

Erik Albæk og Søren Winter

## **Evaluering i Danmark: Rationalitet eller politisk våben?**

Evalueringsudviklingen tolkes ofte i et rationelt lys og sættes i forbindelse med velfærdsstatens udvikling, men så må det undre, at evalueringen kom så relativt sent til Danmark. Mere frugtbart kan evaluering opfattes som et våben i politiske konflikter, og så kan både dens sene ankomst og kraftige udvikling i det sidste årti hænge sammen med et relativt lavt, men i 1980'erne stigende konfliktniveau. Udbuddet af evalueringskyndige har dog også udviklet sig, men kvaliteten er stadig ikke tilfredsstillende.

### **Evalueringsudviklingen<sup>1</sup>**

Hvis man med evaluering mener en systematisk, som regel forskningsbaseret, vurdering af offentlige foranstaltningers og indsatsområders design, implementering og effekt, så er evaluering - eller evalueringsforskning - en forholdsvis ny aktivitet i dansk sammenhæng. Selv om der også tidligere er gennemført evalueringer i Danmark - men som oftest under anden betegnelse - så er evalueringsforskning et fænomen, der sporadisk blev introduceret i dansk politik og forvaltning i slutningen af 1970'erne, og som først begyndte at vinde fodfæste i 1980'erne.

Den danske udvikling er ikke enestående, men nok bemærkelsesværdig derved, at en række lande, som ellers i mange henseender minder om Danmark, på et betragteligt tidligere tidspunkt indførte evaluering som almindelig forvaltningspraksis og på væsentlig mere systematisk vis og i betydeligt større omfang, end vi kender det herhjemme. Det drejer sig først og fremmest om USA, der er evalueringens moderland, men også lande som Vesttyskland, Holland og England (Rist, 1990). Til gengæld er der sket en kraftig vækst i evalueringen i Danmark i løbet af 1980'erne og 1990'erne. Det er interessant at søge en nærmere forklaring på, hvorfor evaluering dukker op, og i hvilken takt den udvikles. Den komparative metode kan være velegnet til belysning af denne problemstilling, og vi vil søge at opstille nogle hypoteser med udgangspunkt i evalueringsudviklingen i navnlig USA og Danmark.

Et forklaringsforsøg må tage udgangspunkt både i evalueringens udbuds- og efterspørgselsside. Udbudssiden er ikke mindst vigtig. Således var bestanden af kandidater med viden om og erfaring med den type samfundsvidenskabelig teori og metode, der ville sætte dem i stand til at udføre evalueringer, yderst beskeden her i landet, da den danske velfærdsstat blev udbygget. Og den var stadig beskeden, da evalueringsforskningen blev populær i USA i 1960'erne og 1970'erne. Centraladministrationen domineredes fortsat af jurister, for hvem evalueringstankegangen ikke er nogen fortrolig størrelse, og det var i det hele taget forholdsvis let for samfundsvidenskabelige kandidater at finde anden form for beskæftigelse end ved evaluering. Ved de danske universiteter er der

kun i meget lille omfang, og først fra slutningen af 1980'erne, blevet undervist i evalueringsmetoder. Udbuddet af kandidater med evalueringskendskab kan skabe sin egen efterspørgsel efter evaluering, når de selv bliver ansat i den offentlige forvaltning.

Efterspørgslen påvirkes dog nok ikke kun af udbuddet, men også af karakteristika ved det politiske system og de politiske problemområder. Amerikanere søger ofte at forklare evalueringens fremvækst i USA ud fra et *rationalistisk* perspektiv i forening med velfærdsstatens udvikling (Albæk, 1988; Rossi og Freeman, 1989; Vedung, 1991). Evalueringen knytter sig til en rationalistisk tradition, der skønt den ikke opfatter den politiske og administrative verden som rationel, så dog alligevel forsøger at gøre de politiske og administrative beslutningsprocesser mere rationelle.

Hvor programbudgettering var et fremadsskuende forsøg på mere rationelt at planlægge sig til beslutninger og løsninger, der effektivt kan realisere de politiske mål, kan evaluering opfattes som en mere tilbageskuende aktivitet, der undersøger, om de politiske mål nu også opnås i praksis, og som kommer med forslag til en politikrevision, der kan sikre en bedre målopfyldelse. Hvor programbudgetteringen og planlægningssystemerne var instrumenter til mere rationelle beslutninger i første omgang, blev evalueringen midlet til rationelt at sikre *feed-back*.

Ud fra dette rationalistiske perspektiv forøges behovet for evaluering med velfærdsstatens vækst. Mens natvægterstatens opgaver var relativt simple, bevæger den moderne velfærdsstat sig ind på mere og mere komplekse områder, hvor sammenhængen mellem offentlige foranstaltninger og effekten på de sociale omgivelser er langt mere usikker. Ved evaluering kan virkningerne kortlægges, og man kan lære af erfaringer og etablere en mere effektiv indsats.

Evaluering skulle således kunne bidrage til en videnskabeliggørelse af politisk problemløsning. "The Moon and the Ghetto" er titlen på en amerikansk policy-analysebog (Nelson, 1977) og giver udtryk for den opfattelse, der var ret udbredt i 1960'erne og 1970'erne, at når man kan sende en mand til månen med anvendelse af naturvidenskab og ingeniørkunst, kan man også løse ghettoproblemerne i de amerikanske storbyer ved hjælp af samfundsvidenskab. Tidsmæssigt er der da også en klar sammenhæng, idet evalueringen i USA for alvor udviklede sig i kølvandet af de store velfærdsprogrammer i 1960'erne og 1970'erne.

Mens det i en amerikansk sammenhæng kan forekomme oplagt, at der er en sammenhæng mellem velfærdstatens og evalueringens udvikling, problematiseres denne sammenhæng straks ved en sammenligning af Danmark og USA. For selv om Danmark i umindelige tider har været et foregangsland, hvad velfærdsstatsforanstaltninger angår, og længe har haft en af de største offentlige sektorer i verden, er evalueringen først kommet meget sent igang her. Dette forhold gør det oplagt at lede efter andre forklaringer.

Politologer har angrebet det rationalistiske evalueringsperspektiv på to fronter. Dels har implementeringsforskningen vist, at de "interventioner" eller politiske foranstaltninger eller indsatser - som evalueres - ikke med rimelighed

kan betragtes som black boxes, som de klassiske evalueringsforskere gjorde det. Offentlige foranstaltninger gennemføres og administreres nemlig ved hjælp af politiske processer, som ofte forvrider den faktiske indsats i forhold til den planlagte (McLaughlin, 1985; Winter, 1990b). Dels har "knowledge utilization" forskningen påvist, at evalueringer sjældent anvendes instrumentelt i den politisk-administrative beslutningsproces som forudsat i det rationelle idealbillede, men snarere anvendes som politisk ammunition eller bruges oplysende eller begrebsmæssigt (Weiss, 1980; Albæk, 1988; Vedung, 1991). Politologer, som ynder en "alt er politik"-opfattelse, undres næppe meget herover, men "knowledge utilization"-litteraturen rejser alligevel et paradoks: Hvis politikere og embedsmænd ikke bruger evaluering instrumentelt, hvorfor bruger de så så mange penge på at gå i byen for at købe evalueringsprojekter?

Anvendes i stedet for et rationalistiske perspektiv et *konflikt- og forhandlingsorienteret* perspektiv på evaluering, gives der andre mulige forklaringer på den forskellige evalueringsudvikling i Danmark og USA. Evaluering kan således være et middel i politiske konflikter. Selv om politik ikke foregår rationelt på aggregeret niveau, har rationalistiske argumenter for det første en særlig legitimeringskraft. Når politiske reformforslag ledsages af evalueringsprojekter kan det være et led i en sådan legitimering af forslagene.

For det andet er der også mulighed for, at både tilhængere og modstandere af en reform kan få lavet hver deres evaluering, som kan legitimere hver parts synspunkter. Og selv om der kun laves én evaluering, kan formentlig både tilhængere og modstandere hente argumenter som politisk ammunition fra evalueringsresultaterne. For det tredje kan evaluering i forlængelse heraf fungere som et kontrolinstrument, der sikrer ikke kun tilhængere, men også modstandere en legitim adgang til at følge og blande sig i den måde, hvorpå en reform administreres, og en adgang til evt. at forlange politikken revideret. På en måde gør evalueringen de politiske beslutninger mindre definitive og sikrer aktørerne en fortsat access og kontrol med politikken og dens gennemførelse (Winter, 1983).

Et rationalistisk perspektiv på evaluering og dens rolle i velfærdsstatens udvikling forudsætter enten en betydelig homogenitet i de politiske forsamlinger, hvad angår politiske værdier og mål, eller meget stærke og homogene flertal. Det konflikt- og forhandlingsprægede evalueringsperspektiv indebærer derimod, at evalueringer vil blive mere efterspurgte, desto mere det politiske system er præget af konflikt og usikre flertalsdannelser.

Ud fra dette perspektiv sættes den forskellige efterspørgsel efter evaluering i USA og Danmark i et helt andet lys. Den sene evalueringsudvikling i Danmark kan således skyldes en langt større konsensus om velfærdsstatens udbygning og stærkere og mere stabil flertalsdannelse end i USA. Omvendt hænger væksten i dansk evaluering siden begyndelsen af 1980'erne - som vi skal se på senere - måske sammen med et øget konfliktniveau og en langt mere usikker flertalsdannelse.

Udbygningen af velfærdsstaten i USA har været præget af et langt højere konfliktniveau end i Danmark. Evalueringer har derfor ikke i samme grad som

i USA været set som et instrument, som en lovgivnings fortalere har kunnet benytte til at overbevise dens skeptikere, eller som en given offentlig foranstaltningens modstandere kunne bruge til at drage dennes resultater eller succes i tvivl.

Men ikke alene den brede opbakning bag den danske velfærdsstatslovgivning har været med til at mindske konfliktniveauet i dansk lovgivning. Også den politiske struktur giver færre muligheder for og incitamenter til konflikt. Den amerikanske magtadskillelse og føderalisme har i forening med den svage partidisciplin tendens til at institutionalisere konflikt og skabe mere autonome grupper af politiske aktører, som benytter evalueringer som et instrument i den politisk-administrative beslutningsproces. Eksempelvis ser den amerikanske kongres det som en af sine fornemste opgaver nøje at kontrollere og kritisere præsidenten og dennes ministre og embedsmænd, et forhold, der yderligere accentueres af, at præsidenten og flertallene i kongressens to kamre ofte tilhører forskellige partier. Evalueringsafdelingen i den amerikanske kongres' rigsrevision, the General Accounting Office, udgør således et af verdens største evalueringscentre.

I modsætning hertil har det klassiske parlamentariske princip og den stærke partidisciplin haft tendens til at mindske konflikten mellem lovgivende og udøvende magt i Danmark, og regeringens støttepartier har traditionelt gjort, hvad de kan, for at beskytte den mod kritik. Der har derfor ikke i Danmark været samme basis for en parlamentarisk efterspørgsel efter evalueringer som i USA. Og danske regeringer har ikke haft samme behov for at benytte evalueringer til at legitimere lovgivning over for den parlamentariske forsamling. Som tidligere omtalt har dansk velfærdsstatspolitik traditionelt været præget af bred politisk konsensus, og større, betydningsfuld lovgivning er som regel blevet til i et samarbejde mellem regering og opposition, idet det er normen at forsøge at opnå brede, midt-søgende kompromisser. Det har også betydet, at tidligere oppositionspartier kun sjældent ændrer tidligere vedtagne foranstaltninger, når de selv får regeringsmagten.

Hertil kommer, at den politiske konsensus og partidisciplin i vid udstrækning genfindes på de øvrige niveauer i det danske politisk-administrative system, hvor danske amter og kommuner ikke har samme incitament til at benytte evalueringer i deres konflikter med de statslige myndigheder, som både de kommunale og statslige niveauer i USA har haft det i forhold til den føderale regering. Derforuden gør man i den danske tilvejebringelse af service til borgerne ikke i nær samme udstrækning brug af ikke-offentlige organisationer, og følgelig har behovet for at udøve kontrol med sådanne organisationer ved hjælp af evalueringer været mindre.

Som nævnt kan evaluering betragtes som en feedback eller tilbagekoblingsmekanisme, der sikrer, at informationer om offentlige foranstaltningers gennemførelse, virkemåde og effekter når tilbage til relevante politisk-administrative beslutningstagere. Her har man i Danmark haft alternative tilbagekoblingsmekanismer, som kan have mindsket behovet for at tage et så forholdsvis omkostningsfyldt og "kunstigt" instrument som evalueringer i anvendelse.

Landets størrelse og homogenitet betyder, at den politisk-administrative afstand er ganske kort sammenlignet med andre lande. Ligeledes sikrer den lange tradition for inddragelse af interesseorganisationer og berørte parter i den offentlige beslutnings- og implementeringsstruktur, at der løbende gennem tætte netværk tilflyder beslutningstagerne informationer om offentlige foranstaltningers virkemåde fra mange og forskellige kilder. Erfaringsopsamlingen er således blevet præget af de forskellige interessenters erfaringer.

Generelt har tiltroen til, at samfundsmæssige problemer kan løses gennem velfærdsforanstaltninger været større i Danmark end i USA. Hertil kommer, at danskernes nok generelt er mere tilbøjelige til at acceptere og følge offentlig regulering og til at mene, at lovgivning implementeres som tilsigtet efter vedtagelsen i Folketinget. Under velfærdsstatens opbygning blev danskerne da heller ikke konfronteret med så åbenlyse fiaskoer, som amerikanerne oplevede i forbindelse med gennemførelsen af nogle af 1960'ernes og 1970'ernes storstilede velfærdsprogrammer. Også sådanne forhold kan have bevirket, at efterspørgslen efter evalueringer har været betydelig mindre eller helt fraværende i Danmark.

Endelig skal nævnes, at amerikanske velfærdsstatsforanstaltninger i deres design generelt har været bedre egnede til at blive underkastet evaluering end danske. Allerede i forbindelse med nogle af the Great Society's store velfærdsprojekter i 1960'erne opererede man i USA med tidsbegrænsede såkaldte "programmer", der meget målrettet tilsigtede at afhjælpe klart afgrænsede sociale gruppers specifikke behov og problemer. I modsætning hertil har den danske velfærdsstatslovgivning sjældent været tidsbegrænset, og den har haft mere omfattende målgrupper. Der har ofte været tale om større samlede lovgivningsområder. Det er først i de senere år, denne type programtænkning synes at have vundet indpas i dansk sammenhæng, endog i en grad så man eksempelvis i forbindelse med Det teknologiske Udviklingsprogram overtog den amerikanske terminologi - og den amerikanske uvane med at benytte for uindviede temmelig uforståelige forkortelser: TUP.

### **Anvendelsesorienteret samfundsforskning og dens organisering**

Der findes endnu ingen selv nok så rudimentære opgørelser over evalueringsforskningens art, struktur og omfang i dansk sammenhæng. Man kan imidlertid få et vist fingerpeg herom ved at betragte evalueringsforskningen som et led i en bredere bestræbelse på at øge den samfundsvidenskabelige forsknings umiddelbare nytteværdi i den politisk-administrative beslutningsproces, dvs. på at gøre samfundsforskningen mere "anvendelsesorienteret". Undertiden benyttes også "sektorforskning" som betegnelse for denne type forskning på grund af forskningens nære tilknytning til det politisk-administrative systems enkelte sektorer.

Den anvendelsesorienterede sektorforskning, herunder evalueringsforskningen, udføres kort fortalt på sektorforskningsinstitutterne, på de højere læreanstalter, i den offentlige forvaltning selv, og i de senere år også i konsulentfirmaer. Sektorforskningsinstitutterne er for langt hovedpartens vedkommende

oprettet specifikt med det formål at gennemføre forskning af umiddelbar nytte-værdi for politiske beslutningstagere. Af sådanne institutter kan nævnes Statens Byggeforskningsinstitut, SBI, oprettet i 1947 og med nære forbindelser til Boligministeriet; Socialforskningsinstituttet, SFI, oprettet i 1958 og med nære forbindelser til Socialministeriet og Arbejdsministeriet; og Amternes og Kommunerne Forskningsinstitut, AKF, som etableredes i 1975, og som har et tæt samarbejde med de to store kommunale organisationer, Kommunernes Landsforening og Amdsrådsforeningen, der også øver indflydelse på finansieringen via repræsentation i Den kommunale Momsfond, der yder basistilskud til AKF. Også Danmarks Pædagogiske Institut, DPI, Udviklingscenteret for Folkeoplysning og Voksenundervisning og Institut for Grænseregionsforskning udfører evalueringsopgaver. I grænseområdet mellem de offentlige forskningsinstitutioner og de private konsulentfirmaer findes også Center for Alternative Samfundsanalyser, CASA, der etableredes i 1986.

Tidligere blev der kun udført meget lidt anvendelsesorienteret samfundsvidenskabelig forskning på de højere læreanstalter. Den langt overvejende del af forskningen blev udført af fastansatte professorer eller lektorer. I 1970'erne skete der en ændring, som betød, at mere og mere forskning blev finansieret af kilder uden for universiteterne. Ofte ville forskningens sponsor være et ministerium eller en anden offentlig instans, der havde behov for at få belyst en specifik problemstilling eller evalueret en given offentlig foranstaltning. Ofte, men ikke eksklusivt, blev yngre, løst-ansatte forskere tilknyttet denne type forskning, og der opstod på daværende tidspunkt tydeligvis et vist interessefællesskab mellem sponsorerne og de yngre forskere: Førstnævnte havde problemstillinger, de ønskede belyst; sidstnævnte havde behov for at få finansieret forskning - og dermed deres løn - da udsigten til at få fastansættelse mere og mere fortegnede sig ude i en fjern horisont, efterhånden som økonomien strammede til inden for undervisningssystemet. Op gennem 1980'erne har der flere steder udviklet sig en forholdsvis fast tradition for at udføre sektorsponsoreret anvendelsesorienteret samfundsforskning på de højere læreanstalter. Dette er ikke mindst tilfældet på de danske universitetscentre, hvor den nære tilknytning til det regionale erhvervsliv og politiske system synes at have virket fremmende for denne udvikling. Som hovedregel entreres der på sektorforskningsinstitutterne ikke med de enkelte forskere, men med selve instituttet, mens det modsatte tit gør sig gældende på de højere læreanstalter.

Der findes ingen opgørelser over omfanget og organiseringen af den samfundsforskning, der udføres af den offentlige forvaltning selv. Af eksempler på offentlige myndigheder, der selv har udført evalueringer, kan nævnes Finansministeriet, Rigsrevisionen og Danida. Ligeledes findes der ingen informationer om konsulentfirmaernes rolle i denne sammenhæng, men det er oplagt, at disse har sat sig på en del af det ekspanderende evalueringsmarked i slutningen af 1980'erne og begyndelsen af 1990'erne.

På sektorforskningsinstitutterne og de højere læreanstalter vides den anvendelsesorienterede samfundsforskning at være øget betragteligt op gennem 1970'erne. I begyndelsen af 1980'erne blev der på initiativ af Statens sam-

fundsvidenskabelige Forskningsråd gennemført en undersøgelse, hvis formål var at kortlægge udviklingen i den anvendelsesorienterede samfundsforskning i Danmark. Undersøgelserapporten konkluderede, at den samfundsvidebaskabelige forskning i perioden 1976 til 1982 i alt øgedes med 14 procent, målt i form af fuldtidsbeskæftigede forskere. Det skyldtes så godt som udelukkende en vækst i den eksternt finansierede forskning, dvs. forskning finansieret af sponsorer uden for forskningsinstitutionerne. Selv om den eksternt finansierede forskning ikke generelt kan klassificeres som anvendelsesorienteret forskning eller som sektorforskning, gælder det dog for en betydelig del heraf. Rapporten konkluderede, at 40 pct. af al samfundsvidebaskabelig forskning i Danmark kunne defineres som anvendelsesorienteret i 1982 (Overgaard, 1984).

Det er imidlertid vanskeligt at sige, hvor meget af sektorforskningen, der kan klassificeres som evalueringsforskning. Nærværende artikel skal da også betragtes som en informeret, men trods dette impressionistisk, oversigt over evalueringsforskningens historiske udvikling i Danmark, baseret på erfaringer gjort af to observatører med nære kontakter til danske evalueringsmiljøer.

### **Udviklingen frem til 1980**

Som allerede nævnt kom evalueringsforskningen i gang betydeligt senere i Danmark end i USA og i mange andre vestlige lande, hvis forvaltningshistorie ellers har mange træk tilfælles med den danske. Hvis vi et øjeblik vender tilbage til denne artikels indledende evalueringsdefinition og nedtoner definitionens forskningsbaserede element, så har man længe i dansk forvaltningstradition kendt til systematiske vurderinger af den offentlige sektors funktionsmåde, som med lidt god vilje kan minde lidt om evaluering. Et eksempel er traditionel revision, som netop går ud på systematisk at undersøge, om offentlige instanser har benyttet flere midler end bevilget eller til andre formål end bevilget. Indtil for ganske nylig indskrænkede danske revisionsmyndigheder imidlertid deres virke til netop denne aktivitet. Offentlige foranstaltningers målopfyldelse, produktivitet og effektivitet blev i det store og hele ikke viet opmærksomhed.

Der findes andre eksempler på kontrol eller tilbagemelding af denne art i dansk forvaltningshistorie. I umindelige tider har borgerne haft adgang til at klage til højere myndigheder i det administrative hierarki eller til uafhængige ankeinstanser over den behandling, de har fået i det offentlige system. Uafhængige ankeinstansers afgørelser i sådanne enkeltsager offentliggøres ofte og får dermed en generel præjudicerende effekt i forhold til de myndigheder, i forhold til hvilke klagen er relevant. Det generelle billede i forbindelse med klagebehandling er imidlertid, at klageafgørelser ikke giver anledning til systematisk feedback om lovgivningens konsekvenser, og nok så vigtigt: Informationerne er yderst sjældent blevet samlet og brugt på systematisk vis, når en offentlig politik og dens implementering har skullet revurderes.

Siden 1955 har borgerne også kunnet klage til Folketingets ombudsmand over den behandling, de har fået hos offentlige myndigheder. Selv om ombudsmandens arbejde tager udgangspunkt i enkeltsager, sker det, at han

også udtaler sig om mere generelle administrative principper og beder forvaltningen ændre den måde, en lovgivning er implementeret på. Tamilsagen er et eksempel herpå.

Når regeringen har ment, at eksisterende lovgivning ikke længere fungerer på tilfredsstillende vis, har der været tradition for at lade embedsmænd udfærdige et udkast til ny lov eller til revision af den eksisterende. Alternativt er der nedsat et udvalg eller en kommission til at overveje behovet for at ændre gældende lovgivning og regler. Typisk vil berørte parter i den offentlige sektor og interesseorganisationerne, eksperter og undertiden politikere være repræsenterede i sådanne udvalg. Som udgangspunkt for forslag til lovændring foretager udvalgsrapporterne ofte en vurdering af den hidtidige lovgivnings effekter og implementering. Indtil 1980 findes der kun få eksempler på, at sådanne vurderinger baseredes på egentlige evalueringsresultater. I stedet byggede man på de erfaringer og den viden, de berørte parter kunne tilføre udvalget gennem deres udvalgsrepræsentanter. I den udstrækning, der blev benyttet videnskabeligt baseret analyse, skete det typisk i form af data fra Danmarks Statistik og i nogle tilfælde i form af tilgængelig forskning, der beskrev problematiske forhold på området.

Indtil slutningen af 1970'erne blev der gennemført få egentlige policy evalueringer i Danmark. Inden for visse fagdiscipliner, som eksempelvis det lægevidenskabelige, psykologiske og pædagogiske område, har man dog i længere tid kendt til social eksperimentering og evaluering. Men her indgik evaluering som et led i udviklingen af de faglige og videnskabelige discipliner og blev ikke benyttet til en systematisk vurdering af offentlige foranstaltninger eller bredere indsatsområder. Enkelte policy evalueringer blev dog gennemført. Eksempelvis foretog Socialforskningsinstituttet fra 1960 til 1978 i gennemsnit 1 - 2 evalueringer om året<sup>2</sup>. Instituttet var også i denne periode det ledende forskningsinstitut med hensyn til gennemførelsen af anvendelsesorienteret samfundsvidenskabelig forskning generelt.

At der ikke har været gennemført megen policy evaluering i Danmark før 1980, betyder ikke, at der ikke har eksisteret feedback-mekanismer, der har givet offentlige beslutningstagere viden om den offentlige sektors funktionsmåde. Men generelt er sådan information blevet indsamlet på betydelig mere usystematisk vis, end det normalt kræves i forbindelse med evalueringspraksis. De er for eksempel indsamlet i form af arbejds erfaringer, offentlige debatter og kontakter til borgere, interesseorganisationer og lovgivere. Det er dog vores generelle indtryk, at der ved vedtagelse af ny lovgivning eller lovrevisioner meget ofte har været et meget usikkert vidensgrundlag for, hvordan lovgivningen hidtil har virket (Wilhelm, 1985).

## **1980 og frem**

Siden begyndelsen af 1980'erne er der sket en betydelig vækst i antallet af evalueringer udført både i den offentlige forvaltning selv, på forskningsinstitutioner og i de senere år også hos konsulentfirmaer. Interessen for evaluering blev yderligere styrket hen mod tiårets slutning, hvor kredsen af evalueringsinteres-



serede offentlige myndigheder udvidedes mærkbart - et forhold, der også manifesterede sig i en betydelig øgning i midler afsat til evaluering af offentlige foranstaltninger og indsatsområder.

Generelt har såvel statslige som kommunale danske myndigheder kun i begrænset omfang selv gennemført evalueringer. Mest udbredt har denne praksis været hos statslige myndigheder med tværgående, koordinerende funktioner. Således er Budgetdepartementet begyndt at inkludere visse informationer om myndighedernes præstationer og produktivitet i budgettet, og departementet har publiceret nogle mindre evalueringer af især produktivitet og inden for forskellige offentlige områder. Departementet er klart interesseret i en større anvendelse af produktivitetsmålinger i den offentlige sektor.

Det gælder også Administrations- og Personaledepartementet, som ligeledes hører under Finansministeriet. Det har publiceret flere produktivitsanalyser og vejledninger om sådanne undersøgelser for andre offentlige myndigheder. Siden Administrationsdepartementets oprettelse i 1973 har departementet desuden foretaget en form for evalueringer af forskellige policy områder, de såkaldte turnusgennemgange. Eksempelvis har man evalueret Civilforsvarets ressourceforbrug, det sociale ankesystem og arbejdsløshedskassernes administration. Indtil midten af 1980'erne udførte man typisk én til to gennemgange om året. Disse munder normalt ud i en analyse af mulighederne for at øge effektiviteten på det pågældende område ved at reducere omkostningerne. Turnusgennemgange består primært af organisationsanalyser af forskellig art, og de benyttede metoder har ikke været særligt avancerede eller systematiske. Hovedparten af analyserne offentliggøres, men effekten heraf har været vanskelig at spore, da departementets rekommendationer som oftest møder stærk institutionel modstand hos de berørte myndigheder. Siden midten af 1980'erne har departementet lagt større vægt på at være konsulent for ministerierne snarere end ekstern kritiker og økonomisk vagthund.

Med inspiration i den metodiske udvikling af forvaltningsrevision, som har fundet sted i eksempelvis USA, Canada og Sverige, er Rigsrevisionen i de senere år begyndt at supplere den traditionelle revision med analyser, som relaterer omkostninger til præstationer og effekter. Produktivitsstudier tillægges stadig større vægt, og der er gennemført enkelte, men ikke videre sofistikerede evalueringer, som fokuserer på effekter og omkostninger. Imidlertid er det fortsat den traditionelle revision, der dominerer Rigsrevisionens arbejde. Som følge af Rigsrevisionens uafhængige stilling i forhold til regeringen har nogle af dens analyser været i stand til at stimulere en bredere debat i offentligheden og i Folketinget.

Også Indenrigsministeriet har vist en større interesse for produktivitsanalyser. Ministeriet har nu i nogle år offentliggjort en årlig rapport om nøgletal for forskellige kommunale funktioner. Dette tillader blandt andet, at der kan foretages grove sammenligninger af kommunale forvaltningers produktivitet. Mens ministeriet endnu havde sundhedsområdet tilknyttet sit ressort, tog det også initiativ til at sammenligne produktiviteten på danske hospitaler.

En vigtig undtagelse fra reglen om, at det især er de tværgående, koordine-

rende myndigheder, der selv gennemfører evalueringer, er Danida, der i det hele taget udmærker sig ved at være den offentlige myndighed, der tidligst og mest systematisk har gennemført omfattende evalueringsinitiativer inden for sit område. Danida har endog oprettet en særlig evalueringsafdeling og et netværk af eksterne konsulenter til at evaluere et vidt spektrum af hjælpeprojekter. Den forholdsvis tidlige og omfattende brug af evaluering hænger for det første sammen med Danidas placering i et internationalt miljø, hvor systematisk evaluering af hjælpeprojekter til tredje-verdens lande betragtes som en naturlig rutine, således som det eksempelvis er tilfældet i Verdensbanken. For det andet accepterede Danida på grund af den internationale tilknytning program- og projektankegangen tidligere end de fleste andre danske myndigheder. Bistandsprojekternes målrettede og tidsbegrænsede indsats i forhold til som oftest klart afgrænsede sociale gruppers specificerede behov og problemer gør dem til ideelle evalueringsobjekter. De anvendte evalueringsmetoder har imidlertid generelt været svage, og evalueringerne baseres på projektbesøg af en enkelt til nogle få ugers varighed. Evalueringerne har indtil fornylig ikke været videre kritisk anlagt og sjældent været publiceret. Danidas evalueringspraksis er da også af Rigsrevisionen blevet tildelt en del kritiske bemærkninger (Wandel, 1988).

På sektorområderne er det yderst sjældent, at offentlige myndigheder selv udfører evalueringsundersøgelser. I stedet enters der med sektorforskningsinstitutter, konsulentvirksomheder og i nogen grad med de højere læreanstalter, ikke mindst universitetscentre. I slutningen af 1980'erne havde Socialforskningsinstituttet således forøget antallet af gennemførte evalueringer fra 1-2 til 3 om året. I 1991 var tallet steget til omkring en halv snes. AKF gennemførte i slutningen af 1980'erne omkring fem evalueringer og produktivetsstudier pr. år. I 1991 var antallet af evalueringer steget til ca. 15<sup>3</sup>.

De fleste evalueringer har fundet sted inden for de store, tunge sektorområder, det vil sige social-, arbejdsmarkeds-, og undervisningsområdet. Men også inden for andre sektorer er det blevet mere almindeligt at gennemføre omfattende evalueringer. Det gælder eksempelvis sundhedsområdet, det teknologiske område og landbrugsområdet. Evalueringerne finansieres fortrinsvis af de statslige myndigheder og undertiden af de to store kommunale organisationer, om end enkelte kommuner også har finansieret evalueringer. Ligeledes er EF i stigende grad blevet sponsor for danske evalueringer - blandt andet af EF's socialfondsstøtte til arbejdsløshedsbekæmpelse i Danmark. Der er tale om et marked, der ikke mindst har bragt danske konsulentfirmaer på banen.

Den voksende arbejdsløshed, der fulgte i kølvandet på oliekrisen i 1973/74, har været forsøgt afhjulpnet gennem en række forskellige policy-instrumenter, de fleste imidlertid med temmelig beskedne effekter. Arbejds- og Undervisningsministerierne har været ansvarlige for disse tiltag og har i tidens løb finansieret mange samfundsvidenskabelige studier af arbejdsløshedsproblemer og evalueringer af beskæftigelsesforanstaltninger.

Socialministeriet var kun involveret i få evalueringer i perioden 1960 til 1979, som for størstedelens vedkommende blev udført af Socialforskningsin-

stituttet. En omfattende evaluering af hele det sociale systems funktionsmåde blev dog udført i slutningen af 1960'erne som en del af beslutningsgrundlaget for reformens gennemførelse (blandt andet Rold Andersen, 1970). Ligeledes gennemførtes i begyndelsen af 1980'erne en omfattende evaluering til at vurdere reformens resultater (Plovsing, 1985).

De senere års betydelige vækst i antallet af evalueringer på det sociale område synes at relatere sig til en øget brug af forsøgsordninger, således som tilfældet også har været inden for andre sektorer, eksempelvis sundheds-, undervisnings- og arbejdsmarkedsområdet, det kommunale område og inden for miljø og teknik. Det forekommer undertiden, som om forsøgstankegangen blev det administrative modeløsen fra midten af 1980'erne som erstatning for 1970'ernes undertiden grandiose planlægningsinitiativer, skarpt forfulgt - og ofte støttet - af decentraliseringstankegangen. Forsøgsordningerne har typisk været centralt initierede, men lokalt implementerede, og som oftest, men ikke undtagelsesfrit, har de statslige myndigheder støttet den lokale eksperimentering finansielt. Forsøgsordninger har karakter af overordnede policy programmer, inden for hvilke en bred vifte af forskellige eksperimentelle projekter kan finde sted. De afsatte forsøgsmidler har undertiden været af betydelig størrelse. Blandt de større forsøgsprogrammer kan nævnes: Socialstyrelsens Forsøgsmidler, SOFUS, fra midten af 1980'erne, det Sociale Udviklingsprogram, SUM, der blev initieret af Socialministeriet i 1988 med en samlet statslig forsøgspulje på 350 millioner kroner; Folkeskolens 7-punktsprogram, forsøg med beskæftigelses- og uddannelsesforanstaltninger for ledige, Det teknologiske Udviklingsprogram, TUP.

I forbindelse med mange af disse forsøgsordninger har der udviklet sig en tradition for, at der særskilt afsættes undertiden endog betydelige midler til evaluering. Således blev der givet ca. 15-20 millioner kr. til evaluering af SUM, svarende til 5-6 pct. af den samlede bevillingssum (Winter, 1990a; Jensen, 1992). I 1980'erne synes der at have udviklet sig en form for tæt sammenkobling mellem offentlige forsøgsordninger og evaluering, som kan minde om den tilsvarende kobling i den amerikanske forsøgs- og evalueringstradition, der netop havde sit udspring i 1960'ernes store eksperimentelle velfærdsprogrammer.

Evaluering af offentlige forsøgsordninger foretages dog ikke undtagelsesfrit. Et interessant eksempel på en centralt initieret, lokalt implementeret forsøgsordning, i forbindelse med hvilken der hverken blev afsat statslige midler til ordningens gennemførelse eller til en større evaluering heraf, er det danske frikommuneforsøg, der initieredes af Indenrigsministeriet i 1986 (Albæk, under udgivelse A).

Det er imidlertid ikke alene i forbindelse med offentlige forsøgsordninger, der kan spores en stigning i evalueringsaktiviteten. Et eksempel på introduktionen af en efterhånden betydelig evalueringsaktivitet på et område, hvor der ikke tidligere har været tradition herfor, er det forsknings- og undervisningspolitiske område. Undervisningsministeriet og de forskningspolitiske råd har stået som initiativtagere til sådanne evalueringer. Der er eksempler på, at hele

forskningsområder er blevet evalueret, for eksempel sundhedsforskningen ("SOFIE"), det bioteknologiske område (Hansen, 1991), jordbrugsforskningen og miljøforskningen. Sidstnævnte evaluering har været medvirkende til igangsættelse af Det strategiske Miljøforskningsprogram, som er et meget omfattende forskningsprogram, der både indeholder grundforskning og mere anvendelsesorienteret forskning, herunder også med evalueringselementer. Ligeledes synes en evaluering af den samfundsvidenskabelige forskning, herunder sektorforskningen, at være under optræk.

Desuden er hele forskningsinstitutter blevet evalueret, for eksempel Sociologisk Institut ved Københavns Universitet, ligesom de videregående uddannelsers styreform generelt har været underkastet evaluering. Uddannelsesinstitutionerne er blevet opfordret til at indføre undervisningsevaluering, og visse steder er undervisningskvaliteten i folkeskolen og gymnasierne blevet evalueret; der er blevet oprettet et særligt center under Undervisningsministeriet til at forestå undervisningsevaluering af de videregående uddannelser. I nogle af de nævnte eksempler på evaluering inden for det undervisnings- og forskningspolitiske område er udenlandske eksperter blevet bedt om at være behjælpelige med at gennemføre evalueringerne. Inddragelsen af udenlandske eksperter har officielt dels været motiveret i et ønske om at få en uvildig vurdering af dansk forskning og undervisning, dels at få introduceret internationale evalueringsstandarder, eksempelvis OECD standarder, der gør det muligt meningsfuldt at sammenholde danske med udenlandske forhold.

### **Hvorfor øget evalueringsaktivitet?**

Er den øgede danske evalueringsaktivitet blot et udtryk for, at man i Danmark har (gen)opdaget og kopieret en administrativ mode, der i mange andre lande efterhånden er blevet fast administrativ rutine? Mange administrative tiltag introduceres oplagt - og ofte med god grund - som kopier af udenlandske nyskabelser. Et godt eksempel er de nordiske frikommuneforsøg, der i midten af 1980'erne med ganske få års mellemrum blev indført i alle nordiske lande bortset fra Island (Ståhlberg, 1990). Også kommunesammenlægningsreformerne i 1960'erne og 1970'erne skete under gensidig international inspiration. Det gælder også i betydelig grad budgetlægningssystemer. De borgerlige regeringer siden 1982 har også, men stort set forgæves søgt at kopiere amerikanske og engelske privatiseringstiltag.

Også evalueringsforskningen hørte man om i Danmark allerede i 1960'erne og 1970'erne, så hvorfor først i 1980'erne kopiere en så gammel original? Hvis vi går tilbage til denne artikels indledning, vil vi kunne finde nogle af forklaringerne på, hvorfor den øgede evalueringsaktivitet ikke alene kan betragtes som en sent ankommet modedille - selv om det ikke skal skjules, at megen evaluering også er et spørgsmål om mode. Der er i perioden siden 1960'erne sket en række ændringer i det danske politisk-administrative system, som i dag tillader evalueringsforskningen at spille en mere oplagt rolle heri end tidligere.

Der er for det første ingen tvivl om, at det politiske konfliktniveau er steget. Den danske velfærdsstat er i dag under et pres, som man ikke kendte tidligere.

Det skyldes flere forhold. Flere og flere opgaver er i tidens løb blevet defineret som offentlige ansvarsområder, hvilket i sig selv har været med til at øge de offentlige udgifter, men også yderligere til at øge befolkningens forventninger til og krav om offentlig service. Ændringer i befolkningssammensætningen har øget udgifterne til nogle områder (for eksempel ældreområdet), mens det tilsvarende har været vanskeligt at reducere udgifter på andre områder, hvor den demografiske udvikling ellers synes at gøre en reduktion oplagt (for eksempel folkeskoleområdet). Understøttelsen til arbejdsløse og renterne for et tidligere statsligt overforbrug gør et markant indhug i de offentlige kasser. Samtidig oplever man heller ikke længere samme stigning i indkomsterne og dermed i skattegrundlaget.

Der er med andre ord opstået en situation, hvor der er større konflikt både om de offentlige udgifters størrelse og deres fordeling mellem formål og myndigheder. Det er derfor i dag vanskeligere at minimere politiske konflikter ved hjælp af inkrementelle budgetforøgelser end tidligere. Konfliktniveauet imellem og inden for de enkelte sektorer er da også øget.

Hertil kommer, at den politiske konsensus i Folketinget synes at være mindsket under 1980'ernes udgiftsbevidste, men parlamentarisk svage mindretalsregeringer, der gang på gang har måttet lide politiske nederlag, uden at den selv eller oppositionen har villet eller turdet tage den traditionelle konsekvens heraf i form af mistillidsvotum og/eller udskrivelse af nyvalg. Det har dog medført en øget kontrolaktivitet fra Folketingets side. De mange kommissionsdomstole og dommerundersøgelser, som søger at kortlægge de retlige aspekter af konkrete administrative begivenhedsforløb er et eksempel herpå. Rigsrevisionens overflytning til Folketingets regi samt overvejelserne af at give Folketinget et lovvurderingsråd, der er uafhængigt af regeringsapparatet, er andre eksempler.

I konfliktladede perioder kan evalueringer som nævnt være et nyttigt politisk instrument for mange aktører, og det synes oplagt, at aktørernes placering i det politisk-administrative system reflekteres i deres interesse for konkrete evalueringsformer. De tværgående, koordinerende ministerier med et overordnet økonomisk ansvar har især haft opmærksomheden rettet mod produktivitet- og effektivitetsmålinger, mens sektorministerierne i særlig grad har fokuseret på proces- og målopfyldelsesevalueringer. Men også sektorministeriernes kontrol- og styringsfunktioner afspejler sig undertiden i deres konkrete evalueringsinteresse, således som det eksempelvis kommer til udtryk i flere af Undervisningsministeriets evalueringsskemaer.

For det andet har der både udviklet sig en større skepsis med hensyn til de offentlige opgaveløsningers effektivitet på flere områder. Der er kommet langt flere offentligt kendte eksempler på, at velfærdsstatens politiske indgreb ikke altid fungerer efter hensigten. Indsatsen mod arbejdsløsheden er et eksempel. Miljøbeskyttelsen af de danske vandløb, søer og havområder et andet og behandlingen af narkotikamisbrug et tredje.

Politikfiaskoer og en øget ressourcebevidsthed er blevet ledsaget af et stigende ønske om at forsøge sig med nye løsninger, omstilling og fornyelse inden

for den offentlige sektor. Dette blev ikke mindst sat højt på den politiske dagsorden af den første Schlüter-regering, da den overtog regeringsmagten i 1982 og lancerede en kampagne for at modernisere den offentlige sektor, en kampagne, der flere gange siden er forsøgt revitaliseret. Men omstilling og fornyelse møder som oftest institutionel modstand uanset forslagsstillernes politiske farve, da vante rutiner og tankemåder, velerhvervede rettigheder og gældende magtkonstellationer nærmest per definition forstyrres.

En måde at mindske anticiperet modstand på er netop ved at præsentere ideer til fornyelse og omstilling gennem forsøg - noget tentativt og tidsbegrænset, der i princippet fortsat er til forhandling. Forsøgsordninger har yderligere den fordel, at de kan virke motiverende: Det er for de enkelte medarbejdere og andre som oftest mere spændende at deltage i et forsøg - noget nyt og uprøvet, hvis udfald blandt andet afhænger af ens eget engagement - end det er blot at implementere en revision af gældende regler og praksis (Winter, 1983), og forsøgsinitiativer og -deltagelse er en ny karrierevej. De nordiske kommunale chefer har da karakteristisk nok også selv betragtet murbrækker- og motivationsfaktorerne som nogle af de vigtigste begrundelser for at iværksætte forsøgs- og omstillingsprojekter i deres kommuner (Adamsen *et al.*, 1990; Albæk, under udgivelse B). Forsøgsideen har nu vundet sådan indpas i den administrative tankegang, at lovgivningen på flere områder har indført generelle forsøgsparagraffer, der letter adgangen til at dispensere for gældende regler i forbindelse med forsøgsvirksomhed (Albæk, under udgivelse A).

De fleste forsøgsordninger har karakter af et generelt forsøgsprogram, under hvilket der kan gennemføres en bred vifte af konkrete forsøgsprojekter, som er forholdsvis overskuelige at evaluere. Man har derfor i 1980'erne udviklet en forsøgs- og evalueringssituation, der kan minde om forsøgs- og evalueringstraditionens indførelse i den amerikanske forvaltning i forbindelse med 1960'ernes store eksperimentelle velfærdsprogrammer. Der er da heller ingen tvivl om, at forsøgstraditionen i Danmark har virket ansporende for evalueringsskningens vækst på samme måde, som det var tilfældet i USA.

Men der er samtidig væsentlige forskelle i den historisk-politiske kontekst, i hvilken forsøgstankegangen fandt fodfæste i de to lande, hvilket også afspejler sig i evalueringsskningens karakter. Hvor man i forbindelse med 1960'ernes store velfærdseksperimenter i USA i vid udstrækning trådte jomfruelig grund og gennem egentlige, videnskabeligt tilrettelagte og afprøvede eksperimenter forsøgte at opnå indsigt i alternative foranstaltningers effekter, har de danske forsøgsordninger i højere grad været motiveret i et ønske om at nedbryde og ændre vante rutiner og tankegange gennem en betydelig decentraliseret og løst prøv-og-fejl logik uden præcise forhåndsforventninger om, hvad forsøgene skulle munde ud i.

Det skal imidlertid nævnes, at evaluering hverken er en nødvendig eller en tilstrækkelig betingelse for, at ideer fra forsøgsordninger dissemineres og benyttes, hvilket illustreres af det danske frikommuneforsøg. Kun ganske få af frikommuneforsøgsprogrammets konkrete projekter er blevet evalueret, mens en generel evaluering af programmet har mødt en stærk ministeriel modstand

helt fra start. Alligevel kunne ni ministerier i 1991 rapportere, at frikommuneforsøget havde inspireret til lovændringer i ikke færre end 24 tilfælde (Albæk, under udgivelse A). Det er i øvrigt interessant, at forsøgevalueringer og andre evalueringer er vundet frem i løbet af 1980'erne, trods den borgerlige regerings betydelige modstand mod evaluering i starten. Den mente, at det drejede sig om handling - ikke forskning - og at erfaringsudvekslingen om forsøgene kunne ske uden evaluering.

For det tredje skal det til sidst nævnes, at også evalueringsforskningens udbudsside har ændret sig, idet produktionen af samfundsvidenskabelige kandidater er vokset markant op gennem 1970'erne og 1980'erne, hvilket både har stimuleret administrationens efterspørgsel efter evaluering og forøget udbuddet, da en del af disse kandidater har fundet beskæftigelse på forskningsinstitutioner og i konsulentfirmaer, der har reageret på de forbedrede markedsbetingelser for udførelse af evalueringsforskning.

### **Evalueringsens forskningsmæssige karakter og kvalitet**

Som nævnt er evalueringsforskningen internationalt meget omfattende. Som forskningsdisciplin er det især de metodologiske aspekter, der samler, og der er en særdeles omfattende metodologisk litteratur. Som internationale eksempler herpå kan nævnes bøger af Rossi og Freeman (1989) og Evert Vedung (1991). Danske eksempler er Kjeld Møller Pedersen (1979), Benedicte Madsen (1984) og Lis Adamsen m.fl. (1986). Der er derimod ikke nogen egentlig international evalueringsteori. Som nævnt har især politologer dog bidraget med teoriudvikling inden for to tilgrænsende områder, nemlig for det første implementeringsforskningen (Winter, 1985; 1990b; Vedung, 1991), der har et betydeligt slægtskab med procesevaluering, men som har et langt mere eksplicit teoretisk udgangspunkt end procesevalueringen, der ofte gennemføres ved hjælp af rent deskriptive eller fænomenologiske metoder. For det andet er der en vis teoriudvikling i forskningen vedrørende anvendelse af evaluering i den politisk-administrative beslutningsproces (Weiss, 1980; Albæk, 1988).

Internationalt er der imidlertid også en del evalueringsforskere, der arbejder med teoridannelser, som ikke drejer sig om evaluering i sig selv, men som omfatter teori inden for de forskellige substantielle problemområder, den enkelte forsker arbejder med, for eksempel behandling af narkotika- eller spiritusmisbrug, pædagogik, arbejdsløshed og immigration. Teorier om de kausale forhold på det pågældende problemområde vil ofte være særdeles frugtbare ved opstilling af evalueringsdesignet og ved tolkningen af resultaterne (Rossi og Freeman, 1989).

Det meste af den danske evalueringsforskning er imidlertid relativt svagt udviklet, både metodisk og teoretisk. Til effektundersøgelserne er der som regel blevet anvendt det svageste evalueringsdesign, nemlig den simple efterundersøgelse, som måler deltagergruppens situation efter foranstaltningens afslutning, og som ikke giver nogen sikkerhed for, at en eventuel ændring i deltagerens situation skyldes den pågældende foranstaltning. Når evalueringer således viser, at en - ofte beskeden - del af deltagerne i beskæftigelses- eller

uddannelsesforanstaltninger for ledige kommer i ordinær beskæftigelse bagefter, ved man ikke, hvor mange der ville være kommet det alligevel.

Når danske evalueringsundersøgelser næsten aldrig bruger det stærkeste effektevalueringdesign, det klassiske eksperiment med tilfældigt udvalgt eksperiment- og kontrolgrupper, hænger det både sammen med etiske betænkeligheder ved en sådan forskelsbehandling og med, at evalueringerne som regel har vedrørt politiske foranstaltninger, hvor hele populationen har et retskrav på at deltage. For eksempel kan man ikke lave nogen ordentlig kontrolgruppe ved evaluering af arbejds- eller uddannelsestilbud til langtidsledige, fordi ordningen har omfattet alle. En anden mangel ved mange danske evalueringsundersøgelser er anvendelse af svage statistiske metoder. Der er som regel anvendt simple krydstabeller, hvor der kun kontrolleres for ganske få variable, skønt man ofte må have en teoretisk forventning om, at flere faktorer har betydning. Når effekten af forskellige foranstaltningstyper sammenlignes, er der i flere undersøgelser ikke blevet kontrolleret for den forskellige deltager-sammensætning. I flere evalueringer i 1980'erne er det således blevet hævdet, at private jobtilbud giver en bedre beskæftigelsesvirkning end offentlige beskæftigelsestilbud, uden at det er kontrolleret, hvor meget af denne forskel der skyldes, at kun de "bedste" ledige udvælges til private arbejdstilbud.

I de seneste år har man dog set danske evalueringer, der betjener sig af kvasi-eksperimentelle metoder, og som i højere grad anvender multivariate statistiske metoder, som multipel eller logistisk regression (Jensen *et al.*, 1991; Pilegaard Jensen *et al.*, 1991; Thaulow og Anker, 1992). Der er således kommet evalueringer, der enten opererer med kontrolgrupper (for eksempel Jensen *et al.*, 1991) eller i det mindste kan klarlægge forskellen i effekten af forskellige foranstaltningstyper over for forskellige typer af deltagere (for eksempel Aarkrog *et al.*, 1991; Pilegaard Jensen *et al.*, 1991; Thaulow og Anker, 1992). Der er også fra flere forskningsmiljøer fremsat forslag om forsøg med tilknyttede evalueringer med kontrolgruppedesign.

Ud over de nævnte evalueringsdesigns og -metoder, som lægger hovedvægten på kvantitative metoder, bør det nævnes, at der også i Danmark findes en del procesevaluering, der i langt højere grad anvender kvalitative metoder. Formålet hermed er ikke kun eller primært at måle effekten af den offentlige indsats, men at belyse og vurdere den proces, hvormed indsatsen føres ud i livet. Størstedelen af disse analyser er rent beskrivende, eller har et erklæret fænomenologisk perspektiv (for eksempel Flyvberg, 1991).

Ligesom det meste af den øvrige danske samfundsvidenskabelige sektorforskning er den teoretiske forankring af evalueringsforskningen gennemgående meget beskeden (Winter, 1990a; 1991; Hansen, 1992). Det gælder for det første - i lighed med den internationale praksis - selve de anvendte evalueringsperspektiver, hvad enten undersøgelserne anvender et eksperimentelt, kvasi-eksperimentelt, simpelt efterundersøgellesdesign eller procesevaluering. Der er uhyre sjældent opstillet en teoretisk ramme, der styrer undersøgelserne. Derimod er der - ligeledes i overensstemmelse med en international tradition - forsøgt en større grad af teoristyring i de få danske politologiske implemente-



ringsanalyser (for eksempel Jørgensen, 1985; Dalsgaard, 1985; Skou Andersen, 1989; Winter, 1985; 1990b; Pilegaard Jensen *et al.*, 1991), som på dette punkt adskiller sig fra størstedelen af den beslægtede procesevaluering. I lighed med den internationale situation er teoriudviklingen inden for implementeringsforskningen dog relativ beskedent.

For det andet bygger danske evalueringer sjældent på teori om de sociale processer og årsagssammenhænge på de problemområder, som de evaluerede foranstaltninger sigter mod at påvirke. Her kan danske evalueringsforskere afgjort lære noget af den internationale evaluerings- og anvendelsesorienterede samfundsvidenskabelige forskning. De danske evalueringsstudier er som regel meget enkeltstående, og graden af videnskummulation er beskedent. Der er dog gjort enkelte forsøg på mere sammenfattende analyser (for eksempel Adamsen og Fisker, 1986; Adamsen *et al.*, 1990/91; Winter, 1988). Ligeledes relaterer evalueringsundersøgelserne sig kun i meget beskedent omfang til international litteratur, ligesom danske evalueringsresultater sjældent publiceres internationalt (Winter, 1990a; 1991; Hansen, 1992). Der er ganske vist taget visse skridt til at forbedre sektorforskningen, ikke mindst gennem etablering af ph.d. forløb.

De hidtidige forhold reducerer imidlertid efter vores opfattelse kraftigt værdien af den danske evalueringsforskning. Desværre findes der ikke nogen større, generaliserbare undersøgelser af anvendelsen af evalueringer i den politisk-administrative beslutningsproces i Danmark. Der foreligger, så vidt vides, kun enkelte undersøgelser, der berører anvendelsen af evalueringsresultater i forbindelse med et par forsøgsordninger, ungdomsgarantiforsøget i 1980-82 (Winter, 1983) og en række sociale forsøg (Adamsen *et al.*, 1990), samt vedrørende det bioteknologiske forskningsprogram (Hansen, 1991). Resultaterne heraf afviger ikke markant fra de amerikanske resultater fra "knowledge utilization"-litteraturen, hvoraf den dominerende del som nævnt hævder, at evalueringsresultater sjældent anvendes direkte og instrumentelt, men oplysende og som politisk argumentation og ammunition (Weiss, 1980; Albæk, 1988). Men selv om en anvendelse i den politiske argumentation ved første øjekast kan forekomme et pauvert resultat, kan evalueringer dog i den sidste ende påvirke de politisk-administrative beslutninger ved at indsnævre argumentationsrummet ved at reducere antallet af argumenter og løsninger, der kan holde vand i en offentlig debat. Det er for eksempel vanskeligt at argumentere for, at UTB-kurser skal være obligatoriske for langtidsledige, når gentagne evalueringer viser, at kun omkring 10 pct. af deltagerne kommer i normalt arbejde bagefter (Aarkrog *et al.*, 1991; Pilegaard Jensen *et al.*, 1991; Thaulow og Anker, 1992).

Det skal dog nævnes, at spørgsmålet om instrumentel brug er omstridt (Rossi og Freeman, 1989; Vedung, 1991), og at nogle forskere hævder, at der finder en vis instrumentel brug sted, navnlig på lavere organisatorisk niveau og lokalt i samspillet mellem forskere og lokale brugere. Vi ved desværre for lidt om, hvordan evalueringsresultater anvendes i Danmark. Det kan ikke udelukkes, at det fungerer anderledes end i USA.

Tager man imidlertid udgangspunkt i den dominerende amerikanske opfattelse, forekommer det oplagt, at både en oplysende og argumentatorisk, overtalende brug af evalueringsresultater fremmes, hvis undersøgelserne ikke blot anvender solide videnskabelige metoder, men også er forankret i teori og begreber, som tillader politiske og administrative aktører at se problemer, politisk-administrative processer, mekanismer og barrierer samt løsninger i et nyt og frugtbart lys (Winter; 1990a; 1991). Selv om undersøgelser viser, at samfundsvidenskabelige forskningsresultater faktisk anvendes på denne måde (Weiss, 1980), kan man naturligvis ikke forvente, at brugerne af sig selv stiller krav om teoretisk baserede undersøgelser, og det gælder vel især - når de som i Danmark - ikke er forvænt med teoretisk perspektiverede undersøgelser. Brugernes evalueringsspørgsmål er ofte formuleret særdeles snævert og beskrivende i relation til den konkrete foranstaltning, der skal evalueres, hvilket bestemt ikke stimulerer en mere generaliserbar forskning, der har en mere blivende værdi. Der er også eksempler på, at evalueringsspørgsmålene - som ved SUM-evalueringen - kan være formuleret meget udflydende, og hvor det faktisk kan være svært at se, hvad evalueringen skal bruges til (Winter, 1991).

For at sætte brugerne i stand til at formulere mere relevante forskningsspørgsmål, er der efter vores opfattelse behov for en bedre og uafhængig forskningsmæssig rådgivning af myndighederne omkring evalueringsopgavernes afgrænsning og udformning. Bedre evalueringsforskning er imidlertid langt fra kun brugernes ansvar. Den danske sektorforskning bør selv blive langt bedre til at udnytte de muligheder, der trods alt er, for at give evalueringerne og den øvrige sektorforskning et indhold, der både er mere teoretisk relateret, og som kan stimulere en begrebsmæssig eller oplysende brug (Winter, 1990a; 1991). Forskningsinstitutionerne bør også i højere grad give brugerne et modspil, der kan medføre en mere interessant forskning, og de bør være mere kritiske med hensyn til, hvilke opgaver de påtager sig.

## **Konklusion**

Evaluering som en forskningsdisciplin er udviklet relativt sent i Danmark. Det er først rigtigt sket i løbet af 1980'erne, og udviklingen har for alvor taget fart omkring 1990. Udviklingen hænger formentlig både sammen med et relativt begrænset udbud i 1960'erne og 1970'erne af samfundsvidenskabelige kandidater med kendskab til evalueringsmetoder og med et politisk system, der indtil begyndelsen af 1980'erne var relativt homogent og konsensusorienteret. Der var ikke så stærke konflikter mellem Folketing og regering, og evaluering var ikke så nærliggende som politisk legitimation og ammunition. Måske hænger den større brug i dag sammen med ændringer i disse parametre. Evalueringsforskningens teoretiske og internationale forankring har været relativt beskednen i Danmark, og der har været anvendt relativt svage evalueringsmetoder, der dog er under udvikling.

## Noter

1. Nærværende artikel er en opdateret og udvidet version af Erik Albæk og Søren Winter, "Evaluation in Denmark: The State of the Art" i Ray C. Rist (ed.), *Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations*, New Brunswick: Transaction Books, 1990. Vi takker Inge Mærkedal, Socialforskningsinstituttet, og Olaf Rieper, AKF, for kommentarer til et manuskriptudkast. Forfatterne bærer dog alene ansvaret for artiklens indhold.
2. Optælling baseret på Socialforskningsinstituttets publikationslister og årsberetninger. Det skal anføres, at det ud fra publikationslister er overordentligt vanskeligt at foretage en optælling af evalueringsrelevante publikationer, blandt andet fordi ordet "evaluering" ikke anvendes på dette tidspunkt. De fleste titler vedrører undersøgelser af forskellige sociale forhold blandt udvalgte målgrupper, for eksempel de ældres levevilkår. Der kan være evalueringslementer indeholdt i flere af denne type publikationer.
3. Optælling baseret på AKF's publikationslister.

## Litteratur

- Aarkrog, Vibe, Jørgen Brems, Bodil K. Hansen, Kim Mørch Jacobsen og Bjarne Wahlgren (1991). *UTB. Uddannelsesstilbud for langtdslidige*, København: Danmarks Lærerhøjskole og Udviklingscenteret for folkeoplysning og voksenundervisning.
- Adamsen, Lis et al. (1986). *Vejledning i evaluering*, København: AKFs Forlag.
- Adamsen, Lis og Jesper Fisker (1986). *Socialt forsøgsarbejde i boligområder - en tværgående analyse af 20 udvalgte forsøgsprojekter i Danmark*, København: AKFs Forlag.
- Adamsen, Lis, Jesper Fisker og Kirsten Jørgensen (1990/91). *Forsøgsstrategi - Samfundsmæssige konsekvenser og fremtidsperspektiver*, bind I og II, København: AKFs Forlag.
- Albæk, Erik (1988). *Fra sandhed til information: Evalueringsforskning i USA - før og nu*, København: Akademisk Forlag.
- Albæk, Erik (under udgivelse A). "The Danish Case: Rational or Political Change", in Harald Baldersheim og Krister Ståhlberg (eds.), *Free Communes: The Nordic Approach to Local Government Reforms and Innovation*.
- Albæk, Erik (under udgivelse B). "Holdning til omstilling og forsøg blandt de nordiske kommunaldirektører", i Larry Rose (red.), *Kommuner og kommunale ledere i Norden*.
- Dalsgaard, Lene (1985). *Jøbskabelse - Brikken, der mangler*, Ålborg: ATA-projektet, AUC.
- Flyvbjerg, Bent (1991). *Rationalitet og magt. Et case-baseret studie af planlægning, politik og modernitet*, København: Akademisk Forlag.
- Hansen, Hanne Foss (1991). "Programevaluering: Videnskab eller politik?", *Politica*, 23. årg., nr. 1, pp. 84-100.
- Hansen, Erik Jørgen (1992). "Kampen om forskningen - elementer i en analyse til forståelse af sektorforskningens elendighed", *Dansk Sociologi*, 3. årg., nr. 3, pp. 4-17.
- Jensen, P., P.J. Pedersen, N. Smith og N. Westergaard-Nielsen (1991). *Arbejdsmarkedsuddannelserne - Evaluering af effekten på arbejdsløshed og løn*, Center for Arbejdsmarkedsøkonomi, Handelshøjskolen i Århus og Aarhus Universitet.
- Jensen, Mogens Kjær (1992). *SLUT-SUM. En sammenfatning af projekterfaringerne fra Socialministeriets Udviklingsprogram*, København: Socialforskningsinstituttet.
- Jørgensen, Henning (1985). *Arbejdsmarkedsnævnene i arbejdsmarkedspolitikken*, Ålborg: ATA-projektet, AUC.
- McLaughlin, Milbrey W.(1985). "Implementation Realities and Evaluation Design", pp. 96-120 in E.L. Shotland og M.M. Mark, *Social Science and Social Policy*, London: Sage.
- Madsen, Benedicte (1984). *Eksperimentet som paradigme for evalueringen af sociale programmer*, Århus: Psykologisk Institut, Aarhus Universitet.
- Nelson, R.R. (1977). *The Moon and the Ghetto: An Essay on Public Policy Analysis*, New York: Norton.
- Overgaard, Hans Chr. (1984). *Forskning som tjener*, Udvalget vedrørende sektorforskning, Statens samfundsvidenskabelige Forskningsråd.
- Pedersen, Kjeld Møller (1979). *Effektmålingens teori og metode*, København: AKFs Forlag.
- Pilegaard Jensen, T. S. Winter, J. Manniche og P. Ørberg (1991). *Indsatsen for langtdslidige - en undersøgelse af administration og effekt af arbejds- og uddannelsesstilbud*, København: AKFs Forlag.
- Plovsing, Jan (1985). "Socialreformens idealer og praksis", *Politica*, 17. årg., nr. 4, pp. 502-519.

- Rist, Ray C. (ed.) (1990). *Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations*, New Brunswick: Transaction Books.
- Rold Andersen, Bent (1970). *Socialreformundersøgelserne, bd. II. Borgerne og tryghedssystemet*, København: Socialforskningsinstituttet, publikation nr. 44.
- Rossi, P.H. og H.E. Freeman (1989). *Evaluation. A Systematic Approach*, 4. ed., Newbury Park, London, New Delhi: Sage Publications.
- Skou Andersen, Michael (1989). "Miljøbeskyttelse - et implementeringsproblem", *Politica*, 21. årg., nr. 3, pp. 312-28.
- Ståhlberg, Krister (red.) (1990). *Frikommuneforsøk i Norden*, Åbo: Åbo Akademi.
- Thaulow, I. og N. Anker (1992). *Hvordan virker indsatsen for de ledige*, København: Socialforskningsinstituttet.
- Vedung, Evert (1991). *Utvärdering i politik och förvaltning*, Lund: Studentlitteratur.
- Wandel, Jens-Chr. (1988). *Evalueret set ud fra to organisationsteoretiske perspektiver - et case-studie af Danida*, speciale ved Institut for Statskundskab, Aarhus Universitet.
- Weiss, C.H. (with M.J. Buchvalas) (1980). *Social Science Research and Decision-Making*, New York: Columbia University Press.
- Wilhelm, Preben (1985). "Lovgivning i blinde", *Politica*, 17. årg., nr. 4, pp. 488-501.
- Winter, Søren (1983). "Styring via forsøg? - Ungdomsgarantiforsøget som eksempel", pp. 91-135 i Bøje Larsen (red.), *Nye styremåder i den offentlige sektor*, København: Danmarks Forvaltningshøjskoles skriftserie, Jurist- og Økonomforbundets Forlag.
- Winter, Søren (1985). "Iværksættelsesbarrierer", *Politica*, 17. årg., nr. 4, pp. 467-487.
- Winter, Søren (1988). "Denmark", pp. 23-50 in Amin Rajan and Kenneth Walsh (eds.), *Policy Measures for Combating Long-Term Unemployment in the European Communities since the 1984 Council Resolution*, Falmer, Brighton: Institute of Manpower Studies, University of Sussex, report for the EEC Commission.
- Winter, Søren (1990a). "Den anvendte forskning bør ændre kurs", *AKF-nyt*, nr. 2, pp. 25-29.
- Winter, Søren (1990b). "Integrating Implementation Research", pp. 19-38 in D.J. Palumbo and D.J. Calista (eds.), *Implementation and the Policy Process: Opening Up the Black Box*, New York/London: Greenwood Press.
- Winter, Søren (1991). "Summarisk forskning: Forskning, politik og SUM", i *SUMma summarum, Tidsskrift for social udvikling*, Nr. 2, pp. 32-41.