

af Seniorforsker cand. polit. Kamma Langberg ph.d.
& Forskningsleder dr.merc. Peter S. Mortensen

Benchmarking på offentlige forskningsinstitutioner – et ledelsesværktøj?

En kommentar til en kommentar...

I sidste nummer af *Dansk Sociologi* kom der en kritik af en rapport, der er udsendt af Analyseinstitut for Forskning om benchmarking. Den kommentar vil vi hermed gerne kommentere. Kritikken har lighedspunkter med den kritik, som er kommet frem fra anden side og den er dermed et eksempel på en fejllæsning, der kan tolkes på flere måder. Vi tolker umiddelbart kritikken på to niveauer:

for det første anvender rapporten en metode, DEA-metoden, der nok er svær tilgængelig for ikke-økonomer. Vi har derfor lavet en række forskellige grafiske fremstillinger af teknikken, der skulle kunne gøre den mere forståelig¹.

for det andet er der tilsyneladende forskere og andre, der ikke ønsker, at ledelsesforholdene på de danske universiteter skal diskuteres på nye betingelser

Når det er sagt, må det også siges, at vi har mødt helt andre holdninger fra mange andre: mange forskere kritiserer ledelsesforholdene og der er miljøer, der har inviteret os til et samarbejde om netop *Benchmarking*.

Offentlig forskningsledelse - og ledelsesværktøjer

I begyndelsen af 1990'erne var der to overordnede diskussioner om forskeres vilkår: den ene pegede på, at mange forskere inden for sektorforskningen ikke havde tilstrækkeligt med råderum til forskning og den anden at der var problemer med ledelsen på universiteterne. Der blev lavet en ledelsesreform på universiteterne, hvorefter lederne fik en større ledelseskompetence, og der blev lavet en ny sektorforskningslov. Det indebar bl.a., at sektorforskningen systematisk har fået implementeret en stillingsstruktur², der ligner universiteternes så meget, at det

skulle være muligt for forskere at bevæge sig mellem universitetssektoren og sektorforskningen.

Kritikken af universiteternes ledelsesstruktur er imidlertid ikke forsvundet. Den interne kritik er gået i flere retninger: nogle steder har der været kritik af, at ledelsen ikke mere var 'demokratisk' (= at alle medarbejdere ikke blev indkaldt jævnlige til møder for derigennem at kunne øve indflydelse); andre steder er kritikken, at ledelsen ikke tager ledelsesansvaret på sig. Den eksterne kritik har været koncentreret på to områder: dels en meget konkret kritik fra erhvervslivet om, at det er vanskeligt at indgå samarbejdsaftaler med universiteterne og dels en mere diffus kritik gående på, at ledelsesformerne er forældede. Det er nødvendigt at stille krav til universiteterne, så der bliver 'value for money'. Denne sidste diskussion ligner den diskussion, der generelt set har været i forhold til de offentlige arbejdspladser.

Kravet om forskningsledelse bliver ofte mødt med, at det er vanskeligt at vurdere forskning: forskere kan arbejde med noget i mange år, der pludseligt bliver brug for, og som pludselig er betydningsfuldt. Desuden kan forskere uden for det konkrete felt ikke vurdere forskningen, så hvem skal afgøre, om en artikel hvert femte år er mere end tilstrækkeligt, eller om rapporter er lige så meget værd som bøger osv. I modsætning hertil bliver yngre forskere til stadighed målt på deres produktion, når de søger stillinger.

Afledt af denne holdning har forskere i tidens løb afvist at informere om, hvad de bruger deres forskningstid til, når de f.eks. er blevet spurgt til forskningsstatistikken. Vi mener, at det er helt legitimt, at de, der betaler, også får en viden om, hvad der betales til, dvs. de danske skatteborgere og det danske samfund. Vi mener, at det er relevant at spørge, om samfundet får noget for forskningspen-

gene. Dermed bliver det også relevant at diskutere, hvordan der sikres en forskningsproduktion – og hvordan den skal opgøres.

Der er en tradition for, at det er forskere, der *leder* forskningsprojekter, forskningsafdelinger og forskningsinstitutioner. Inden for nogle sektorer er det også (tidligere) forskere, der har det administrative hovedansvar. Umiddelbart kan man pege på to begrundelser, der går igen for dette:

Forudsætningen for at kunne give et fagligt feedback er en forståelse for problemstillingerne, dvs. at lederen selv er ekspert på et højt niveau. Forudsætningen for at lederen kan få den respekt, der er nødvendig for at fungere som leder, er faglig kompetence.

Forskningslederrollen har en lang række ens aspekter uanset sektor. Forskningslederen skal sørge for, at medarbejderne får feedback, forskningslederen skal repræsentere projektet/institutionen udadtil, og forskningslederen skal sikre projekternes økonomiske og menneskelige ressourcer.

En vigtig opgave for forskningslederne er at vurdere enhedens samlede præstation (performance): om instituttet/forskningsprojektet set under et har en forskningsproduktion, der er i overensstemmelse med det forventede, både vurderet mht. kvaliteten og omfanget, dvs:

en fagkyndig evaluering
en evaluering af, om ressourcerne udnyttes godt nok - eller om det kan gøres bedre.

De to dimensioner hænger nøje sammen; det giver ikke mening med en fagkyndig evaluering uden at tage hensyn til de ressourcer, der har været til rådighed.

Inden for både den private sektor og sektorforskningen er forskningsledelse noget, der tages hånd om - noget der diskuteres. I disse organisationer er det helt åbenbart, at ikke alle forskere kan fungere som ledere. Det ser ud til, at det i disse organisationer er indlysende, at nogle forskere har 'talent for at lede' og dermed kan blive gode ledere, hvis disse talenter udvikles f.eks. gennem træning eller kurser. Ligeledes kan de fleste forskere blive gode projektledere, hvis der undervises i projektledelse.

For offentligt ansatte forskningsledere svarer problemet om udnyttelsen af ressourcerne til problemet i resten af den offentlige sektor. En ledelse kan gribe dette problem an på flere måder, bl.a. ved at bruge en række *ledelsesværktøjer*. Inden for forskningssektoren har to metoder/koncepter været en del omtalt:

Kvalitetsledelse (Total Quality Management / Business Excellence)³
Benchmarking

Analyseinstitut for Forskning har derfor udviklet en model til brug ved benchmarking af forskningsmiljøer – forstået som et ledelsesværktøj.

Benchmarking - et ledelses-værktøj

Den bærende ide i den udgave af benchmarking-konceptet, som vi arbejder med, er en intern opgørelse over egne ressourcer og produktion med en efterfølgende sammenligning med andre for at lære af disse andre. Der er altså tale om en *læringsproces*, hvor formålet er at blive bedre ved at se på egne forhold og lære af andres erfaringer.

Benchmarking bygger således på åbenhed, tillid og selverkendelse. Læringsprocessen er bygget systematisk op, her udtrykt i en række faser møntet på

større benchmarking-projekter med flere organisationer involveret:

1. Valg af enheder, der har noget til fælles med hensyn til det, der ønskes benchmarket.
2. Valg af kvantitative indikatorer og metode - i samarbejde med enhederne.
3. Måling af de valgte kvantitative indikatorer.
4. Identifikation af enheder, der er bedst inden for konkrete områder (*Best Practices*).
5. Afdækning af forskelle mellem *Best practice*-enheder og de andre vha. kvalitative metoder.
6. Individuelle revisioner af strategier, visioner og arbejdsmåder samt opstilling af nye mål og milepæle.

Baggrunden for at anvende både kvantitative og kvalitative mål er, at det er vanskeligt at måle, hvad der kommer ud af forskningsprocessen: viden, formidlingen af viden og effekten af formidlingen. Der må i stedet anvendes *indikatorer*, således at det efter punkt 3 og 4 er nødvendigt at undersøge nuancerne bag indikatorerne vha. kvalitative sammenligninger.

De kvantitative indikatorer i en benchmarking-proces for forskningsinstitutioner

Ved de kvantitative sammenligninger er et af de største problemer, at mange forskere tror, at en sådan sammenligning *kun* kan finde sted, hvis der laves en kvalitativ vurdering af de forskellige former for forskningsproduktion. Det er forkert, idet der i dag findes velgennemprøvede metoder til at lave sammenligninger uden kvalitative opdelinger - og rent praktisk er det nemmere at lave sammenligningerne uden prædefinerede vægte. Det ændrer dog ikke ved, at det fælles

valg af indikatorer, som skal repræsentere mængden af forskning, vil påvirke den indbyrdes placering.

Inden for det samfundsvidenskabelige område udtrykkes resultatet af forskningen primært gennem publikationer, der kan opdeles i kategorier som artikler i tidsskrifter med og uden referee, bøger (hele og bidrag), lærebøger, notater, rapporter, ph.d.- og doktorafhandlinger samt populærformidling. Desuden sker der en mundtlig formidling ved konferencer o.lign., der eventuelt dokumenteres i papers. Endelig kan produktionen afspejles i forskeruddannelses-aktiviteter, samarbejdsrelationer, mens patenter nok ikke er relevant for dette område.

På ressourcensiden tages der typisk ikke udgangspunkt i antal forskere, fordi den tid, den enkelte forsker har til rådighed til forskning, varierer⁴. I stedet anvendes forskningsårsværk – som i forskningsstatistikken – som man på en række forskningsinstitutioner har relativt detaljerede oplysninger om pr. forsker fra opgørelser over fordelingen af forskning, undervisning, rådgivning og administration. Andre steder må medarbejderne skønne over tidsforbruget. Da de krav, der stilles til adjunker og professorer, ikke er de samme, kan man vælge at opdele ressourcerne efter dette eller lade det være en del af den efterfølgende kvalitative del af benchmarking-processen.

Hver enheds placering samt identifikationen af Best Practices sker nu ved at sammenholde ressourceforbrug og produktionsmål. Kravene til denne beregning er:

Den skal kunne håndtere mindst en form for input (forskningsårsværk)

Den skal kunne håndtere flere former for output (produktionsmål)

Der skal ikke på forhånd vælges vægte til sammenvejning af hverken input eller output

Den skal kunne anvendes til at sammenligne mange institutioner

DEA-metoden⁵ er en matematisk metode, der er skræddersyet til denne beregning. Vægtene fastsættes individuelt for hver deltagende enhed, hvilket betyder, at *de enkelte enheder stilles 'bedst muligt' i sammenligningen med de andre*. Når placeringen for en enhed skal bestemmes, beregnes der altså et vægtsæt for denne enhed, der sikrer den bedst mulige placering. Hvis det ikke er muligt at fastsætte vægte, der giver enheden en bedre placering end de øvrige (når de vurderes med samme vægtsæt), så bliver enheden *ikke Best Practice*. Det forventes så, at enheden kan lære noget ved en sammenligning med andre. Det vil de kvalitative sammenligninger afsløre.

De kvalitative sammenligninger i benchmarking-processen

Efter afdækningen af forskelle i produktionen som målt ved de valgte indikatorer må man gå nærmere ind og finde begrundelser for forskellene. Det gøres ved at sammenligne Best Practice-enhederne med de andre på forhold som:

Indikatorerne (niveau og art)

Miljøets forskningsprofil, struktur, organisation, demografi mm.

Andre arbejdsopgaver for miljøet: undervisning, administration, udredningsarbejde

Udvikling over tiden: indikatorernes udvikling og fremtidigt potentiale

Her kunne et eksempel på forklaringer være forskelle mellem to universitetsinstitutioner, hvor det ene endnu er under opbygning, mens det andet institut har en velfungerende og international anerkendt forskergruppe, der i flere år har arbejdet med en konkret problemstilling, der flere år efter projektets afslutning 'ka-

ster artikler af sig'.

Begrundelserne for forskellene vil dele sig i to grupper, hvor den første omfatter forhold som det nævnte eksempel, og der derfor ikke kan gøres – eller ønskes gjort – noget ved. Den anden gruppe omfatter forhold, hvor der ses muligheder for at ændre på et eller andet. Det kan være på den overordnede strategi f.eks. vedrørende formidlingspolitik, det kan være på arbejdsgange eller rutiner, f.eks. bedre feedback i workshops eller teamstruktur med projektledere eller det kan være på forskningsledelsen, f.eks. brug af medarbejdersamtaler. Fordelen ved sammenligningerne er, at den kan være 'øjenåbner' for medarbejdere, der ellers er socialiseret ind i en konkret måde at arbejde på og derfor har vanskeligt ved at se, at arbejdet kan gøres anderledes.

Kritik af benchmarking-analyser.

Analyseinstitut for Forskning (AFSK) har i flere år arbejdet med benchmarking-begrebet i forbindelse med benchmarking af forskning. Det første offentliggjorte bidrag til debatten findes i et notat fra 1998, hvor begreberne nærmere afklares⁶. Herefter har flere af instituttets medarbejdere diskuteret begrebet i forskellige sammenhænge, og i 2001 kom så rapporten⁷, der for alvor satte benchmarking på dagsordenen.

Desværre har der været en række misforståelser af, hvad der er sigtet og indholdet i denne rapport. Vi har nærmest fået det indtryk, at ganske mange har sprunget teksten i rapporten over og er gået direkte til tabellerne, ligesom vi har fået indtryk af, at DEA-metoden er vanskelig tilgængelig. Det sidste har vi forsøgt at råde bod på ved at udsende nogle forklarende notater, der alle findes på instituttets hjemmeside. Herudover har dele af debatten været præget af fejllæs-

ninger og misforståelser - hvor Finn Hansson og Frode Frederiksen: *Forskningsvognning eller forskningsudvikling*⁸ (F&F) kan anvendes som eksempel på disse fejllæsninger:

F&F's artikel gør sig skyldig i fire fundamentale misforståelser i forhold til rapporten, som her skal klart afvises:

Rapporten er *ikke* et bidrag til afklaring af problemer med forskningsevaluering.

DEA-metoden præsenteres *ikke* som benchmarking processen, men som en mulig metode i forbindelse med de kvantitative opgørelser.

DEA-metoden indebærer *ikke* i sig selv kvalitative vurderinger af f.eks. publikationstyper.

Ressourcemålet *forskningsårsværk* er *ikke* fastsat ud fra en bevidstløs brug af 40-50-10 fordelingen mellem forskning, undervisning og administration.

En læsning af rapportens indledning vil kunne vise de to første punkter og en læsning af rapportens afsnit 3 og 4 de to sidste.

Forskningsevaluering og DEA-metode

F&F's artikel starter med at fastslå, at rapporten skal ses som en evaluering:

"Rapporten vil således forstås som et bidrag til evaluering af forskning...."

Det sker på baggrund af et citat fra indledningen, der er taget ud af sammenhængen, og det er samtidig et af de meget få steder i rapporten, hvor ordet evaluering bliver nævnt. Tages sammenhængen med ind i betragtning, står der, at DEA-metoden *kan* anvendes ved Benchmarking, der igen *kan* anvendes ved en evaluering. Rapporten skal ikke forstås som - og hævder heller ikke nogen steder - at der er tale om et bidrag til

forskningsevaluering:

Følgende tre afsnit/sætninger står fremhævet med fed skrift i den indledning, som citatet er trukket ud af:

Det er her vigtigt, at metoden kun finder de enheder, som er forbilleder, når resultatet af de udvalgte målepunkter anvendes. Andre målepunkter kan betyde, at andre enheder bliver forbilleder. Undersøgelsens resultater kan derfor ikke uden videre generaliseres

og lidt senere:

Den empiriske undersøgelse i denne publikation giver et eksempel på denne metodes anvendelse. Da der ikke er foretaget kvalitative analyser til at be- eller afkræfte resultatet eller forklare evt. forskelle, må resultatet af denne undersøgelse ikke overfortolkes.

Det sidste, der er fremhævet med fed:

Igen er det vigtigt at nævne, at værdi af andre inputs og outputs kan ændre resultatet, og at denne analyse kun er ment som et eksempel på nogle relevante variabler.

Kvalitative vurderinger.

F&F mener, at der ved anvendelsen af DEA-modellen finder en kvalitetsvurdering sted af forskningsproduktionen. Det er - som det fremgår af afsnit 3 i rapporten og af afsnit 3 i denne artikel - faktisk det modsatte, der sker. Der fastlægges ikke vægte ud fra andre kriterier, end at der vælges et individuelt vægtsæt pr. institut, så samtlige institutter bliver stillet bedst muligt.

Ressourcemålet.

Der står i F&F's artikel:

Men heller ikke input-beregningerne er en enkel sag. For at nå frem til et enkelt tal til modellen antager rapporten, at antal forskningsårsværk gange 40% er det bedst mulige input-mål. Baggrunden må være at arbejdstiden for forskere på universiteter beskrives ud fra en norm på 40-50-10 (forskning-undervisning-administration) som dog aldrig er anerkendt af forskningsministeriet. Problemet med at bruge dette grove gennemsnit er, at der med de mange nye typer stillinger (ph.d., forskningsassistenter, -lektorer,-professorer mm.) er kommet meget store lokale variationer mht. forskningstid ind i billedet, hvorfor brugen af en gennemsnitlig procent kun har mening, hvis man antager at der er samme struktur i stillingsmønsteret på alle institutter (Barré 2001: 2262).

Dette er helt forkert. Forskningsårsværkene kommer fra et udtræk af mikrodata fra den officielle danske forskningsstatistik, hvor institutledelsen for hver af 14 personalegrupper angiver antal personer og det antal forskningsårsværk, de har præsteret. Faktisk står der på rapportens side 39-40, at forskningsårsværkene fra forskningsstatistikken er blevet sammenlignet med oplysninger fra årsberetningerne med udgangspunkt i normen 40-50-10. Kun et sted var der så store afvigelser mellem tallene, at der er lavet to beregninger for det pågældende institut. Hertil kommer, at netop forskelle i stillingsstrukturen er brugt som eksempel i afsnit 6 i rapporten, hvor institutternes placeringer korrigeres for forskelle i omverdenen. Det fremgår her bl.a., at pro-

fessorandelen har en positiv indflydelse på placeringen.

Afsluttende kommentar

Når man læser den type kritik, som F&F's artikel er et eksempel på, så er det nærliggende at overveje, om der er noget helt andet på spil: måske handler debatten slet ikke om DEA-metoden eller benchmarking, men om at enhver diskussion om 'value for money' i den offentlige forskningsverden skal lægges på is. Er der tale om en automatisk reaktion: nu kommer økonomerne med deres uforståelige modeller. Er der tale om, at ledelsen på universiteterne ikke skal have økonomiske redskaber i hånden, og/eller at man ønsker at fastholde et monopol på en bestemt type diskussion af begrebet evaluering? At nye former for ledelse inspireret af positive erfaringer fra andre offentlige og private institutioner ikke skal ind på universiteterne, fordi 'den frie universitetsforskning' skal forsvares?

Her er det relevante modspørgsmål, om 'den frie universitetsforskning' virkelig skulle være truet af nye ledelsesformer, hvor forskere f.eks. bliver afkrævet svar på, hvad de bruger tiden til? Det omvendte er nok nærmere tilfældet: jo flere historier, studerende kan fortælle om offentligt ansatte forskere, der ikke forsker, og en ledelse, der ikke leder, jo større er sandsynligheden for, at samtlige universitetsforskere kommer under pres – og det vil så også gælde de forskere, der gør en forskel. Vi må derfor opfordre alle forskere til at gå positivt ind i udviklingen af værktøjer, der kan fremme forskningens kvalitet og omfang.

Noter:

1. Nogle findes på vores hjemmeside, andre fås ved henvendelse til kl@afsk.au.dk
2. se f.eks. Kindtler, Kallehauge og Langberg: *Sektorforskningens roller og rammebetingelser* AFSK 1998/7, og Kallehauge og

Langberg: *Forskere på sektorforskningsinstitutioner* AFSK 1999/6

3. Der er et delprojekt under REMAP-projektet *Forskningsledelse under Forandring* med Kamma Langberg som projektleder, der fokuserer på Kvalitetsledelse på danske forskningsinstitutioner. Det forventes, at der udsendes en rapport i 2002 fra AFSK om dette.

4. se f.eks. Tabel 4.1 og 4.2 i Langberg og Lauridsen: *Universitetsforskernes arbejdsvilkår og holdninger til forskningen og forskeres vilkår*, AFSK rapport 2001/5

5. DEA står for Data Envelopment Analyse. Nærmere beskrivelser af metoden findes på www.afsk.au.dk

6. Udviklingskontrakter mellem Forskningsministeriet og Universiteter. Belyst med inddragelse af erfaringer fra Benchmarking og Evalueringer. Notat 3, 1998

7. Bettina Damm: Et forsøg på benchmarking. En analyse af de samfundsvidenskabelige og juridiske institutter ved de højere læreranstalter Rapport 2001/6