



Voiko tieteenaloja verrata toisiinsa?

Antti Arjava

Kun tiede on ammattimaistunut ja sen rahoitus lisääntynyt, tutkijoilta odotetaan yhä enemmän näyttöjä tuloksista. Siksi arvioinnit lisääntyvät. Perinteisesti yksilöitä ja yksiköitä on arvioitu oman alansa sisällä. Tieteenaloja voi kuitenkin verrata myös toisiinsa. Se voidaan tehdä monin eri tavoin ja päätyä erilaisiin tuloksiin. Tuloksista riippumatta vertailu johtaa hyödyllisiin pohdintoihin tieteellisen tutkimuksen olemuksesta sekä alojen eroista ja yhtäläisyyksistä. Tulostavastuun kurimukselta voisi hypätä ulos vain se, joka uskaltaisi luopua vaatimasta lisää resursseja.

Helsingin yliopisto suoritti vuonna 1999 kaikki tieteenalat kattavan tutkimuksen tason arvioinnin. Alkuperäisenä tarkoituksena oli toteuttaa hanke yhdessä muiden kotimaisten korkeakoulujen kanssa, niin että kutakin yliopistollista laitosta verrattaisiin saman tai lähialojen muihin laitoksiin. Periaatteellisten ja käytännöllisten vaikeuksien vuoksi yhteistyö ei toteutunut, vaan Helsingin yliopisto sai toteuttaa arvioinnin yksin. Siitä seurasi, että eri alojen laitoksia jouduttiinkin itse asiassa vertaamaan toisiinsa. Laitosten konkreettisten tulosten ja tieteenalojen yleisen tason arviointi kietoutuivat toisiinsa paljon monimutkaisemmin kuin esimerkiksi Yhdistyneen kuningaskunnan kansallisissa laatuarvioinneissa, joita on tehty 1980-luvulta lähtien.

Kuvaan ensin lyhyesti Helsingissä tehtyä arviointia ja sen mahdollisia virhelähteitä. Siitä saatujen kokemusten pohjalta esittelen kysymyksiä, jotka heräävät arvioinnin tuloksia tutkiessa. Ensinnäkin voidaan pohtia, onko olemassa yleisiä menestystekijöitä, jotka selittävät yksittäisten laitosten tutkimuksen korkeaa laatua yli tiedekuntarajojen. Toisaalta joudutaan tarkastelemaan kokonaisten tieteenalojen toimintakulttuureita ja laatueroja maailmanlaajuisessa mittakaavassa.

Joustavat kriteerit

Koska Helsingin arvioinnissa tuloksen haluttiin vaikuttavan suoraan tiedekuntien määrärahoihin, piti vertailun yhteismitallisuuteen kiinnittää sitäkin enemmän huomiota. Ongelma päätettiin ratkaista vertaamalla kutakin laitosta saman tieteenalan kansainväliseen tasoon. Sitä varten yliopiston noin 120 ainelaitosta jaettiin 24 paneeliin (ks. taulukko). Paneelien arviointijoina toimi yhteensä 124 asiantuntijaa, joista noin neljännes oli kotimaasta, toinen neljännes muista pohjoismaista ja loput muualta maailmasta. Näistä eniten oli saksalaisia, brittejä ja amerikkalaisia, mutta paneelisteja tuli myös esimerkiksi Puolasta, Unkarista ja Japanista.

Tehtävänä vertailu saman tieteenalan kansainväliseen tasoon saattaa kuulostaa selkeältä. Paneelit joutuivat kuitenkin ensin ratkaisemaan, mikä on sama tieteenala: yksi oppiaine, laitos, koko yhden paneelin kattama alue vai jokin vielä suurempi kokonaisuus? Useimmat paneelit päätyivät siihen, että paneelin sisällä tulisi käyttää samaa vertailuasteikkoa. Toisin sanoen sama arvosana kahdelle laitokselle kertoi suoraan niiden tutkimuksen olevan toisiinsa nähden samalla tasolla. Mutta ainakin yksi paneeli päätyi päinvastaiseen ratkaisuun. Se antoi kahdelle laitokselle saman arvosanan, vaikka ei katsonut niiden olevan keskenään samantasoisia: vain niiden etäisyys oman alansa kansainvälisestä kärjestä oli sama.

Edellä olevasta seuraa, että vaikka sijoittuminen tiettyyn paneeliin ei varmasti yksinään ratkaissut minkään laitoksen arvostelua, se ei ollut yhdentekevä asia. Paneelijaolla on toinenkin merkitys arvioinnin uskottavuuden kannalta. Tieteen kuten taiteen perusongelma Suomen kaltaisessa maassa on vaikeus löytää riippumattomia asiantuntijoita, jotka ovat valmiita kriittisesti arvioimaan omia kollegoitaan alan ulkopuoliselle yleisölle. Esimerkiksi tiedotusvälineet usein saavat puristetuksi avointa kritiikkiä vain tutkijoilta, jotka ovat ilmiiridoissa kritiikin kohteen kanssa - tämäkään ei ole terve lähtökohta oikealle laatuarvioinnille. Ulkomaalaisten käyttö ei automaattisesti takaa luottavampaa lopputulosta, varsinkin jos he tietävät arvioinnin suoraan vaikuttavan oman alansa resurssointiin Suomessa. Siksi laajempi paneeli on lähtökohtaisesti uskottavampi: koska asiantuntijoita on enemmän, he joutuvat vertaamaan useampia yksiköitä, ja yhden jäsenen on vaikeampi vedättää omaa alaansa ylöspäin. Oman alan nostaminen ei ollut täysin tuntematon riski Helsingin arvioinnissa, mutta kokonaisuutena tilanne pysyi hyvin hallinnassa, eikä useimpia paneelle voi syyttää ainakaan lepsuilusta.

Jos halutaan toisistaan suuresti poikkeavista tieteenaloista jollakin tavoin vertailukelpoisia arvioita, on arviointikriteerien joustavuus kaikkein tärkein edellytys. Hyvin yksityiskohtaiset kriteerit antavat näennäisen kuvan objektiivisista mittareista mutta todellisuudessa kohtelevat aloja eri tavoin, koska niiden toimintakulttuurit poikkeavat toisistaan. Joustavuus antaa epäilemättä mahdollisuuden hieman pelata arvioinnissa (esimerkiksi ilmoitettavaa materiaalia valittaessa), mutta se on parempi kuin että pelaaminen siirtyisi itse tieteelliseen työhön ja toisi siihen piirteitä, jotka eivät ole yksittäisen alan kannalta mielekkäitä. Kriteereissä ei myöskään saa vaatia mitään tiettyä tapaa saavuttaa tuloksia. Jos tehokkain tapa tehdä hyvää tutkimusta on maata sohvalla odottaen inspiraatiota, arvioinnin tulee tämän menetelmäkysymyksen suhteen olla täysin neutraali.

Kukin laitos ilmoitti arviointiin vuosilta 1994-98 keskimäärin kaksi julkaisua tutkijaa kohti. Arviointikriteerit olivat sikäli yksinkertaiset, että paneelin tuli lähinnä määrittää, kuinka suuri osa ilmoitetuista julkaisuista oli erinomaista, hyvää tai tyydyttävää kansainvälistä tasoa. Esimerkiksi korkein arvosana 7 edellytti, että julkaisuista yli puolet oli erinomaista ja käytännöllisesti katsoen kaikki muut hyvää kansainvälistä tasoa. Matalimman arvosanan 1 sai laitos, jonka julkaisut eivät yltäneet edes tyydyttävälle kansainväliselle tasolle. Erinomaiselle kansainväliselle tasolle työ luokiteltiin, jos se laatunsa puolesta olisi voitu julkaista oman alansa johtavissa lehdissä tai parhailla kustantajilla, joilla valintakriteerit olivat tiukimmat. Vastaavasti määriteltiin hyvä ja tyydyttävä kansainvälinen taso.

Tästä eteenpäin kunkin paneelin oli sovellettava kriteereitä omalle alalleen parhaiten sopivalla tavalla. Esimerkiksi biotieteiden paneeli katsoi, että sen piiriin kuuluvilla aloilla "johtavia lehtiä" oli 25 % tieteellisistä julkaisuista. Lääketieteen



paneeli taas rajasi huomattavasti pienemmän osuuden lehdistä "johtaviksi". Sama paneeli totesi myös, että lääketieteessä kaksi julkaisua tutkijaa kohti on löysä vaatimus ja johtaisi yliopiston asteikkoa käytettäessä liian helposti korkeisiin arvosanoihin; siksi se sovelsi asteikkoa tätä tiukemmin. Voikin ymmärtää, että kun lääketieteessä julkaistaan nopeassa tahdissa lyhyitä artikkeleita, niitä ei voi rinnastaa suoraan monisatasivuisiin monografioihin. Paneeleille jäi ohjeiden soveltamisessa suuri vastuu, ja yleisesti ottaen se onnistui hämmästyttävän hyvin.

Etukäteen varauduttiin kuitenkin siihen, että joillakin paneeleilla saattaisi olla kiusaus antaa kohtuuttoman korkeita arvosanoja. Sen takia kriteereihin lisättiin ehto, ettei ylintä arvosanaa 7 saanut antaa, ellei laitos kuulunut Euroopassa alansa laitosten joukossa parhaaseen kymmenekseen. Vastaavasti arvosanaan 6 vaadittiin kuuluminen parhaan neljänneksen sisään. Tämän uskottiin siis tiukentavan arvostelua. Todellisuus osoittautui päinvastaiseksi. Paneelien näkemysten mukaan Euroopassa on kaikilla aloilla niin paljon huonoja laitoksia, että parhaan neljänneksen joukkoon pääseminen ei ole kovin vaikea suoritus. Useimmissa tapauksissa lisäkriteeri jäi näin ollen vaille merkitystä, koska sen saavuttaminen ei oikeuttanut korkeampaan arvosanaan, ellei ensijainen, julkaisuihin perustuva kriteeri täyttynyt.

Kohdeltiinko kaikkia tasapuolisesti?

Noin puolen vuoden työn jälkeen kaikki paneelit olivat antaneet kirjalliset lausuntonsa. Yksittäisten laitosten arvosanat vaihtelivat yhdestä seitsemään, eli koko asteikko tuli käyttöön. Laitosten arvosanoista muodostuneet tiedekuntien keskiarvot sen sijaan olivat suhteellisen lähellä toisiaan. Koko yliopiston keskiarvo 5,5 oli varsin hyvä siihen nähden, että paneelien mielikuvissa Euroopan keskiarvo liikkui yleensä kolmen ja neljän välillä. Sitä alempia arvosanoja laitoksista sai vain kourallinen.

Teologinen tdk 5,8
Lääketieteellinen tdk 5,7
Matemaattis-luonnontieteellinen tdk 5,7
Humanistinen tdk 5,7
Valtiotieteellinen tdk 5,1
Kasvatustieteellinen tdk 5,0
Maatalous-metsätieteellinen tdk 4,9
Oikeustieteellinen tdk 4,8
Eläinlääketieteellinen tdk 4,5

Parhaaksi tiedekunnaksi selviytyi ehkä joidenkin yllätykseksi teologinen. Kolme suurta tiedekuntaa oli käytännössä sen kanssa samalla tasolla, pienen eron jälkeen tulivat muut. Erityisesti suurten monitieteisten tiedekuntien ja valtiotieteellisen tiedekunnan arvosanoja voi pitää sikäli luotettavina, että ne muodostuivat monien eri paneelien arvosanoista: mahdolliset tulkintaerot tasoituivat keskiarvoa laskettaessa.

Tiedekuntien arvosanoja tarkastellessa huomaa heti, ettei laadun rajalinja kulkenut ainakaan humanististen ja luonnontieteellisten alojen välillä. Ilmeisesti myöskään kriteerit eivät systemaattisesti suosineet niistä kumpiakkaan. Luonnontieteelliset alat vaikuttivat kyllä helpommilta arvioida. Sen sijaan kärkeen sijoittuneissa tiedekunnissa on ehkä yhteistä se, että niissä perustutkimuksen osuus on suurempi soveltavaan tutkimukseen verrattuna. Tällainen tulos ei olisi täysin yllättävä ottaen huomioon, että ennen arviointia keskusteltiin erityisen paljon kriteerien sopivuudesta soveltaviin tieteisiin. Osoittaako tulos siis, että kriteerejä luotaessa epäonnistuttiin? Onnistuneen arvioinnin tunnusmerkkinä ei tietenkään voi pitää sitä, että kaikki arvioitavat saavat saman tuloksen. Mutta tällaisen arvioinnin luonteeseen kuuluu, ettei väitettä väärästä kriteereistä voi sitovasti todistaa oikeaksi tai vääräksi. Tosin jokaisesta tiedekunnasta löytyi erinomaisia yksiköitä tai vähintään laitosten osia. Voittajiksi selvisivät ne tiedekunnat, joissa tutkimusintensiivisten laitosten osuus oli suurempi. Tämän perusteella ei näytä siltä, että valitut kriteerit olisivat tehneet huippuarvosanojen saavuttamisen joillekin tiedekunnille mahdottomaksi. Ns. kansallisiin aloihin kiinnitettiin alusta lähtien erityishuomioita. Esimerkiksi suomen kielen laitosta oli määrä verrata ruotsin kielen laitokseen Tukholmassa tai ranskan kielen laitokseen Pariisissa. Kansallisia aloja ja suomenkielistä tuotantoa haluttiin tiettyssä mielessä suojella mutta ei ylisuojella. Kielimuurin problematiikka on sama kuin esimerkiksi kansainvälisen kaupan rajoituksissa. Monissa tapauksissa paikalliset olosuhteet aidosti vaativat kansallisia määräyksiä. Ei ole kuitenkaan ilmeistä, miksi jonkin maan kansallinen sähkövirta vaatisi pistokkeeseen kolme piikkiä, kun naapurissa riittää kaksi. Silloin on syytä pelätä, että kilpailun rajoituksilla tietoisesti tai tiedostamatta suojellaan omaa tehotonta toimintaa. Tämän ongelmakentän avoin tarkastelu olisi monille tieteenoiloille eduksi.

Mutta koska Helsingin yliopiston arviointikriteereissä kysyttiin vain, missä työ olisi voitu julkaista - eikä siis missä se oli todellisuudessa julkaistu - suomen kielen käyttö sinänsä ei estänyt mitään laitosta saavuttamasta korkeaa arvosanaa. Kansallinen ala par excellence, suomen kielen laitos, sai arvioinnissa korkeimman arvosanan, koska paneeli ei tuntenut Euroopassa sitä korkeatasoisempaa äidinkielen laitosta. Ja esimerkiksi teologian koulutusallalla vain neljännes artikkeleista julkaistaan Suomen Akatemian tilastojen mukaan ulkomaisissa refereesarjoissa, mutta silti tiedekunta sijoittui arvioinnissa ensimmäiseksi.

Kaikessa on vaikea olla hyvä

Arvioinnin keskeisiä tavoitteita oli kertoa, missä hyvää tutkimusta tehdään. Pettynneiden lisäksi yliopistolla oli paljon laitoksia, jotka nyt ensimmäistä kertaa saivat julkista tunnustusta vuosikausien hiljaisesta työstä. Tuloksista voi lukea yllättäviäkin havaintoja. Esimerkiksi paljon puhuttu kriittinen massa ei Helsingin arvioinnin valossa painanut paljonkaan. Monet korkeimman arvosanan saaneista yksiköistä kukoistivat vain muutaman aktiivisen tutkijan voimin, eikä ainoastaan humanistisissa tieteissä. Toki pienet laitokset saivat myös heikkoja arvosanoja, kun taas suurlaitosten arvosanat luonnostaan hakeutuivat lähemmäs keskiarvoa. Jos kuitenkin halutaan etsiä jotain yhteistä tekijää, joka selittäisi laitosten sijoittumista arviointiasteikolla, se ilmeisesti liittyisi jollain tavoin kansallisten intressien ongelmaan. Näyttää nimittäin siltä, että keskimäärin arvioinnissa menestyivät parhaiten ne laitokset, jotka näkivät tutkimustyönsä keskeisenä yleisönä kansainvälisen tiedeyhteisön. Niille laitoksille, jotka mielsivät työnsä tulosten hyödyttävän eniten kotimaista yhteiskuntaa, oli huomattavasti vaikeampaa vakuuttaa arviointipaneeleita työnsä korkeasta laadusta. Tämä ei

koskenut ainoastaan tyypillisiä humanistisia kansallisia aloja, vaan sama linja kulkii jokseenkin kaikkien tiedekuntien läpi. Mikäli tämä havainto pitää paikkansa, se ei tietenkään ole hyvä uutinen. Se kuvastaa sitä, kuinka vaikeaa millekään laitokselle on täyttää kaikki ne odotukset, joita yhteiskunta nykyään kohdistaa yliopistoille. Paitsi ihmiskuntaa, korkeakoulujen tulisi hyödyttää myös isänmaata ja lähialuettaan. Jos strategisesti keskittää ponnistukset jälkimmäisiin tavoitteisiin, saattaa kansainvälinen kärki jäädä saavuttamatta. Kaikki alat ovat tässä samojen kysymysten edessä. Oikeustieteilijät hoitavat sivutyönään lainvalmistelua, kun taas humanistien ja yhteiskuntatieteilijöiden odotetaan osallistuvan yhteiskunnalliseen keskusteluun kaikissa medioissa. Lääketieteilijät parantavat ihmisiä, ja luonnontieteilijöillä on omat houkutusensa yritysmaailmassa. Tällaisten tehtävien olemassaolo on vaikea hyväksyä syyksi olla tekemättä perustutkimusta, koska silloin tutkimusvelvollisuutta ei jäisi itse asiassa juuri kenellekään. Yhteiskunnallisesta aktiivisuudesta ei arvioinnissa tietenkään annettu minkäänlaisia miinus pisteitä, pikemminkin päinvastoin, mutta sillä ei voinut korvata oman tiedeyhteisön arvostaman tutkimuksen vähäisyyttä. Samantapainen optimointiongelma näyttäisi vallitsevan tutkimuksellisen laadun ja maisterituotannon välillä. Tutkimuksessa parhaiten menestyneet laitokset ovat usein niitä, jotka eivät juuri voi yleillä maisteriluvuilla. Tässä valintatilanne saattaa kuitenkin olla näennäinen. Vaikka jokainen laitos voi varmasti kohentaa tulostaan paremmalla opintojen ohjauksella, laitosten omien toimien vaikutus ei voi selittää alojen välisiä lähes paradigmaattisia eroja. Valmistumisprosenttia ei nosteta kahdestakymmenestä yhdeksäänkymmeneen gradupiorejä perustamalla. Valmistuneiden määrää säätelee enemmän alan työllisyystilanne tai jokin muu opiskelijoiden motivaatiotekijä. Paljon maistereita valmistuu erityisesti laitoksilla, joilla on selvä ammattikuvaa yliopiston ulkopuolella. Vahva ammattisuuntautuneisuus puolestaan korreloi negatiivisesti opiskelijoiden tutkimusmotivaation kanssa, ja se heijastuu helposti koko alan tutkimukselliseen imagoon.

Yliopistossa on joitakin laitoksia, jotka ovat pystyneet hoitamaan monet eri tehtäväalueet kiitettävästi, vaikkapa filosofian ja psykologian laitokset. Avuksi ovat sekä onnelliset olosuhteet että monipuolisesti pätevät henkilökunta. Sikäli kuin varsinaista menestysreseptiä on olemassa, se on ehkä jonkinlainen voimien jakaminen, niin että osa tutkijoista keskittyy perustutkimukseen, osa vastaa yhteiskunnan lyhytjänteisempiin vaatimuksiin. Parhaassa tapauksessa ihminen voi vaihtaa roolia uransa eri vaiheissa. Kaiken tekeminen yhtäaikaa on vaikeaa sille enemmistölle tutkijoista, joka tarvitsee säännöllisen yön ja palasen perhe-elämää.

Tieteenalojen erot

Tieteenalojen välisiä eroja voi lähestyä toisestakin suunnasta. Oheisen mallikuvan palkit kuvaavat teoreettisena lähestymistapana yhden tieteenalan tutkimusta (laitoksia, tiedekuntia) maailmanlaajuisesti. Maailman parhaat laitokset sijoittuisivat kriteereillämme palkin yläpäähän (arvosana 7) ja heikoimmat laitokset alapäähän (arvosana 1). Pystyakseli kuvaa alojen absoluuttista tutkimuksellista tasoa, mitattaisiinpä sitä tutkimuksen intensiteetillä, inhimillisellä suorituskyvulla tai jollakin muulla täysin teoreettisella ja abstraktilla suurella.

Helsingin arvioinnissa pyrittiin siis määrittämään, mihin osaan palkkia ko. Helsingin yliopiston laitos sijoittui. Tässä arvioinnissa palkkien asemaa toisiinsa nähden ei edes yritetty selvittää. Sen sijaan keskityttiin arvioimaan yksittäisen laitoksen sijaintia oman alansa palkin sisällä. Se ei tarkoita, ettei alojen välisiä eroja voisi olla olemassa. Päin vastoin toimintakulttuurin ja saavutuksien erot tulevat varmasti ilmi jokaiselle kuuntelijalle, kun suuri joukko eri alojen laitoksia esittelee työtään.

Teoriassa eri tieteenalojen vertailu (so. palkkien sijainnin määrittäminen) voitaisiin toteuttaa niin, että tarpeeksi laaja-alaiset asiantuntijat kattaisivat ketjuna limittäin tieteen kirjon yhdestä reunasta toiseen. Näin matematiikkaa ja fysiikkaa voitaisiin arvioida osittain yhdessä, fysiikkaa ja kemiaa taas tarkastella fyysikaalisen kemian kautta toisiinsa verrannollisina. Kemiasta päästäisiin biokemian kautta biotieteisiin, niistä maatalous- ja metsätieteisiin, lääketieteeseen ja eläinlääketieteeseen. Loikkaus luonnontieteistä ihmistieteisiin onnistuisi ehkä maantieteessä, kansanterveystieteessä ja tilastotieteessä. Sen jälkeen olisi taas yksinkertaisempaa edetä poikkitieteellisten asiantuntijoiden avulla kautta humanististen ja yhteiskuntatieteellisten alojen kentän. Jos voidaan luottaa siihen, että jokaisessa yksittäisessä vertailussa on päästy yhteismallisuuteen, myös niistä syntyvä ketju antaisi periaatteessa totuudenmukaisen kuvan vertailtavien laitosten keskinäisestä tasosta. Tällainen koe olisi mielenkiintoinen, vaikka se tyydyttäisi varmaankin enemmän inhimillistä uteliaisuutta kuin mitään varsinaista käytännön tarvetta.

Mistä johtuu, että joillakin aloilla inhimillisiä voimavaroja uhrataan intensiivisemmin uuden tiedon etsintään? Syitä on monia, joista yksi mutta ei ainoa on kova kansainvälinen kilpailu. Periaatteessa olisi täysin mahdollista juosta 5000 metrin maailmanennätys pelkästään kansallisissa kisoissa kilpaillen, mutta jostain syystä ennätysjätkä tekevät yleensä ne, jotka ovat kohdanneet alansa parhaimmiston monissa kilpailuissa ympäri maailmaa. Mitä sitten voisi tehdä, jos kansainvälisiä kilpakumppaneita on vaikea löytää? Tai jos jouduttaisiin valitsemaan Suomen parasta urheilijaa ja vastakkain olisivat maailman paras pesäpallolijja ja maailman paras formulakuljettaja? Heidän asettamistaan paremmuusjärjestykseen on vaikea perustella sitovasti. Kriteeriksi ei varmasti käy, että toiselle on ulkopuolisella rahoituksella hankittu alle kallis menopeli, kun toisen aseena on puinen maila. Mutta intuitiivisesti useimmat varmaankin ajattelisivat, että autoilijan suorituskyyky on joutunut kovempaan testiin, koska kilpailijoita on niin paljon enemmän ja panokset suuremmat. Samasta syystä jalkapallon maailmanmestaruutta voisi pitää noin 1000 kertaa arvokkaampana saavutuksena kuin jääkiekon maailmanmestaruutta, vaikka sitä ei Suomessa useinkaan uskoisi.

Toisaalta alat elävät myös historiallisessa kehityksessään eri vaiheita. Jotkin oppiaineet, kuten opettajankoulutus, ovat maailmanlaajuisesti vasta aksettään saaneet akateemisen statuksen ja hakevat tieteellisen kulttuurinsa mallia. Onko kilpailuasetelma silloin kohtuuttoman helppo? Paavo Numenkaan tuloksilla ei tänään päästäisi edes kansalliseen kärkeen, ja hänen menestyksensä perustui ensi sijassa siihen, että kilpakumppanit ulkomaila eivät olleet oppineet harjoittelemaan. Pidämme silti häntä edelleen erinomaisena urheilijana, koska hän omassa historiallisessa yhteydessään pystyi parempiin suorituksiin kuin kukaan muu.

Joillekin aloille tekninen kehitys on äkkiä antanut lisävauhtia, niin että ne etenevät uudesta huimaavasta löydöstä toiseen. Etenemisen vauhti ei kuitenkaan välttämättä ole suoraan verrannollinen matkan intellektuaaliseen

vaativuuteen. Geenienkin metsästys voi olla varsin mekaanista toimintaa, kun tekniikan on ensin omaksunut. Tällaista tutkimusta voisi verrata suuriin löytöretkiin. Silloinkin riitti, että retkeilijällä oli taitoa ja uskallusta lähteä purjehtimaan kotisatamasta. Ei ollut niinkään ratkaisevaa, mihin suuntaan kokkansa käänsi, jossakin vaiheessa tuli kaikkialla uusi manner vastaan. Useimmat tieteenalat ovat jo kauan sitten jättäneet löytöretkien ajan taakseen. Ne joutuvat koluamaan tutumpia seutuja, joista uuden löytäminen vaatii jo huomattavaa äylyllistä ponnistusta. Kääntöpuolena on se keskinkertainen tutkimus, joka ei jaksakaan kiinnostaa ketään, kun sitä ei pelasta edes uusi ja kiihottava aihe.

On vielä aloja, joilla mahdollisuuksia varsinaisiin innovaatioihin on yksinkertaisesti vähemmän, mutta useat oikeustieteen haarat. Useimmissa maissa tieteellinen perustutkimus ei ole keskeinen osa akateemista oikeustieteellistä kulttuuria. Ja esimerkiksi kirjallisuudentutkimuksessa kansainvälinen käytäntö vaatii professorilla pikemmin taitavasti kirjoitettua esseistikkää kuin tutkimuksia sellaisilla argumenteilla, jotka ovat tuttuja luonnontieteistä, historiasta tai kielitieteestä.

Kaikesta tästä seuraa, että oikeudenmukaisten kriteerien asettaminen eri tieteenaloille on aina tasapainoilua. Yhteiskunnan avokätinen resurssointi ja kova kansainvälinen kilpailu nostavat alan tasoa mutta tekevät myös yksittäiselle laitokselle vaikeammaksi yltää alansa kärkeen. Helsingin yliopiston arviointikriteerit pyrkivät ottamaan tämän huomioon ja takaamaan eri aloille mahdollisimman samanlaiset mahdollisuudet korkeimpiin arvosanoihin mutta myös jonkinlaisen perusvertailun alojen välille. Lopputuloksena saatu arvosanojen jakauma on kompromissi, josta ei voi suoraan lukea sellaisia asioita kuin laitosten keskimääräinen lahjakkuustaso, tutkijoiden henkilökohtaisista ominaisuuksista puhumattakaan.

Onko vertailusta hyötyä?

Se että tieteenaloja voidaan verrata keskenään, ei tietenkään automaattisesti merkitse, että niitä myös kannattaa vertailla. Ei ole eduksi, jos muutaman vuoden kuluttua maailman kaikki tutkijat vain arvioivat toisiaan. Arviointeja voidaan vastustaa joko yleensä tai sitten voidaan ajatella, että arvioinnit tulisi rajata kunkin tieteenalan sisälle, jotta alojen välisten vertailujen ylivoimaiset ongelmat vältettäisiin.

Tehokkuuden korostaminen on johtanut yhteiskunnassa yleisesti ongelmiin, jotka saattavat pitkällä tähtäyksellä ylittää siitä saatavat hyödyt. Arviointi-ideologian taustalla on tietenkin ajatus, että laatua voidaan mitata ja sillä on väliä. Tämä ajatus voidaan asettaa kyseenalaiseksi, mutta tieteessä sillä on aina ollut laajaa kannatusta. Yliopistoihin kilpailu ja inhimillisen suorituskyvyn kehittäminen on kuulunut viimeistään 1800-luvulta lähtien, ja olisi vaikea kuvitella yliopistoa, jossa tutkijoiden urakehitys ei perustuisi tulosten rehelliseen mittaamiseen. Tutkijanuralla keskeiset paineet syntyvätkin henkilökohtaisesta kilpailusta. Itse asiassa useimmat laitosten arviointiin liittyvät periaatteelliset ja käytännölliset ongelmat koskevat myös professorien virantäyttöjä. Arvioinnin työläisyys on johtanut siihen, että professuureja täytetään yhä enemmän kutsusta. Tämä on johtanut epäilyksiin, että puolueettoman arvioinnin puuttuessa valtarakenteet johtavat epäreiluun suosintaan, esimerkiksi miesten suosintaan naisten kustannuksella. Nyrkkisääntönä voikin sanoa, että mitä vähemmän arviointeja, sitä parempi on niiden henkilöiden ja laitosten asema, jotka ovat menneisyydessä saavuttaneet itselleen etuja, olivatpa ne alunperin ansaittuja tai ansaitsemattomia.

Laitosten arvioinnissa korostuu henkilökohtaisen kilpailun sijasta hurskas odotus ihmisten yhteistyöstä. Tärkeintä eivät ole numeeriset lopputulokset, joiden tarkkuuteen voidaan aina kohdistaa enemmän tai vähemmän oikeutettuja epäilyksiä. Olenaisempaa on ajattelu, että tekemällä parempaa tutkimusta yhteisö sijoittuu arvioinnissa paremmin. Toisaalta vaikka yksilöarvioinnin ainakin periaattessa tulisi johtaa siihen, että tutkijat oman etunsa vuoksi tekevät parempaa tutkimusta, ei ole takeita siitä, että koko yhteisön arviointi johtaa toimintakulttuurin muutokseen - silloinkaan kun se ulospäin näyttäisi tarpeelliselta. Tieteelliset yhteisöt, kuten laitokset, voivat ryhminä olla paremmin puskuroituja ulkopuolista kritiikkiä vastaan.


Ihmisillä on monilla elämänoilla luonnollinen taipumus pyrkiä rajaamaan omat pelikenttensä, joilla he voivat määritellä säännöt ja valita tuomarit omasta keskuudestaan. Kohtuullisessa mitassa se on hyväksyttävää, koska se antaa helpommin mahdollisuuden päteä ja tekee ihmisistä onnellisempia. Mutta se voi johtaa myös vahingollisten kuplien syntymiseen. Pienessä maassa sellaiset kuplat voivat elää pitkäänkin, jopa sukupolven ajan. Lopulta ne yleensä puhkeavat, viimeistään silloin kun uusi sukupolvi alkaa kritisoida edeltäjiään osoittaakseen omaa pätevyytään.

Keskeistä kaikenlaisessa vertailussa on tietenkin se, että arviointikriteereissä ei ole mitään sellaista, mitä hyvältä yliopistolliselta laitokselta ei muutenkin odotettaisi. Kilpailussa menestyminen ei saa edellyttää sellaista toimintaa, joka ei yliopiston perustehtävien kannalta olisi mielekästä ja kestäisi eettistä tarkastelua. Vähemmän vaarallista on, jos joku sinänsä hyvänä pidetty toimintamuoto jää kriteerien ulkopuolelle: kaikkea ei ole pakko arvioida. Yleensäkin tulee varoa vaatimasta arviointisijoilta tyhjentävää lausuntoa laitosten kaikesta toiminnasta ja olemassaolon oikeutuksesta. Tutkimuksen arvioinnin tehtävä on antaa tietoa tutkimuksen laadusta. Se ei voi kertoa koko totuutta mistään laitoksesta tai tieteenalasta - ainoastaan yhden kohtuullisen luotettavan osan totuudesta.

Kirjoittaja on akatemiaturkija ja klassillisen filologian dosentti, joka organisoi tutkimuksen arvioinnin Helsingin yliopistossa 1999. Kirjoitus perustuu alustukseen Suomen Kulttuurirahaston järjestämällä kulttuurinpuolustuskurssilla 14.-15.3.2002.

Kuvat ja kuvatekstit:
Arviointipaneelit. Helsingin yliopiston arvioinnin toteutus ja tulokset esitellään tarkemmin osoitteessa <http://savotta.helsinki.fi/researcheval>

- 1 MATHEMATICS, COMPUTER STUDIES, AND STATISTICS
- 2 PHYSICS AND ASTRONOMY
- 3 CHEMISTRY
- 4 BIOLOGICAL SCIENCES
- 5 GEOLOGY, METEOROLOGY, AND GEOGRAPHY
- 6 PRE-CLINICAL MEDICINE
- 7 CLINICAL MEDICINE AND DENTISTRY
- 8 PUBLIC HEALTH

- 
- 
- 
- 
- 
- 
- 9 PHARMACY
 - 10 VETERINARY SCIENCE
 - 11 AGRICULTURAL SCIENCES AND FOOD SCIENCE
 - 12 FORESTRY
 - 13 ECONOMICS
 - 14 LAW
 - 15 SOCIAL SCIENCES
 - 16 PSYCHOLOGY
 - 17 EDUCATION
 - 18 POLITICS AND COMMUNICATION STUDIES
 - 19 PHILOSOPHY
 - 20 LANGUAGES
 - 21 ART RESEARCH
 - 22 THEOLOGY
 - 23 HISTORY
 - 24 ASIAN AND AFRICAN STUDIES, CLASSICS, CULTURAL RESEARCH