti – češ da jih je mogoče uporabljati kot veljaven algoritem v vseh okoliščinah enako – se zanemarija naslednje okoliščine:

Prvič, pri kazalcih znanosti ne gre za enostavno odslikavo kompleksne realnosti, temveč za konstrukcijo specifičnega modela realnosti. Če je realni svet znanosti sestavljen iz sodb, vrednot, označb, rangiranih nagrad, kvalitativnih predstav, potem tudi kvantitativni kazalci znanosti, ki naj bi "odslikavali" to realnost, ne morejo uiti tej svoji začetni subjektivni oz. kontekstualni pogojenosti, pa čeprav se jih morda želi prikazati kot od subjektivnih odločitev povsem neodvisne objektivne odslikave realnega znanstvenega sveta.

Tudi podatki, na katerih temeljijo kvantitativne analize znanosti, so že "obremenjeni" s subjektivnimi sodbami znanstvenikov. Kvantitativna analiza sicer lahko zamegli ali celo prikrije, ne more pa prekiniti zveze med subjektivnim in objektivnim v znanosti. Poleg tega vedno večja zahtevnost statističnih izračunov in analiz sama po sebi nič ne pomeni, če nima ustreznega teoretskega temelja. Tu ne gre toliko za kriterij zanesljivosti kot predvsem veljavnosti kvantitativnih mer znanosti.

² Vzemimo primer ZDA, kjer so se tako družbenoekonomski kot znanstveno in tehnološki kazalci začeli najprej uporabljati. Ko je bil leta 1921 v ZDA prvič izdan Indeks, je prišlo do njegove naraščajoče uporabe ne samo znotraj ekonomske, temveč tudi tarifne in drugih politik. S tem se je spremenil tudi njegov notranji ustroj in življenjski tok: ni bil več samo analitični instrument, namenjen opisovanju kompleksne ekonomske realnosti, temveč je postal uporabno sredstvo v političnem komuniciranju.

Drugič, do končne rešitve kvantitativnih merjenj znanstvenega napredka (kakovosti) ni mogoče priti že zato, ker ne bomo nikoli dosegli absolutne standardizacije kvantitativnih kazalcev znanosti. Res pa je, da je danes pred javnostjo – tega dejstva se znanstvena politika zelo dobro zaveda – mnogo lažje upravičevati odločitve s kvantitativnimi podatki, pa četudi ne povsem zanesljivimi, saj jim zaradi lažje berljivosti javnost tudi prej verjame. Kvantitativni kazalci znanosti niso enake vrste kot strogo kontrolirane spremenljivke v laboratorijskih eksperimentalnih proučevanjih. Največkrat se opirajo na sekundarne podatke, ki so bili ustvarjeni v podatkovnih bazah, katerih primarni cilj ni bil uporaba za namene znanstvene politike.

Tretjič, razvoj in uporaba kvantitativnih kazalcev znanosti nista mogoča samo na temelju ponavljajočih se empiričnih (statističnih) analiz iste vrste, ampak zahtevata poglobljene teoretske premisleke o tem, kaj pojmi znanstvena kakovost, produktivnost ali vpliv v procesu vrednotenja znanosti sploh pomenijo; skratka, upoštevati je treba tako kriterij zanesljivosti kot kriterij veljavnosti. Vsaj od Kuhna naprej se tudi v okviru filozofije (metateorije) znanosti vedno bolj postavlja pod vprašaj kategorija kumulativnega znanstvenega napredka kot od cele vrste kognitivnih in družbenih dejavnikov neodvisnega procesa, ki vodi k “približevanju” znanstveni resnici. Zdi se, da spremenjeni epistemološki pogled na napredek v znanosti, po katerem pojma resnice (napredka) kot takega ni mogoče vzeti za kriterij vrednotenja v znanosti, še ni v celoti razumljen ne samo med uporabniki, temveč tudi in predvsem ne med samimi proizvajalci kvantitativnih kazalcev znanosti. Ali se lahko približamo pojmu znanstvenega napredka s čim večjo funkcionalno povezanostjo posameznih kazalcev znanosti? Odgovor je negativen, kajti vedno smo v položaju, ko merimo hipotetične konstrukte, ki so ne samo povezani, ampak se med seboj tudi prekrivajo. Čeprav so nekateri scientometriki – na primer Martin in Irvin (1983) – poskušali razviti metodologijo, s katero naj bi prišli do neodvisnih ocen o kakovosti, pomembnosti in vplivu objavljenega dela, ki naj bi v seštevku pokazale, koliko je to delo prispevalo k znanstvenemu napredku, so bila takšna prizadevanja že vnaprej obsojena na neuspeh. Znanstveni napredek namreč ni nikoli v celoti izmerljiv.

Znanstvena publicistika kot osnova za kvantitativne in kvalitativne ocene v znanosti

Pri kvantitativnih kazalcih, ki naj bi merili napredek v znanosti, se neprestano srečujemo z vprašanjem, kako izmeriti pomembnost in kvaliteto znanstvenega dela. Najpogostejši primer je merjenje znanstvenih dosežkov na temelju znanstvenih objav. Kljub temu, da ob hitrem razvoju posameznih znanstvenih področij vedno večjo težo dobivajo neformalni načini znanstvenega komuniciranja, se osnovni “profil” znanosti oziroma njenih posameznih disciplinarnih področij še vedno prvenstveno oblikuje na osnovi njihove publicistične dejavno-

sti.³ Z ekonomsko terminologijo bi znanstvene objave lahko opredelili kot rezultat procesov znanstvene produkcije in distribucije. V znanstveni objavi je udejanjen pomemben del raziskovalnih aktivnosti, tako da tudi tu lahko sledimo določenim kvantifikacijam. Seveda pa tu ne gre – če ostanemo pri analogiji z ekonomijo – predvsem za neposredne številčne vrednosti, saj ni ključnega pomena objava sama po sebi, temveč njena vsebina. Iz tega sledi, da publicistična dejavnost ni zadostni, temveč samo nujni pogoj za razvoj (napredek) in s tem povezano kakovost v znanosti. Kaj s preštevanjem publicističnih enot v znanosti sploh merimo – znanstveno aktivnost, znanstveno produktivnost ali znanstveni napredek? Teh razsežnosti ni mogoče nedvoumno ločiti med seboj.⁴

Če z znanstveno aktivnostjo razumemo obseg vloženega dela znanstvenikov, s produktivnostjo obseg znanstvene aktivnosti glede na nekatere druge količine (vložen denar, čas itd.) in z napredkom produktivnost glede na doseganje zastavljenih ciljev, postane očitno, da je mera znanstvene aktivnosti najbolj enostavna in v tej enostavnosti tudi najlažje merljiva razsežnost kompleksnega pojava znanosti, če seveda upošteva razlike med posameznimi znanstvenimi disciplinami. Obstoj razlik so dokazale številne bibliometrične študije in analize. Njihov skupni zaključek je, da na področju družboslovnih in humanističnih znanosti v primerjavi z naravoslovnimi in tehničnimi znanostmi obstaja večja fragmentacija glede vrste literature, v kateri se objavlja, hkrati pa se bistveno večji pomen daje knjižnim objavam. Govorimo lahko o različnih publicističnih habitusih znanstvenikov v različnih disciplinah.

Diana Hicks (1999) je v najnovejših primerjalnih bibliometričnih raziskavah po posameznih deželah (Avstralija, Španija, Nemčija, Nizozemska) ugotovila, da znanstveniki na področju naravoslovno-tehničnih znanosti večji del svojih rezultatov objavljajo v člankih v znanstvenih revijah in v zbornikih znanstvenih kongresov. V nasprotju z njimi znanstveniki s področja družbenih in humanističnih ved veliko objavljajo tudi v knjigah in strokovnih (ne samo znanstvenih) revijah. Tudi za Slovenijo je že bila opravljena podobna bibliometrična študija (Mali 1996). Rezultati bibliometrične analize disciplinarnih vzorcev publicistične produktivnosti raziskovalnih skupin na Slovenskem so pokazali, da tudi tu obstaja razlika v publicističnem habitusu med skupino naravoslovno-tehničnih in skupino družboslovno-humanističnih znanstvenikov. Glede tega se Slovenija ne razlikuje od drugih znanstvenih okolij.

³ Z vidika ocenjevanja rezultatov znanstvenega dela je treba upoštevati, da med neformalnim in formalnim (publicističnim) tipom znanstvenega komuniciranja obstaja specifična kategorija "sivih" publikacij. Gre za različne vrste pol-publikacij, ki sicer niso brezpogojno dostopne, vseeno pa presegajo zgolj z medosebnimi stiki vzpostavljena komunikacijska omrežja. Pri vrednotenju znanstvenega outputa se jim običajno ne pripisuje pomembnejše vloge (glej Whitley 1984; Krohn in Kueppers 1989).

⁴ Podobne pojmovne težave se pojavljajo že pri poskusih razmejevanja znanstvene uspešnosti in učinkovitosti. Po Fischu in Danielu se pojma nanašata na razmerje med "inputom" in "outputom" v znanosti (Fisch in Daniel 1986). Toda, ali ni pri merjenju znanstvenih dosežkov oziroma učinkovitosti mogoče imeti tudi podatkov o vloženih finančnih sredstvih, čeprav – gledano povsem tehnično – predstavljajo samo kazalce znanstvenega "inputa", za kazalce znanstvenega "outputa"? Drugi pol istega metodološkega problema predstavljajo razmere, ko se posamezni strukturni kazalci (npr. znanstvena kocičiranja) uporabljajo kot "output" mera znanosti.

Ena izmed pomembnih razlik med družboslovjem in humanistiko ter naravoslovno-tehničnimi znanostmi je tudi narava "preglednih znanstvenih člankov". Razliko med "izvirnimi" in "preglednimi" znanstvenimi članki je v družboslovju težko smiselno določiti (npr. v pomenu večje ali manjše izvirnosti ali znanstvenosti) zaradi specifične narave družboslovnega raziskovanja. Če z raziskovanjem razumemo empirično zbiranje podatkov z namenom, da bi preskušali, sistematično oblikovali ali spodbujali oblikovanje teorij, ki nam dejansko šele pomagajo razumeti in pojasniti rezultate empiričnega opazovanja, potem moramo reči, da najpomembnejši in prevladujoči del družboslovnega *znanstvenega* dela ni "raziskovanje" (research), pač pa "teoriziranje" (theorizing) oz. sekundarno raziskovanje (secondary research).⁵ Dihotomija empirično/teoretično je v angleščini poudarjena z razlikovanjem med "družboslovnim raziskovanjem" (social research) in družboslovno teorijo (social theory). Še izrazitejši poudarek na teoriji najdemo na primer v Tönniesovi trihotomiji "teorija - aplikacija - empirija", v kateri je teorija samozadostna oz. "samoreferenčna" in tudi aplikacija pomeni dedukcijo iz "čistih teoretskih pojmov", torej v temelju teoretsko raziskovanje, in je le empirija induktivno raziskovanje. V istem času je veliki ameriški pragmatist John Dewey razločno pokazal, da empirična dejstva sama za sebe v družbenih vedah ne povedo nič ali vsaj ne dovolj: "Ugled matematičnih in fizikalnih znanosti je precejšen, in prav ja tako. Toda razlike med dejstvi, ki so, kar so, neodvisno od človeške želje in prizadevanja, in dejstvi, ki so do določene mere to, kar so, zaradi človekovega interesa in smotra in se z njima spreminjajo, ne moremo odpraviti z nobeno metodologijo. Iskreneje ko se sklicujemo na dejstva, pomembnejša je razlika med dejstvi, ki pogojujejo človeško delovanje, in dejstvi, ki so pogojena s človeškim delovanjem. Družboslovna znanost postaja psevdoznanost do mere, do katere zanemarjamo to razliko" (Dewey 1927/1954, 7).

Obstajajo različni postopki in načini (t. j. teoretični in empirični), po katerih prihajamo do našega védenja o družbenih pojavih. V vsaki znanstveni disciplini sočasno obstajajo različne paradigme, kot jih je opredeljeval Kuhn, ki se med seboj razlikujejo glede temeljnih ciljev, predmetov in metod raziskovanja ter specifične uporabnosti izsledkov, in prav te *paradigmatske razlike* so dejansko najpomembnejši in trajni "predmet raziskovanja". Temu je treba dodati, da od najzgodnejših začetkov družboslovnih znanosti v antiki njihov pomemben del pripada *normativnim* in/ali *čistim* teorijam, katerih empirično preskušanje bi bilo sploh nesmiselno, saj so njihov predmet "čisti teoretski pojmi" (Tönnies), "kontrafaktični ideali" (Habermas) ali "znanstvene utopije" (Marx). V danes nedvomno prevladujočem pozitivističnem razumevanju (družboslovne) znanosti, je empirično raziskovanje "zgolj" metodološko rigorozna aplikacija teorije; operacionalistično razumevanje samozadostnosti empiričnega raziskovanja iz začetka tega stoletja je že zdavnaj povsem izgubilo znanstveno veljavo. Najmanj smiselno je končno poskus, da bi

⁵ Gibaldi in Achtert v večkrat ponatisnjenem Handbook for Writers of Research Papers (1977/1988, 1.1) opredeljujeta primarno raziskovanje kot "študij predmeta z neposrednim opazovanjem in preiskovanjem, npr. izvedba ankete ali laboratorijskega eksperimenta", medtem ko sekundarno raziskovanje "zahteva proučevanje študij, ki so jih o predmetu opravili drugi", in poudarjata, da "številni akademski članki ter poročila in predlogi ... temeljijo na sekundarnem raziskovanju."

“stopnjo znanstvene izvirnosti” še ožje določili samo z uporabo kvantitativnih metod v (empiričnem) raziskovanju (po čemer naj bi prednjačila ekonomija), saj so se v mnogih okoliščinah *kvalitativne* metode za zbiranje podatkov v empiričnem raziskovanju izkazale za bolj produktivne.

Izvirnost je takorekoč po definiciji lahko samo paradigmska, saj pomeni “rešitev” problema, ki “ni odvisna od kakega vzora, predloge” (SSKJ), ki se torej razlikuje od znanega po neki bistveni (paradigmski) razsežnosti (npr. nove teorije, nove metode) – v nasprotju z *novostjo*, ki jo (po definiciji) prinaša vsaka informacija (npr. *vsako*, še tako trivialno dejstvo, ki ga odkrijemo v empiričnem raziskovanju). V jeziku publiciranja oz. kazalcev raziskovalnega dela gornja opozorila pomenijo, da so najpomembnejši, torej v najzlahtnejšem pomenu “izvirni znanstveni” članki v temelju in najprej vedno “pregledni”, saj novo vedenje nujno izhaja iz primerjalno ugotovljenih (za izbrano problematiko bistvenih) razlik med že razvitimi paradigmi in teorijami. Drugače rečeno: znanstveno *izvirne* so ideje, ki (lahko) nastopajo kot zasnove empiričnega raziskovanja, interpretacije raziskovalnih izsledkov (“dejstev”) ali so opredmetene v specifičnih raziskovalnih metodah (npr. metode za analizo podatkov), ne pa sama uporaba raziskovalnih metod in sami izsledki, čeprav sicer slednji vedno predstavljajo novo *vedenje*. Razlikovanje med “izvirnimi” in “preglednimi” znanstvenimi članki je torej smiselno le, kolikor so slednji razumljeni izključno v ožjem pomenu *kritičnih razprav* oz. poglavljenih ocen (review essay).

Standardizirana struktura znanstvenih člankov v naravoslovnih, medicinskih, tehniških in biotehniških vedah (po shemi IMRAD - “Introduction, Methods, Results And Discussion” oz. “uvod, materiali in metode, rezultati, razprava in viri”) se je v družboslovju uveljavila samo v ožjem delu poročanja o izledkih empiričnih raziskav, medtem ko v teoretskih člankih, kjer se večinoma izvirne ideje najprej pojavijo, struktura člankov bistveno odstopa od empiričnih in niti ni standardizirana, saj, kot poudarjata Gibaldi in Achtert (1977/1988, 1.1), “nobene konvencije o pripravi rokopisa ne morejo nadomestiti živahnega in inteligentnega pisanja.”⁶

Težave, ki jih povzročajo disciplinarne razlike, in splošne težave z (omejeno) veljavnostjo postanejo še izrazitejše ali že kar brezizhodne, če kazalec “število publikacij” uporabimo za merjenje znanstvenega *napredka*. V tem primeru namreč ne gre več samo za merjenje znanstvene *aktivnosti*, ampak tudi za merjenje znanstvene *kakovosti*. Izenačevanje produktivnosti in kakovosti v znanosti v pomenu, da je več tudi bolje, se je na področju merjenja znanosti sicer močno uveljavilo. Četudi v znanosti pogosto obstaja tesna zveza med količino in kakovostjo, ima

⁶ *Ob enaka protislovja kot pri neutemeljenem razlikovanju med izvirnimi in preglednimi članki, naletijo tudi pri poskusih razlikovanja med monografijo (“znanstvenim delom”) in knjigo nasploh (ne ali manj znanstvenim delom?); mnoga najpomembnejša družboslovna dela novega veka po tej delitvi ne bi bila “znanstvene monografije”, pač pa “druge knjige”. Razlikovanje med “monografijo” in “knjigo” je sploh nedosledno; tipologija dokumentov Comarc govori o “monografijah in drugih zaključenih delih” (poleg monografij še učbeniki, ostalo učno gradivo, priročniki, slovarji, leksikoni, umetniška dela itd.), znanstveno monografijo pa opredeljuje kot “specializirano knjigo”. Knjiga je torej monografija s specifičnimi tehničnimi (število strani, vezava, statistična evidenca) in ne vsebinskimi značilnostmi. Vsekakor pa je kot “znanstvene monografije” skladno z definicijo Comarc treba upoštevati samo tiste, ki so izšle kot knjige in torej izpolnjuje dogovorjene mednarodne bibliotekarske standarde.*

lahko v postopkih ocenjevanja znanstvenikov upoštevanje zgolj števila njihovih objav prej negativne kot pozitivne učinke. Lahko vpliva na strategijo znanstvenega publiciranja, ki ni vedno v korist dejanskemu napredku v znanosti. Vzemimo primer neskrupuloznega uveljavljanja načela "publish or perish". Nekateri sociologi znanosti opozarjajo, da je uveljavljanje tega načela v praksi neredko vodilo znanstvenike k prirejanju rezultatov raziskovanja (S. Hornbostel 1997; P. Weingart in M. Winterhager 1984). Z namenom, da bi "iztržili" čim večje število objav, so znanstveniki minimalno prirejali rokopis za petkratne ali celo desetkratne objave. Če v sicer od vsega začetka ambivalentni družbeni funkciji objavljanja – kdo bi lahko danes trdil, da je edini interes znanstvenikov spoznavanje resnice? – motiv pridobivanja znanstvenega ugleda in s tem povezanih nagrad prevladuje nad pridobivanjem in širjenjem novih spoznanj, se stopnja kognitivne redundance v sistemu znanosti samo še povečuje. Temu pojavu smo bili priča že v času, ko zunanji pritiski za čim večjo znanstveno produktivnost še niso bili tako močni.

Po študiji O.K. Maya (1968) naj bi od šestih znanstvenih člankov na področju matematike samo eden dejansko posredoval novo, vsi ostali pa naj bi predstavljali samo variacije in ponavljanja že znanega. Ekstremne oblike možnih deformacij, ki izhajajo iz načela "publish or perish", naj bi predstavljali odkriti primeri znanstvenih goljufij in prevar. Abelson, od leta 1962 do 1985 odgovorni urednik najbolj ugledne naravoslovne revije *Science*, je na temelju dveh znanih primerov znanstvenih goljufij (Darseejev primer na področju kardiologije in Slutskyev primer na področju radiologije) opozoril, da so se zadeve razkrile ravno zaradi pretirane znanstvene produktivnosti Darseeja in Slutskyja. V zadnjem primeru je bil novi članek objavljen v povprečju na vsakih deset dni, četudi naj bi bilo na tem področju v tistem času povprečje 2-3 objavljenih člankov na leto. Zato je Abelson predlagal, naj bi v vrednotenjih znanstvenih del pri znanstvenih napredovanjih upoštevali samo omejeno število znanstvenih objav.⁷

Ko govorimo o razmerju med količino in kakovostjo v znanosti, se je treba zavedati, da gredo objave praviloma skozi recenzijsko sito znanstvenih revij, kar je ena izmed oblik izvedenskega ocenjevanja; tako objave kot citati v znanosti so torej povezani s predhodno kvalitativno izvedensko oceno. To pomeni, da se z izvedenskim ocenjevanjem srečujemo na vsakem koraku znanstvenega delovanja in se njihov vpliv čuti v vseh fazah raziskovalnega dela. Treba je sicer priznati, da ni veliko empiričnih raziskav pristranosti delovanja izvedenskega ocenjevalnega sistema. Opravljene raziskave pa so opozorile na dve vrste deformacij: na socio-kognitivne pristranosti in pristranosti, ki izhajajo iz slučajnih napak.⁸ V krog prvih možnih pristranosti bi lahko navedli težave delovanja izvedenskega sistema v majhnih znanstvenih skupnostih, kamor je treba šteti tudi slovensko.

⁷ V zvezi s številom objav kot mere znanstvene produktivnosti, se neizogibno postavi vprašanje primerljivosti posameznih publikacij glede na njihovo znanstveno pomembnost. Še najbolj posrečeno je to dilemo izrazil De Solla Price (1963, 40) z naslednjim retoričnim vprašanjem: "Who dares to balance one paper of Einstein on relativity against even a hundred papers by John Doe, Ph.D., on the elastic constant of the various timbers (one to a paper) of the forests of Lower Basutoland?"

⁸ O tem zlasti veliko pišejo ameriški avtorji; glej npr. Abrams, 1991; Cole in Cole, 1981; Mazuzan, 1992.

Dobra stran izvedenskih ocen je v tem, da izvedenci pri presoji uspešnosti in relevance v večji meri upoštevajo kontekst. Slaba stran pa je v tem, da bolj prihaja do izraza subjektivna narava presoje, kar prinaša nevarnost, da zaradi nezadostnih informacij ali preveč poudarjenega izvedenčevega jaza (ega) pride do napak v oceni. Kvantitativna merjenja znanosti (vzemimo primer štetja objav) sicer zadovoljijo kriterij ponovljivosti in regularnosti, ne morejo pa specificirati veljavnosti. Zanesljivost merjenja ne more zagotoviti veljavnosti merjene znanstvene kakovosti. Gre za problem, ki ga je v analogiji s Heisenbergovim načelom nedoločljivosti ugotavljal že Francis Narin, ko je zapisal: "Evalvacijske študije, ki so najbolj relevantne za merjenje dejanske stopnje prispevka k znanosti, izkazujejo največjo stopnjo nezanesljivosti glede na objektivnost, tiste, ki so najbolj objektivne, pa največjo stopnjo nezanesljivosti glede na relevantno" (Narin 1978, 36).

Uporabo kvantitativnih kazalcev v ocenjevanju znanstvene aktivnosti raziskovalcev je načelno vedno smiselno kombinirati z izvedenskim (kvalitativnim) ocenjevanjem njihovih najpomembnejših del. Kvantifikacija ocenjevanja znanstvene uspešnosti in ustvarjalnosti ne bi smela imeti za edini cilj doseganje "mehanične objektivnosti". Kvantitativni kazalci ne morejo v celoti nadomestiti individualnega (subjektivnega) ocenjevanja kakovosti *raziskovalnih projektov*, predvsem pa ne ocenjevanja uspešnosti *raziskovalcev*, ker veljavneje merijo usposobljenost raziskovalne skupine za izvajanje projekta oz. verjetnost, da bo raziskovalna skupina realizirala predlagani projekt, kot pa ("splošno") kakovost raziskovalca ali raziskovalne skupine. Njihova uporabnost je omejena na sprejemanje odločitev, s katerimi želimo nadzorovati in usmerjati npr. razvoj raziskovalnih področij (*ex ante*), mnogo manj smiselna pa je taka kvantifikacija v ocenjevanju kakovosti (odličnosti) že opravljenega raziskovalnega dela (*ex post*). Ocenjevanje znanstvenega napredka postane zanesljivo šele v okviru dolgoročnih *ex-post* analiz.

Kljub pomanjkljivostim in slabostim izvedenskih ocen v znanosti, se njihova moč kaže v tem, da so zgodovinsko prispevale k ohranjanju standardov raziskovanja. Razlog za to je v tem, da gre v podsistemu znanosti v vseh oblikah komuniciranja bolj kot kjerkoli drugje v družbi za specifično izmenjavo "znanstvenih dobrin". V tej izmenjavi se vsak posamezni znanstvenik lahko (hipotetično) znajde v vsaki izmed naslednjih treh vlog: kot avtor dela, ki se ga ocenjuje, kot ocenjevalec dela drugih znanstvenikov in kot uporabnik rezultatov predhodnega in sedanjega znanstvenega dela. Ker je torej glede na takšno zasedbo vlog vsakdo, ki deluje v sistemu znanosti, bistveno odvisen od kakovosti standardov, je družbeni sistem znanosti uspel ohraniti takšna pravila igre, ki so (dolgoročno) vzdrževala ustrezno raven njene profesionalizacije.⁹

⁹ V sociološki teoriji gre za primere "zapornikove dileme". V okviru te "cirkularnosti" gre za naslednjo (hipotetično) situacijo: jaz recenziram X, X recenzira Y, Y recenzira Z, ta zadnji pa recenzira mene. Če bi v tem zaprtem krogu eden izmed znanstvenikov, ki se nahaja v vlogi recenzenta - zato da bi si čim hitreje pridobil ugled - dal profesionalnemu kolegu slabšo oceno kot jo zasluži, vsi ostali pa bi bili v ocenjevanju korektni, potem bi ta recenzent vsaj na kratek rok lahko imel korist. Glede na to, da takšni strategiji lahko sledijo vsi znanstveniki, saj se pojavljajo v izmenični vlogi, namreč enkrat kot tisti, ki ocenjujejo, drugič pa kot tisti, ki se jih ocenjuje, dolgoročno od tega, za znanost ne bi imela nobene koristi. Sistem znanosti bi zaradi lastnega neupoštevanja profesionalnih pravil uničil samega sebe. Zato znanstveniki sami skrbijo, da ne pride do erozije profesionalnih vrednot.

Danes bi bilo torej nesmisleno vztrajati pri institucionalni delitvi uporabe kvantitativnih kazalcev znanosti in kvalitativnih izvedenskih ocen; kot da se nujno potrebna avtonomija znanosti lahko ohrani samo z uporabo kvalitativnih ocen in da vsak poseg od zunaj – v obliki uporabe bolj formaliziranih in standardiziranih metod ocenjevanja (kvantitativnih kazalcev) – ogroža svobodo znanstvenikov. Res pa je, da se je treba v praktični uporabi kvantitativnih kazalcev čim bolj izogibati njihovi napačni uporabi ali celo zlorabi. Le tako se bo mogoče izogniti še vedno močno navzočemu nezaupanju znanstvenikov do kvantitativnih kazalcev znanosti. Po drugi strani je še naprej ključnega pomena etični vidik uporabe in interpretacije kvantitativnih indikatorjev znanosti – ne samo zato, ker njihova uporaba pogosto vzbuja lažna pričakovanja, da z njihovo pomočjo lahko presežemo vse, kar je subjektivnega v znanosti (pa to ni mogoče), temveč tudi zato, ker se v zadnjem času vedno bolj uporabljajo za ugotavljanje uspešnosti posameznih znanstvenikov. Ravno v zvezi s tem pa se pojavlja največ metodoloških problemov.

Uporaba kvantitativnih kazalcev v slovenskem družboslovju

Podobno kot za druge znanosti, so kazalci raziskovalne uspešnosti v družboslovju zajeti v (1) Merilih za volitve v nazive visokošolskih učiteljev, znanstvenih delavcev in sodelavcev Univerze v Ljubljani s posebnimi interpretacijami članic (Ekonomska fakulteta, Fakulteta za družbene vede, Pedagoška fakulteta, Filozofska fakulteta, Pravna fakulteta, Fakulteta za šport); (2) Merilih za volitve v nazive visokošolskih učiteljev, znanstvenih delavcev in sodelavcev Univerze v Mariboru ter (3) Merilih za ocenjevanje raziskovalnih projektov znanstvenega sveta MZT za področje družboslovnih ved. Veljavnost kazalcev raziskovalne uspešnosti v slovenskem družboslovju bi bilo smiselno presojati v treh razsežnostih: (1) v mednarodni primerjavi (še zlasti glede na tolikokrat poudarjano nacionalno-kulturno specifičnost - slovenskega - družboslovja); (2) v medpodročni primerjavi (družboslovje - humanistika - medicina – tehnika...) ter (3) v znotrajpodročni primerjavi (razlike med posamičnimi disciplinami in znanstvenimi polji).

Primerjava kazalcev raziskovalne uspešnosti v družboslovju, ki so uveljavljeni v Sloveniji, z drugimi državami oz. z mednarodnimi standardi doslej ni bila opravljena. Vse dosedanje spremembe (ali zahteve po spremembah) kazalcev v družboslovju so izhajale predvsem iz (ne)primerljivosti z naravoslovjem, tehniko in medicino ter temeljile na "prilagajanju" sistema kazalcev iz v tem pogledu razvitejših disciplin v Sloveniji. To sicer ne pomeni, da je rezultat nujno slab, je pa odsotnost mednarodne primerljivosti nedvomno sama po sebi šibkost. "Prilaganje" pomeni predvsem dvoje: (1) težnjo po izenačevanju objav v "reprezentativnih" slovenskih revijah z objavami v mednarodnih indeksiranih revijah; (2) poudarjanje pomembnosti knjižnih objav v primerjavi s članki, pri čemer pogosto *v izhodišču ni razlike* med objavo monografije v Sloveniji in pri mednarodno uveljavljenih založbah. Skupni imenovalac in utemeljitev teh prizadevanj je poudarjanje nacionalne specifičnosti (dela) družbenih ved in njihovega pomena za kulturno in politično (državno) samobitnost slovenskega naroda. Ne da bi

odrekli znanosti tovrstno vrednotno določenost (ki nedvomno obstaja), pa se zdi, da poudarjanje nacionalnega pomena *na račun* mednarodne primerljivosti znanstvene uspešnosti disciplinarnemu razvoju prej škoduje kot koristi predvsem zato, ker v Sloveniji v mnogih (družboslovnih) disciplinah zaradi premajhnega števila raziskovalcev takorekoč fizično ni mogoče vzpostaviti sistema kolegialnega ocenjevanja. Po drugi strani so pretirana prizadevanja za uveljavljanje "nacionalne specifičnosti" posameznih disciplin, in s tem njihove izolacije v odnosu do mednarodne znanstvene skupnosti, tudi v škodo mednarodni uveljavitvi Slovenije kot kulturno specifične entitete – torej prav tistemu, s čimer naj bi bila utemeljena mednarodna neprimerljivost.

Doslej tudi ni bila opravljena izčrpna primerjalna analiza, ki bi ne le predstavila, ampak tudi pojasnjevala razlike med posameznimi družboslovnimi disciplinami. Področje družboslovja je v primerjavi z vsemi drugimi znanstvenimi področji pri nas verjetno najbolj heterogeno in se kaže v (1) različni stopnji disciplinarne izoblikovanosti (pri čemer je družboslovje nasploh med vsemi področji najmlajše) ter s tem (ne)uveljavljenosti trdnih kriterijev ocenjevanja uspešnosti znanstvenega dela ter v (2) razdvojenosti med prizadevanji za uveljavljanje mednarodnih kriterijev vrednotenja raziskovalnega dela na eni strani ter uveljavljanjem narodne oz. kulturne specifičnosti (in posledično mednarodne neprimerljivosti) na drugi strani. Ta razdvojenost ni značilna le za Slovenijo, ampak tudi za marsikatero drugo državo (npr. Nizozemska, Belgija, Italija).

Pogosto se razlike med disciplinami še vedno povezujejo z (ne)uporabo kvantitativnih metod zbiranja podatkov v empiričnem raziskovanju, kar naj bi bil ključni kriterij "znanstvenosti". Ta kriterij, ki očitno ni uporab(lj)en v humanistiki, je nadvse problematičen tudi v družboslovju. Pogosto se namreč prav v družboslovju kaže paradoks, da (pretirana?) težnja po visoki *zanesljivosti* raziskovanja (ki jo omogoča – ali vsaj omogoča preverjati – prav uporaba kvantitativnih metod zbiranja in analize podatkov) vodi v manjšo *veljavnost* raziskovanja, čeprav veljavnost načelno predpostavlja zanesljivost. Vsaj v delu družboslovja, ki poudarja pomen teoretske veljavnosti empiričnega raziskovanja (v nasprotju z empirizmom, ki ne pozna razlike med teoretskimi in operacionalno oz. empirično definiranimi pojmi), je primat kvantitativnega nad kvalitativnim empiričnim raziskovanjem že zdavnaj odpravljen. To se konec koncev izraža tudi v spoznanju, da kvantitativni kazalci v ocenjevanju znanstvenega dela ne morejo nadomestiti kvalitativnih ocen oz. recenzij, ampak sta si obe vrsti orodij komplementarni.

Znotrajpodročna primerjava kazalcev v družboslovju je predstavljena s tabelo, iz katere je razvidno, da Ekonomska fakulteta, Fakulteta za družbene vede in Pedagoška fakulteta¹⁰ v skupino najpomembnejših mednarodnih revij uvrščajo samo tiste, ki jih indeksira philadelphijski Institute for Scientific Information (SSCI, SCI, AHCI), medtem ko Filozofska fakulteta, Fakulteta za šport in Pravna fakulteta kot povsem enakovredne obravnavajo tudi "osrednje" slovenske revije. Univerza v

¹⁰ Konec leta 1998 je senat Pedagoške fakultete sicer sprejel predlog sprememb habilitacijskih pogojev, ki je v osnovi povzemal pogoje, določene na Filozofski fakulteti - toda na predlog svoje habilitacijske komisije, je senat Univerze v Ljubljani predlog zavrnil.

Mariboru je sprejela enotno klasifikacijo revij, ki ne upošteva specifičnosti posameznih področij (oziroma je posamezne fakultete niso operacionalizirale).

Še pomembnejše, kot so razlike v razvrščanju revij, so razlike med disciplinami oz. fakultetami Univerze v Ljubljani pri operacionalizaciji posebnega pogoja iz Meril za volitve v nazive visokošolskih učiteljev, znanstvenih delavcev in sodelavcev Univerze v Ljubljani glede "načina, ki mu stroka priznava mednarodno pomembnost ali pomembnost za narodno in državno samobitnost in kulturo" (12. člen). Za izvolitev v posamezni naziv mora imeti kandidat/ka določen delež objav, ki izpolnjujejo pogoj "pomembnosti" po mednarodnih *ali* nacionalnih kriterijih. Operacionalizacije tega pogoja so po disciplinah oz. fakultetah zelo različne (v tabeli so posebni pogoji, določeni na posameznih fakultetah, predstavljeni s krepkim tiskom). Samo dve fakulteti (Fakulteta za družbene vede in Pedagoška fakulteta) omejujeta vrhunsko pomembnost na revije v SSCI in drugih mednarodnih bibliografskih bazah; Ekonomska fakulteta in Fakulteta za šport za vrhunske štejeta tudi vse revije v lastni knjižnici oz. na lastnem ad hoc seznamu, medtem ko na Filozofski fakulteti in Pravni fakulteti med "najpomembnejše objave" sodijo članki v revijah po lastnem ad hoc seznamu (vključno s slovenskimi revijami), prispevki v zbornikih konferenc, poglavja v knjigah in knjige... Visoka šola za socialno delo povzema specifična določila Fakultete za družbene vede, Visoka upravna šola pa določila Pravne fakultete. Iz operacionalizacij je očitno, da izpolnjevanje istega posebnega pogoja univerzitetnih Meril – tj. objav, ki jim "stroka priznava mednarodno pomembnost ali pomembnost za narodno in državno samobitnost in kulturo" – na različnih fakultetah pomeni zelo različne vrste objav: od FF in PF, kjer je ta pogoj izpolnjen že z objavami v slovenskih revijah in knjigah, do PeF in FDV, kjer je raziskovalna izvrstnost določena z objavami v uveljavljenih mednarodnih revijah, ki so uvrščene v indeks citiranja pri ISI ali v najpomembnejše sisteme abstrahiranja oz. bibliografske podatkovne baze za posamezne discipline, ki jih gojita fakulteti. Nobena družboslovna fakulteta kazalcev *znanstvene izvrstnosti* ne omejuje na indekse citiranosti (torej na objavo člankov v revijah SSCI) in nobena tudi ne vključuje citiranosti raziskovalcev v SSCI med raziskovalne kazalce! Razlike v merilih ocenjevanja med fakultetami ljubljanske univerze so dobro ponazorilo za splošno tezo o podloženosti kvantifikacij s subjektivnimi (izvedenskimi) ocenami; v tem primeru pa žal tudi dokazujejo, da izvedenske ocene niso nujno nepristranske.

Čeprav je bilo doslej največ razprav in polemik v zvezi z raziskovalnimi kazalci namenjenih razlikam v znanstveni pomembnosti slovenskih in tujih *revij*, pa so za vzpostavitev enotnega sistema raziskovalnih kazalcev za področje družboslovja pomembnejše razlike v vrednotenju znanstvenih revij in monografij. V Merilih... Univerze v Ljubljani je razmerje med najvišje vrednoteno periodično in monografsko objavo 8:20. Pri tem je predvsem vsaj nenavadno, da je določanje vrednosti objavljenega članka (od 1 do 8 točk) vsaj relativno natančno določeno, medtem ko je za skoraj trikrat "vrednejše" monografske objave vrednotenje prepuščeno povsem subjektivni volji ocenjevalca ali celo ocenjevanca oz. samoocenjevalca. Nikakor pri tem ne gre prezreti, da pri ocenjevanju znanstvene uspešnosti tudi v razvitejših družboslovnih okoljih knjižne objave pri mednarodnih založbah veljajo

več kot objave člankov v še tako uglednih revijah, toda res je tudi, da je redkokje mogoče tako hitro in lahko izdati knjigo kot v Sloveniji (s finančnimi subvencijami in brez uredniških recenzij). Pomembnost knjižne objave je mogoče kvantitativno meriti samo ex post, npr. s citiranostjo knjige, toda težava pri velikem številu slovenskih knjižnih del bi najbrž bila, da bi bil indeks blizu nič. Za veljavno ocenjevanje torej ostanejo na voljo samo še izvedenci.

Predlog "točkovanja bibliografije", ki ga je kot začasni instrument pri ocenjevanju prijav raziskovalnih projektov sprejel znanstveni svet za družbene vede pri MZT za leto 1998, določa razliko med vrednostjo člankov in monografij celo v razmerju 5:30 ("znanstveni članek v mednarodno uveljavljeni reviji" in "znanstvena knjiga pri mednarodno priznani založbi") oz. celo v razmerju 2:20 ("druga revija" in "druga založba"). Razmerje 2:20 v prid monografije lahko celo presega golo kvantitativno razmerje med količino besedila za članek in monografijo! Poleg tega je poglavje v npr. slovenski knjigi enakovredno članku v najvišje vrednoteni mednarodni reviji! Teh nekaj primerov jasno kaže, da v družboslovju sistema raziskovalnih kazalcev ni mogoče postaviti brez določitve specifikke monografskih in periodičnih objav in – predvsem – postavitve posebnih pravil za vrednotenje monografskih objav. Kljub očitnim nedoslednostim, je pri razvrščanju revij vsaj v nastavkih operacionaliziran kriterij razvrščanja – mednarodna priznanost revije, medtem ko pri monografijah, ki so bistveno više vrednotene kot članki, ni nobenega razvidnega kriterija za ocenjevanje kakovosti monografije.

Vrh vsega ni natančno določljiva niti meja med znanstveno in strokovno (zlasti) knjižno objavo. Če bi iz *znanstvenih* kazalcev povsem izključili *strokovno* knjigo, se bi najbrž dogajalo, da se bi neupravičeno povečevalo število "znanstvenih" knjig. Razloga za vključevanje strokovnih knjig med znanstvene kazalce in s tem za spodbujanje objav strokovnih knjig sta za Slovenijo predvsem dva: (1) slovenske strokovne knjige kot oblika prenosa znanja iz tujine v Slovenijo in posebej v prakso imajo pomembno vlogo v širjenju kulturne osnove znanosti, in (2) mednarodne založbe za pisanje strokovnih knjig izbirajo mednarodno uveljavljene strokovnjake. Oboje, torej pisanje strokovnih knjig za slovensko in mednarodno javnost, je treba spodbujati, ker je tako v interesu razvoja znanosti nasploh kot v interesu promocije slovenske znanosti v svetu in doma. Ne nazadnje (najbrž s podobnimi razlogi) tudi Ministrstvo za znanost subvencionira izdajanje ne le znanstvenih, ampak tudi strokovnih knjig in revij.

Kvantitativne ocene pomembnosti knjižnih in periodičnih objav in razmerja med njimi, ki jih predlaga v svoji študiji skupina pri SAZU (Adamič in sod. 1999), so ta hip najbolj obetaven poskus, da bi kvantifikacijo znanstvenih kazalcev v Sloveniji napravili vsaj dovolj zanesljivo in primerljivo med področji. Skupina strogo omejuje izbiro na kazalce *raziskovalne* uspešnosti. Pri člankih izenačuje "izvirne" in "pregledne" in jih razvršča v pet skupin: (1) znanstveni članki v revijah, ki jih indeksirajo zbirke SCI, SSCI in A&HC (20 točk), (2) članki v revijah zunaj SSCI, ki jih vključuje Current Contents (10 točk); (3) članki v revijah, ki so vključene v druge specializirane bibliografske baze (6 točk); (4) znanstveni članki v osrednjih oz. pomembnih slovenskih revijah, ki niso v ključene v skupine 1-3 (3 točke) ter (5) znanstveni članki v drugih revijah (1 točka). Pri knjižnih objavah pa je tudi tej

skupini pošla imaginacija: (vsaka) "znanstvena knjiga, izdana pri mednarodni znanstveni založbi" je ocenjena s 40 točkami in pri "nemednarodnih znanstvenih založbah" z 12 točkami, kar je dvakratnik vrednosti ustrezne (t. j. mednarodne ali domače) periodične objave (poglavje v knjigi pa je ocenjeno s 25-odstotnim deležem, tj. 10 oz. 3 točkami, pri čemer vsota ocen za vse prispevke oz. poglavja posameznega avtorja v knjigi ne more presegati polovice točk, s katerimi je ocenjena knjiga). Previdnost tu gotovo ni odveč, na kar opozarja zaletavost Ministrstva za znanost, ki je z dekreti in brez kriterijev preocenjevalo knjižne objave iz "znanstevnih" v "strokovne" v bazi COBIB in potem od avtorjev zahtevalo, naj dokažejo, da ministrstvo nima prav, ter na koncu za razsodnike postavilo predsednike področnih svetov – spet brez vsakršnih kriterijev.

Nekaj naštetih primerov opozarja, da je "mehanična objektivnost" kljub vsem pomislekom lahko neprimerno manj arbitrarna od "mehanične *subjektivnosti*" pri ocenjevanju znan(stveno)sti. Ugotovitve so skladne z rezultati že omenjenih empiričnih raziskav (zlasti v ZDA), ki opozarjajo na večjo verjetnost pristranskega izvedenskega ocenjevanja v majhnih znanstvenih skupnosti. Zaradi jezikovne ekskluzivnosti lahko to za Slovenijo postane še toliko bolj pereč problem, zato je nujno čimprej uveljaviti enoten sistem kvantitativnih kazalcev znanosti, ki bo spodbujal razvoj kakovosti in internacionalizacijo znanstvene ustvarjalnosti v Sloveniji.

RAZRED/ TOČKE	EF	FDV	FF	PeF	FS	PF
1 do 8	SCI, SSCI	SSCI	revije po seznamu publikacij FF; SSCI in AHCI; "revije, ki jim druge hum. in družb. vede priznavajo kriterij mednarodne pomembnosti in odmevnosti	(S)SCI, AHCI	mednarodne revije po seznamu: 36 revij, od tega ena slovenska, 10 iz SSCI	A: 5 slovenskih revij in 81 tujih revij; B: "spominski zborniki, izdani v tujem jeziku in z mednarodno udeležbo"; C: "zborniki kongresov, simpozijev ipd., izdani v tujem jeziku in z mednarodno udeležbo"
2 do 6	International Bibliography of Economics	Sociological Abstracts, Political Science Abstracts, Communication Abstracts		17 mednarodnih bibliografskih baz	mednarodno priznane nacionalne revije, uvrščene v mednarodne bibliografske zbirke	
3 do 4	Spremljane v J. of Economic Literature ter Bibliographie der Wirtschaftswissenschaften	ostale domače in tuje revije z recenziranjem, mednarodno izmenjavo in povzetki v tujem jeziku	"tuje in domače revije, če niso zajete drugje"			članki, ki niso v 1. skupini
4 do 2	CEK, nacionalne revije po seznamu	ostale revije		nacionalne revije; revije z uredniško recenzijo		"referati v publikacijah I./C"

LITERATURA

- Abrams, P. 1991. The Predictive Ability of Peer Review of Grant Proposals: The Case of Ecology and the NSF. *Social Studies of Science*, 21, 211-224.
- Adamič, Š., J. Dular, V. Kaučič, D. Miklavčič, F. Nekrep, S. Splichal. 1999. *Vrednotenje raziskovalne uspešnosti v Sloveniji*. Ljubljana: SAZU.
- Cole, S. in Cole, J. 1981. Chance and Consensus in Peer Review. *Science*, 214, 881-886.
- Dewey, John. 1927/1954. *The Public and Its Problems*. Athens: Swallow Press.
- Gibaldi, J., W. S. Achtert. 1977/1988. *MLA Handbook for Writers of Research Papers*. New York: The Modern Language Association of America.
- Fisch, R. in H. Daniel. 1986. Zur Einfuehrung: Messung von Forschungsleistungen - wie, wozu und mit welchen Implikationen? V: *Messung und Foerderung von Forschungsleistung*, ur. R. Fisch in H. Daniel. Konstanz: Universitaetsverlag Konstanz GmbH, 11-21.
- Gibbons, M. et al. 1994. *The New Production of Knowledge*. London: Sage.
- Hacking, I. 1995. *The Taming of Chance*. Oxford: Oxford University Press.
- Hicks, D. 1999. The Difficulty of Achieving Full Coverage of International Social Science Literature and The Bibliometric Consequences. *Scientometrics*, 44, 193-215.
- Hornbostel, S. 1997. *Wissenschaftsindikatoren*. Opalden: Westdeutscher Verlag GmbH.
- Krohn, W. IN g. Kueppers. 1989. *Die Selbstorganisation der Wissenschaft*. Frankfurt/M.: Suhrkamp Verlag.
- Mali, F. 1996. Bibliometrična analiza raziskovalnih rezultatov znanstvenikov v Sloveniji, Teorija in praksa, 33, 955-970.
- Martin, B.R. 1996. The Use of Multiple Indicators in the Assessment of Basic Research. *Scientometrics*, 36, 343-362.
- Martin, B.R. in J. Irvine. 1983. Assessing Basic Research. Some Partial Indicators of Scientific Progress in Radio Astronomy. *Research Policy*, 12, 61-90.
- May, O.K. 1968. Growth and Quality of Mathematical Literature. V: *Institute for Scientific Information Service*, 59, 363-71.
- Mazuzan, G.T. 1992. Good Science Gets Funded - The Historical Evolution of Grant Making at the National Science Foundation. *Knowledge: Creation, Diffusion, Utilization*, 14, 63-90.
- Narin, F. 1978. Objectivity versus Relevance in Studies of Scientific Advance. *Scientometrics*, 1, 35-41.
- Porter, T. M. 1995. *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton, NY: Princeton University Press.
- Price de Solla, D. 1963. *Little Science, Big Science*. New York: Columbia University Press.
- Siverston, G. 1997. Ethical and Political Aspects of Using and Interpreting Quantitative Indicators. V: *Evaluating Science and Scientists*, ur. M. S.Frankel in J. Cave. Budapest: Central European University Press, 212-220.
- Van Der Meulen, B.R.J. 1997. The Use of S&T Indicators in Science Policy: Dutch Experiences and Theoretical Perspectives From Policy Analysis. *Scientometrics*, 38, 87-101.
- Van Raan, A.F.J. 1988. *Handbook of Quantitative Studies of Science and Technology*. Amsterdam: North-Holland.
- Weingart, P. in M. Winterhager. 1984. *Die Vermessung der Forschung - Theorie und Praxis der Wissenschaftsindikatoren*. Frankfurt/M.: Campus Verlag.
- Whitley, R.D. 1983. From the Sociology of Scientific Communities to the Study of Scientists' Negotiations and Beyond. *Social Science Information*, 22, 681-720.