Vedr. den Bibliometriske Forskningsindikator (BFI).

Den 30. juni, 2012

**Høringssvar fra faggruppe 68 (Forskningsanalyse og Videnskabsstudier) vedrørende udkast af 24. april 2012 til evaluering af BFI, udgivet af NIFU med Gunnar Sivertsen og Jesper Schneider som forfattere.**

I. Generelle kommentarer:

*1. Forskningens resultater er mere end blot publikationer*

Rapporten sammenfatter en positiv holdning til den fortsatte brug og udvikling af modellen, men primært baseret på holdninger fra dekanerne på de danske universiteter. Det første og vel vigtigste hovedspørgsmål, som Styrelsen for Videregående Uddannelser og Uddannel­ses­støtte VUS (tidligere FI) har formuleret til evalueringen, er om BFI har indfriet positive hhv. negative incitament forventninger.

Her har rapportens forfattere følt sig alt for bundet af opgavens formulering og kun få steder antydet, at BFI modellen og dens norske forbillede har nogle alvorlige indbyggede begrænsninger. Det drejer sig om at BFI modellen alene måler forskningsperformance og kvalitet i ud fra en snæver og ganske traditionel opfattelse af moderne videnskab.[[1]](#footnote--1) Rapporten nævner et enkelt sted den nye model for performancemåling som er undervejs i England som et resultat af mange utilsigtede bivirkninger af hvad man i England erfarede i årene fra 1986, hvor den første RAE blev gennemført. Når man i England, som har været foregangsland for budgetfordeling efter performance målt på artikler, nu er på vej til at anvende en anden og langt mere kompleks model, Research Excellence Framework, hvor både forskningens samfundsmæssige betydning samt forskningsmiljøernes kvalitet indgår med en væsentlig vægt, så er det beklageligt, at evalueringsrapporten ingen steder tager en diskussion om BFI modellens generelle begrænsninger op. Rapporten kunne med fordel have tematiseret disse begrænsninger langt tydeligere end tilfældet er, eksempelvis således, at forskningsaktiviteter som rapporter og andre formidlende publikationer, ofte i lokalt sprog, samarbejde med organisationer og virksomheder, netværksarbejde og meget mere synliggøres og ikke som det er tilfældet nu er defineret ude. Over tid kan begrænsningen i BFI få alvorlige konsekvenser ved at styre forskernes incitamenter væk fra en række aktiviteter, som alle der kender forskningsverdenen ved er uundværlige - og som for at gøre ironien helt tydelig - er aktiviteter som der forskningspolitisk har været fokus på de seneste år.

*2. Kontraintenderede effekter*

En anden negativ incitament funktion som kun berøres kort i rapporten er erfaringerne fra RAE i England med forskeres yderst "kreative" håndtering af performancemåling, hvor ’game playing’ og ’salami slicing’ i vidt omfang hørte til dagens orden. At den danske BFI model helt skulle kunne undgå sådanne utilsigtede effekter er usandsynligt. Det er derfor problematisk, at spørgsmål om utilsigtede incitamentstrukturer i modellen i så ringe omfang inddrages i evalueringen.

*3. Antallet af tidskrifter på højniveau (niveau 2).*

Vi har noteret, at evalueringsrapporten ukritisk videreformidler, at visse universitetsledere giver ud tryk for den opfattelse, at der er for mange tidskrifter på niveau 2.[[2]](#footnote-0) Vi mener ikke dette gælder generelt, og vi er i tvivl om det overhovedet gælder for nogen faggrupper. Vi mener tværtimod, at der er plads til for få tidskrifter som højniveau, og at der konstant har været problemer internt i faggrupperne med at nå til enighed om hvilke tidskrifter (ud af faggruppens samlede ‘ønskeliste’ for niveau 2-tidskrifter alene ud fra en vurdering af kvalitet, en liste som ofte repræsenterer indbyrdes ret forskellige og svært sammenlignelige videnskabelige specialer) kunne “få lov til” at stå som niveau 2. Niveaudelingen får her en delvist arbitrær karakter, når 20%-kriteriet skal overholdes. Dette har noget at gøre med faggruppestrukturen som vi kommenterer i næste punkt.

*4. Konfliktende kriterier for niveau-deling virker negativt for diversiteten af forskningsfelter*

Det har i faggruppernes niveaudeling for en lang række grupper været et genkommende problem, at de to nævnte kriterier for optagelse af et tidskrift på niveau 2 — det ene gående på videnskabelig kvalitet (tidskriftet skal anses for ubestridt internationalt førende, etc.), det andet gående på en ydre, kvantitativ grænse (max 20% af verdensproduktionen af artikler indenfor faggruppens samlede antal specialer må ikke overstiges) — ikke nødvendigvis stemmer overens, således at ses kun på kvalitet og dækning i forhold til videnskabelige specialer, vil max-grænsen evt. overskrides, og omvendt, hvis det ikke tillades faggruppen at sprænge grænsen, vil flere forskningsfelter indenfor en faggruppe mangle niveau 2 tidskrifter at publicere i. Faggruppeinddelingen med kun 67 grupper er her den usynlige, tredje parameter, som har en uheldig tendens til at stille en række videnskabelige specialer ringere.[[3]](#footnote-1)

*5. Det snævre kommissorium, metoden og det spinkle datagrundlag er fatalt for evalueringens troværdighed*

Det er et meget snævert kommissorium VUS har udstukket og kritisabelt, at man ikke har sørget for at sammensætte evalueringen som et bredere panel med andre kompetencer end bibliometri, når det handler om at vurdere effekterne på forskningen helt alment, og ikke alene mht. de bibliometriske parametre. Et tværdisciplinært og uafhængigt evalueringspanel havde været ønskeligt.

Man kan undre sig over at en evaluering af et sådant redskab som BFI helt går uden om at inddrage holdninger fra de direkte berørte brugere, dvs. forskerne på institutionerne. Dette bliver sat i relief, når der på s. 37 er et helt afsnit, som hedder “Mottagelse af indikatoren” (afsnit 4.5). Her refereres til “forskernes” holdninger til BFI, men grundlaget for dette er udelukkende interviews med dekaner og institutledere – ikke med et udvalg af forskere. Eller når der på side 34 står, at forskerne på CBS er mere interesserede i internationale lister over toptidsskrifter. Det var der for få år siden en heftig debat om og der var langt fra enighed blandt forskerne på CBS om sådanne listers betydning. Det fortsætter på de følgende sider, hvor rapporten evaluerer virkninger af indikatoren i de faglige miljøer – alene på baggrund af udsagn fra dekaner og andre ledere. Det svarer til at man eksempelvis i en evaluering af MUS samtalesystemet alene havde interviewet ledere for at evaluere virkningen!

Det skal dog siges, at det for så vidt er et problem, som rapportens forfattere ikke har det fulde ansvar for, da dette også ligger hos opdraggiveren i Bredgade, VUS/FI, som besluttede at evalueringen udelukkende skulle bygge på udsagn fra ledelsen på universiteterne. Det er måske i overensstemmelse med VUS’ opfattelsen af enstrenget ledelse, at forestille sig at de altvidende dekaner og institutledere kan give dækkende svar for hvordan deres forskere ser på BFI, men det er en udeladelse der ikke imponerer professionelle evaluatorer. Ved helt at udelade udsagn om BFI fra de faglige udvalg og fra menige forskere er der på forhånd lagt en skævhed ind i resultatet, og den gennemgående positive holdning til BFI modellen som evalueringen refererer har af disse grunde ingen høj gyldighed i den virkelige verden. I forlængelse af kritikken af kommissoriets snæverhed og VUS’ styring af metodevalg, noterer vi os, at der andetsteds er rejst kritik af Gunnar Sivertsens roller som både løbende konsulent for ministeriet, medlem af Fagligt Udvalg og ansat hos den ‘kunde’ (NIFU) som modtog evalueringsopdraget; her gør ministeriet sig unødigt sårbar overfor beskyldninger om at have evalueret sig selv.[[4]](#footnote-2)

II: Specifikke kommentarer

*6. Datakvaliteten og den lille tidshorisont*

Datakvaliteten trænger til at blive grundigere og mere kritisk belyst, især for de danske publiceringskanaler. De tal, evalueringen benytter, og som ifølge rapporten er analyseret af Jesper Schneider (hvor to år er slået sammen idet 2008 data er mangelfulde) kan der næppe siges meget ud fra. Det er ret evident, at publikationsdata fra 2009 reelt er forskning, der kan være iværksat et eller flere år forinden, og på et tidspunkt hvor BFI ikke fyldte meget i bevidstheden hos forskerne.[[5]](#footnote-3) De første årgange af data fra BFI bliver først interessante hvis man fx har en 10-årige periode; så 2009 er nærmere et udtryk for hvordan publicerings­adfærden var før BFI. Det kræver naturligvis en helt anden tidshorisont hvis en analyse skal laves på denne måde. Og så belyser en sådan analyse jo kun makroforhold.

*7. Evalueringens retorik*

Som nævnt er det invaliderende for evalueringen, at der kun er talt med ledelsen, og ikke er undersøgt hverken et udsnit af menige forskeres holdninger, eller de faglige rådvidere over VUS, dvs. faggruppemedlemmer, som har et mere grundigt kendskab til BFI og dens problemer. Selvom dette fremgår af rapporten, virker det som om den retorisk søger at dække over denne mangel. En ret stor del af rapporten er viet referater fra interview med de enkelte universiteters ledelser. I fremstillingen sker en glidning *fra* – i metodeafsnittene - en eksplicitering af, at der kun tales med de ansatte ledere, ikke forskerne selv, *til* at læseren, i gennemgangen af disse samtaler, foregøgles at få formidlet almene opfattelser på universiteterne af BFI. Et eksempel: I opsummeringen af afsnit 4.5 (s. 35) hedder det

Kjennskapet til og diskusjonen om BFI blant forskerne varier med i hvilken grad universitetene selv benytter indikatoren lokalt og med hvilket hovedområde man tilhører. Diskusjonen av indikatoren har vært størst i HUM og SAM, hvor oppfatningen har endret seg fra å være overveiende negativ til mer varierende eller positiv. I NAT/TEK og SUND har interessen for BFI blant forskerne vært liten eller fraværende, særlig ved universiteter uten lokal bruk av indikatoren. En viss diskusjon har det likevel vært om nivåinndelingen.

Læseren bibringes det indtryk, at de humanistiske forskere nu er mere positive i holdningen til BFI end førhen. Dette er qua datamaterialet helt udokumenteret, og der er tværtimod mange tegn på, at ledelsen (dekaner osv.) og forskerne har helt forskelligt perspektiv på BFI.

Tilsvarende omtaler rapporten disse dekaner, prodekaner, direktører m.v. som “forskningsledere”, hvad der igen er misvisende for ikke at sige manipulerende, idet en forskningsleder ved at dansk universitet typisk er en lektor eller professor, der selv forsker og leder en gruppe forskere indenfor et bestemt forskningsfelt.

Rapporten hævder (bl.a. s. 10) at “indikatoren har bidratt til å sette mer fokus på forskning som et felles ansvar i fagmiljøene”, men dette er dels en uklar påstand, dels ikke dokumenteret tilstrækkeligt.

En anden men ikke uvæsentlig retorisk effekt er samforfatterskabet, annonceret på forsiden, bestående af Gunnar Sivertsen og Jesper Schneider. Sidstnævnte er kendt i Danmark som en dygtig ekspert og forsker i bibliometri, og Schneider har tidligere i år har været med til at udgive en artikel[[6]](#footnote-4) om BFI, som leverer en samlet og selvstændig analyse, er særdeles kritisk, og som på flere punkter går imod evalueringen. Det virker derfor besynderligt at han står som medforfatter til hele rapporten, hvis hans rolle alene (jf. kommissoriet) har været at analysere enkelte data.

*8. BFI som internt fordelingsinstrument.*

Vi finder, at evalueringsrapporten mere kritisk burde forholde sig til, at flere universiteter bruger den bibliometriske forskningsindikator internt som ledelsesinstrument (fx Syddansk Universitet og Aalborg Universitet bruger indikatorens tal til centralt at allokere midler ud til fakulteter og i enkelte tilfælde videre til institutterne).[[7]](#footnote-5) Dette er i modstrid med en tidligere anbefaling fra Fagligt Udvalg, som i brev[[8]](#footnote-6) af 29/9-2009 skrev til faggrupperne:

“Fagligt Udvalg opfordrer alle faggrupperne til at medvirke til at advare mod for nærværende at bruge de bibliometriske data til andre formål. Det er Fagligt Udvalgs anbefaling, at den bibliometriske forskningsindikator ikke bruges internt på institutionerne til måling af forskningsproduktionen for mindre enheder, institutter og lignende.”

Hvis en stiltiende tilslutning til en mere omfattende brug af BFI internt på lavere niveauer har fundet sted hos Fagligt Udvalg og i ministeriet, er dette ikke kommunikeret videre til universiteterne, men vi tvivler på, at der foreligger en ny målsætning fra ministeriet/VUS herom. Dette aktualiserer på sin side et af de hovedkritikpunkter, som analysen af Schneider og Aagaard (se note 6) gør gældende: En manglende afklaring med målsætningen med indikatoren.

M.v.h.

faggruppe 68

Claus Emmeche, Finn Hansson, Tinne Hoff Kjeldsen, Helge Kragh og Peter Øhrstrøm

1. Se kritikken heraf i Ernø-Kjølhede, E., & Hansson, F. (2011). Measuring research performance during a changing relationship between science and society. *Research Evaluation*, *20*(2), 130-142. doi:10.3152/095820211X12941371876544 [↑](#footnote-ref--1)
2. I følge-emailen til faggrupperne fra VUS af 12/6-2012 bliver det til at ” Universiteterne rejser dog også flere kritikpunkter, herunder at der er for mange tidsskrifter på det høje niveau to.” [↑](#footnote-ref-0)
3. problemet er yderligere belyst i notatet "Mareridt, damage control, eller forskningsrelevante kvalitetskriterier? - Notat om faggruppernes forbehold overfor den bibliometriske forskningsindikator efter niveaudelingsprocessen og indtastning af tidskriftlisterne pr. 15/9-2009" (http://faggruppe68.pbworks.com/f/Faggruppernes-forbehold-v2.doc). [↑](#footnote-ref-1)
4. Den 18/6-2012 i dagbladet *Information,* "[Forskere: Inhabil konsulent har blinde pletter](http://www.information.dk/303765)" og "[Evalueringsrapport kritiseres for at være politisk bestillingsarbejde](http://www.information.dk/303764)"; *FORSKERforum* ([her](http://www.forskeren.dk/?p=2319)); d. 19/9 lederen "Selvevaluering" ([her](http://www.information.dk/303858)); den 20/6 svarer direktøren for NIFU på kritikken på bloggen *Forskningsfrihed?* ([her](http://professorvaelde.blogspot.dk/2012/06/nifu-direktr-svarer-pa-kritik-af.html)).  [↑](#footnote-ref-2)
5. Rapporten er selv inde på dette, jævnfør eksempelvis s. 47: ”Konsekvensen er at det nok er for tidlig å besvare det ene av denne evalueringens hovedspørsmål: *I hvilket omfang har indførelsen af den bibliometriske forskningsindikator indfriet de positive incitamentsvirkninger, og i hvilket omfang har indikatoren haft negative incitamentsvirkninger?*” [↑](#footnote-ref-3)
6. Aalborg Universitetsforlag udgav d. 10. april 2012 bogen *Dansk Forskningspolitik Efter Årtusindskiftet* (red. af Kaare Aagaard & Niels Mejlgaard) med kapitlet "Stor ståhej for ingenting. Den danske bibliometriske indikator" af Jesper W. Schneider & Kaare Aagaard (s. 229-260). [↑](#footnote-ref-4)
7. Også ved KU og RUC er der ifølge rapporten (s. 33) “enkelte enheter på lavere nivåer som selv har valgt å videreutvikle data fra BFI til lokal ledelsesinformasjon, til sammenligninger med andre universiteter på samme fagområde, til økonomiske omfordelinger, som informasjonsgrunnlag for medarbeidersamtaler og i noen få tilfeller som et moment ved tildeling av individuelle forskningsmidler eller lønnsøkninger.” [↑](#footnote-ref-5)
8. “Høring om den samlede niveaudeling” fra FI (v. K. K. Christensen, sagsnr. 07-020539, dok.nr. 1057384), hvor det også hed: ”Der er således mange grunde til at være yderst varsom med at anvende de bibliometriske data til andre formål end det, de er udarbejdet til, nemlig at fordele basisbevillinger til universiteterne. Til det formål er indikatoren acceptabel i sin nuværende form – også fordi der i første omgang kun fordeles meget beskedne beløb.” [↑](#footnote-ref-6)