

Tiedeindikaattorit ja tutkimuksen arviointi

TERTTU LUUKKONEN-GRONOW

Korkeakoulujen arviointi, erityisesti tiedeindikaattorien avulla, on tällä hetkellä Suomessa varsin ajankohtainen kysymys. Käsitellen seuraavassa tavallisimpia tuotosindikaattoreita, joita on sovellettu tai esitetty sovellettaviksi perustutkimuksen ja erityisesti korkeakoulujen tutkimustoiminnan arviointiin, ja pohdin niiden käyttökelpoisuutta tähän tarkoitukseen.

Käsite »tiedeindikaattorit» on yleistynyt viime vuosina tiedepoliittiseen käyttöön. Tiedeindikaattorit suppeasti käsitettyinä tarkoittavat perustutkimusta arvioivia kvantitatiivisia mittareita, jotka on tulkittu käsitteellisesti tutkimusta koskevien teoreettisten käsitysten, olettamusten tai mallien puitteissa. Laajasti käsitettynä ne tarkoittavat koko tutkimus- ja kehittämistyötä ja teknologiaa koskevia indikaattoreita. Suppea merkitys vastaa tarkoitustamme, kun käsittelemme korkeakoulujen arviointia. Tiedeindikaattorikäsitteen juuret ovat 1930-luvulla syntyneessä sosiaaliindikaattorien laatimisperinteessä. Tiedeindikaattorikäsite vakiintui Yhdysvaltain National Science Foundationin alettua käyttää sitä joka toinen vuosi ilmestyvissä raporteissaan »Science Indicators» maan tutkimuksen tilasta ja »terveydestä» (vuodesta 1972) ja OECD:n ryhdyttyä käyttämään käsitettä 70- ja 80-lukujen vaihteessa.

Puhuttaessa tiedeindikaattoreista tutkimuksen

arviointiin yhteydessä tarkoitetaan tavallisesti tutkimuksen tuloksellisuutta ja tuotosta mittaavia indikaattoreita. Vertailun vuoksi todettakoon, että esim. tutkijoiden keski-ikä ja vaihtuvuutta kuvaavat luvut voivat olla indikaattoreita tutkijakunnan uusiutumisesta ja tutkimuksen potentiaalisesta uudistumisesta, mutta nämä eivät ole tuotosindikaattoreita, vaikkakin ne mittaavat erästä ulottuvuutta tutkimusjärjestelmän »terveydestä». Tutkimuksen arviointikeskustelu on painottunut voimakkaasti tutkimusjärjestelmän edustaman »mustan laatikon» tuottamiin lopputuloksiin. Niiniluoto on äskettäin (1987) esittänyt käsityksen, jonka mukaan tällainen ajattelu näkee tieteellisen tutkimuksen kilpaurheiluun verrattavana toimintana. Tämä ei ole kovin liioiteltu kuva. Tiedeindikaattoreita on sovellettu kansallisten tutkimusjärjestelmien tuotosten keskinäiseen vertailuun kansainvälisten kilpaurheilujen tapaan nationalistisessa hengessä.

Tavallisimmat tuotosindikaattorit

Perustutkimuksen tuottavuutta koskevat indikaattorit voidaan jakaa kolmeen ryhmään:

- mittarit tai tunnusluvut, jotka pohjautuvat tieteellisen yhteisön antamiin tunnustuspalkintoihin
- tutkimusaktiivisuutta kuvaavat tiedot
- bibliometriset indikaattorit.

Tunnustuspalkinnot

Hagstrom tulkitsee klassisessa tieteen sosiologian teoksessaan *The Scientific Community* (1965: 16) palkitsemisen olevan osa tieteen sosiaalista kontrollia, jonka avulla taataan tutkijoiden normien mukainen käyttäytyminen. Keskeisenä tähän käyttäytymiseen kuuluu tutkimustulosten saattaminen laajemman tutkimusyhteisön käyttöön. Palkitseminen tapahtuu tunnustuksen (recognition) muodossa. Merton (1973) on puolestaan todennut tunnustuksen olevan tapa palkita originaalin tutkimustuloksen esittänyttä tutkijaa »hyvästä tutkijanroolin suorituksesta» (s. 293). Eri-laisia tunnustuksen antamisen muotoja Hagstrom luokittelee olevan pääasiassa kahdenlaisia: institutionalisoitunut tunnustuksenanto ja elementäärinen (elementary) eli tutkijoiden keskinäisessä vuorovaikutuksessa tapahtuva tunnustuksenanto (mt., 23—33). Nyt käsittelemämme tunnustuksen muodot kuuluvat lähinnä Hagstromin ensimmäiseen kategoriaan eli institutionalisoituun tunnustukseen ja siinä kollektiivisiin kunnianosoituksiin. Kollektiivisuus tarkoittaa tiedeyhteisön kollektiivisten muodostumien eli tieteellisten seurojen tai vastaavien antamia palkintoja ja kunnianosoituksia. Näitä ovat tieteellisten seurojen ja tutkimusinstituutioiden antamat mitalit ja palkinnot. Usein niihin liittyy seremoniallinen antaminen, jonka yhteydessä palkinnon saaja pitää kunnialuennon. Myös yliopistot järjestävät kunnialuentoja (honorific lectures). Tieteelliset seurat tai säätiöt järjestävät myös symposiumeja, joihin kutsutaan ansioituneita tutkijoita esitelmöimään ja keskustelemaan keskenään. Tällaisten symposiumien esitelmät julkaistaan tavallisesti kirjoina. Tieteelliset aikakauslehdet julkaisevat johtavien tutkijoiden laattimia katsausartikkeleita. Tunnettujen tutkijoiden kunniaksi laaditut ja julkaistut kunniakirjat, *Festschriften*, ovat eräs muoto kollektiivisista kunniapalkinnoista. Oleellista Hagstromin mukaan on, että näihin liittyy seremoniallinen tiedonjakaminen (luentojen, julkaisujen muodossa) ja kunniamaininta tai -palkinto. Tunnustuksen osoituksina voidaan pitää myös erilaisia luotamustehtäviä tai kunniajäsenyyksiä, kuten arvostettujen tieteellisten seurojen tai akatemioiden jäsenyys, arvostettujen tieteellisten julkaisujen päätoimittajuus, tai vähäisempiä palkintoja ja

kunniatehtäviä. Kaikkiin näihin ei liity seremoniallisuutta, vaan esim. tieteellisen julkaisun päätoimittajuus antaa mahdollisuuden vallankäyttöön, gatekeeper-asemaan tieteellisessä yhteisössä, mikä sisältää mahdollisuuden normien valvontaan ja tulkintaan. Tällaisen aseman saaminen on tunnustus tutkijan pätevydestä ja arvostuksesta kollegojen parissa ja siksi luettavissa myös tunnustuspalkintoihin.

Luonteenomaista tieteellisten yhteisöjen antamille tunnustuspalkinnoille ja arvostuksenosoituksille on, että ne muutamia poikkeuksia lukuunottamatta ovat tieteenalakohtaisia ja -siddonaisia. Tämä tekee tunnustuksen ja palkintojen luokittelun vaikeaksi ja yhteismitattomaksi tieteenalojen välillä. Näkyvin ja arvostetuin tieteellinen palkinto, Nobel-palkinto, lienee ainoa, jonka asema suhteessa muihin palkintoihin on selvä (ks. Zuckerman, 1967).

Ei ole ongelmatonta pisteyttää tunnustus-palkintoja edes yhden alan sisällä. Jonkinlainen pisteyttäminen on kuitenkin välttämätöntä, muutoin arvostettu kansainvälinen tieteellinen palkinto saa saman painon kuin esim. kotimaisen seuran antama »Vuoden paras gradu -palkinto». Rajanveto olisi myös tehtävä tieteellisiksi katsottavien tunnustus-palkintojen ja kunniatehtävien ja muiden asiantuntijatehtävien ja toimintojen välillä. Raja tässä helposti hämärtyy. Ongelmallista tunnustus-palkintoja ja muita vastaavia tunnustuksen osoituksia koskevien tietojen indikaattorikäytössä on myös se, ettei meillä ole tarkempaa tietoa tällaisten palkintojen yleisyydestä, jakautumisesta tutkijoiden kesken (muutoin kuin että jakautumat ovat todennäköisesti hyvin vinoja) tms. mikä auttaisi johtopäätösten tekoa tietojen pohjalta. Nyt voimme tehdä vertailuja vain kulloinkin kyseessä olevan joukon sisällä.

Tieteellisten yhteisöjen antamat tunnustus-palkinnot ovat parhaiten sovellettavissa tieteenalan sisäiseen, laadulliseen käyttöön, kuten tähän asti on tapahtunut, sillä palkintojen arvottaminen vaatii tieteenalan sisäistä asiantuntemusta. On myös ajateltavissa, että tieteenalan ulkopuolinen voi asettaa kysymyksiä, esim. kun yliopiston laitokselta puuttuvat kokonaan kaikki tunnustus-palkinnot ja arvostuksen osoitukset. Tämä voi johtaa jatkotutkimuksiin asiantilan syistä. Kuitenkin tällainen käyttö on laadullista ja hyvin karkealla tasolla luokittelevaa eikä itse asiassa

vastaa indikaattorikäsitteen sisältämää kvantitatiivisen mittarin ajatusta.

Tutkimusaktiivisuutta koskevat indikaattorit

Tutkimusaktiivisuutta koskevat indikaattorit ovat poikkeus muista tässä esiteltävistä indikaattoreista siinä suhteessa, että ne mittaavat nykyisten toimintojen intensiivisyyttä ja itse asiassa kuvaavat »mustan laatikon» sisällä tapahtuvaa toimintaa. Tällaisia indikaattoreita on sovellettu erityisesti yliopistojen laitosten tutkimustoiminnan mittaamiseen. Ne perustuvat tutkimuksen organisaatiota koskevan tutkimuksen tuloksiin ja mittaavat mm. tutkimuksen kommunikaatiota, tutkimuksen resursseja tai tutkimuksen sosialisatiota koskevia ilmiöitä. Näitä ilmiöitä kuvaavien seikkojen indikaattorikäyttöä ei ole kuitenkaan pohdittu tieteen tutkimuksen kirjallisuudessa ja indikaattorikäyttö vaikuttaa kovin ad hoc-pohjaiselta. Esimerkkejä aktiivisuusindikaattoreista ovat:

- yliopistojen laitosten saama ulkopuolinen tutkimusrahoitus
- laitosten tutkijoiden kanssakäyminen ulkomaisten tutkijoiden kanssa, jota mittaavat laitoksen tutkijoiden vierailut ulkomaille ja vastaavasti ulkomaisten tutkijoiden vierailut laitokseen ja
- tutkimuksen jälkikasvun ja potentiaalisen uudistumisen turvaaminen: tätä mittaavat esim. jatko-opiskelijoiden määrät, jatko-opiskelijoiden saamien tutkimusvarojen määrä ja jatkotutkintojen määrä.

Monet tutkimusaktiivisuutta koskevista indikaattoreista mittaavat samanaikaisesti tutkimusprosessin eri vaiheita edellytyksistä tuotoksiin. Esimerkki tällaisesta indikaattorista on korkeakoulujen ulkopuolisen tutkimusrahoituksen suuruus. Tutkimusrahoitus on edellytys tutkimuksen tekemiselle ja siten panostekijä. Ulkopuolisen rahoituksen saaminen tieteellisiltä toimikunnilta, säätiöiltä tai yrityksiltä osoittaa samalla korkeakoulun laitoksen tutkimusaktiivisuutta ja aikaisempia tuotoksia. Vastaava voidaan todeta tutkijoiden välisestä kanssakäymisestä kansainvälisellä tasolla. Tutkijoiden välinen kanssakäyminen on merkittävä tutkimuksen aineeton resurssi, mutta jotta runsasta kanssakäymistä voisi syntyä ja kanssakäymisessä olisi jo-

tain vaihdettavaa, edellytyksenä ovat aikaisemmin suoritettu tutkimus ja siinä saavutetut tulokset.

Tutkimusaktiivisuutta koskevat indikaattorit ovat myös monimerkityksisiä ja -ulotteisia. Runsa ulkopuolinen tutkimusrahoitus yrityksiltä ja viranomaisilta ilmaisee tutkimuksen suuntautuneisuutta lyhyen tähtäimen ongelmanasetteluihin; vastaavasti runsas ulkopuolinen rahoitus säätiöltä tai tieteellisiltä toimikunnilta merkitsee perustutkimusorientoituneisuutta. Siksi pelkkä ulkopuolisen rahoituksen määrä ei yksin kerro kovin paljon. Ulkopuolisen rahoituksen tarve, suuruus ja suuntautuneisuus vaihtelevat tieteenalasta toiseen ja indikaattorin vertailu tieteenalojen kesken johtaa sen vuoksi harhaisiin johtopäätöksiin. Tämä koskee myös muita yksittäisiä aktiivisuusindikaattoreita. Yksittäisten aktiivisuusindikaattorien mittaamat asiat ovat keskenään varsin erilaisia eikä niiden keskinäistä painoarvoa voida helposti yhdistää samaksi mittariksi. Tosin esim. Ruotsissa Karolinska Institutissa on näin tehty.

Tutkimusaktiivisuutta koskevien indikaattorien käytölle ei ole olemassa selvää tulkintaohjetta. Mitä enemmän sen parempi, ei välttämättä päde kaikkien näiden toimintaa kuvaavien indikaattorien osalta, toisin kuin esim. ensin mainittujen tiedeyhteisön palkitsemisten suhteen. Esimerkiksi ei ole varmasti toivottavaa, että kaikilla aloilla tuotetaan jatkuvasti kasvavaa tohtoreiden määrää, joille ei sitten ole kuitenkaan riittävästi tutkijanpaikkoja. Siksi indikaattorien tulokinnan ongelma nousee varsin keskeiseksi.

Bibliometriset indikaattorit

Käsite bibliometriset indikaattorit viittaa joukkoon kvantitatiivisia mittoja, jotka pohjautuvat tieteellisen kirjallisuuden ominaisuuksien laskeamiseen. Pritchardin klassisen määritelmän (1969) mukaan kyseessä on matematiikan ja tilastollisten menetelmien soveltaminen kirjoihin ja muihin tieteellisen tutkimuksen kirjallisiin kommunikaatiovälineisiin. Bibliometriikka kehittyi alunperin tieteellisten kirjastojen hankintojen avuksi selvittämään kirjastokokoomien riittävyttä ja auttamaan kirjallisuuden etsimisessä. Bibliometriikka tarjoaa monipuolisen tutkimusvälineen myös tieteellisen kommunikaation tutkijoille. Puhun tässä yhteydessä bibliometrinen menetelmien

soveltuvuudesta tutkimuksen arvioinnin tarpeisiin ja näiden menetelmien käyttökelpoisuuden punnitseminen tästä näkökulmasta ei tuota välttämättä samoja päätelmiä niiden soveltuvuudesta muihin tarpeisiin. Tämä erottelu on syytä pitää mielessä.

Bibliometriset menetelmät ja indikaattorit ovat herättäneet voimakkaita tunteita sekä puolesta että vastaan, mutta erityisesti vastaan, kun on puhuttu niiden soveltamisesta tutkimuksen arviointiin. Olettaisin, että tärkeä syy niiden vastustamiselle on pelko yksioikoisten tulkintojen teosta niiden perusteella ja eri tieteen- ja tutkimusalojen joutumisesta tällöin eriarvoiseen asemaan. Indikaattorien käyttäminen edustaa merkittävää valtaa, sillä ne oletettavasti ohjaavat käyttäytymistä tulevaisuudessa. Käsittelen seuraavassa joi-takin bibliometrinen, mutta erityisesti viiteanalyysiin pohjautuvien indikaattorien tulkinnan ongelmia, joilla on merkitystä niiden sovellutuksille arviointikäyttöön.

Julkaisujen määrä

Tieteellisten julkaisujen kvantiteettia, määrää, pidetään varsin yleisesti tutkijoiden tuottavuuden (productivity) indikaattorina. Tällöin tuottavuudelle ei anneta laadullista painotusta. Tästä syystä erimielisyyttä ei vallinne siitä, että julkaisujen määrä yksin on riittämätön indikaattori tutkimuksen tuotoksista, ja julkaisujen määrää käytetään tavallisesti yhdessä muiden indikaattorien ja tutkimuksen laadun arviointien yhtedessä. Julkaisujen määrän on katsottu kuitenkin mittaa-van jossain määrin myös laatua, mikäli tällöin rajoitutaan julkaisuihin referee-järjestelmää käyttävissä julkaisukanavissa. On myös ehdotettu julkaisujen painottamista julkaisukanavan näkyvyyden tai arvostuksen mukaan, jota ns. impact factorit mittaavat (perustuvat tieteellisessä aikakauslehdessä julkaistujen artikkelien saamaan keskimääräiseen viitemäärään lyhyellä tähtäimellä) (Garfield, 1973, Martin & Irvinen mukaan, ks. Martin & Irvine, 1983; ja Suomessa Helsingin Yliopiston lääketieteellisen tiedekunnan henkilökohtaisten ylimääräisten professorien työryhmä 1986). Tällaiset ehdotukset eivät kuitenkaan ota huomioon sitä, että yksittäisten aikakauslehtien sisällä on erittäin suuri variaatio artikkelien saaman viitemäärän suhteen, ja aika-

kauslehtien impact-luvulla on suhteellisen pieni ennustusarvo artikkelien viitemäärästä (korrelaatio impact-lukujen ja artikkelien viitemäärien välillä omassa tutkimuksessani, joka koski pohjoismaista syöpätutkimusta oli 0.37, Luukkonen-Gronow, 1987 b).

Ongelmallisinta julkaisujen määrän käyttämisessä indikaattorina on erityyppisten julkaisujen keskinäinen painottaminen ja arvostaminen. Eri-laisia yrityksiä painottamiseen on tehty, mutta täysin subjektiivisin perustein ilman empiirisiin tutkimuksiin pohjautuvia arviointeja (ks. Bayer & Folger, 1966). Esimerkiksi tieteellisessä aikakauslehdessä ilmestyneet artikkelit on jossain tutkimuksissa katsottu monografian arvoiseksi, jos niiden lukumäärä on 18, toisissa taas jos niiden lukumäärä on 4, 3 tai 6. Joissain tutkimuksissa tämä luku on vielä jaettu tekijöiden määrällä (ks. ma.). On muistettava myös, että julkaisurakenteet ja vastaavasti erityyppisten julkaisujen keskinäinen verrattavuus vaihtelevat tieteenaloittain eikä sen vuoksi voitane kuvitella, että voitaisiin luoda yhtä, yhteneväistä julkaisujen määrän indikaattoria Suomen korkeakoululaitoksen tarpeisiin.

Viiteindikaattorit

Viiteanalyysi pohjautuu tieteellisissä julkaisuisa tehtyihin viittauksiin aikaisempaan tieteelliseen kirjallisuuteen. Viiteanalyysi perustuu olettamuksiin, joiden mukaan nämä viittaukset ilmaisevat jonkinlaista vaikutussuhdetta, esimerkiksi viittaava artikkeli on saanut vaikutteita viitatulta artikkelilta. Se olettaa myös kaikkien viitteiden olevan samanarvoisia. Hagstromin mukaan (1965) toisten tutkijoiden työn arvon tunnustaminen viittausten muodossa on ensimmäinen esimerkki tieteen institutionaalisen tunnustuksen muodoista. Kysymys on yksittäisten tutkijoiden antamasta tunnustuksesta vastakohtana aikaisemmin käsiteltyihin kollektiivisiin tunnustuksen-osoituksiin. Cole & Cole (1973) toteavat viitteiden mittaavan useita seikkoja. Näitä ovat tutkimuksen *laatu*, toisten tutkijoiden työn *tunnustaminen*, tutkimuksen *leviäminen* (diffusion) ja tutkimuksen *hyväksikäyttö*.

Viitteiden käyttöä tutkimustuotosten laadullisena indikaattorina on perusteltu lukuisilla tutkimuksilla, joiden mukaan viiteluvut korreloivat

muiden suorituksia kuvaavien mittojen tai arviointien kanssa, esim. kunniapalkintojen ja vertaisarvioiden kanssa (Cole & Cole, 1973). Esimerkki viitelukujen ja muiden arviointien tulosten vastaavuudesta on myös se että Nobel-palkintojen saajat kuuluvat harvalukaiseen eniten viitattujen tutkijoiden joukkoon, joka käsittää 0.1 % kaikista tutkijoista.

Viitteiden käyttöä tutkimuksen laadun tai aikaisemman työn vaikutuksen mittarina on kritisoitu vetoamalla tutkimuksiin, joissa sisältöä ja tekstikontekstia analysoimalla on todettu viitosten sisällön ja tarkoituksen vaihtelevan. Esimerkiksi osa viitteistä tarjoaa lisäinformaatiota käsiteltävälle asialle, osoittaa lukeneisuutta tai yrittää vakuuttaa lukijan kirjoittajan väitteiden paikkansapitävyydestä. Viitteiden käytön syistä ja motiiveista tarvittaisiin tällaisten tekijöiden tärkeyden selvittämiseksi kuitenkin lisätutkimuksia, jotka pohjautuvat sekä viittaavien tutkijoiden että artikkeleiden yleisön eli niitä lukevien tutkijoiden haastatteluihin (vrt. Brooks, 1986). Tällaisilla lisätutkimuksilla voitaisiin selvittää eri tieteenalojen eroja viittauskäyttäytymisessä. Nyt varsin useat viitteitä koskevat tutkimukset koskevat valikoituja tieteenaloja, erityisesti fysiikkaa.

Viitteiden käyttöä vastaan vaikutuksen tai tuotoksen laadullisten tekijöiden indikaattorina on esitetty myös muunlaista kritiikkiä. Se sisältää esimerkkejä tapauksista, joissa väitteiden määrää ei voida tulkita osoitukseksi tutkimuksen myönteisestä vaikutuksesta muuhun tutkimukseen. Usein mainitaan esimerkkeinä *kielteiset viitteet* ja *itseviittaukset*. Kolmas esimerkki sisältää ajatuksen *runsaasti viittaamisesta joillain kriteereillä rajatun sisäpiirin* julkaisuihin.

Esimerkki *kielteisistä viitteistä* sisältää ajatuksen, että korkea viitemäärä ei ilmaise tutkimuksen vaikutusta tai laatua, jos se perustuu kielteisiin viitteisiin. Tämän esimerkin merkittävyys kokonaisviitteiden kannalta on kuitenkin kyseenalainen. Chubinin ja Moitran (1975) tutkimuksessa, jossa tekijät analysoivat joukkoa (N = 43) suurenergiafysiikan alaan kuuluvia artikkeleita ja näiden esittämien väitteiden sisältöä, ei yksikään viite ollut viitattuun teokseen nähden täysin kielteinen, osittain kielteisten väitteiden määräkin oli vain 5 % kaikista viitteistä. Moravcsik ja Murugesan (1975) saivat puolestaan samaa tie-

teenalaa koskevassa 30 artikkelin aineistossa kielteisten väitteiden osuudeksi 14 %. MacRoberts ja MacRoberts (1984) ovat todenneet, että tieteellisessä kirjallisuudessa esiintyy vain vähän kielteisiä viitteitä, koska kritiikki tiedeyhteisön sosiaalisista syistä puetaan valeasuun ja mm. positiivisilla kommentilla varustettuna pyritään hälventämään kritiikin purevuutta. Brooks (1986) on omassa useita tieteenaloja kattaneessa, mutta suppeassa aineistossaan havainnut vahvistusta em. MacRoberts & MacRobertsin väitteille. Tämä keskustelu ei kuitenkaan kykene ratkaisuun kysymystä kielteisten väitteiden merkittävydestä viitelukujen tulkinna. Periaatteellisesti voidaan myös todeta, että jos tutkija tekee erheen, joka aiheuttaa runsaasti kielteisistä kommentointia, se voi selvittää asioita ja ohjata jatkotutkimuksia ja tällä tavalla edistää tieteellistä tietoa.

Itseviitteiden aiheuttama ongelma on varmasti kielteisiä viitteitä merkittävämpi. Riippuen viiteanalyysin soveltamistavasta, itseviittaus aiheuttama väitteiden määrän kasvu on otettava analyysissä huomioon, esim. poistamalla itseviittaukset. Tämä ei tosin ole kovin helppoa eikä halpaa. Jo itseviittauksen määrittely voi olla vaikeaa aloilla, joilla yhden artikkelin kirjoittajakunta koostuu jopa kymmenestä kirjoittajasta ja alan tutkijat tekevät keskenään runsaasti yhteistyötä, mutta eri töissä erilaisilla kombinaatioilla. Itseviittauksen avulla ei kuitenkaan viittauksen määrää saatane kovin runsaasti kasvatettua. Itseviittauksen määrän on joissain tutkimuksissa havaittu olevan noin kymmenen prosenttia kaikista viittauksista. Kun viittauksen jakautumat ovat erittäin vinoja ja yksittäisten tutkimusten tai tutkijoiden väliset viitemäärien erot moninkertaisia, kymmenen prosentin molemmilla puolilla pyörivät itseviittauksen määrät eivät vertailuja heittäne kovinkaan paljon.

Runsas viittaaminen sisäpiirin julkaisuihin voi tarkoittaa eri asioita eri yhteyksissä. Cole & Cole (1973) ovat määritelleet sisäpiirin eminenttien tutkijoiden joukoksi ja vertailleet sitä, missä määrin runsaasti viitatuut tutkijat ovat saaneet viitteitä toisilta runsaasti viitatuilta tutkijoilta. Heidän tulostensa mukaan väitteiden alkuperässä ei ollut merkittäviä eroja. On myös tuotu esille, että kokeneet tutkijat suosisivat viitteissä omia oppilaitaan tehdäkseen näitä tunnetuksi muiden tutki-

joiden parissa (Ks. Zuckerman, 1987). Tällaisen mahdollisen yliviittaamisen vaikutuksia ei ole empiirisesti tutkittu. Toisaalta nuorten tutkijoiden osalta toimii myös vastatendenssejä, koska senioritutkijat monilla aloilla esiintyvät oppilaidensa ja nuorempien kollegojensa töissä ensimmäisenä kirjoittajana, mikä vähentää nuorten saamaa tunnustusta työstänsä. Joissain suomalaisissa puhenvuoroissa on esitetty myös, että viitteiden määrä voitaisiin nostaa keinotekoisesti samalla alalla toimivien tutkijoiden muodostamien ystäväverkostojen avulla. Tätäkään seikkaa ei ole missään tutkittu. Sen sijaan viiteverkostoja tutkimalla on identifioitu tutkimusalueita. Käsitykseni on, ettei viitemäärien »keinotekoista» nostamista ystävien avulla ole kovin helppoa tehdä, ainakaan siinä määrin, että sillä olisi reaalista vaikutusta viitteiden kokonaismääriin. Kilpailu samalla alalla toimivien tutkijaryhmien välillä on siksi kova, että »ystäväpalveluna» tehdyt viitaukset lienevät suhteellisen harvinaisia.

Viime vuosina on alan tutkijoiden parissa keskusteltu viitteiden käsitteellisestä merkityksestä ja useissa tutkimuksissa tultu tuloksiin, joiden mukaan viitteet mittaavat tutkimuksen *tosiasialista vaikutusta* (engl. impact) muuhun tutkimukseen lyhyellä aikaviiveellä (Martin & Irvine, 1983) tai erityisesti tutkimuksen *eturintamaan* (Moed et al., 1985). Martinin ja Irvinen (1983) mukaan viitteillä mitataan tutkimuksen tosiasialista vaikutusta *tieteen edistymiseen* (scientific progress). Viitteet eivät mittaa tutkimuksen laatua sinänsä, mikä tarkoittaa esimerkiksi sitä kuinka hyvin tutkimus on tehty, onko se virheetön, tai matemaatiikassa kuinka »esteettinen» päättely on yms. Kyseessä ei ole myöskään tutkimuksen *tärkeiden* mittaaminen, sillä tarkoittaa tutkimuksen potentiaalista merkitystä muulle tutkimukselle, edellyttäen että tieteessä vallitsee esteetön kommunikaatio. Tämä erottelu torjuu osan viiteanalyysiä kohtaan esitetystä kritiikistä, esimerkiksi viitteiden kyvyttömyyden rekisteriöidä merkittävää työtä ennen aikaisten löydösten tapauksessa tai tutkijan aseman vaikutuksen viitemääriin (Matteus-efekti, johon palataan myöhemmin). Osittain on kuitenkin kysymys ehkä käsitteellisestä leikistä, sillä tieteiden edistäminen on myös eräs tutkimuksen laadun tavallisesti käytetyistä kriteereistä. Jotta tutkimus käsitetään laadullisesti hyväksi, ei riitä, että se on virheet-

tömästi ja metodisesti oikein suoritettu, vaan tutkittujen ongelmien merkittävyys vaikuttaa laadun arviointiin. Lienee tarpeen myös korostaa, mielenkiinto tutkimuksen arvioinnissa kohdistuu tavallisesti tutkimuksen laatuun. Siksi liiallinen käsitteellinen erittely viiteanalyysin merkityksistä vaikuttaa ehkä haittaavasti sen käyttöön ja tulkintaan. Tutkimuksen laatu vaikuttaa työn haitsemiseen ja vastaanottoon, mutta tämä haitseminen ei tapahdu sosiaalisessa tyhjiössä, vaan tiede on Mertonin (1968) sanoja lainaten »sosiaalisesti väliintynyt ja todennettu tietojoukko» ja »ainoastaan työ, jonka muut tutkijat tosiasiallisesti havaitsevat ja käyttävät hyväkseen, merkitsee jotain». Siten viitteet voitaisiin ymmärtää *tiedeyhteisön sosiaalisesti tulkitsemaksi laaduksi*.

Esimerkkinä viitteiden riippuvuudesta tiedeyhteisön sosiaalisista tekijöistä on viiteanalyysin herkkyys tutkimuksen kommunikaatorakennetta koskevien tekijöiden vaihtelun suhteen. Olen omassa tutkimuksessani, joka koskee lääketieteen osa-alueita, syöpätutkimusta, havainnut, että näkyvissä, kansainvälisissä lehdissä julkaistut artikkelit ovat muita artikkeleita runsaammin viitattuja (Luukkonen-Gronow, 1987 b). Todennäköinen selitys havainnolle on se, että näkyvillä lehdillä on muita lehtiä ankarampi karsinta. Ne ovat arvostettuja ja siksi niille on runsas artikkelitarjonta. Ilmeistä on kuitenkin myös, että näkyvät lehdet antavat niissä julkaistuille artikkeleille paremman näkyvyyden ja siten lisäävät todennäköisyyttä, että nämä artikkelit huomataan. Lehden maine antaa kaikille siinä julkaistuille artikkeleille tietyn laatutakuun. Järvinen ja Pietiäinen (1988) ovat tutkineet suomalaisessa ornitologian alan lehdessä *Ornis Fennica* julkaistujen artikkelien samaa vastaanottoa viitteiden muodossa ja he ovat havainneet, että ulkomaisien viitteiden kertymisessä on usean vuoden viive (jopa yli 10 vuotta) verrattuna kotimaasta tuleviin viitteisiin, mikä osoittaisi lehden heikkoa näkyvyyttä ja saatavuutta ulkomaisille tutkijoille ja tämän tekijän vaikutusta lehdessä julkaistujen artikkelien vastaanottoon.

Merton (1968) on kuvannut Matteus-efektillä ilmiötä, jossa tieteellinen kontribuutio havaitaan helpommin, jos sen tekee arvostettu tutkija verrattuna tuntemattomaan tutkijaan. Tällaisten arvostettujen tutkijoiden töihin viitataan Mertonin

olettamuksen mukaisesti runsaammin kuin heidän panoksensa tieteen edistymiselle edellyttäisi. Mm. Cole & Cole (1973) ovat testanneet Mertonin hypoteesia ja todenneet, että Matteus-efektillä on jokin verran vaikutusta artikkelien leviämiseen eli sille kuinka pian niihin viitattiin; tämä efekti toimii myös siten, että tutkijan instituution arvonannolla on tähän vaikutusta; kuitenkin »parhaiden» (viitemäärillä mitattuna) artikkeleiden ollessa kyseessä Matteus-efektillä on vähemmän vaikutusta ja vähemmän merkittävät artikkelit hyötyivät siitä eniten (Cole & Cole, 1973: 194—295). Oma olettamukseni on, että tällainen Matteus-efekti ei toimi vain tukijakohtaisesti, vaan myös aikakausjulkaisujen osalta. Eksponentiaalisesti kasvavassa tieteellisten julkaisujen tulvassa tutkijat joutuvat tekemään valintoja seuraamiensa julkaisujen suhteen. Mertonin mukaan tutkijan maine on eräs tätä valintaprosessia helpottava tekijä. On ilmeistä, että myös julkaisujen maine vaikuttaa siihen, mitä julkaisuja tutkijat säännöllisesti seuraavat, ja sitä kautta vaikuttaa myös siihen, mitä artikkeleita he helpoimmin huomaavat ja joihin he mahdollisesti itse myöhemmin viittaavat. Julkaisujen seuraamiseen vaikuttaa toki julkaisujen saatavuus. Julkaisujen tilauskustannusten kasvaessa ja kirjastomäärärahojen jäädessä jälkeen kirjastot joutuvat myös tekemään entistä tiukempia valintoja sen suhteen, mitä aikakausjulkaisuja tilataan. Tässä julkaisujen impact-luvut toimivat kirjastojen tilauksia ohjaavina ja siten tätäkin kautta näkyvien julkaisujen asema voi entisestään vahvistua.

Vaikka voitaisiin periaatteellisesti hyväksyä se, että viitteet mittaavat tutkimuksen tosiasiallista vastaanottoa ja vaikutusta, viiteanalyysin soveltaminen eri tieteenaloilla ja erilaisten maiden tutkimukseen aiheuttaa merkittäviä ongelmia. Viiteanalyysiin kuuluu, että sen avulla mitataan tutkimuksen tosiasiallista vaikutusta kansainvälisesti, alan tutkimuksen eturintamaan. Tämän periaatteen yksioikoinen soveltaminen ei käy yhtäläisesti kaikille tieteenaloille. Vastaavasti kansainväliset viitetiedostot, joiden systemaattisessa keräämisessä Institute for Scientific Informationilla on käytännössä monopoliasema, perustuvat etupäässä aikakauslehtien julkaisemaan artikkeliaineistoon. Aikakauslehtien muodostaessa viitetiedostoissa valta-aineiston viiteanalyysi käytännössä pohjautuu olettamuksiin tieteellisen aikakaus-

lehden keskeisestä asemasta tieteellisen kommunikaation välittämisessä. Näin on asianlaita selvimminkin ilmeisesti lääketieteessä ja monilla luonnontieteen aloilla. Sen sijaan esim. teknisten tieteiden tai yhteiskuntatieteiden aloilla myös muuntotyypiset tuotokset tai julkaisut ovat merkittäviä, esimerkiksi kirjat tai erilaiset raporttijulkaisut. Nämä kaksi olettamusta, kansainvälisyys ja aikakauslehtien keskeinen asema, rajoittavat viiteanalyysin sovellutuksia.

Koska viiteanalyysi mittaa tutkimuksen tosiasiallista vaikutusta kansainvälisesti, pienten maiden osalta se antaa tietoa myös näissä suoritetun tutkimuksen kansainvälistymisen asteesta. Ongelmana pienten maiden osalta on kuitenkin se, että niitä koskevat viiteaineistot saattavat jäädä kovin vähäisiksi tutkimuksen suhteellisen pienen volyymin vuoksi. Sitäpaitsi vain osa tästä tutkimuksesta julkaistaan sellaisilla foorumeilla, että se huomataan kansainvälisesti. Edellä käsiteltyä tästä aiheutuvia ongelmia eli viitteiden viivästymistä vähemmän näkyviin julkaisuihin ja toisaalta mahdollisen Matteus-efektin vaikutusta näkyvien julkaisujen aseman korostamisessa.

Vaikka hyväksyttäisiinkin periaatteessa se, että tieteellisten tutkimustulosten merkittävyys on sosiaalisesti määräytynyt, viiteanalyysin käytössä arvioinnin tarpeisiin joudutaan juuri tästä syystä tulkinnallisiin ongelmiin. Viiteanalyysiä sovellettaessa ei nimittäin täsmälleen tiedetä, mitä itse asiassa menetelmällä mitataan.

Viiteanalyysin soveltaminen ei ole ongelmaton myöskään käytännön syistä. Analyysi on erittäin työlästä, mikrojen käytöstä huolimatta työvoimaintensiivistä ja kallista. Analyysiä ei voida toistaiseksi soveltaa mekaanisesti. Erilaisia tutkimuskokonaisuuksia koskevien viitetietojen määrä kasvaa koko ajan, sillä alalla tehdään varsin intensiivistä tutkimustyötä. Siitä huolimatta juuri kulloinkin kyseessä olevien tarpeiden mukaisia vertailutietoja (eim. syöpätutkimuksen samat keskimääräiset viitteet kansainvälisesti) ei yleensä ole saatavilla. Kukin tutkimus tehdään hieman eri lailla käyttäen erilaisia seurantajaksoja, mitkä molemmat seikat vaikuttavat viitteiden määriin. Tämän vuoksi vertailujen tekeminen eri tutkimustulosten välillä on vaikeaa. Viiteanalyysin käytännön suorittamiseen liittyy myös teknisiä, erilaisista virhetekijöistä aiheutuvia ongelmia, joista syistä analyysiä suositellaan

tavallisesti sovellettavaksi suurempiin kokonaisuuksiin kuin yksittäiset tutkijat tai artikkelit. Kaikista ongelmista huolimatta voidaan todeta, että kovin suurille kokonaisuuksille ei ole kvantitatiivisen viiteanalyysin lisäksi muuta käytännöllisesti mahdollista keinoa arvioida tutkimuksen »sosiaalisesti tulkittua laatua».

Kuka tulkitsee indikaattoreita. Kysymys vallasta

Olen edellä tuonut esille lukuisia ongelmia erilaisten indikaattorien tulkinnassa ja käytössä. Siksi lienee itsestään selvää, etteivät tiedeindikaattorit ole ratkaisu Suomen korkeakoululaitoksen tai perustutkimuksen arvioinnin ongelmaan. Tämä ei kuitenkaan merkitse etteikö indikaattorien tuoma tieto olisi hyödyllinen elementti arviointitoiminnassa. Järkevintä indikaattorien käyttöä olisi mielestäni tilanne, jossa indikaattorit antavat taustatietoa ja lähtökohtatietoa muilla menetelmillä, lähinnä vertaisarvioinnilla suoritettavalle arvioinnille (ks. Luukkonen-Gronow, 1987 a). Vertaisarviointi tarkoittaa tiedeyhteisön omien jäsenten suorittamaa arviointia, joka on institutionalisoitunut tieteessä monin tavoin, esimerkiksi virantäytöissä ja tieteellisten julkaisujen refereerijärjestelmässä. Suomen Akatemian toimikuntajärjestelmä on eräs muoto institutionalisoitunutta vertaisarviointia tutkimusvarojen jakamiseksi. Suomen Akatemian suorittamat tieteentutkimuksen arvioinnit ovat varsin uusi sovellutus vertaisarvioinnin periaatteista kokonaisten tutkimusalojen arviointiin.

Monet näistä indikaattoreista ovat parhaiten sovellettavissa tutkimusalueille, jotka ovat luonteeltaan kansainvälisiä. Indikaattorien järkevintä käyttöä olisi kysymysten asettaminen siitä, miksi esimerkiksi yliopiston laitoksella ei ole lainkaan kansainvälisiä kontakteja, miksi julkaisurakenne on havaitun kaltainen, miksi yhteenkään julkaisuun ei ole viitattu tms. Olen jo useaan otteeseen tuonut esille kysymyksen indikaattorien tulkinnan tärkeydestä ja ongelmallisuudesta. Tämä tulkintavaihe ei ole mekaanista ja toteutettavissa vain opetusministeriön virkamiesvoimin. Useiden indikaattorien kohdalla ei ole olemassa selvää tulkintaohjetta, vaan indikaattoreiden tulkinta riippuu siitä miten erilaiset asiat halutaan arvostaa. Tähän sisältyy selvä vallankäytön elementti. Tulkinnassa on tarpeen turvautua jonkinlaiseen *vertaisten* (engl. peers) käyttöön, ja siten

indikaattorien käyttö tulisi yhdistää vertaisarvioinnin sovellutuksiin.

LÄHTEET

- Bayer, A.E. & Folger, J. Some Correlates of a Citation Measure of Productivity in Science. *Sociology of Education*, 39 (1966): s. 381—390.
- Brooks, T.A. Evidence of Complex Citer Motivations. *Journal of The American Society For Information Science*, 37 (1986): 1, s. 34—36.
- Cole, J.R. & Cole, S. *Social Stratification in Science*. Chicago & London: University of Chicago Press, 1973.
- Chubin, D.E. & Moitra, S.D. Content Analysis of References: Adjunct or Alternative to Citation Counting. *Social Studies of Science*, 5 (1975): s. 423—441.
- Garfield, E. What Scientific Journals Can Tell Us About Scientific Journals. *IEEE Transactions on Professional Communication* PC-16, no. 4 (1973): s. 200—202.
- Hagstrom, W.O. *The Scientific Community*. New York & London: Basic Books Inc. 1965.
- Järvinen, O. & Pietiäinen, H. Citation Patterns of Papers Published in *Ornis Fennica*. *Ornis Fennica*, 65 (1988): s. 31—36.
- Luukkonen-Gronow, T. Scientific Research Evaluation: A Review of Methods and Various Contexts of Their Application. *R. & D Management*, 17 (1987): 3, s. 207—221 (1987 a).
- Luukkonen-Gronow, T. Publish in a Visible Journal or Perish? A Bibliometric Analysis of Nordic Cancer Research. Paper Prepared for the Workshop »The Relations Between Qualitative Theory and Scientometric Methods in S & T Studies», Amsterdam, December 10—11, 1987 (1987 b).
- MacRoberts, M.H. & MacRoberts, B.R. The Negational Reference: or the Art of Dissembling. *Social Studies of Science*, 14 (1984): s. 91—94.
- Martin, B.R. & Irvine, J. Assessing Basic Research. *Research Policy*, 12 (1983): s. 61—90.
- Merton, R.K. The Matthew Effect in Science. *Science*, 159 (1968): s. 56—63.
- Merton, R.K. Priorities in Scientific Discovery, teoksessa *The Sociology of Science*. Chicago & London: University of Chicago Press (1957) 1973, s. 286—324.
- Moed, H.F. et al., The Use of Bibliometric Data for the Measurement of University Research Performance. *Research Policy*, 14 (1985): s. 131—149.
- Moravcsik, M.J. & Murugesan, P. Some Results of the Function and Quality of Citations. *Social Studies of Science*, 5 (1975): s. 86—92.
- Niiniluoto, I. Tiede, hyöty ja kilpailu. Teoksessa K. Mäkelä (toim.) *Tieteen vapaus ja tutkimuksen etiikka*. Helsinki: Tammi 1987, s. 90—113.
- Pritchard, A. Statistical Bibliography or Bibliometrics. *Journal of Documentation*, 25 (1969): s. 348—349.
- Zuckerman, H. Nobel Laureates in Science: Patterns of Productivity, Collaboration, and Authorship. *American Sociological Review*, 32 (1967): 3, s. 391—403.
- Zuckerman, H. Citation Analysis and the Complex problem of Intellectual Influence. *Scientometrics*, 12 (1987) 5—6, s. 329—338.