

Forskningsevaluering: Den danske situation i internationalt lys

Af Hanne Foss Hansen

1. Indledning

Forskningsevaluering er på en og samme tid en klassisk og en relativt ny aktivitet i forskersamfundet. Den klassiske form for forskningsevaluering har fokus på personer, enkeltprodukter og forskningsplaner. Ansøgere vurderes i forbindelse med stillingsbesættelser. Afhandlinger vurderes i forbindelse med tildeling af Ph.D. og doktorgrader. Manuskripter til artikler vurderes af referee's og tidskriftsredaktioner, og ansøgninger om projektstøtte vurderes af f.eks. forskningsråd.

Klassisk forskningsevaluering er baseret på anerkendte forskeres læsning af andre forskeres arbejder og oplæg. Og den er koblet til klare beslutningssituationer. Der siges ja eller nej til ansættelse, kompetencetildeling, gradstildeling, publicering eller bevilling. Klassisk forskningsevaluering indebærer, når der arbejdes i udvalg, som ved f.eks. stillingsbesættelser og gradstildelinger, en eksplicit diskussion af kvalitetsbegrebet. Endelig er der i

klassisk forskningsevaluering indbygget en faglig kontrol af de der evaluerer. Indstillinger om stillingsbesættelser og gradstildelinger godkendes af dekaner og eventuelt fakultetsråd. Indstillinger om publicering godkendes af tidsskriftredaktioner, og også i forbindelse med beslutninger om bevillinger i f.eks. forskningsråd er der typisk flere beslutningsniveauer, før der gives endeligt grønt lys.

Den nye form for forskningsevaluering, som vi har set den vokse frem i Danmark ved siden af den klassiske siden starten af 1980'erne, har fokus på større enheder f.eks. institutter, centre, forskningsprogrammer, discipliner eller den samlede indsats inden for et givent forskningstema.

Den nye form for forskningsevaluering har baggrund i en ændret relation mellem samfund og forskersamfund. Tre forhold synes at gøre sig gældende. For det første er forventningerne til forskningens resultater øget. Forskning skal bidrage til teknologisk og økonomisk udvikling samt til løsning

af samfundsmæssige problemer, f.eks. bekæmpelse af forurening, behandling og forebyggelse af sygdomme, udbedring af indeklimaproblemer o.s.v. For det andet er også skepsisen mod visse typer forskning og forskningsprocesser øget. Der stilles i stigende grad spørgsmålstegn ved, hvilke risici forskning indebærer. Er det f.eks. forsvarligt at lave forsøg med udsættelse af gensplejsede planter i naturen? Eller er det forsvarligt at gennemføre lægevidenskabelige forsøg med screening i relation til forskellige former for kræft? Endelig synes forskningens legitimitet i samfundet svækket. Eksperterne er ofte uenige. Er kosttilkud f.eks. gavnligt eller skadeligt? Er stress årsag til mavesår eller mavesår årsag til stress? O.s.v.

Såvel øgede forventninger, øget skepsis som svækket legitimitet synes at bidrage til at øge offentlighedens og det politisk-administrative systems interesse for forskningen og dens udvikling. Styringspresset synes øget og forskersamfundets autonomi indsnævret i forhold til tidligere. Forskningspolitik har udviklet sig til et selvstændigt politikfelt med eget ministerium. Flere ressourcer end tidligere kanaliseres til forskning, og samfundet forlanger rimeligt nok dokumentation for, at investeringen giver udbytte. Evaluering, som også anvendes i stigende omfang i andre dele af samfundet, betragtes af mange som et virkemiddel til at sikre samfundet udbytte af investeringen.

Forskerne og forskningsmiljøerne betragter ofte evaluering som et nødvendigt onde. Evaluering tager tid; tid der går fra den tid, der er til rådighed til forskning. Men evaluering betragtes også som et virkemiddel til legitimering og dermed som et element i en strategi for tiltrækning af yderligere ressourcer til egen forskergruppe, eget institut, egen disciplin, eget forskningstema og/eller eget videnskabsområde.

Den nye form for forskningsevaluering indlejres på denne baggrund såvel i forskningspolitik i snæver forstand (fag- og institutionspolitik) som i forskningspolitik i bred forstand. Udviklingen er ingenlunde et isoleret dansk fænomen. Trenden er international. Måske ser vi udviklingen af et internationalt jetset af forskningsevaluatore?

Evalueringemetodisk adskiller den nye form for forskningsevaluering sig på en række dimensioner fra den klassiske forskningsevaluering. Den er i ringe grad (ofte slet ikke) baseret på læsning af egentlige forskningsprodukter. Den er i stedet baseret på statistik vedrørende publikationer, patenter, ressourcer, forskerkompetence, alders- og kønsfordeling, mobilitet m.v. samt på besøg, de såkaldte "site visits". Til tider anvendes også bibliometriske analyser f.eks. i form af citationsanalyser.

Den nye form for forskningsevaluering gennemføres af paneler sammensat af forskere, der er aktive, inden for det berørte område. I Danmark bruges altovervejende udlændinge. Panelerne er ikke sammensat således at deltagernes kompetencer overlapper hinanden, således som det er tilfældet i den klassiske form i forbindelse med bedømmelsesudvalg f.eks. ved professoratsbesættelser. Panelernes faglige profil er derimod komplementært sammensat, fordi deres vurderinger skal dække et større fagligt spænd end tilfældet er i den klassiske forskningsevaluering. Resultatet er dels en tyndere faglig dækning på evaluatorside og en formodentlig heraf følgende mindre dybdegående faglig vurdering, dels ringere muligheder for at panelets deltagere kan krydskontrollere hinandens vurderinger.

Endelig er det karakteristisk for den nye form for forskningsevaluering, at den oftest indgår i en uklar beslutningssituation, en beslutningssituation hvor dagsorden og mulige beslutningsudfald skabes undervejs i processen, modsat den klassiske evaluering hvor disse er startpunkt for processen. Sammenlignes den nye form for forskningsevaluering med den klassiske form er det således klart, at det fag- og forskningspolitiske aspekt i evalueringprocessen er øget, mens det rent faglige kvalitetsaspekt er nedtonet. I den nyligt afholdte høring om den pågående evaluering af udvalgte dele af dansk samfundsvidenskab kom dette tydeligt frem, da en deltager i et af evalueringspanelerne efter en kritik af panelets arbejde sagde: "Don't take evaluations too seriously. It is just opinions. The panel is in no position to judge."

Med udgangspunkt i metodemæssige og organisatoriske overvejelser kan der altså opstilles en kritik af den nye form for forskningsevaluering. Mens vi fra en række empiriske studier af klassisk forskningsevaluering har kendskab til de her anvendte kvalitetskriterier, processens organisering, forekomsten af bias m.v.¹, er empiriske studier af moderne forskningsevaluering desværre langt mere sparsomme.² Der er et stort behov for systematisk viden om disse evalueringers organisering, proces, bias'es o.s.v. samt ikke mindst for yderligere viden om evalueringernes betydning og evalueringsrapporternes og -resultaternes efterfølgende anvendelse.

Selvom den nye form for forskningsevaluering som anført kan siges at adskille sig fra den klassiske forskningsevaluering på en række dimensioner, er der på ingen måde tale om en entydig ting. Forskningsevaluering kan organiseres på mange måder og en række metoder kan anvendes. I det følgende diskuteres en række forskellige evalueringsformer, og der gives konkrete eksempler. De forskellige former for forskningsevaluering implicerer ikke blot et syn på forsknings- og vidensbegrebet, jfr. diskussionen i Hjørlands artikel andetsteds i dette tidsskrift, de implicerer også et syn på forskningsorganisationer samt virkemidler og procesrettelæggelse relateret til disses udvikling. I sidste instans implicerer de et menneskesyn.

Diskussionen af evalueringsformer vil tage udgangspunkt i en enkel typologi. Typologien sonder mellem forskellige evalueringsformer på baggrund af forskelle i disses målsætninger og "ejerskab". Den danske tradition for og overordnede diskussion af forskningsevaluering vil blive sat i relief via redegørelser for erfaringerne i England og Holland. Artiklen afsluttes med påpegning af en række principielle designvalg i udformningen af evalueringspraksis og evalueringssystemer.

2. Forskningsevaluering: En typologi

Der er flere forsøg på typologisering af fænomenet forskningsevaluering. I et kort debatindlæg argu-

menterer Aage Bøttger Sørensen (1994) skarpt imod, hvad han kalder "den pædagogiske evalueringsmodel", og advokerer i stedet for anvendelse af en evalueringsmodel, der benytter sig af "effektive og konsekvente sanktioner". Sondringen er enkel. Desværre tyder meget på, at den er for enkel. Resultaterne af empiriske studier af udviklingen i den offentlige sektor understreger, at hverken intern læring eller eksternt pres, f.eks. etableret via evaluering koblet med sanktioner, i sig selv sikrer forandring og kvalitetsudvikling. Kvalitative forandringer sker derimod som resultat af komplicerede samspil mellem eksternt pres og intern læring.

Jeg har derfor i stedet valgt at strukturere diskussionen i det følgende på basis af en lidt mere udbygget typologi, jfr. figur 1. Typologien er udviklet af Trow (1994). Eksemplerne står for egen regning.

Typologiens opdeling er baseret på to dimensioner. Den ene dimension vedrører formålet med at gennemføre forskningsevaluering. For det første kan forskningsevaluering have til formål at kontrollere om en given indsats udføres som den bør. Om forskningsproduktiviteten og/eller -kvaliteten er på et rimeligt niveau, i praksis ofte defineret som "på internationalt niveau". For det andet kan forskningsevaluering have til formål at støtte udviklingen af en given indsats, med andre ord at give feedback til forskere, forskergrupper, institutter og discipliner samt formulere forslag til tiltag, der kan forbedre indsatsen fremover.

Kontrolformålet vedrører primært fortiden. Kontrollen synliggør fortidens helte og syndere. Støtteformålet derimod er mere direkte orienteret mod fremtiden, mod at udvikle forskningen og forskningsmiljøerne.

Den anden dimension vedrører ejerskabet til forskningsevalueringen. Ejerskabet kan være forankret internt eller eksternt. Internt ejerskab er defineret ved at evalueringen er initieret internt på en forskningsinstitution (f.eks. på et universitet eller anden højere læreanstalt eller på et sektorforsknings-

	Intern	Ekstern
Kontrollerende	forskningsvognning	rating
Støttende	institutevaluering	disciplinevaluering

Figur 1: Forskellige typer af forskningsevaluering

institut) og gennemføres på dennes ansvar. Eksternt ejerskab er modsat defineret ved at evalueringen er initieret eksternt af en offentlig myndighed, f.eks. et ministerium eller et forskningspolitisk organ, og gennemføres på dette ansvar.

Der kan rejses en række kritikpunkter mod typologien. For det første kan typologien karakteriseres som en ex-ante typologi. Hermed menes at typologien udelukkende specificerer formålene med givne former for forskningsevaluering, men ikke nødvendigvis den efterfølgende anvendelse af evalueringernes resultater. Man kan f.eks. tænke sig, at resultater fra forskningsevalueringer hvis formål fra starten har været støttende efterfølgende (forsøges) anvendt til kontrolformål. Det er i denne sammenhæng vigtigt at være opmærksom på, at resultater fra forskningsevalueringer, som allerede nævnt, indgår i forskningspolitiske processer. Der kan både være tale om processer, der afspejler "kampe" internt i forskersamfundet, og processer der udspilles i relationen mellem forskersamfund og samfund.

For det andet kan man anføre, at andre formål end kontrol og støtte kan være aktuelle, og at typologien derfor burde udvides. Forskningsevaluering kan f.eks. tænkes at have til formål at legitimere en given forskningsindsats eller forskningsinstitution; og/eller forskningsevaluering kan have til formål at øge gennemsigtigheden og måske efterfølgende konkurrencen forskningsinstitutionerne imellem. Dernæst kan man selvfølgelig anføre, at forskellige aktører kan have forskellige formål med at gennemføre en given forskningsevaluering. Et

ressortministerium for et sektorforskningsinstitut kan måske ønske at kontrollere, mens instituttet selv ønsker faglig feedback og støtte. Konflikter om målsætninger for evalueringer kan føre til konflikter og kompromiser om disses udformning og tilrettelæggelse. Endelig kan man anføre, at en ting er de officielle målsætninger, noget andet ofte de reelle. Evalueringer kan garneres med sympatiske parademålsætninger, som kan dække over mere illegitime hensigter.

Også i relation til den anden dimension kan der fremføres kritikpunkter. Først og fremmest kan man anføre, at sontringen mellem intern og ekstern er uklar. Der er mange niveauer i forskningssystemet, og ofte er de overlappende i den forstand, at de samme forskere optræder med forskellige "kasketter" på flere niveauer. En evaluering af et forskningsinstitut, der initieres af et fakultet, er ifølge definitionen intern, men opfattes måske af forskerne på det pågældende institut som ekstern.

Trods de nævnte kritikpunkter er det min opfattelse, at modellen er et enkelt og frugtbart redskab til at diskutere karakteristikaene ved de forskellige former for forskningsevaluering, der anvendes i Danmark såvel som i udlandet. Som sådan vil den blive benyttet i det følgende.

2.1. Forskningsevaluering som led i intern kontrol

Et dansk eksempel på internt initieret, kontrollerende forskningsevaluering er den såkaldte *forskningsvognning*. Forskningsvognning er en form for forskningsevaluering, hvor universitetsledelsen på

forskellige niveauer, typisk institut- eller fakultets-niveau, anvender publikationstælling med henblik på at kontrollere, hvorvidt forskerne har levet op til deres forskningsforpligtigelse. Mange universiteter har udformet regelsæt, der specificerer de minimumskrav de fastansatte forskere skal leve op til. Typisk forlanges et antal publikationer over en nærmere defineret periode. Forskningsvognning udgør i denne forstand en produktivitetskontrol.

2.2. Forskningsevaluering som led i intern støtte

Danske eksempler på internt initierede, støttende forskningsevalueringer er de forskergruppe- og institutevalueringer, der i stigende omfang foregår på både universiteter og sektorforskningsinstitutter. Der er blandt andet eksempler på sådanne processer på naturvidenskab på Københavns Universitet, samfundsvidenskab på Aalborg Universitet, Handelshøjskolen i København, Arbejdsmiljøinstituttet, Statens Planteavlsvorsøg og Statens Bygge-forskningsinstitut.

På det erhvervsøkonomiske fakultet på Handelshøjskolen i København har man f.eks. igangsat en evalueringsprocedure, der betyder at alle institutter på fakultetet over en årrække evalueres (Det erhvervsøkonomiske Fakultet 1994 & 1995). Der nedsættes et panel med 2 anerkendte udenlandske forskere for hvert enkelt institut. På basis af et baggrundsmateriale udarbejdet af instituttet (en slags selvevalueringsrapport), besøg samt eventuelt interviews og seminarer udarbejder panelet en evalueringsrapport typisk indeholdende vurderinger af faglig kvalitet, faglig profil, organisering samt forskeruddannelse. Rapporten følges op af diskussioner og udmeldinger fra det enkelte institut om fremtidige forsknings- og rekrutteringsplaner m.v.

Trods de fælles rammer synes evalueringsprocesserne at forløbe relativt forskelligt på de forskellige institutter, ligesom de tilskrives og får forskellig betydning på institutterne efterfølgende (Borum og Foss Hansen, 1995). Faktorer som det enkelte instituts kultur og traditioner, de aktuelle lokale dagsordener samt de udpegede evaluators evalueringer

erfaringer påvirker den enkelte evaluering design, procesforløb og den efterfølgende effekt. Der er både eksempler på evalueringer, som ikke synes at få meget andet end symboleffekt og eksempler på evalueringer med omfattende effekter på omorganisering, faglig profilering samt rekrutteringsplaner.

Herudover kan man som eksempel på en internt støttende evaluering lignende aktivitet nævne den monitorering som Københavns Universitet har gennemført vedrørende den bioteknologiske forskning på universitetet (Københavns Universitet, 1996). Her er tale om et omfattende beskrivelses- og analysearbejde, der indgår som basis for universitetets strategiarbejde. Den bioteknologiske forskningsprofil, publikationsaktivitet, ressourcesituation samt impact målt på omfanget af citeringer analyseres med basis dels i oplysninger fra forskerne selv, dels i kørsler i internationale citationsdatabaser.

2.3. Forskningsevaluering som led i ekstern kontrol

I Danmark findes kun enkeltstående eksempler på evalueringer gennemført som led i ekstern kontrol. Et omdiskuteret eksempel med dramatiske konsekvenser er den på foranledning af det daværende Direktoratet for de Videregående Uddannelser i Undervisningsministeriet af Aage Bøttger Sørensen udarbejdede vurdering af danske sociologers publikationer (Aage B. Sørensen, 1988).

Herudover findes en del analyser, der ledsages af forslag om udvikling af eksterne kontrolsystemer. Her kan f.eks. nævnes Jennergren og Obels (1985-86) nu 10 år gamle analyse af 22 økonomiske institutter baseret på den såkaldte DEA-metode samt Gøtzsche m.fl.'s (1995) analyse af dansk lægevidenskabelige forskning. Led i egentlige evalueringer kan disse analyser dog ikke siges at have været.

For at finde gode eksempler på institutionaliserede evalueringssystemer baseret på ekstern kontrol må vi derfor kigge til udlandet. Et sådant eksempel er

det engelske *ratingsystem*. I dette system er der siden 1986, i starten hver tredje år, nu hver fjerde år, givet karakter til alle universitære forskningsinstitutter, der ønsker at deltage i evalueringen. I 1992 tildeltes over 2500 forskningsinstitutter hver *en karakter* på en skala fra 1 til 5, hvor 5 var bedst. Karaktererne, der gives uden kommentarer, begrundelser eller anden form for feedback, blev uddelt af 62 altovervejende nationalt sammensatte faglige paneler og en række underpaneler med i alt omkring 450 medlemmer.

Til grund for karaktergivningen lagde panelerne en rapport fra hvert institut. På en given skæringsdato opgjorde rapporten statistiske data vedrørende aktive forskere ansat på skæringsdatoen, publikationslister for de foregående fire år, opgørelser over omfang og kilder for eksterne ressourcer ligeledes for de forgangne fire år, redegørelse for forskningsplaner samt henvisning til maksimalt 2 publikationer og 2 andre offentligt tilgængelige forskningsoutput (f.eks. patenter) pr. forsker. Panelerne tilrettelagde selv deres arbejdsform. Nogle paneler valgte at inkludere anvendelse af bibliometriske analyser, f.eks. citationsanalyser.

Karaktergivningen anvendes efterfølgende som basis for ressourcefordeling. Institutter der ikke "melder sig" til evalueringen har derfor ingen muligheder for at få andel i de basisressourcer, der stilles til rådighed for forskning.

Evaluering- og ressourcefordelingssystemet har en række konsekvenser, som blot skal nævnes kortfattet her:

Systemet fastholder en skæv ressourcefordeling. 21 universiteter med 22% af den samlede studentebestand modtager 64% af de samlede forskningsmidler.³ Herudover synes systemet at favorisere store institutter på bekostning af de mindre (Martin m.fl., 1993), hvilket på nogle fagområder fører til koncentration.

Systemet påvirker på mange måder adfærden i forskningssystemet. Forskernes publiceringsadfærd tilpasses ønsket om, at kunne fremvise flest mulige

publikationer. Inden for samfundsvidenskab og humaniora giver det særlige problemer, fordi bøger ikke længere er en attraktiv publiceringsform. Institutterne går på jagt efter hinandens publiceringstunge forskere. "Poaching" (=krybskytteri) eller headhunting udbredes, og der udvikles en forskningsmæssig elite med særligt gode arbejdsforhold. Udlandsophold, gæsteforskerbesøg m.v. planlægges ikke ud fra faglige hensyn men ud fra strategiske hensyn knyttet til skæringsdatoen for opgørelsen.

Sammenfattende må man sige, at evalueringssystemets organisering samt dets kobling til ressourcefordeling skaber produktivitetsudvikling men næppe kvalitetsudvikling. Filosofien systemet bygger på er kontrol og sanktioner, et tælle-veje system. Eksemplet viser, at evalueringssystemer, der benytter effektive og konsekvente sanktioner, jfr. den tidligere omtalte dikotomi, ikke sikrer kvalitetsudvikling. Det er i øvrigt interessant at notere sig, at man i England tilsyneladende er på vej væk fra systemet, dels i erkendelse af dets svagheder og uønskede bieffekter, dels fordi systemet med dets fokusering på fortiden ikke matcher et aktuelt politisk ønske om i ressourcefordelingen at prioritere samfundsmæssig relevans. År 2000 bliver formodentlig sidste runde med det nuværende system.

2.4. Forskningsevaluering som led i ekstern støtte

Eksempler på eksternt initierede, støttende forskningsevalueringer er de område- og disciplinerevalueringer, der pågår både i Holland og i Danmark. Først et par ord om den hollandske model.

Det nuværende hollandske forskningsevalueringssystem er indført i 1992. Det baserer sig på at alle områder, typisk defineret som discipliner, evalueres ved regelmæssigt tilbagevendende evalueringer. Tanken er, at et område skal evalueres hvert 5. år. Man startede i 1993 med evaluering af 4 områder som en form for pilotprojekt. De øvrige 21 områder skal evalueres første gang i perioden 1994-

1997. Herefter startes forfra. Man har således endnu ikke erfaringer med, hvordan turnustankegangen kommer til at virke i praksis.

Evalueringerne foregår efter følgende procedure: 1) Der indsamles baggrundsmateriale fra institutterne inden for området i form af beskrivelser af forskningsprofil, nøgledata vedrørende finansiering, bemanning, publikationsvirksomhed m.v.; 2) Der udpeges et panel bestående af 5-7 medlemmer, alle internationalt anerkendte eksperter inden for området; 3) Panelet gennemfører interviews og/eller besøger institutterne samt gennemfører i nogle tilfælde bibliometriske analyser, f.eks. citationsanalyser; 4) Hvert institut gives fire kommenterede karakterer på en skala fra 1 til 5. Karaktererne gives for henholdsvis videnskabelig kvalitet, videnskabelig produktivitet, videnskabelig og eventuel samfundsmæssig relevans samt for langsigtet levedygtighed. Herudover kan andre forhold, f.eks. ledelse vurderes, hvis dette er ønskeligt. Blandt karaktererne er nogle således rettet mod fortidens resultater, andre mod fremtidigt potentiale. Det er ikke kun fortiden, der tæller, som i England.

Evalueringsrapporterne distribueres bredt. Der er ingen mekanistisk kobling til ressourcefordeling, som i det engelske system. Tværtimod har regeringen forpligtiget sig til ikke at omfordele ressourcer før forskningsinstitutionerne har haft 3-5 år til at forbedre eventuelt kritisable forhold. Det hollandske system er således baseret på en fremadrettet proces-tænkning.

De landsdækkende danske forskningsevalueringer er mindre systematiske og mere pragmatiske end de hollandske. Alligevel giver det mening at tale om en institutionaliseret dansk model, hvis styrker og svagheder det er vigtigt at diskutere i en situation, hvor etableringen af et nyt overordnet forskningspolitisk organ, Danmarks Forskningsråd, muliggør et kursskift.

I Danmark startede forskningsevalueringerne meget forsøgsvis på initiativ af den daværende Undervisningsminister Ritt Bjerregaard og med Planlægningsrådet for Forskning (PRF) som

ansvarlig for gennemførelsen. Internationalt blev der lagt mærke til Danmark, da vi satte "videnskabelige lægmænd" til at vurdere forskningen inden for områderne engelsk og matematik. Erfaringerne var blandede, og forsøget forblev et forsøg (se f.eks. Ståhle, 1985).

Senere udvikledes, hvad der kan kaldes den danske model, først i regi af PRF (evalueringen af miljøforskning, Planlægningsrådet for Forskning, 1989) senere af Forskningspolitisk Råd (FPR; evalueringerne af jordbrugsforskning, Danish Council for Research Policy, 1992, sundhedsvidenskabelig forskning, Danish Council for Research Policy 1993, samt den igangværende evaluering af samfundsvidenskabelig forskning).

På nogle punkter ligner den danske model den hollandske. Der indsamles et ofte omfattende baggrundsmateriale hos forskere og forskningsmiljøer. Der nedsættes et panel for hvert område bestående af internationalt anerkendte forskere. Der gennemføres besøg og evalueringsrapporterne spredes bredt.

Der er imidlertid også væsentlige forskelle de to lande imellem. Evalueringerne er i Danmark sat mindre i system. Valget af områder sker ad hoc, der er ingen samlet plan og ingen tanke om gentagne evalueringer med års mellemrum af de samme områder. Der er ikke i Danmark faste vurderingskriterier, men der udarbejdes et kommissorium ad hoc for hver evaluering. Også hvad angår sekretariatsbetjening er den danske model ad hoc. Et (delvist) nyt sekretariat opbygges hver gang, hvilket synes at svække metodebevidsthed og -udvikling.

Hertil kommer, at der i den danske model lægges relativt stor vægt på vurderingen af organisatoriske forhold. Oftest etableres et særligt panel for vurderingen af organisatoriske spørgsmål, selvom også dette altovervejende er sammensat af eksperter inden for det pågældende fagområde, der evalueres, og ikke af f.eks. eksperter i forskningsorganisation og -ledelse. Faktisk synes de organisatoriske vurderinger fra tid til anden at overskygge de fagli-

ge vurderinger. Muligvis giver det mening at tale om en målglidning undervejs i processen fra kommissorium til evalueringsrapport.

Endelig er der en vigtig forskel hvad angår ejerskab mellem evalueringerne i Holland og i Danmark. I Holland har en organisation kaldet VSNU (på engelsk: Association of Universities in the Netherlands) ansvaret for driften af evalueringssystemet. I Danmark har ansvaret som nævnt været tillagt skiftende overordnede forskningspolitiske organer. De hollandske evalueringer kan således siges at ligge mere til venstre (i retning af internt ejerskab) i figur 1 end de danske.

Den senere tids kobling af de danske evalueringer til udarbejdelsen af nationale strategier (for jordbrugsforskningens vedkommende se Udvalget om udarbejdelse af forslag til en national strategi for den offentlige jordbrugsforskning, 1994, for den sundhedvidenskabelige forsknings vedkommende Det Nationale Strategiudvalg for Sundhedsvidenskab, 1995) understreger den danske pragmatisme. En ny dagsorden, der muligvis for nogle aktører kan give evalueringerne større mening, blev sat/kom forbi. Evalueringerne var imidlertid på ingen måde planlagt efter denne dagsorden; kommissorierne ikke tilpasset dette formål.

3. Principielle overvejelser omkring forskningsevaluering

Af beskrivelserne af de forskellige evalueringsformer ovenfor kan udtrages en række principielle dimensioner og valg, som alle evalueringsaktiviteter, -systemer og -politikker bør forholde sig eksplícit til. De to første dimensioner kan henføres til sondringerne i figur 1, de øvrige er supplerende:

- 1) Formålsdimensionen: Sigter evalueringen mod at gennemføre kontrol og/eller sigter den mod at igangsætte og støtte en udvikling?
- 2) Ejerskabsdimensionen: På hvilket niveau i forskningssystemet forankres evalueringsansvaret? Opleves forankringen, og dermed ansvaret for

design og opfølgning, som internt eller eksternt af de evaluerede?

Både dimension 1 og 2 relaterer sig til forhold som tillid og accept. Kontrol opfattes, hvis det er noget der går ud over opnåelse af en minimumsnorm, ofte som et signal om mistillid. Eksternt ejerskab, herunder evaluatorens udpeget eksternt uden forslagsindkaldelse eller høring, kan vanskeliggøre accept af evalueringens resultater. Manglende accept af en evalueringens resultater fører ofte til manglende implementering af dens anbefalinger.

3) Tidsdimensionen: Giver evalueringen kredit til fortidens helte (dokumenterede resultater, merit) eller udpeger den potentialer (unge, lovende forskere eller forskningsmiljøer; kreative ideer; velbeskrevne, relevante forskningsplaner etc.)? Er evalueringen fagligt betragtet ex-post eller ex-ante orienteret? Her gemmer sig grundlæggende forskellige forskningspolitiske betragtningsmåder. Nogle hævder, at forskningsressourcer for at sikre kvaliteten af forskningen udelukkende bør gives til forskningsmiljøer, der allerede har vist hvad de du'r til, andre at en sådan strategi indebærer faglig konservatisme og risiko for indavl af både tanker og personer.

4) Resultatdimensionen: Udmøntes evalueringen i endimensionel karaktergivning, flerdimensionel karaktergivning eller udpegning og beskrivelse af eventuelle problemer, der bør tages hånd om (monitorering)? Der stilles mange konkurrerende og til tider også modsætningsfyldte krav til forskere og forskningsmiljøer; hvis der skal evalueres med sigte på karaktergivning, bør karaktergivningssystemet reflektere flerheden af krav.

5) Publikumsdimensionen: Anlægger evalueringen internt videnskabelige kriterier og/eller samfundsmæssige relevanskriterier?

6) Effektivitetsdimensionen: Anlægger evalueringen produktivitets- og/eller kvalitetskriterier?

7) Afgrænsningsdimensionen: Fokuserer evaluering

gen udelukkende på evaluering af forskning eller inkluderes andre tilgrænsende aktiviteter såsom forskeruddannelse, undervisning, politisk-administrativ rådgivning etc.? Inkluderer evalueringen evaluering af forskningsorganisation og -ledelse?

8) Konsekvensdimensionen: Skal evalueringen f.eks. mekanistisk udløse ressourcer, som i det engelske system, eller være dekoblet fra ressourcfordeling? Litteraturen om metoder til forskningsevaluering påviser, at såvel bibliometriske som peer review baserede metoder er så problemfyldte, at de kun vanskeligt kan bære mekanistisk konsekvensdragnet. Samtidig vil evalueringer uden konsekvenser tendere mod symbolske processer og måske tidsspilde. Der må findes en mellemvej.

9) Procesdimensionen: Er evalueringen et engangskud eller tænkt ind i en proces (opfølgning, turnus)?

10) Etikdimensionen: Vurderes de evaluerede på egne præmisser eller eksterne præmisser? Her gemmer sig flere forhold:

a) Faglighed: f.eks. kan der være problemer og urmeligheder i at vurdere flerfaglige og tværfaglige forskningsmiljøer på disciplinære kriterier;

b) Tidsaspektet: forskning og forskningsmiljøer udvikles ikke fra den ene dag til den anden. Forskning tager tid. Derfor kan der være problemer i i 90'erne at evaluere resultaterne af forskning initieret i 80'erne på kriterier udledt af 90'ernes aktuelle forskningspolitiske kæpheste.

4. Afslutning og anbefalinger

Beskrivelserne og diskussionerne ovenfor af forskellige former for forskningsevaluering samt de dimensioner, der ligger til grund for disse former, har vist, at design af forskningsevalueringer og udformning af institutionaliserede evalueringssystemer er både kompliceret og udfordrende.

I Danmark kompliceres situationen i disse år for universitetsforskningens vedkommende yderligere af opsplittningen af, konkurrencen mellem og manglen på samarbejde mellem forskellige dele af det politisk-administrative apparat. Med dannelsen af det nye Danmarks Forskningsråd er tiden måske moden til en mere principiel debat om disse forhold?

Skal jeg slutte med en anbefaling til Danmarks Forskningsråd må det være: Etablering af en evalueringspause. Mens "våbenhvilen" råder, bør der reflekteres over, hvordan evalueringssopgaven bedst forankres og løses fremover. Set med mine briller er der behov for udvikling af et system der på en og samme tid, dels via eksternt pres sikrer at evaluering pågår, dels forankrer evalueringprocesserne internt på forskningsinstitutionerne, så et design der matcher lokale problemstillinger sikres. Man kan forestille sig en struktur med et overordnet metaevaluerende organ, som har ansvar for evaluering af evalueringsskemaerne lokalt. Samtidig med at den egentlige evaluering sker lokalt. En sådan struktur må formodes at kunne afpolitiseres og fagligt opkvalificere evalueringsskemaerne.

Følges en sådan model, sættes der spørgsmålstegn ved, om Danmarks Forskningsråd skal være et egentligt evalueringsskema og -udførende organ, sådan som Forskningspolitisk Råd har været det. I stedet tegnes konturerne af et organ, der besidder evalueringskompetence, herunder følger udenlandske erfaringer på området, og som derfor kan være inspirerende og overvågende i forhold til andre organer og organisationers rutiner.

Noter

- 1 Det er ikke muligt fyldestgørende at omtale og referere til denne litteratur her. Der kan henvises til oversigten i Foss Hansen og Holst Jørgensen, 1995, kapitel 5.
- 2 Der er i Norden gennemført en større undersøgelse, se Luukkonen og Ståhle, 1993.

- 3 Dette var så at sige også tilsigtet. Systemet blev indført i en situation, hvor antallet af universiteter blev kraftigt forøget, idet de tidligere Polytechnics fik status af universiteter, uden at den samlede pulje til forskningsressourcer øgedes. Systemet skulle i denne situation sikre, at ressourcerne ikke blev spredt for tyndt.

LITTERATUR

Borum, Finn & Hanne Foss Hansen: Beneath the Smooth Surface of Evaluation Rhetoric in John K. Christiansen m.fl. (eds.): *Proceedings of the 13th Nordic Conference on Business Studies*, vol. 1 (321-331). København, 1995.

Christiansen, John K. & Hanne Foss Hansen: *Forskningsevaluering i teori og praksis*. København: Samfundslitteratur, 1993.

Danish Council for Research Policy: *International Evaluation of Danish Agricultural Research*. Copenhagen: Danish Council for Research Policy, 1992.

Danish Council for Research Policy: *International Evaluation of Danish Health Research*. Copenhagen: Danish Council for Research Policy, 1993.

Det erhvervsøkonomiske fakultet: *Forskningsevaluering 1994*. København: Handelshøjskolen i København, 1996.

Det erhvervsøkonomiske fakultet: *Forskningsevaluering 1995*. København: Handelshøjskolen i København, 1996.

The Faculty of Social Sciences, Aalborg University: *IKE Group. Industrial Development and International Competitiveness*. Aalborg, 1996.

Foss Hansen, Hanne & Birte Holst Jørgensen: *Styring af forskning: Kan forskningsindikatorer anvendes?* København: Samfundslitteratur, 1995.

Gøtzsche, Peter C.; Jacob W. Krog og Rasmus Moustgaard: Bibliometrisk analyse af dansk sundhedsvidenskabelig forskning 1988-1992, i *Ugeskrift for Læger*, 157/37, 11. september 1995 (5075-5081).

Jennergren, L. Peter & Børge Obel: *Forskningsevaluering - eksemplificeret ved 22 økonomiske institutter*, i *Økonomi og Politik*, 59. Årgang, nr. 2, 1985-86.

Københavns Universitet: *Strategi for forskning og undervisning i bioteknologi 1996-2002*. Københavns Universitet, 1996.

Luukkonen, Terttu & Bertel Ståhle: *Evaluation of Research Fields - Scientists' Views*. Nord 1993:15. København: Nordisk Ministerråd, 1993.

Martin, Ben; Diana Hicks; E. Nigel Ling & Jim E. F. Skea: *The Effects of Size and Other Factors on the Research Performance of University Departments*. Brighton, SPRU, May, 1993.

National Institute of Occupational Health: *International Evaluation of the National Institute of Occupational Health*. Copenhagen, 1992.

Det Nationale Strategiudvalg for Sundhedsvidenskab: NASTRA, *Forslag til national strategi for sundhedsvidenskab*. Forskningsministeriet 1995.

Planlægningsrådet for Forskningen: *International Review of Danish Environmental Research*. København: Planlægningsrådet for Forskningen, 1989.

Statens Byggeforskningsinstitut: *Evaluering af forskningen 1995*. Hørsholm: Statens byggeforskningsinstitut, 1995.

Ståhle, Bertel: Om utvärdering av forskning och evaluering av evaluering - en nordisk introduktion i Bertel Ståhle (red.): *Utvärdering av forskning. Nordiska erfarenheter*. København: Nord FPR-publikation nr. 2, 1985.

Sørensen, Aage B.: *Danske sociologers publikationer*. Harvard University, September 1988.

Sørensen, Aage B.: Evaluering som middel til organisationsindlæring, i Danmarks Grundforskningsfond: *Beretning og årsregnskab 1994*.

Trow, Martin: *Academic Reviews and the Culture of Excellence*, Kanslerämbetets skriftsserie 1994:1. Stockholm, 1994.

Udvalget om udarbejdelse af forslag til national strategi for den offentlige jordbrugsforskning: *Forslag til en national strategi for jordbrugsforskning. Betænkning nr. 1274*. København, juni 1994.

Hanne Foss Hansen
Lektor, lic. merc. Institut for Statskundskab,
Københavns Universitet