

LIV LANGFELDT & SVEIN KYVIK

Inre spänningar och framtida utmaningar för *peer review*

Forskare ägnar mycket tid åt att bedöma forskning, vilket kan leda till rollkonflikter och andra motsättningar. Bland de framtida utmaningarna för *peer review* finns frågan hur kraven på insyn och samhällsnytta ska kunna mötas samtidigt som den akademiska självständigheten upprätthålls.



EN VIKTIG DEL av de många uppgifter som ingår i forskarrollen är att utvärdera andra forskares arbete. Avsikten är att främja god forskning på bekostnad av dålig. Utvärderarna ger eller nekar kolleger tillgång till forskningsbidrag, publicering och befordran genom en kollegial granskningsprocess kallad *peer review*. Problemet är, enligt den brittiske fysikern och vetenskapsteoretikern John Ziman, att detta kräver ”stora mängder av den mest värdefulla resursen inom vetenskapen: de mest kompetenta forskarnas tid”.¹ Dessutom har forskningsutvärdering ökat i både betydelse och omfattning på senare år, bland annat till följd av att en större andel av resurserna för forskning fördelas via projektbidrag. Krav på bättre kvalitet inom universitetsforskningen, större samhällsrelevans och förbättrad effektivitet (valuta för pengarna) har lett till en våg av utvärderingar med fokus på organisationsnivån.

En rad olika utvärderingsroller och uppgifter

Det finns åtminstone nio skilda funktioner i forskares roll som utvärderare:

1. bedömning av doktorsavhandlingar
2. val av ny personal och befordran av personal
3. fördelning av forskningsbidrag
4. bedömning av manuskript insända för publicering
5. recensioner av böcker och författande av översiktsartiklar
6. bedömning av kandidater för vetenskapliga utmärkelser
7. utvärdering av forskningsorganisationer
8. bedömning av forskning som underlag för policy och beslut
9. bedömning av framtida forskningsstrategier och -prioriteringar

De tre första funktionerna handlar om tillförsel av mänskliga och ekonomiska resurser till forskningssystemet, de fyra följande rör forskningens produkter och de två sista avser framtida strategier. Några av utvärderingsfunktionerna ingår i de arbetsuppgifter som följer av att vara anställd vid ett universitet (bedömningar av avhandlingar och sökande till universitetspositioner), andra utförs som arvoderade bisysslor (utvärderingsarbete för andra organisationer), medan åter andra sker som oavlönat arbete inom det egna ämnesområdet (bedömning av tidskriftsartiklar).

Förväntan att akademisk personal ska medverka i dessa uppgifter är inbäddad i normer, sällan i formella regler eller förordningar. Rollen som utvärderare är visserligen frivillig, men den erbjuder en möjlighet att utöva akademisk makt och är viktig för fördelningen av prestige och trovärdighet inom forskarsamhället.

Det finns en omfattande litteratur om *peer review*, men den behandlar främst granskning av ansökningar om forskningsbidrag och av manuskript som sänts in till vetenskapliga tidskrifter. Studier av övriga bedömarroller är mer begränsad. Nedan presenteras de viktigaste aspekterna av de nio olika rollerna.

1. *Examinator*. I de flesta länder svarar doktorander för en betydande del av den totala universitetsforskningen, och en viktig bedömarfunktion är att granska doktorsavhandlingar. Dels är det ett sätt att certifiera disputeranden som färdigutbildad forskare, dels identifieras lovande forskartalanger. För opponenter och ledamöterna av betygskommittén är avhandlingsgranskning dessutom ett viktigt sätt att hålla sig uppdaterad om ny forskning och unga löftesrika forskare. En undersökning av avhandlingsgranskning tyder på att opponenter och examinatorer ikläder sig en roll som liknar läsare av vilken ny text som helst², men trots det är avhandlingsgranskning en av de mer krävande utvärderingsuppgifterna.

Avhandlingar är i allmänhet långa och en grundlig genomgång förväntas, ofta med både muntlig och skriftlig redovisning.

2. *Personalväljare*. I uppgiften att bedöma sökande till akademiska tjänster ingår exempelvis att avgöra huruvida en sökande uppfyller kraven för att antas som docent eller utses till professor. I denna utvärderarroll kan forskarna påverka en viss institutions framtida vetenskapliga profil, liksom könsfördelningen bland dess personal.³ Detta är processer som kan orsaka konflikter.⁴

3. *Fördelare av forskningsbidrag*. En tredje viktig uppgift är att delta i kommittéer som inrättats för att dela ut forskningsbidrag eller att vara enskild *referee* av ansökningar om sådana bidrag. Eftersom en allt större andel av resurserna för forskning fördelas via forskningsråd, stiftelser och andra forskningsstödjande organisationer där *peer review* ligger till grund för tilldelningen, har betydelsen av denna funktion ökat. Baksidan av detta är den ökande andelen avslag, vilket är frustrerande både för de sökande och för bedömargrupperna. Som en följd av detta har användningen av kollegial granskning för fördelning av forskningsbidrag ifrågasatts, och det har länge funnits förslag om att ersätta den, även av skäl som inte har att göra med att bedömningen förbrukar värdefull forskningstid.⁵ Dels är det svårt, och inte alltid meningsfullt, att bedöma forskning som ännu inte utförts. Dels kan partiska bedömningar ha omfattande inverkan på vilken typ av forskning som finansieras och vilken som blir utan pengar. Om endast vissa tankeskolor är representerade i en beredningsgrupp, kan följderna bli att andra inriktningar missgynnas.⁶ Att nå oväldiga avvägningar i mångdisciplinära bedömningskommittéer är ytterligare en utmaning.⁷ Bedömning av projektansökningar med *peer review* är sålunda en omtvistad form och utgör den bedömningsuppgift som lättast skulle kunna ersättas, exempelvis med direkt tilldelning till institutioner utifrån tidigare resultat.

4. *Referee och redaktör*. Medan en *referee* har till uppgift att bedöma kvalitet och relevans hos manuskript som sänts till en tidskrift för publicering, fattar redaktören det slutliga beslutet grundat på råd från normalt två eller fler personer.⁸ Ett närliggande men mindre formaliserat uppdrag är att bedöma *abstracts* och *papers* till vetenskapliga konferenser. Det finns en omfattande litteratur om *peer review* för publicering och även särskilda konferenser på området.⁹ Partiskhet är en central fråga i den diskussionen.

5. *Recensent*. Den här funktionen innefattar dels bokrecensioner, dels översiktsartiklar om det aktuella kunskapsläget inom ett avgränsat forsk-

ningsfält. Recensenterna fördelar beröm och klander till de berörda forskarna, och de kan också markera sin åsikt genom att negligera artiklar, böcker och andra vetenskapliga bidrag.

6. *Prisutdelare*. Att dela ut vetenskapliga pris och andra utmärkelser betraktas som prestigefyllda uppdrag och kan ta mycket tid. Dessutom kan sådant utvärderingsarbete ge möjlighet att utöva makt och delta på hög nivå i vetenskapspolitiken, vilket framgår bland annat av studier av arbetet bakom de vetenskapliga Nobelprisen.¹⁰

7. *Utvärderare av forskningsorganisationer*. Funktionen att utvärdera forskargrupper, institutioner, program, institut och till och med hela universitet har ökat kraftigt i betydelse under det senaste kvartsseket som svar på "den utvärderande statens" behov.¹¹ Dessa avser inte bara bedömningar av kvalitet, utan kan även omfatta konkreta rekommendationer om hur forskningen kan förbättras och hur den utvärderade enheten ska organiseras. Det finns stora skillnader mellan länder vad avser både dessa utvärderingar syfte och hur de utförs.¹²

8. *Policyrådgivare*. Ytterligare en viktig uppgift är att värdera forskning som underlag för politiska beslut och lagstiftning. Exempelvis utförs omfattande översiktsarbete av forskare anställda på myndigheter inom områden som hälsa och miljö. Vetenskapliga kommittéer av skilda slag betjänar också offentliga organ med denna typ av sammanställningar, liksom enskilda forskare som bistår internationella organisationer av typen FN:s klimatpanel IPCC. Kombinationen *peer review* och praktisk politik medför dock ett antal utmaningar, till exempel möjlig oenighet bland forskare liksom eventuell partiskhet.¹³

9. *Framtidsbedömare*. Under 1980- och 90-talen uppmuntrades statliga forskningsfinansiärer av sina länders regeringar att bli mer strategiska i sin finansieringspolitik, bland annat genom tillskapandet av storskaliga FoU-program riktade mot särskilt lovande nya forskningsområden. De brittiska vetenskapspolicyforskarna John Irvine och Ben R. Martin myntade termen *foresight activities* för de mekanismer och förfaranden som användes för att identifiera grundforskningsområden med stor framtidspotential.¹⁴

Sammantaget kan det verka som om forskare är ständigt sysselsatta med att utvärdera eller bli utvärderade: de är del av en utvärderingsspiral där samma forskning granskas upprepade gånger. Ett projekt kan gå igenom sin första utvärdering (1) när forskarna ansöker om (ett eller flera) ekonomiska bidrag, och senare när (2) *abstracts* till konferenser,



(3) tidskriftsartiklar, eventuella (4) doktorsavhandlingar och (5) bokmanuskript från projektet lämnas in. Dessutom blir (6) publikationer från projektet föremål för granskning när författarna söker akademiska befattningar, när de är (7) nominerade till pris och andra utmärkelser, och när (8) deras institutioner eller (9) forskningsorganen som finansierade deras projekt utvärderas. En eller flera (10) recensioner av böcker från projektet kan publiceras i vetenskapliga tidskrifter, och publikationer från projektet kan komma att (11) utvärderas i översiktsartiklar. Om projektet ger resultat av politisk eller ekonomisk betydelse eller bidrar med intressanta öppningar för ny forskning, kan det också bli en del av (12) politiska beslutsprocesser och strategiska framtidsstudier. Och slutligen återförs resultaten från projektet till steg 1, nämligen när forskarnas tidigare prestationer bedöms sedan de lämnat in nya ansökningar om forskningsbidrag – utvärderingsspiralen fortsätter. Till detta kommer att löpande utvärdering är en integrerad del av forskningsprocessen. Förutom all formell *peer review* pågår därför en kontinuerlig granskning av ansatser, metoder och resultat.

Det finns rader av sätt på vilka *peer review* påverkar forskningen och forskarsamhället, och som en följd härav kan rollen som utvärderare i flera skilda avseenden hamna i konflikt med rollen som forskare.

Motsättning 1: Tid för forskning kontra tid för utvärdering

Den mest uppenbara spänningen mellan utvärderarrollen och forskarrollen är tidskonflikten. Ju mer tid forskare spenderar på utvärdering, desto mindre utrymme finns för forskningsarbete.

Inte så mycket är känt om hur mycket tid som läggs på formella utvärderingsuppgifter. I en studie som vi genomförde 2001 bland alla fast vetenskapligt anställda vid norska universitet undersöktes fyra av de nio utvärderarfunktionerna: granskning av doktorsavhandlingar, bedömning av sökande till lediga tjänster och befordran till högre rang, granskning av manuskript för publicering samt bedömning av forskningsorganisationer.¹⁵ Sammanlagt medverkade cirka 80 procent av den akademiska personalen i åtminstone en av dessa fyra utvärderarroller under ett år, och dessa forskare använde i genomsnitt nära 17 arbetsdagar åt uppgifterna. Den uppgift som de flesta var engagerade i var manuskriptbedömning. Utvärderarrollen tog mest tid för professorer (19 arbetsdagar). Eftersom undersökningen inte innehöll samtliga nio utvärderarroller, är det totala antalet dagar högre än de observerade talen. Om vi lägger till några dagar för granskning av forskningsansökningar till siffrorna ovan, blir resultatet för en genomsnittlig forskare – som tillhör de 80 procent som över huvud taget tog på sig utvärderingsuppgifter – omkring 20 arbetsdagar per år på utvärderingsarbete (runt 25 dagar för professorer).¹⁶ Detta motsvarar fyra till fem arbetsveckor.

Frågan om huruvida utvärderarrollerna har förändrats över tid kan bara besvaras delvis av enkätdata från vår studie. Andelen anställda som deltog i granskning av doktorsavhandlingar och i urval av personal ökade från 46 till 58 procent mellan 1991 och 2000. Det genomsnittliga antalet dagar som användes för dessa funktioner var detsamma vid båda tillfällena, vilket indikerar att den totala tiden för dessa båda syften ökade under 1990-talet. Även för övriga funktioner finns det anledning att tro att det totala trycket på personalen har ökat. Insatser som *referee* har knappast minskat de senaste årtiondena, bland annat beroende på att en allt större del av resurserna för universitetsforskning fördelas efter ansökan, vilket får till följd att allt fler forskare engageras i bedömningsarbetet. Men den granskarfunktion som troligen ökat sin tidsåtgång mest är utvärderingen av forskningsorganisationer. Detta beror på införandet av *new public management* med större tyngd på bedömningar i efterhand.

En annan förändring som framgår av enkätdata är en mycket mer internationell profil på rollen som utvärderare. År 1991 deltog 9 procent av den akademiska personalen vid norska universitet i utvärderingsarbete utomlands, vilket ökade till 22 procent år 2000, och till 38 procent år 2012.¹⁷ Sammanfattningsvis finns det alltså all anledning att förmoda att utvärderarrollen har ökat i både betydelse och tidsåtgång.

Tidskonflikten innebär rimligen att de mest efterfrågade forskarna omöjligen kan ta på sig alla utvärderingsuppgifter som de inbjuds till. De begränsar då sina insatser till vad de uppfattar som de viktigaste uppdragen. Det troliga resultatet är att de högst rankade akademikerna kommer att ta på sig de mest prestigefyllda och maktpåliggande utvärderingsuppgifterna¹⁸, vilket lämnar de mindre prestigefyllda uppdragen till mer juniora forskare.

Motsättning 2:

Expertis kontra opartiskhet

Jäv utgör en allvarlig spänning mellan utvärderarrollen och forskarrollen. De forskare som är mest kompetenta att utvärdera viss forskning är ofta nära kolleger eller konkurrenter till dem som står bakom forskningen¹⁹, men nära kolleger diskvalificeras enligt jävsregler.

Förekomsten av jäv kan vara argument både för försiktighetsåtgärder i syfte att säkerställa vetenskapens självständighet och för att tillämpa något alternativ till kollegial granskning, det vill säga få forskarsamhället att avstå från en del av sitt självstyre i syfte att undvika att "sätta bocken till trädgårdsmästare". Ett exempel: om en forskare med kommersiella intressen i läkemedelsindustrin ombeds granska artikelmanuskript som rör effekter av läkemedel skulle det äventyra vetenskapens autonomi och trovärdighet. Det kan därför hävdas att endast "rena" akademiker bör utföra sådana uppgifter. Å andra sidan, om en forskare deltar i utformningen av en programutlysning hos en forskningsfinansierare kan det diskvalificera personen och dennes medarbetare från att ansöka om bidrag från programmet. Sådana rättighetsförluster står i konflikt med forskares vetenskapliga intressen och kan göra dem mer tveksamma till att ta på sig sådana uppdrag.

Dessa motsättningar kan hanteras genom att begränsa självstyret för enskilda forskningsområden eller minska användningen av *peer review*. Exempel på det förstnämnda är att använda utländska granskare eller

brett sammansatta paneler utan några egentliga specialister inom det aktuella forskningsområdet. Då förlorar inhemska forskare möjligheten att fördela resurser inom det egna forskningsområdet. Ett exempel på det senare fallet är att fördela mer resurser genom kanaler som inte kräver *peer review*, exempelvis att överföra offentliga medel direkt till universiteten baserat på resultatindikatorer eller administrativa beslut, samt genom uppdragsforskning och anbudstävlingar.

Motsättning 3: Neutrala bedömare kontra att främja vissa forskningsintressen

Det finns också spänningar förknippade med dubbla förväntningar avseende utvärderares neutralitet. Å ena sidan förväntas utvärderarna vara neutrala bedömare som utför en opartisk och grundlig granskning. Å andra sidan förväntar sig utvärderarna att kunna påverka vad som anses vara god forskning och hur viktiga resurser inom deras område fördelas. Dessa dubbla förväntningar lämnar ett förhandlingsbart utrymme för kraven på, och betydelsen av, "en neutral bedömare". En sådan kan uppenbarligen inte driva personliga intressen, men hur är det med omstridda uppfattningar grundade i utvärderarnas forskningsfält, forskningsintressen eller "tankeskola"? I vissa avseenden är föreställningen om en vetenskapligt neutral utvärderare meningslös. Forskare har olika vetenskaplig bakgrund och skiljaktiga synpunkter på exempelvis metod- och teorival, och skillnaderna i bedömningar till följd av sådant betraktas rimligen som legitima avvikelser, inte som *bias*.²⁰ Med andra ord förväntas utvärderare inte vara vetenskapligt neutrala, utan de ska avge bedömningar utifrån sitt vetenskapliga gottfinnande. De motsättningar som uppstår har sin grund i att det inte finns några klara gränser för vad som är ett vetenskapligt omdöme, det vill säga ingen tydlig gräns mellan vad som är legitimt för utvärderare att försöka påverka och vad de bör avstå från. Till exempel kan det vara svårt att skilja främjandet av vetenskapligt lämpliga metoder och forskningsperspektiv från bedömningar som stöder utvärderarens egna forskningsintressen.

Spänningar mellan förväntningar om en neutral bedömare och om att kunna påverka utfallet kan vara särskilt störande när det inte finns några tydliga normer för val av utvärderare och när processerna saknar transparens. En del av litteraturen om examination av doktorsavhandlingar

behandlar just bristen på nationella riktlinjer för doktorsexamen²¹ och ger exempel på förfaringssätt som gör det möjligt för störningar av olika slag. Till exempel kan bristen på standarder och öppenhet om hur examinatorer väljs medföra skadliga maktrelationer mellan handledare och examinatorer.²²

Motsättning 4: Avvikande bedömningar kontra enhålliga slutsatser

I sin roll som utvärderare gör forskare olika bedömningar och prioriteringar. Samtidigt syftar *peer review*-processer till att nå enstämmiga slutsatser om fördelning av knappa resurser och prestige.

Det finns två huvudsakliga sätt att hantera dessa motsättningar, dels genom att de sakkunniga granskarna möts personligen och diskuterar fram en kompromiss, dels genom att icke-experter fattar beslut utifrån *peer review*-bedömningar inhämtade från individuella specialister på det aktuella området. Tidskriftsredaktörers beslut om publicering av insända manuskript är ett exempel på det sistnämnda. Ett exempel på det första sättet är bedömningsarbete i beredningsgrupper. Studier visar att en viktig mekanism för att nå enighet i sådana grupper är att överlåta avgörande inflytande till den i gruppen som är mest kompetent på det område som den aktuella ansökningen avser.²³ Även det sätt på vilket processen är organiserad kan påverka utfallet.²⁴

Vid bedömningen av manuskript, bidragsansökningar och de flesta andra *peer review*-tillämpningar i den vetenskapliga världen är avvikande bedömningar inte bara oproblematiske, utan till och med en viktig del av dynamiken inom vetenskapen. Däremot, vid bedömningen av forskning för användning i praktisk politik orsakar skiljaktiga bedömningar ofta problem. Insatserna är höga, och granskningsprocesser som lämnar utrymme för partiskhet eller slumpmässiga resultat tolereras inte. Potentialen för *bias* är mer uttalad när det gäller tillämpningar av forskning än i inomvetenskaplig *peer review*, eftersom granskarnas forskningsperspektiv kan påverka hur risker bedöms.²⁵

Motsättning 5:

Peer review kontra kvantitativa indikatorer

Vid sidan av den utvärderingsspiral som beskrevs ovan och allmänt ökat bruk av *peer review*, ökar också användningen av kvantitativa indikatorer av skilda slag. Dessa kan verka övertygande eftersom de erbjuder enkla svar och är billiga.²⁶ Även när de kvantitativa indikatorerna är aggregerade slutsatser av tidigare *peer review* (det vill säga bibliometriska mått), finns motsättningar mellan dessa indikatorer och kollegial granskning, och de kan ge olika slutsatser. Studier som jämför utfallet av *peer review* och bibliometriska mått har funnit ett visst samband, men i många fall är korrelationen mycket lägre än vad man skulle förvänta sig.²⁷ Även om båda i sista hand bygger på kollegial granskning, direkt eller indirekt, är de baserade på mycket olika logiker. *Peer review* grundas på närkunskap om forskningsfronten inom det område som bedöms²⁸, medan bibliometri räknar artiklar och citeringar i indexerade tidskrifter. Bibliometri kan ge viktig information om en forskares tidigare verksamhet, nätverk och samarbetsmönster, men har ifrågasatts som belägg på vetenskaplig kvalitet. I vissa fall kombineras kollegial granskning och bibliometri, till exempel vid utvärderingar av större forskningsprogram, vilket kan spara tid. I andra fall ersätts *peer review* av kvantitativa mått. När forskningsfinansierande myndigheter försöker öka forskares ansvarstagande genom att erbjuda dem produktivitetsincitament, ersätter de gärna fasta budgetramar för universiteten med finansiering baserad på kvantitativa resultatindikatorer²⁹, snarare än genom forskningsbidrag baserade på *peer review*.

Motsättning 6:

Autonomi kontra ansvar

Även om kollegial granskning är en viktig mekanism för att utöva kontroll i den vetenskapliga världen, kan den motverka ansvarstagande gentemot övriga samhället. *Peer review* innebär en "professionell självreglering som ger forskarna ett mått av frihet från granskning av allmänheten".³⁰ Vetenskaplig kvalitet är en av de minst politiserade aspekterna av vetenskap. Det vetenskapliga samfundet är fritt att självt definiera och bedöma kvaliteten på vetenskaplig forskning. I tider då den vetenskapliga autonomin ifrågasätts³¹ och det finns föga överlappning mellan offentliga och

akademiska forskningsintressen, kommer samhällets krav på att bestämma forskningsagendan sannolikt att stärka forskarnas strävan efter akademisk autonomi. Forskare åberopar *peer review* i sitt eget försvar³², och ett exempel på detta är de utvärderingar av forskningsprogram där bedömarna rekommenderar en ökad andel resurser till forskarinitierad grundforskning.³³

Det vetenskapliga samfundets förmåga att skydda sin självständighet är generellt stark. Behovet av *peer-expertis* för att bedöma den vetenskapliga kvaliteten är uppenbar, och forskare lyckas ofta definiera "kvalitet" som "vetenskaplig kvalitet". Utrymmet för icke-specialister att medverka i utvärdering av forskningen är tydligt begränsat. Det finns dock uppenbara skillnader mellan olika utvärderarroller. Medan uppgiften att bedöma doktorsavhandlingar inte kan utföras av lekmän, kan fördelning av forskningsbidrag göras genom exempelvis direkta anslag baserade på tidigare produktivitet.

Dessutom kan lekmanamedverkan vid bedömning av ansökningar öka motsättningar som rör den akademiska friheten. Beslut om forskningsbidrag baseras ofta på ytterligare aspekter vid sidan av vetenskaplig kvalitet, exempelvis huruvida projektet passar in i önskad fördelning av medel vad avser tema, geografiska och institutionella hänsyn, könsbalans, möjliga samhällsliga konsekvenser och etiska överväganden.³⁴ I allmänhet krävs inte *peer-kompetens* för bedömning av sådana frågor. Inte heller är det självklart vem som bäst kan bedöma samhällseffekterna av föreslagen forskning. Forskare kan hävda att ett projekts samhällsliga förtjänster är osäkrare än de vetenskapliga och att samhällseffekter dessutom är beroende av vetenskaplig framgång, och därmed prioritera vetenskaplig kvalitet framför möjlig samhällslig nytta. Dessutom, om ett forskningsfinansierande organ skulle organisera sin bedömningsprocess så att ickeexperter kan rösta ner experter, skulle det uppfattas som illegitimt och möta opposition från forskarsidan. Även när kollegial granskning definieras som en del av det system som har att tillse att offentliga medel fördelas ansvarsfullt, prioriteras inte ansvarsutkrävande och samhällsnytta av de utvärderande forskarna.³⁵

Rollen som utvärderare av forskning befinner sig under förändring. Under de senaste två decennierna har arbetet blivit alltmer internationellt, vilket bidrar till att både förbättra kvaliteten och undvika jäv. Öppen *peer review* är en annan trend som kan medverka till mer jämlika granskningsprocesser. Vissa tidskrifter har infört öppnare *peer review*-processer,

bland annat öppna inbjudningar att kommentera insända manuskript.³⁶ Nedan diskuterar vi en agenda för framtida studier av rollen som forskningsutvärderare utifrån den ökade globaliseringen och de motsättningar som analyserats ovan.

Val av utvärderare, tidsbrist och skiktningen av vetenskap

Utfallet av en *peer review* beror på vem som gör den. Ändå föreligger föga forskning om hur utvärderare väljs ut, liksom om tidsbrist och tillgången på lämplig expertis. Här finns viktiga forskningsfrågor: Vilka är de mest efterfrågade utvärderarna och varför, vilka är de mest attraktiva utvärderingsuppgifterna och hur mycket tid läggs i dag på olika utvärderingsuppgifter?

Urvalsprocesser och tidsbrist kan påverka arbetsfördelning och stratifiering inom vetenskapssamhället på olika sätt, bland annat den globala arbetsfördelningen. När den totala mängden utvärderingar ökar, kan tidsbrist komma att minska stratifieringen inom vetenskapen på så sätt att de mest önskade utvärderarna blir tvungna att överlåta vissa utvärderingsuppdrag till juniora medarbetare eller forskare från en bredare krets av länder. Införandet av öppen granskning (självurval av granskare) kan förstärka denna tendens. Trots detta finns anledning att anta att forskarna med högst rang kan behålla de politiskt viktigaste uppdragen. Forskningsfrågor här är bland annat hur mycket tid forskare är villiga att lägga på utvärderingsarbete och hur de väljer utvärderingsuppdrag. Exempelvis, i tider när den akademiska autonomin ifrågasätts kanske forskare är mer villiga att utföra utvärderingar för att säkerställa att den vetenskapliga kvaliteten inte komprometteras av andra hänsyn.

Autonomi, indikatorer och lekmän

Andra viktiga frågor är hot mot det vetenskapliga självstyret i form av användning av kvantitativa indikatorer och utvärderingar som inte kräver vetenskaplig kompetens. Låg bevilningsgrad hos forskningsfinansiärer liksom problem med partiskhet och jäv i *peer review* kan leda till åtgärder för att begränsa den tid som forskare ägnar åt att delta i kollegial bedömning och att skriva ansökningar, med följd att mer betoning läggs på kvantitativa indikatorer. Ökat utrymme för samhällelig relevans vid

bedömning av ansökningar och ökade önskemål om mätbara resultat kan leda till utvärderingar genomförda av icke-specialister.

En utvidgad användning av kvantitativa mått vid fördelningen av forskningsmedel kan minska betydelsen av *peer review*. Å andra sidan ökas den aggregerade betydelsen av kollegial granskning vid finansiering baserad på tidigare prestationer. Viktiga nyckeltal vid prestationsbaserad finansiering är nämligen antalet vetenskapligt granskade publikationer och framgång i att erhålla forskningsbidrag baserat på *peer review*. I flera länder utgör också *peer review* underlag för prestationsbaserad finansiering, såsom i det brittiska Research Excellence Framework.³⁷

Kvantitativa mått ger upphov till två uppsättningar av forskningsfrågor om forskare som utvärderare. För det första, i vilken mån är utvärderarna medvetna om den aggregerade betydelsen av sitt bedömningsarbete, i vilken utsträckning inverkar den på deras vilja att ta på sig utvärderingsuppdrag, och i vilken grad påverkar den hänsyn till jäv och användning av utländska granskare? Den andra gruppen av frågor rör forskares åsikter och önskemål rörande indikatorer och expertgranskning. Föredrar forskare att pengar fördelas med hänsyn till kvantitativa mått och incitamentssystem, eller ser de hellre *peer review* av bidragsansökningar? Båda tillvägagångssätten kritiseras ofta av forskare. Kollegial granskning anses vara konserverande såtillvida att metoden diskriminerar tvärvetenskaplig och banbrytande forskning; dessutom är den mycket tidskrävande. Eftersom prestationsbaserad finansiering baseras på aggregerade resultatet av sådan granskning, riskerar den att reproducera svagheter hos *peer review*. Kvantitativa mått kan dessutom producera dysfunktionella incitament, som att prioritera kvantitet före kvalitet. Det är fortfarande inte självklart vad som uppfattas vara det rättvisaste, det minst konservativa och det effektivaste sättet att fördela forskningsbidrag.

Kvantitativa mått antyder ett konkurrenssystem som bygger på rimligt förutsägbara kriterier, medan den ökande förekomsten av lekmän liksom förstärkt betoning på samhällsnytta vid vetenskapliga utvärderingar innebär en tydligare utmaning för den akademiska autonomin. Det finns även forskare som ifrågasätter traditionell *peer review* och föreslår mer öppenhet och bredare urvalskriterier.³⁸

Det finns många argument emot fullt akademiskt självstyre vad gäller att fördela forskningsresurser, liksom det finns argument för ökad betoning på externa hänsyn och lekmannainflytande vid sådana beslut.

Det brukar anses att offentliga medel för forskning ska tjäna offentliga intressen och allmänheten. För att legitimera stora offentliga utgifter för forskning krävs dessutom öppenhet för kritik utifrån och lyhördhet för allmänhetens behov och oro. Vi har också sett ökade krav på medborgarinflytande (demokratisering av vetenskapen), större tonvikt på nytta, ökade önskemål om att kunna utkräva ansvar, och utmaningar i samband med kontroversiella frågor som kännetecknas av stor osäkerhet.³⁹ Att inkludera samhällskriterier och lekmän i bedömningar kan innebära att forskningen bättre tjänar både allmänna behov och behoven hos utsatta grupper.⁴⁰ Icke-akademisk kompetens kan minska *insider bias*⁴¹ och leda till att en bredare uppsättning kriterier beaktas – sociala, ekonomiska, miljömässiga, hälsorelaterade och så vidare.

Eftersom bedömning av forskningsansökningar är ett nollsummespel som definierar forskningsfokus inom ett program eller forskningsfält, uppfattas denna uppgift sannolikt som viktig när det gäller att skydda den vetenskapliga autonomin. Hur införandet av lekmän som bedömare, liksom av bredare relevanskriterier, påverkar bedömningsarbetet och lyhördheten för allmänhetens behov kräver fördjupade studier. Dessutom bör både utvärderares och forskningsfinansierande myndigheters uppfattningar om akademisk autonomi studeras.

Texten är en uppdaterad och något förkortad version av L. Langfeldt & S. Kyvik, "Researchers as evaluators: Tasks, tensions and politics", *Higher Education*, 62 (2), 2011, s. 199–212.

*

1. J. M. Ziman, *Prometheus Bound: Science in a Dynamic Steady State*, Cambridge: Cambridge University Press, 1994, s. 103.

2. S. Johnston, "Examining the examiners: An analysis of examiners' reports on doctoral theses", *Studies in Higher Education*, 22(3), 1997, s. 333–347; s. 340.

3. *Tenure Denied: Cases of Sex Discrimination in Academia*, Washington, DC: American Association of University Women Education Foundation & American Association of University Women Legal Advocacy Fund, 2004.

4. J. C. Hearn & M. S. Anderson, "Conflict in academic departments: An analysis of disputes over faculty promotion and tenure", *Research in Higher Education*, 43(5), 2002, s. 503–529.

5. D. F. Horrobin, "Peer review: A philosophically faulty concept which is proving disastrous for science", *Behavioral and Brain Sciences*, 5(2), 1982, s. 217–218; R.

Roy, "Alternatives to review by peers: A contribution to the theory of scientific choice", *Minerva*, 22(3-4), 1984, s. 316-328.

6. G. D. L. Travis & H. M. Collins, "New light on old boys: Cognitive and institutional particularism in the peer review system", *Science, Technology & Human Values*, 16(3), 1991, s. 322-341.

7. M. Lamont, *How Professors Think: Inside the Curious World of Academic Judgment*, Cambridge, MA: Harvard University Press, 2009.

8. A. C. Weller, *Editorial Peer Review: Its Strengths and Weaknesses*, Medford, NJ: Information Today (ASIS&T Monograph Series), 2001.

9. S. J. Ceci & D. P. Peters, "Peer review: A study of reliability", *Change*, 14(6), 1982, s. 44-48; J. M. Campanario, "Peer review for journals as it stands today – part 1", *Science Communication*, 19(3), 1998, s. 181-211; J. M. Campanario, "Peer review for journals as it stands today – part 2", *Science Communication*, 19(4), 1998, s. 277-306; B. W. Speck, *Publication Peer Review: An Annotated Bibliography*, Westport & London: Greenwood Press, 1993; Weller 2001; C. J. Lee, C. R. Sugimoto, G. Zhang & B. Cronin, "Bias in peer review", *Journal of the American Society for Information Science and Technology*, 64(1), 2013, s. 2-17; K. C. Soh, "Peer review: Has it a future?", *European Journal of Higher Education*, 3(2), 2013, s. 129-139.

10. R. M. Friedman, *The Politics of Excellence: Behind the Nobel Prize in Science*, New York: Times Books, 2001.

11. G. Neave, "The evaluative state reconsidered", *European Journal of Education*, 33(3), 1998, s. 265-284.

12. H. F. Hansen, *Forskningsevaluering: Metoder, praksis og erfaringer*, København: Forsknings- og innovationsstyrelsen (Forskning: Analyse og evaluering 1), 2009.

13. S. Jasanoff, *The Fifth Branch: Science Advisers as Policymakers*, Cambridge, MA: Harvard University Press, 1990, s. 79 ff.

14. J. Irvine & B. Martin, *Foresight in Science: Picking the Winners*, London: Pinter, 1984.

15. S. Kyvik & L. Langfeldt, "Tidkrevende bedømmelsesarbeid", *Forskningsspolitikk*, 27(2), 2004.

16. Ibid., s. 15.

17. S. Kyvik & J. Wiers-Jenssen, "Internasjonalisering av norsk høyere utdanning", *Arbeidsnotat* 2/2014, Oslo: NIFU, 2014.

18. S. Cole, "The hierarchy of the sciences?", *American Journal of Sociology*, 89(1), 1983, s. 111-139; s. 138.

19. D. E. Chubin & E. J. Hackett, *Peerless Science: Peer Review and US Science Policy*, New York: State University of New York Press, 1990, s. 80.

20. L. Langfeldt, *Decision-making in Expert Panels Evaluating Research: Constraints, Processes and Bias*, Oslo: Unipub, 2002, tillgänglig på www.sv.uio.no/isv/forskning/aktuelt/arrangementer/disputaser/arkiv/fulltxt/langfeldt.pdf (hämtat 16 januari 2015), s. 67-69.

21. P. Tinkler & C. Jackson, "Examining the doctorate: Institutional policy and

- the PhD examination", *Studies in Higher Education*, 25(2), 2000, s. 167–180; L. Morley, D. Leonard & M. David, "Variations in vivas: Quality and equality in British PhD assessments", *Studies in Higher Education*, 27(3), 2002, s. 263–273.
22. Morley m.fl. 2002, s. 270–271.
 23. Lamont 2009, s. 117.
 24. L. Langfeldt, "The decision-making constraints and processes of grant peer review, and their effects on the review outcome", *Social Studies of Science*, 31(6), 2001, s. 820–841.
 25. Jasanoff 1990, s. 76–83.
 26. Ziman 1994, s. 103–105.
 27. D. W. Aksnes & R. E. Taxt, "Peer review and bibliometric indicators: A comparative study at a Norwegian university", *Research Evaluation*, 13(1), 2004, s. 33–41; L. Bornman & H. D. Daniel, "Reliability, fairness, and predictive validity of the peer review process for the selection of research fellowship recipients of the Boehringer Ingelheim fonds", i B. M. Kehm (red.), *Hochschule im Wandel: Die Universität als Forschungsgegenstand. Festschrift für Ulrich Teichler*, Frankfurt am Main: Campus, 2008, s. 365–376; P. van den Besselaar & L. Leydesdorf, "Past performance, peer review and project selection: A case study in the social and behavioral sciences", *Research Evaluation*, 18(4), 2009, s. 272–288.
 28. J. R. Ravetz, *Scientific Knowledge and its Social Problems*, Oxford: Clarendon Press, 1971, s. 274.
 29. S. Sörlin, "Funding diversity: Performance-based funding regimes as drivers of differentiation in higher education systems", *Higher Education Policy*, 20(4), 2007, s. 413–440.
 30. E. J. Hackett, "Peer review in science and science policy", i M. S. Frankel & J. Cave (red.), *Evaluating science and scientists*, Budapest: Central European University Press, 1997, s. 51–60; s. 57.
 31. M. Henkel, "Can academic autonomy survive in the knowledge society? A perspective from Britain", *Higher Education Research & Development*, 26(1), 2007, s. 87–99.
 32. Chubin & Hackett 1990, s. 5.
 33. Langfeldt 2002.
 34. Langfeldt 2001, s. 827–829.
 35. B. Van der Meulen, "Science policies as principal-agent games: Institutionalization and path dependency in the relation between government and science", *Research Policy*, 27(4), 1998, s. 397–414; s. 405 ff.
 36. U. Pöschl & T. Koop, "Interactive open access publishing and collaborative peer review for improved scientific communication and quality assurance", *Information Services & Use*, 28(2), 2008, s. 105–107.
 37. Tidigare känt som Research Assessment Exercise; se Research Excellence Framework, www.ref.ac.uk (senast kontrollerad 16 januari 2015).
 38. A. Scott, "Peer review and the relevance of science", *Futures*, 39(7), 2007, s. 827–845.

39. Ibid.; S. Jasanoff, "Technologies of humility: Citizen participation in governing science", *Minerva*, 41(3), 2003, s. 223–244.

40. E. Woodhouse & D. Sarewitz, "Science policies for reducing societal inequities", *Science and Public Policy*, 34(3), 2007, s. 139–150.

41. B. Martin, "Research grants: Problems and options", *Australian Universities' Review*, 43(2), 2000, s. 17–22, tillgänglig på www.bmartin.cc/pubs/00aur.html (hämtat 16 januari 2015).