



Evaluering og magt

Peter Dahler-Larsen

Magtudredningen

Evaluering og magt

Magtudredningen

Folketinget besluttede i marts 1997 at iværksætte en dansk magtudredning eller, som det officielle navn er, *En analyse af demokrati og magt i Danmark*. Projektet ledes af en uafhængig forskningsledelse.

Magtudredningens forskningsresultater publiceres i en række bøger, som udgives på Aarhus Universitetsforlag, og i en skriftserie, som udgives af Magtudredningen.

Lise Togeby
(formand)

Jørgen Goul Andersen

Peter Munk Christiansen

Torben Beck Jørgensen

Signild Vallgård

Peter Dahler-Larsen

Evaluering og magt

Magtudredningen

Evaluering og magt

© Magtudredningen og forfatteren, 2004

ISBN: 87-7934-863-7

Omslag: Svend Siune

Tryk: AKA-PRINT A/S, Århus

Magtudredningen

c/o Institut for Statskundskab

Aarhus Universitet

Universitetsparken

8000 Århus C

Danmark

Magtudredningen@ps.au.dk

www.magtudredningen.dk

Alle rettigheder forbeholdes. Mekanisk, fotografisk eller anden gengivelse af eller kopiering fra denne bog er kun tilladt i overensstemmelse med overenskomst mellem Undervisningsministeriet og CopyDan. Enhver anden udnyttelse er uden forlagets skriftlige samtykke forbudt ifølge dansk lov om ophavsret. Undtaget herfra er korte uddrag til brug ved anmeldelser.

Forord

Der rettes en varm tak til forskningsassistenterne Marie Konge Nielsen og Carsten Strømbæk Pedersen, der har ydet store og væsentlige, selvstændige bidrag i forbindelse med de casestudier, som præsenteres i denne bog.

Peter Dahler-Larsen

Indhold

Kapitel 1. Evaluering og magt – en indledning	7
Kapitel 2. Teoretiske skitser af evaluering og magt.....	15
Kapitel 3. Evalueringernes anvendelsesformer.....	23
Kapitel 4. Evalueringsprocessens komponenter	41
Kapitel 5. Cases: Metode.....	50
Kapitel 6. Cases: Evaluering i praksis	54
Kapitel 7. Magtens hvem: Aktører og rationaler i evaluering.....	78
Kapitel 8. Sprog og virkelighed i evaluering: Magtens hvordan.....	87
Kapitel 9. Konklusion: Evaluering og magt	93
Bilag 1. En note om magtbegrebet	96
Litteratur	99
Om forfatteren	103

Kapitel 1

Evaluering og magt – en indledning

Vores aktuelle samfund nærmest oversvømmes af en veritabel evalueringsbølge. Evaluering er tilsyneladende et populært og tidstypisk styringsinstrument. Evaluering har i de senere år fået en lang række synlige udtryk i samfundet. Lovgivning foreskriver evaluering på områder, som ellers traditionelt set har været friholdt for obligatorisk evaluering. Et eksempel herpå er højskolerne – bannerførere for den danske tradition for folkelig oplysning. Evaluering har manifesteret sig i evalueringsinstitutter, centre osv. på flere forskellige områder. Mange offentlige myndigheder har egne evalueringskontorer eller specialister i evaluering. Omfanget af evalueringsindustrien i form af kurser, konferencer og konsulentvirksomhed er ikke kendt, men ganske betragteligt.

Evaluering fremtræder i dag som en ”naturlig”, forventelig eller direkte obligatorisk proces i tilknytning til offentlig virksomhed. Ganske mange styringskoncepter implicerer evaluering, f.eks. ”kvalitetsmåling” som opfølgning på ”kvalitetsudvikling” eller et sæt af indikatorer som led i ”resultatstyring”. Evaluering er i dag også et resultat af uformelle forventninger. Evaluering fremtræder som det, den amerikanske organisationsforsker James March kalder ”protected discourse”, dvs. en særlig beskyttet pulje af ord, som nærmest pr. automatik aftvinger respekt, og som det er svært at opponere imod. I den henseende kan man tale om, at evaluering er blevet et tidstypisk ritual.

Hvorfor er evaluering vigtig?

Ordet ”ritual” kan antyde noget ceremonielt, en procedure uden indhold og konsekvenser. Evaluering kan givetvis i nogle tilfælde optræde som sådan. Men ritualer kan også have stærke konsekvenser – tænk blot på, at en vielse og en domsafsigelse også er ritualer. Faktisk giver evaluerings status som tidstypisk ritual anledning til at tiltænke evaluering en ganske stor betydning. Evaluering svarer til træk i vores samfundstype, som benævnes ”den refleksive modernitet” (Beck, 1994; Dahler-Larsen, 1998). I den refleksive modernitet har man ikke længere automatisk tiltro til fremskridtet. Den refleksive modernitet er i stedet bieffekternes tidsalder. Uforudsete virkninger, koordinationsproblemer og kritiske reaktioner er almindelige i tilknytning til en offentlig indsats. Mange aktuelle styringskoncepter opererer

derfor med løbende tilbagemeldinger og vedvarende refleksionsprocesser. Her passer evaluering ind. Evaluerings aktuelle status passer med den almindelige kulturelle favorisering af løbende læring og omstillingsparathed. Uden at man heraf kan slutte noget om det enkelte evalueringsforløbs betydning og konsekvenser, ser evaluering ud til som styringside og tankefigur at have stor kulturel medvind.

Evaluering kan skabe tvivl og rejse debat. Evaluering kan henlede opmærksomheden på manglende målopfyldelse og på ineffektivitet. Evaluering kan gøre opmærksom på, at en given offentlig indsats har en lang række bivirkninger ud over de tilsigtede. Evaluering kan fungere som talerør for grupper, for eksempel brugergrupper, der har forventninger til en bestemt offentlig indsats, og som ønsker at blive hørt i den politiske proces. Evaluering kan medvirke til, at viden om problemer ved indsatser eller forskelle i resultater spredes i tid og rum. Evaluering kan dermed medvirke til, at bestemte problemer synliggøres og sættes på en mediemæssig og en politisk dagsorden. Evaluering kan skabe et informationsgrundlag, der sagligt eller usagligt kan bruges som løftestang til at få foretaget konkrete beslutninger. Disse beslutninger kan vedrøre fortsættelse eller nedlæggelse af bestemte politiske initiativer såvel som hele offentlige institutioner. Evaluering kan dermed indgå i spillet om ressourcer, jobs og interesser. I bredere forstand kan evaluering påvirke måden, vi tænker og taler om en offentlig indsats på. Vores billede af skolen præges for eksempel af den måde, skolen evalueres på.

Derfor er der god grund til nærmere at studere evaluering i et magtperspektiv. Der er god grund til nærmere at interessere sig for, hvordan det afgøres, hvad evalueringer handler om, og hvad der kommer ud af dem. Hvem sætter dagsordenen? Er der strid om, hvordan evalueringer skal forløbe? Bliver de brugt til noget? Kort sagt: Hvordan udfolder magt sig i forbindelse med evalueringer? Det er denne bogs hovedspørgsmål.

Evaluering eller ej har betydning i sig selv

At evaluering ser ud til at have en vis samfundsmæssig og kulturel medvind skal ikke forlede os til at tro, at evaluering er lige udbredt på alle områder inden for den offentlige sektor. Områder som uddannelse, ulandsbistand, arbejdsmarkedspolitik og hele det socialpolitiske område ser ud til være præget af intens evalueringsevne (Foss Hansen, 2001: 121), mens f.eks. kirken, forsvaret, kongehuset og domstolene indtil videre ser ud til at have klaret frisag. Disse forskelle er formodentlig ikke tilfældige. Selve beslutningen om at evaluere kan være en signifikant beslutning. Men

den er samtidig tvetydig. På den ene side kan evaluering, specielt i tilfælde af negative resultater, skabe risiko for ubehagelige nedskæringer, omstruktureringer og dårlig omtale. Omvendt kan evaluering også være en løftestang for intern og ekstern opmærksomhed, der kan give udviklingsmuligheder og måske positiv omtale. Hvad man ”ser” i evaluering kan altså være afhængigt af, hvordan man vurderer sine egne interesser, og hvordan man forventer, evalueringen falder ud. Ens egne muligheder for at påvirke evalueringens emne, design og anvendelse kan også spille ind.

Ud fra en observation af, at evaluering finder sted, kan man altså ikke uden videre slutte noget om, hvilke interesser, der eventuelt har været bestemmende. Det samme gælder, hvis evaluering ikke finder sted. På den anden side kan vi heller ikke udelukke, at der kan ligge en væsentlig magt i magten til at undvige evaluering i en tid, hvor evaluering ellers ser ud til at blive stadig mere udbredt.

Hvad er evaluering? – En definition

I takt med vore dages evalueringebølge har evalueringebegrebet en tendens til at blive udflydende.

I enhver undersøgelse af evaluering er det derfor på sin plads at indkredse evalueringebegrebet nærmere. Formelt kan evaluering defineres som en systematisk retrospektiv (dvs. tilbageskuende) bedømmelse af gennemførelse, præstationer og udfald i offentlig virksomhed, og det er en bedømmelse, som tiltænkes at spille en rolle i praktiske beslutningssituationer (Vedung, 1998: 20).

Inden for denne lidt abstrakte definition falder en lang række aktiviteter, som ikke officielt eller uofficielt benævnes ”evaluering” (f.eks. kvalitetsmålinger, brugerundersøgelser og monitorering ved hjælp af indikatorer). Man kan også forestille sig, at en aktivitet, som i praksis benævnes evaluering, reelt ikke er systematisk eller retrospektiv og derfor falder uden for definitionen. Det betyder, at enhver empirisk undersøgelse af evaluering kan få problemer med at få det helt rigtige med, hverken for lidt eller for meget.

I ovenstående definition er det værd at hæfte sig nærmere ved fire særligt betydningsfulde ord, nemlig ”systematisk”, ”bedømmelse”, ”offentlig virksomhed” og ”praktiske beslutningssituationer.”

Systematisk betyder, at evaluering i det mindste ideelt er baseret på en metode, der giver en vis grad af systematik i data. I denne henseende ligner evaluering videnskabeligt arbejde. Det betyder, at nogle af de diskussioner, man kan have om videnskabelige undersøgelser, herunder diskussioner om

metodisk kvalitet, også findes inden for evaluering (se f.eks. Lomborg, 1998).

Ordet ”bedømmelse” viser, at evaluering må involvere værdier. Værdier kan være implicitte eller eksplicitte, men uden dem ingen bedømmelse.

Ordet ”offentlig virksomhed” viser hen til, at evaluering handler om den offentlige sektor eller i bredere forstand ”res publica”, det offentlige anliggende, dvs. det som er fælles for borgerne i et samfund. Offentligheden og demokratiet kan altid gøres til mere eller mindre aktive medspillere i et evalueringsforløb. Der kan evalueres i den parlamentariske styringskædes navn, i Folketingets navn, i almenvellets navn osv.

Mere specifikt kan evaluanden (evalueringsgenstanden) være både gennemførsel (organisation og samarbejde), præstationer (det som gøres af de offentligt ansatte) og udfald (de resultater, der i sidste led opnås med hensyn til uddannelse, beskæftigelse, livskvalitet osv.). Evaluering kan beskæftige sig både med mikrokvalitet (snævert knyttet til en bestemt offentlig indsats) og makrokvalitet (bredere samfundsmæssige effekter).

Det fjerde og sidste aspekt af definitionen er, at evaluering tiltænkes at spille en rolle i praktiske beslutningssituationer. Evaluering har således et anvendelsessigte. Med denne praktiske dimension adskiller evaluering sig fra klassisk videnskabeligt arbejde, som den ellers via sin metode er i familie med. Evaluering sigter mod at informere og begrunde ganske praktiske beslutninger om, at offentlige indsatser skal fortsætte, nedlægges eller reformeres. Bemærk, at evaluering på dette punkt er defineret ved sit sigte snarere end ved sin faktiske funktion. Det afgørende er evalueringens selvforståelse – den vil gerne hjælpe til med at skaffe vidensgrundlaget for informerede beslutninger, og den *tiltænkes* en sådan rolle. Evaluatorene og evalueringsforskningen er imidlertid helt klar over, at evalueringer kun sjældent lever op til deres officielle formål (Albæk, 1988; Weiss & Bucuvalas, 1980). Til gengæld har man de senere år fået øjnene op for en hel række forskellige anvendelser af evaluering, hvoraf flere rækker hinsides det formelle og det tilsigtede (Vedung, 1998; Dahler-Larsen, 1998; Dahler-Larsen & Larsen, 2001; Kirkhart, 2000).

I det følgende vil jeg kort rendyrke to overordnede billeder af evaluering, som repræsenterer to helt forskellige syn på, om evalueringer har tilsigtede og formålstjenlige virkninger.

To billeder af evaluering og magt

Denne bog handler om evaluering og magt, nærmere bestemt hvordan magt udfolder sig i forbindelse med evaluering. Formålet er ikke at udrede

magtbegrebet som sådan – det er blandt andet sket i Magtudredningen (Thomsen, 2000; Christiansen & Tøgeby, 2003). På den anden side har der været et behov for at afklare og eksplicite det magtbegreb, som undersøgelsen bygger på. Jeg har løst dette problem ved at give en kort redegørelse om undersøgelsens magtbegreb som sådan i bilag 1.

I resten af bogen vil jeg beskæftige mig med evaluering i et magtperspektiv. Indledningsvis skitserer jeg to rendyrkede og meget forskellige opfattelser af, hvordan magt udfolder sig i forbindelse med evaluering. Disse to opfattelser markerer samtidig yderpunkterne i det begrebslige rum, som undersøgelsen udfolder sig inden for.

Tabel 1.1. To opfattelser af evaluering og magt

	et instrument	en arena
Evaluering er for magtudøvelse	Finalitet	Emergens
Teoretisk hovedide	Givne	Mindre givne
Aktører	Givne	Mindre givne
Interesser og formål	Høj	Lav
Evalueringsforløbets forudsigelighed	Velafgrænset	Ikke velafgrænset
Evaluerings anvendelse	Ud fra forudliggende intentioner og ressourcer	Ud fra evalueringsforløbet
Anvendelsens bestemmelse	Entydig	Flertydig
Evaluerings betydning	Legetøjstog	Radiobiler
Billede af evalueringsforløbet		

Evaluering som instrument

I den første opfattelse er evaluering et instrument for magtudøvelse. Evaluering udmønter en magt, der er givet forud. Evalueringen er præget af finalitet – formålsrettethed – i forhold til denne bagvedliggende magt. Denne magt består i, at givne aktører med givne interesser og formål gør hvad de kan for, at evalueringen kommer til at tilgodese disse interesser og formål, og det lykkes, fordi de har ”magt” til det.

Evalueringsforløbet følger – som et legetøjstog – fuldstændig de skinner, der på forhånd er lagt ud. Evalueringsforløbet er således forudsigeligt. Evalueringens anvendelse – f.eks. at understøtte en given beslutning – er velkendt og klart afgrænset i tid og rum. Udfaldet er givet på forhånd ud fra de bagvedliggende interesser og formål samt naturligvis de ressourcer, som aktørerne har haft til at sikre tilgodeseelsen af deres interesser og formål. Evalueringens hele mission er så at sige at fuldbyrde det formål, som lå i

kimform allerede inden evalueringen. Ud fra denne tankefigur er det ikke vanskeligt at forklare et givet udfald af en evaluering. Udfaldet kan nemlig forklares udtømmende med henvisning til de bagvedliggende intentioner. Evalueringsforløbet i sig selv tilføjer intet nyt. Evalueringsforløbet er i sig selv ikke kreativt. Der er intet irreversibelt ved selve forløbet. Som i en matematisk funktion kan man entydigt regne frem og tilbage i tid. Evalueringen er en nøjagtig omsættelse af formålet til en anvendelse. Baglæns kan anvendelsen forklares med henvisning til formålet.

Evalueringens betydning er entydig. Det udfald, som evalueringen fik, var det ene udfald, som den kunne få. Selv om denne skitse af et evalueringsforløb er ganske skitseagtig, finder den udbredt genklang i gængse evalueringsopfattelser, af hvilke her skal fremhæves to undertyper.

Den første undertype er den officielle selvforståelse i meget evalueringssarbejde. Der opstilles en formelt formål med evalueringen – det være sig for eksempel ”kontrol” eller ”læring” – og evalueringssarbejdet tilrettelægges derefter, så det så sikkert som muligt fører frem til netop det tilsigtede formål (”intended use for intended users”, som er sloganet hos Michael Quinn Patton (1997), en toneangivende amerikansk evaluator). Hvad enten man sigter efter f.eks. kontrol eller læring foreskrives der altid, at der bruges kræfter og energi på virkelig at afklare formålet. Hvorfor? Fordi man antager, at det opstillede formål virkelig er bestemmende for, hvad der faktisk kommer ud af evalueringen.

At evalueringen først og fremmest, om ikke udelukkende, er et instrument for tilgodeseeelsen af bagvedliggende intentioner er også en styrende idé i en anden undertype af gængse evalueringssforståelser. Denne anden undertype opererer med, at evaluering først og fremmest er et brækjern til at fremme skjulte politiske dagsordener. Tankefiguren er igen, at skal man forstå udfaldet af en evaluering, skal man blot henvide til de intentioner, som findes før evalueringen hos en magtfuld gruppe af aktører. Disse intentioner kan imidlertid kun identificeres begrebsligt, men ikke nødvendigvis empirisk. I modsætning til den første undertype er intentionerne nemlig her skjulte og ganske ofte illegitime. Men der er samme mulighed for at regne frem og tilbage fra formål til funktion. At forklare udfaldet af en evaluering er derfor logisk set en let sag for begge undertypers vedkommende (selv om den selvfølgelig kan byde på praktiske og metodiske problemer, f.eks. at få kortlagt en skjult dagsorden). For begge undertyper af denne forståelse kan man let rekonstruere forløbet – man kan sige at sige blot køre tilbage ad de samme skinner.

Evaluering som arena

En helt anderledes opfattelse af evaluering og magt har man under overskriften evaluering som arena. Her præges en evaluering løbende. Evalueringsforløbet er ikke givet, men det ”bryder frem” af samspillet på arenaen. Det præges af dem, der optræder. Aktørerne er imidlertid ikke givne på forhånd. Nye kan blande sig undervejs. Og – måske mere subtilt – selve aktørernes konstitution indgår i spillet på evalueringsarenaen. Et givet evalueringsforløb kan således være med til at definere den rolle eller identitet, som en aktør kan optræde med. Det kan f.eks. være, at en evaluering fastlægger, hvad det vil sige at være ”bruger” i og med, at evalueringen tilordner ganske bestemte udsagnsmuligheder for dem, der samtidig henregnes til denne kategori. Tilsvarende kan intentioner nydannes og formuleres undervejs. Evalueringen er ikke styret alene af ”bagvedliggende” intentioner og formål. I stedet formes evalueringsforløbet undervejs af dem, der blander sig. Evalueringen præges af, hvad der viser sig at kunne gøres gældende. Ingen aktør kan være helt sikker på, hvad der kommer til at gælde på arenaen.

Der kan være mange sammenstød i evalueringsforløbet. Evalueringsforløbet vil derfor minde mere om en tur i radiobilerne end en tur med et legetøjstog. Turen i radiobilerne præges selvfølgelig delvist af føreren af den enkelte bil, men sammenstødene med de andre biler undervejs giver en høj grad af uforudsigelighed i den faktiske rute. Skal man drive billedsproget lidt vidt, kan man endda sige, at selve ruten ikke er så afgørende. Men kørslen i forhold til de andre biler er både fornøjelsen og udfordringen.

Når evalueringer minder om kørsel i radiobiler, kører de forskellige deltagere ikke på en fælles rute, som tilfældet var i skinnetoget. Nogen vil blive kørt mere ned end andre. De forskellige deltagere vil opleve turen forskelligt, ikke bare på grund af psykologiske opfattelsesforskelle, men fordi de faktisk tager forskellige erfaringer med fra turen. Ud fra dette billede kan en evaluering have ganske forskellige virkninger på forskellige aktører. *Således kan evalueringer ramme forskelligt inden for og uden for en organisation såvel som på forskellige niveauer i organisationen* (Christina Segerholm, oplæg i Odense, 9.10.2003).

Evalueringsforløbet peger ikke frem mod én bestemt evalueringsanvendelse. Med et relativt uforudsigeligt evalueringsforløb kan anvendelsen af evaluering blive begrænset i forhold til det tilsigtede, men virkninger af en evaluering kan også forplantes gennem tid og rum – ikke mindst gennem de andre biler! – på måder, der rækker langt videre end det tilsigtede. Man kan i en sådan situation ikke blot komme tilbage til udgangspunktet ved at

bakke ad den rute, man kom fra. Den kompleksitet i køreturen, som sammenstødene med de andre biler tilføjede, kan ikke reproduceres blot ved at bakke. *Selve evalueringsforløbet kan tilføje noget nyt, noget irreversibelt.*

Man kan formode, at aktører, der involveres i en evaluering, forholder sig til denne ud fra det billede, de har af evalueringens virkninger. Dette billede kan være præget af, men ikke nødvendigvis identisk med, evalueringens officielle formål. Hver aktør kan formodes at forsøge at præge evalueringen i en sådan retning, at den pågældende aktørs ønsker til evalueringsanvendelsen tilgodeses. Den resulterende evalueringsanvendelse kan være et komplekst konglomerat af aktørernes forventninger, aktørernes handlinger på baggrund af både forventninger og ønsker samt af den interaktion, der opstår imellem disse handlinger. I den forbindelse kan man tænke på evalueringsanvendelsen som noget, der *bryder frem igennem evalueringsprocessen*. Evalueringsanvendelsen kan ikke udledes udtømmende af interesser og magtbaser, der er givne *forud* for selve processen.

Evaluering kan dermed være fundamentalt flertydig. Den bliver til det, den gøres til. Det betyder, at ethvert entydigt udsagn om udfaldet af en evaluering for så vidt blot kan være et indslag i et spil om, hvad evaluering kan blive til, men det er ikke i sig selv en analyse.

At evalueringens betydning afgøres undervejs betyder ikke, at alle udfald er lige sandsynlige. Radiobilen påvirkes for eksempel ganske klart af køreernes absolutte og relative hurtighed, vægt og køreevner.

Dagsordenen for en undersøgelse af magt og evaluering er at forsøge at finde ud af, hvordan det gik til, at en evaluering blev præget præcis, som den gjorde. I denne undersøgelse af evaluering og magt vil der nærmere bestemt blive fokuseret på følgende spørgsmål:

1. Hvilke aktører indgår med hvilke rationaler? Her tænkes både på aktører, som "blander" sig, såvel som på evalueringens eventuelle tilskrivning af en aktør-position, som man kan "blende sig" fra. Hvilke ressourcer trækker de forskellige aktører på?
2. Hvordan afgøres ruten på korsveje, hvor evalueringen kan forløbe på forskellig måde? Hvad tages for givet, og hvad udfordres?
3. Hvordan forløber evalueringerne? Ligner evalueringsforløb mest legetøjstoget eller radiobilerne?

Disse spørgsmål vil være ledetråde igennem det teoretiske og empiriske arbejde.

Kapitel 2

Teoretiske skitser af evaluering og magt

Samfundsvidenskaben har leveret forskellige skitser af, hvordan magt udøves. Jeg vil gennemgå udvalgte af disse skitser, som kan medvirke til at indkredse, hvordan evaluering er med i et større magtsystem. Disse forskellige skitser disponeres efter

- a. om hovedvægten lægges på strukturelle relationer eller på kulturelle betydninger
- b. om hovedvægten lægges på, at magten er monologisk (udgår fra et center eller en kilde) versus, at magten er pluralistisk (dvs. at magten sker i samspil mellem flere instanser og/eller principper).

Det giver følgende muligheder:

Tabel 2.1. Fire teoretiske perspektiver på evaluering som magt.

	Monologisk	Pluralistisk
Strukturelle relationer	1: Kontrol og overvågning	2: Kamp mellem instanser
Kulturelle betydninger	3: Evalueringssideologi	4: Kamp mellem værdier

Denne opdeling er tænkt som indledende og overbliksskabende. De fire felter fanger nogle typiske grundtankegange om, hvordan evaluering virker som magtudøvelse. Hver af disse grundtankegange er vist ved en stikordsagtig overskrift. Men skal man have en mere avanceret forståelse af magtudøvelse i evaluering er der ingen garanti for, at man kan nøjes med at holde sig inden for ét af de fire felter. Faktisk viser den opstillede flerhed af perspektiver, at hvert enkelt perspektiv i sig selv giver et meget begrænset udsyn. Lad os imidlertid først præsentere dem et for et.

Kontrol og overvågning

Har man et monologisk og strukturbaseret syn, vil man tænke på samfundet som et stort system, der er præget af kontrol. I den positive version er kontrollen et legitimt led i en demokratisk styringskæde (Vedung, 1998). Folket holder øje med politikerne, der holder øje med embedsmænd, der holder øje med fagfolk, der udfører nogle offentlige aktiviteter for borgere og bru-

gere. Hvert led i kæden har et legitimt krav på at udøve kontrol. Evaluering er en mekanisme, der er med til at holde styringsstrukturen på plads. Evaluering checker, om det hele virker.

Det er imidlertid heller ikke vanskeligt at komme i tanke om meget negative versioner af samfundet som en kontrolstruktur. Ingen har analyseret denne samfundsvision bedre end Foucault i sin analyse af Benthams panoptikon (Foucault, 1980; Bauman, 2002). Bentham var optaget af at bygge et effektivt fængsel med så få personaleressourcer som muligt. I hans ideal-fængsel skal alle celler derfor kunne ses fra et enkelt punkt, det panoptiske punkt. Fangerne kan imidlertid ikke se deres vogtere direkte. Den enkelte fængselsfange skal vænne sig til evigt og altid at være overvåget. Alene fangens bevidsthed om fangevogterens blik virker disciplinerende. Når først dette princip er installeret, er det for så vidt ligegyldigt, om der i hvert enkelt øjeblik faktisk også findes fangevogtere på det panoptiske punkt. Fangerne disciplinerer sig selv, for så vidt som deres adfærd er en funktion af hele det strukturelt baserede overvågningssystem, som de lever i. Parallellt til evalueringssamfundet er sigende, hvis man tænker magt på denne måde. Vi lever i et evalueringssamfund, hvor vi må disciplinere os selv, så vi løbende kan leve op til at blive evalueret. Evalueringer møder os overalt. Vi vænner os til alt fra overvågningskameraer til tidsregistreringer, præstationsmålinger, kvalitetsundersøgelser og brugerundersøgelser af vores arbejde. I et velstruktureret kontrolsystem leder man forgæves efter moral, konsensus, værdier, prædikener og legitimering – som man ellers ofte finder helt afgørende for, at et samfund hænger sammen. Ligeledes er det nyteløst at appellere til ”fælles værdier.” Kontrolsystemet fungerer fint uden nogen fælles vedtagelse eller opbakning. I den forstand er det selve overvågningen, der fanger os. Det er måske mindre afgørende, hvad der måles, registreres og vejes. Mediet er så at sige budskabet. Magten løsgøres fra enkeltpersoner. Magten er, som Foucault siger (1980: 156), ”en maskine, som ingen ejer.”

Kamp mellem instanser

Tænker man pluralistisk og strukturelt vil man pege på, at samfundet består af en lang række institutioner og organisationer med forskellige udgangspunkter og interesser. Ganske vist er nogle institutioner i stand til at sætte andre under evaluerende overvågning, f.eks. når et ministerium opretter et evalueringsinstitut til at bedømme forskellige uddannelser og skoler. Alligevel når intet kontrolsystem fuldstændigt at dominere, og ingen offentlig institution er kun omspundet af ét kontrolsystem. Kontrolsystemerne er i

stedet talrige, overlappende og delvist modstridende (Wirth, 1986). Kontrollen kommer ikke kun oppefra og udefra, men afbalanceres af mekanismer indefra og nedefra. En skole styres kun ganske partielt og indirekte af et evalueringsinstitut. Andre kontrolmekanismer etableres via lovgivning, kommunalpolitik, brugerbestyrelse, skoleledelse, mediernes og lærernes faglige organisation. Mange instanser giver deres besyv med, og mange kan foretage evaluering. Mange er i position til at sikre begrænsninger i, hvilke evalueringer af hvilke skoler, klasser og lærere, der foretages, og hvilke, der offentliggøres, og i hvilken form. Der er også mange appelinstanser, ikke mindst pressen. I et strukturelt og pluralistisk perspektiv vil man interesse sig for, hvilke magtbaser, de forskellige instanser har i en sådan evalueringskamp. Magtbaserne ses i høj grad som strukturelle positioner karakteriseret ved forskellige ressourcer, forhandlingsstyrke, adgang til information, vetoret osv. I et sådant perspektiv er en vigtig kamp retten til at evaluere andre samt retten til at sikre, at andre ikke evaluerer ens egen instans. Forhandlinger om udformningen af eventuel selvevaluering vil også være et vigtigt punkt inden for dette perspektiv, hvor kampen om evaluering ses som kampen om strukturelt herredømme over evaluering.

Evalueringsideologi

Tænker man i kulturelle baner snarere end strukturelle, vil man se magtudøvelse som et spørgsmål om sprog, symboler og betydninger. Tænker man monologisk på kulturen, ser man ethvert samfunds mest aktuelle betydninger som præget af en overordnet verdensforståelse eller ideologi.

Et sådant syn på vor tids evalueringsbølge præger dem, der i vor tids evalueringsbølge ser en dominans af værdier angående konkurrence, mistillid, effektivitet og målbarhed i den tænkning, som styrer den offentlige sektor. I stedet for tillid til, at den enkelte offentlige ansatte gør sit bedste for at fremme de menneskelige værdier, der bør præge den offentlige ansattes møde med den enkelte borger/bruger, fremmes i dag en mistillid, der fører til onde cirkler af måling/registrering, mistillid, manglende motivation og mere overvågning, som stjæler tid fra det egentlige arbejde. Det målbare prioriteres som følge deraf højere end de bløde og menneskelige værdier, der ikke kan måles, lyder kritikken. Evalueringsbølgens udbredelse går således hånd i hånd med udbredelse af en bestemt management-kultur. Evaluering og den tilhørende managementkultur er for så vidt blot forlængelser af de grundlæggende træk i det moderne samfund, der lægger vægt på bureaukrati, dokumentation, målbarhed, opsyn, forudsigelighed og kalkulerbarhed.

Andre ser langt mere positive værdier i vore dages evalueringskultur. I vore dages post-konventionelle kultur (Ziehe, 2003) står det meste til forhandling, og ingen autoritet kan automatisk regne med at læne sig op ad overleverede traditioner og privilegier. I stedet må en autoritet kunne begrunde sine handlinger. Evaluering kan tjene en funktion her. Evaluering kan virke som en begrundelse for at benytte en fremgangsmåde frem for anden, for eksempel i skolen. Men så snart evalueringer peger på, at andre fremgangsmåder er bedre, må den offentlige institution følge med. I den positive variant er vore dages evalueringskultur således et krav om åbne, begrundede og demokratiske præmisser for udøvelsen af offentlig virksomhed.

Kamp mellem værdier

I et kulturelt perspektiv præget af pluralisme er evaluering et møde mellem forskellige værdier. Evaluering er et medium for tilbagemelding, koordinering og pejling mellem forskellige synspunkter. Samfundet er ikke længere præget af en overordnet værdimæssig paraply eller et overordnet billede af sine egne visioner (Castoriadis, 1982). Ideen om fremskridt har mistet sin opbakning, ikke mindst fordi ”fremskridtet” har vist sig at medføre talrige bivirkninger og risici. I en tid uden fælles normer og værdier, som samtidig er præget af stor kompleksitet, hvad angår utilsigtede virkninger af snart sagt enhver handling, bliver det ekstra nødvendigt for den enkelte organisation at sikre sig løbende tilbagemeldinger om samarbejdspartneres, brugeres og andre interessenters reaktioner og/eller oplysninger om virkninger af organisationens egne handlinger. Evaluering gør en udvidet kreds af interessenter relevant (Albæk, 2001). Evaluering øger i den forbindelse den offentlige organisations sensitivitet over for værdier, som den ikke selv rummer. Evaluering får den offentlige organisation til at forholde sig til andres oplevelser af dens bivirkninger. Evaluering afstemmer forskellige interessenters forventninger og vurderinger.

I den positive version handler evaluering om det åbne, dialogprægede, i bedste fald helt herredømmefri møde mellem forskellige værdier. I den negative version handler evaluering om konstante forhandlinger uden fælles værdier, hvor den der er bedst organiseret, har flest ressourcer, har det største sekretariat eller den dygtigste pressemedarbejder, ser bedst ud på tv, eller filer dygtigst med data går af med sejren – for en tid. I hvert fald ser det ud til, at produktion af viden i dag, inklusive evalueringsviden, er kraftigt påvirket af magtfaktorer som økonomiske ressourcer, mandskab og adgang til medierne.

Muligvis er der et gran af sandhed i hver af de fire ovenstående skitser af evaluering som magtudøvelse, der henholdsvis er strukturel og kulturel, henholdsvis monologisk og pluralistisk. I hvert fald kan man møde afskygninger af dem alle i dagens debat.

Tilhængere af det strukturelle perspektiv har en tilbøjelighed til at lægge vægt på institutionelle indretninger og medier til overvågning; kulturperspektivet understreger mere mentalitet, værdier og normer. Tilhængere af monologiske perspektiver fokuserer på, hvordan et princip eller en logik dominerer over alle andre, mens det pluralistiske perspektiv understreger udvekslingen og samspillet mellem forskellige synspunkter. Tilhængere af et monologisk perspektiv vil derfor ofte forvente, at evalueringsforløb er instrumenter for en enkelt dominerede magt, hvis proces minder om legetøjstogets bane, mens pluralister ser evalueringer mere som kørsel i radiobil.

Den enkelte evaluering som magtudøvelse

Den enkelte evaluering udøver i grundlæggende forstand magt ved at give en udfordring til evaluanden (altså evalueringsgenstanden). Det kan være en bestemt form for praksis, en indsats eller en institution. Udfordringen lyder: Kan det dokumenteres, at denne praksis er god nok?

Evalueringen antaster dermed den selvfølgelighed, hvormed den pågældende praksis ellers ville blive foretaget. Evaluering viser, at den pågældende praksis er kontingent. Kontingent betyder, at den er situationsbettinget, afhængig af noget (f.eks. anerkendelse), og at den således kunne være anderledes. Det er let at se, at evaluering og kontingens hænger sammen. Institutioner, der (endnu) ikke er blevet kontingente, kan man ikke evaluere. Lærere i Japan udsættes ikke for evaluering på samme måde som danske lærere. Det skyldes, at der er knyttet så meget autoritet til lærerrollen i Japan, at man ikke kan problematisere den enkelte lærers måde at udøve sin funktion på. Ingen kan på den offentlige arena stille spørgsmålet: Gør læreren det godt nok? Med andre ord er der endnu ikke tilstrækkeligt med kontingens omkring lærerens funktion til, at den kan evalueres. Også herhjemme har vi institutioner, der ikke er så kontingente, at de kan evalueres, for eksempel folkekirken og kongehuset. Men grænsen for, hvad der er kontingent, er ikke stabil. I de senere år er der f.eks. stillet lovgivningsmæssige krav om, at højskolerne, der tidligere var nærmest automatiske og direkte bærere af alle positive værdier i den folkelige oplysningstradition nedarvet fra Grundtvig, skal foretage evaluering. En ganske vigtig magtmæssig funktion af evaluering er således at antaste noget, der ellers ville fremstå

som nærmest naturgivet. Evaluering gør opmærksom på, at evalueringsgenstanden skal bevise sit værd, ellers kan der rejses tvivl om dens eksistensberettigelse eller i hvert fald om dens nuværende måde at fungere på.

Men evaluering skal helst foretage denne antastelse på en relativt uantastelig måde. Evaluering skal selv helst stå på et grundlag, der ikke samtidig sættes til debat på samme måde som det, der udfordres. Evaluering må for eksempel benytte videnskabelige metoder, således at evalueringsresultaterne ikke virker tilfældige, opfundne eller konstruerede til lejligheden. Evaluering virker bedst, hvis der er opbakning til processen og metoden. Evaluering må derfor benytte velbegrundede evalueringskriterier og en forsvarlig metode. Evalueringsfeltet rummer en lang række overvejelser om de grundlæggende begrebslige dimensioner i enhver evaluering (såsom viden, værdier og anvendelsen (Shadish, Cook & Leviton, 1991)), alt sammen for at styrke velbegrundetheden i evalueringens fremgangsmåde. En sådan "velbegrundethed" er ikke endimensional. Mange forhold kan gøre en forskel. Der kan være mere eller mindre legitimitet i en evaluerings proces, metode og kriterier.

En evaluering kan f.eks. benytte en evalueringsmetode, som ikke nyder anerkendelse hos dem, der evalueres, men alligevel sætte sig igennem, fordi den udføres med hjemmel i lovgivningen af en autoritativ instans, for eksempel et evalueringsinstitut. Samtidig er der imidlertid i dag i mange vidensprocesser, herunder evalueringsprocesser, krav om at mange interessenter inddrages som legitime partshavere (Albæk, 2001). Der forventes ganske ofte inddragelse og dialog, også selv om legitime autoriteter står bag. Således ser man ikke mindst i dag, at mange evalueringsprocesser indeholder selvevaluering som en mere eller mindre dominerende komponent: de evaluerede skal også høres. Selvevaluering eller ej – i mange tilfælde gøres der en betragtelig indsats for at fremstille en evaluering som saglig, begrundet og legitim i mange interessenters øjne. Med andre ord er idealet i evaluering, at godt nok skal en bestemt praksis antastes, men det skal helst ske ud fra et uantasteligt grundlag.

Når det så er sket, er idealet, at der tages informerede beslutninger om den praksis, der evalueres. Aktiviteten kan blive fortsat efter at være blevet godkendt via evaluering. Indsatsen kan blive justeret som følge af evalueringens anbefalinger. Aktiviteten kan også blive nedlagt. Under alle omstændigheder er ideen, at evaluering udgør et slags kvalificerende filter, som bringer den evaluerede aktivitet op på et højere vidensniveau. Man søger at optimere effekter, mindske bieffekter, tage højde for brugerreaktioner osv. Man søger at udluge det ineffektive. Man søger at iværksætte orga-

nisatorisk læring som følge af evaluering. Med andre ord bliver praksis et skridt mere uantastelig, i hvert fald for en tid. *Idealet i evaluering er med andre ord ud fra en relativt uantastelig position at antaste en praksis for gennem informeret beslutning at gøre denne praksis et skridt mere uantastelig.*

Evaluering etablerer således en grænse mellem det antastelige og det uantastelige og opererer derefter ovenpå sondringen. Set fra et magtperspektiv er det selvfølgelig bedst at stå på det uantasteliges side og antaste noget andet. Det vanskelige ved grænsedragningen mellem det antastelige og det uantastelige er imidlertid, at evaluering netop ikke er en traditionel og dogmatisk praksis, som kun kan gøres på en måde, men i stedet rummer forskellige modeller og metoder, der i hver enkelt situation må søge at sætte sig igennem som mest uantastelige evalueringspraksis givet de nærmere politiske, organisatoriske og praktiske omstændigheder. Evaluering foregår ikke uden for den sociale virkelighed, men prøver som andre aktuelle former for videnskabelse at finde en passende vej frem givet en lang række kriterier, der både omfatter videnskabelighed, social acceptabilitet og nytte (Gibbons et al., 1994). Ganske ofte mislykkes de forsøg på grænsedragning, som en given vidensproduktion, herunder evaluering, forsøger sig med.

Den aktuelle polemik omkring Bjørn Lomborg er et eksempel herpå. Bjørn Lomborg søger at udfordre megen eksisterende miljøforskning ud fra metoder, han selv finder uantastelige. Andre antaster hans metode og indbringer sågar hans videnskabelige praksis for et organ, der skal vurdere hans metodes lødighed. Da udvalgene for videnskabelig uredelighed – ud fra en formodet uantastelig position og procedure – fælder sin dom, er der straks andre, der antaster udvalgets position og procedure. Der igangsættes i øvrigt en evaluering af det institut, som Bjørn Lomborg leder, ligesom andre miljøinstitutters arbejde ønskes evalueret for at finde ud af, om disses metoder relativt set er lige så antastelige som Bjørn Lomborgs. Den ene vidensproduktion afløser den anden, og konstant udfordres de andre vidensproduktioners grænse mellem det antastelige og det uantastelige. Det betyder naturligvis ikke, at alle vidensproduktioner har lige stor chance for at fremstå som uantastelige og at sætte sig igennem som gældende. Det betyder blot, at vidensproduktion såsom evaluering ikke på nogen let måde kan holde sig selv uantastet, mens den antaster noget andet. Det er nemmere at antaste end at holde sig selv uantastelig. Denne udvikling ser i stigende grad ud til at være et uomgængeligt vilkår for vidensproduktion (Beck, 1997). Den skyldes ikke blot, at produktion af viden, herunder evaluering, i dag må forholde sig til en lang række praktiske, organisatoriske, politiske

og kulturelle godkendelseskriterier, men også at den videnskabelige arbejdsform i dag har vendt sig mod sig selv. I dag bruges videnskabelig metode i høj grad til at udfordre videnskabelig vidensproduktion (Beck, 1997). At vidensproduktionen således i dag i højere grad end tidligere er kontingent (afhænger af sine arbejdsbetingelser), betyder imidlertid hverken, at al rationalitet har forladt vidensproduktionen (Longino, 2002) eller at alle vidensproduktioner har lige store chancer for at fremstå som acceptable og troværdige. Det betyder blot, at en vidensproduktions grænsedragning mellem det antastelige og det uantastelige ikke på forhånd kan regne med at være absolut, sikker og stabil.

I det lys kan vi formulere, hvad det vil sige at udøve magt igennem evaluering: Der udøves magt igennem evaluering, når det via evaluering lykkes at holde en bestemt grænsedragning mellem noget antasteligt og noget uantasteligt fast længe nok til, at den er socialt produktiv. Med socialt produktiv menes, at noget gøres gældende. Der etableres måder at tænke og tale på i relation til den offentlige indsats. Der krystalliseres bestemte definitioner af ”problemer” og ”løsninger”. På den baggrund kan der tages beslutninger og etableres prioriteringer. Samtidig sker markeringer af, hvem de relevante involverede parter er, i hvilken egenskab de kan sige noget, og hvad deres rettigheder og forpligtelser er i relation til evaluanden og måske i relation til andre forhold i verden.

De to billeder af evaluering, som blev skitseret indledningsvist – med evalueringsforløb som legetøjstog versus evaluering som radiobiler opererer med to helt forskellige grader af stabilitet i grænsedragningen mellem det antastelige og det uantastelige. Når et evalueringsforløb forløber jævnt og forudsigeligt, er det et tegn på at en bestemt sontring mellem det antastelige og det uantastelige er stabil. Når sontringen udfordres løbende, bliver fundamentet under en evaluering udfordret, og det giver et mindre lineært forløb. Hvis også udfordringerne udfordres, bliver evalueringsforløbet uforudsigeligt, og evalueringen får virkninger, der har meget lidt at gøre med det oprindelige oplæg (hvis et sådant ellers findes).

Kapitel 3

Evalueringernes anvendelsesformer

Som nævnt i indledningen har evaluering et anvendelsessigte. Per definition udarbejdes evalueringer for at spille en rolle i handlingssituationer. Udviklingen af metoder og modeller inden for evaluering har i høj grad været drevet af dette hensyn til praktisk anvendelse (Hellstern, 1986). Samtidig har tilhængere og udøvere af evaluering haft en tilbøjelighed til at fokusere på ganske udvalgte former for anvendelse af evaluering, især de positive og tilsigtede – altså hvad man kan kalde de uantastelige måder at anvende evaluering på. Til gengæld har man måske haft en udpræget blind plet (Mark & Henry, 2002) i forhold til anvendelsesformer, som er mere problematiske og diskutabile. Det betyder ikke, at man inden for evalueringskredse ikke har diskuteret forekomsten af politisk misbrug af evalueringer. Det har man nemlig gjort (Christie & Alkin, 1999). Men det har vedvarende været ud fra en ide om, at misbrug var let at kende, når man mødte den – eller med forrige kapitels ord: det har været ud fra en klar sondring mellem det antastelige og det uantastelige.

I dette kapitel drøftes syv begreber om evalueringers anvendelsesformer, som er udviklet inden for forskningen i evaluering. Hver af dem giver et bud på, hvori indflydelsesmulighederne for evaluering ligger. Hver af dem forholder sig til spørgsmålet om det antastelige versus det uantastelige.

Det understreges, at de følgende syv anvendelsesformer er analytiske begreber. Ved hjælp af hvert af disse begreber kan man se aspekter af virkeligheden. Flere af de syv begreber svarer i øvrigt til, hvad mange populært mener, er evalueringers funktion helt generelt.

Men samtidig med, at de hver især kaster lys over aspekter af virkeligheden, efterlader de andre dele i mørke. Hvert af de syv begreber fanger et typisk mønster i evalueringsanvendelse, som kan udledes ved hjælp af teori, systematiske undersøgelser og/eller mere casebaserede erfaringer fra evalueringspraksis. I det følgende gennemgås hver af disse så at sige på dets egne præmisser. Det er imidlertid kun som begreber, at de kan adskilles. I virkelighedens verden er det meget sandsynligt, at flere af disse anvendelsesformer er filtret sammen.

Evaluering som kontrol

Den første anvendelsesform er evalueringer som kontrol. Typisk bruger overordnede myndigheder evaluering til at kontrollere underliggende myndigheder. Typisk er det centrale spørgsmål, om de underordnede dele af systemet har leveret ydelser, der lever op til de formelt fastsatte målsætninger i systemet som helhed. Den typiske beslutningssituation er relativt endimensionel: Skal indsatsen fortsætte eller nedlægges, skal den have flere penge eller færre penge, og skal der i øvrigt uddeles positive eller negative sanktioner.

I et kontrolsigte ønsker man typisk at beskrive relativt komplicerede aktiviteter og udfald heraf ved hjælp af relativt simple kriterier. Dette er forudsætningen for, at der er god styringsøkonomi i et givet kontroltiltag. At beskrive noget komplekst ved hjælp af noget komplekst er sjældent ressourcemæssigt fornuftigt. I et kontrolperspektiv kan der således være organisatoriske grunde til at vælge relativt simple evalueringskriterier, også selv om andre kan mere, at kriterierne ikke er fuldstændigt dækkende og valide udtryk for kompleksiteten i de processer og ydelser, der evalueres.

Ikke overraskende har kontrolevalueringer et ret dårligt image hos de kontrollerede. Det er ikke rart at få en kompleks indsats belyst ved hjælp af simple kriterier, især når det kan føre til sanktioner.

Brugen af evalueringer til kontrol er imidlertid ikke afhængig af at undgå en indvending af denne type (selv om kontrollanten selvfølgelig ofte vil argumentere, at evalueringskriterierne trods alt er ganske rimelige). Legitimiteten i brugen af evalueringer til kontrol ligger i stedet i den funktion, som evaluering har i forhold til hele forvaltningssystemet. Evert Vedung (1998) peger således på, at evaluering i forhold til det repræsentative demokratis styringskæde kan tjene hele fire legitime kontrolfunktioner: Borgere kan holde øje med politikerne; politikerne kan holde øje med embedsmændene; embedsmændene kan holde øje med fagfolkene, der udfører en bestemt indsats. Og endelig kan brugere og borgere holde direkte øje med de fagfolk, som de modtager ydelser fra.

Kontrolanvendelsen kan kendes på, at der sendes en positiv eller negativ sanktion tilbage fra en overordnet instans i styrekæden til en underordnet instans. Sondringen mellem det antastelige og det uantastelige er i princippet ganske klar. Det antastes, om de forskellige led i kæden faktisk leverer det, som er blevet lovet. Evalueringen finder sted på et uantastet grundlag, nemlig de demokratiske forestillinger om den parlamentariske styringskæde. Således har evaluator at holde sig til et evalueringsopdrag og nogle kriterier, der er konsistente hermed. I praksis kan der naturligvis være forskel-

lig grader af antastelighed i forhold som evaluators kompetence, fortolkning af opdraget, valg af metoder og kriterier og så videre, ligesom nogle interessenter kan operere med legitimitetskriterier, der rækker langt videre end den parlamentariske styringskæde. Men kernen i ideen med at bruge evalueringer til kontrol opererer med ganske veldefinerede opdelinger mellem hvad der kan antastes, og hvad der ikke kan.

Det betyder også, at der i forhold til magtudøvelsen egentlig ikke er noget problem i denne form for evalueringsanvendelse. Magtudøvelsen er i hvert fald ikke illegitim. Der er kun tale om illegitim magtanvendelse, hvis nogen sniger sig til at påkalde sig en plads i det politiske systems styrekæde, som der reelt ikke er hjemmel til, altså opfinder et mandat, som der ikke er dækning for. Men i så fald har kontrol-anvendelsen af evaluering allerede forladt sit udgangspunkt.

Evaluering som læring

Sigtet med evaluering som læring er ikke at holde opsyn og uddele sanktioner, men at forbedre indsatsen på et område. Der er derfor typisk tale om andre metoder og data end kontrolevaeringernes. Der er ikke så stor interesse for målinger i forhold til standardiserede kriterier, men mere for procesbeskrivelser og kontekstnære metoder. Der er også typisk tale om en helt anden mentalitet hos de involverede parter. Hvor kontrolevaeringerne typisk leder til, at man viser sine bedste sider og skjuler sine dårlige, så godt man kan, forudsætter læringsevalueringer en åben og tillidsfuld fremlægning især af ens egne svageste sider, for det er her, at læringspotentialet er størst. Ikke mindst derfor opererer læringsorienteret evalueringsarbejde ofte med deltagerorienterede og dialogprægede metoder. Det gøres også for at øge sandsynligheden for, at evalueringens resultater opleves som relevante og brugbare for de personer, der er ansvarlige for at udvikle indsatsen.

Også læringsanvendelsen af evaluering opererer med en klar sondring mellem det antastelige og det uantastelige. Hidtidig praksis er antastelig og både kan og bør gøres til genstand for refleksion. Til gengæld er selve ideen om læring uantastelig. Læring er per definition altid til det bedre. Læring repræsenterer et trin opad i et indsigts- eller kompetencehierarki. Man bliver gennem læring altid bedre til at løse bestemte opgaver eller klare "fremtidens krav og udfordringer" (som den, der bruger læringsbegrebet åbenbart er bekendt med).

Evaluering til læring opererer således ofte med en tavs antagelse om, hvad det er, der skal læres i forhold til og dermed hvordan indsigts- og kompetencehierarkiet er ordnet (før man kan begynde at stige opad). Denne

tavse antagelse kan være rodfæstet hos evaluator eller være mere eller mindre konsensusbaseret blandt de vigtigste interessenter. Afgørende for hele læringsbegrebet er imidlertid, at læring altid er noget positivt.

Af samme grund skildres modstand mod læring som regel som uhenigtsmæssig og uheldig. Modstand mod læring skyldes en psykologisk brist eller en mangel på kognitiv kapacitet. Ligeledes tales i dag ganske ofte om organisatorisk læring; i så fald består barrieren mod læring i uhenigtsmæssige kommunikations- og samarbejds mønstre på det organisatoriske plan. Problemet er i begge tilfælde, at man beklageligvis ikke er i stand til at lære. Der er ikke tale om decideret uvilje. Læringsproblemer henføres således ikke til en strid om værdier mellem evaluator og den lærende. På den måde "udglatter" læringsanvendelsen eventuelle værdikonflikter og taler i stedet om øget kompetence, udvikling, læring, forandringsparathed osv. (Om evalueringens udglattende funktion se endvidere Schwandt, 2003).

Kontrolanvendelse og læringsanvendelse opstilles ofte som diametrale modsætninger. Der kan da også som skitseret være ganske god grund til at understrege de typiske forskelle mellem dem både i sigte, design og faktisk anvendelse.¹

Imidlertid er der også store fællestræk mellem kontrol- og læringsanvendelsen. I begge tilfælde er det kendt og velafgrænset, hvem der anvender evalueringen (henholdsvis et overordnet led i styrekæden eller de, der faktisk udfører en indsats). Anvendelsen er ligeledes klart afgrænset i tid og rum. Det handler om som umiddelbar følge af evalueringen at fortsætte, nedlægge eller udvikle den pågældende indsats. Således opererer begreber om begge anvendelsesformer med, at anvendelsen er forudsigelig ud fra evalueringens formål. Effekterne af evaluering er intet andet, end hvad man på forhånd sagde, de ville være. Vi har altså i disse anvendelsesformer først en afklaring, dernæst en anvendelse. Forholdet mellem formål og funktion er i begge tilfælde 1:1. Derfor kan både tilhængere af kontrol og tilhængere af læring enes om, at det er vigtigt at sikre en afklaring af formålet med evaluering, inden man går i gang. Denne forskrift er kun meningsfuld under den antagelse, at der er en snæver sammenhæng mellem formålet og den faktiske anvendelse. Denne antagelse deles af begge anvendelsesforståelser.

I den henseende er både kontrol og læring instrumentelle anvendelsesformer. Med instrumentel anvendelse menes, at evalueringernes faktiske virkninger på en given praksis er identiske med, hvad man siger, er evalueringernes officielle formål. Man får den funktion, man har bestilt. Man kan bruge formålserklæringen til at forudsige virkningerne med.

I kontrolperspektivet kan man ikke se andre virkninger end den ønskede kontrol. I læringsperspektivet kan man ikke se andre virkninger end den ønskede læring. Til gengæld får man i hvert perspektiv nøjagtig det, som man forventer. Funktion er lig med sigte.

Strategisk og taktisk evalueringsanvendelse

Begreber om instrumentel evalueringsanvendelse har været kritiseret for at overse, i hvor høj grad evalueringer er politiske fænomener. Ud fra en opfattelse af organisationer som præget af interesser, ideologier og magtkampe, er derfor udviklet begreber om hhv. strategisk og taktisk evalueringsanvendelse. I begge disse tjener evalueringer til at fremme skjulte, uofficielle formål. Således bruges evalueringsresultater strategisk til at enten torpedere eller hvidvaske bestemte aktiviteter, institutioner eller faggrupper. Der kommer ingen reel ny viden frem. Evalueringen bruges kun som ammunition til at fremme positioner, der allerede var indtaget, før evalueringen blev gennemført. Den strategiske evalueringsanvendelse kan udføres med større eller mindre finesse i det kunstneriske udtryk. Der kan skrives et mere eller mindre selektivt evalueringskommissorium. Evaluator kan vælges med større eller mindre omhu. Evaluators arbejdsbetingelser undervejs kan påvirkes. Informationer kan skjules eller formidles til de rette. Udvalgte evalueringsresultater kan blæses ud af proportioner. Sluttelig torpederes uønskede indsatser og institutioner. Alt dette er imidlertid ikke overraskende, men ganske forudsigeligt, idet det kan bruges til at fremme den underliggende, men skjulte politiske dagsorden.

Den taktiske evalueringsanvendelse tjener også skjulte formål, men her er det evalueringsprocessen snarere end evalueringsresultaterne, der bruges. Når evaluering bruges taktisk, afviser man en kritisk diskussion med henvisning til, at en bestemt aktivitet evalueres. Man bruger evaluering som en politisk "syltekrukke", hvor ubekvemme sager gemmes væk eller henlægges til bedre tider.

I den strategiske og taktiske evalueringsanvendelse opereres der også med klare begreber om det antastelige og det uantastelige. Hele udgangspunktet for evaluering er nemlig antasteligt, idet evaluering bruges til at fremme en skjult politisk dagsorden. Det betyder, at det er svært at identificere den strategiske eller taktiske evalueringsanvendelse på noget ydre kendetegn. Det er f.eks. ikke i sig selv forkasteligt at bruge en negativ evaluering til at nedlægge en indsats med, som reelt er dårligt fungerende. Evalueringsanvendelsen er kun strategisk, hvis man alternativt alligevel havde ønsket at nedlægge indsatsen, selv om evalueringsresultaterne havde været

positive. En klar identifikation af en strategisk evalueringsanvendelse forudsætter således data om den kontrafaktiske situation (altså en situation, der faktisk ikke indtræffer i praksis) (Vedung, 1998). Alternativt kunne man identificere en strategisk evalueringsanvendelse på, at en bestemt politisk aktør havde forkastelige motiver, altså at vedkommende spekulerede i at bruge evaluering til et politisk formål. Men det er empirisk vanskeligt at få indsigt i en politisk aktørs indre motivation. Hvordan kan man med sikkerhed vide, hvad en politisk aktør egentlig er ude på? Et tilsvarende ræsonnement gør sig gældende ikke blot for den strategiske, men også den taktiske evalueringsanvendelse. At evalueringer tager tid er givet, men det er illegitimt at spekulere i at nøle. Det er spekuleringen i tidsforbruget snarere end tidsforbruget i sig selv, der gør en evalueringsanvendelse til taktisk. Det betyder i praksis, at den strategiske og den taktiske evalueringsanvendelse er ganske vanskelige at identificere i praksis. Med andre ord er det let at have mistanke om, at andre bruger evalueringer strategisk og taktisk, men at eftervise en sådan påstand forskningsmæssigt er temmelig krævende.

Selv om grundlaget for evalueringen pr. definition er antasteligt i den strategiske og taktiske evalueringsanvendelse – hvor det var uantasteligt for kontrol og læring – har alle disse fire evalueringsanvendelsesformer alligevel en del tilfælles. Der er fortsat god sammenhæng mellem intention og udfald, mellem evalueringens formål og dens faktiske funktion. Hvis man skal forstå evalueringernes faktiske funktion, så kan den logisk (men ikke empirisk) udledes af bestemte aktørers bagvedliggende intentioner, selv om de strategiske og taktiske intentioner er mere gustne og altså antastelige end tilfældet er, når sigtet er mere ”renfærdigt” såsom kontrol eller læring.

Yderligere et fællestræk er, at kontrol, læring, strategisk og taktisk evalueringsanvendelse alle typisk opererer med, at evalueringens udfaldsrum typisk omfatter handlinger i forhold til evaluanden (den evaluerede indsats). Evaluanden bevares, nedlægges, justeres – eller i det mindste beskyttes, mens evalueringen står på. Evalueringens udfaldsrum følger evaluandens afgrænsning i tid og rum. Det kan naturligvis ikke udelukkes, at den politiske aktør har interesser, der rækker videre end evaluanden. Ligeledes kan en evaluering indgå i en politisk studehandel om ganske andre anliggender. Men slutter man ud fra politiske aktørers adfærd til et begreb om strategisk eller taktisk evalueringsanvendelse, vil man ofte gøre det ud fra den politiske aktørs interesser i evaluanden.

Oplysende evalueringsanvendelse

Den amerikanske evalueringsforsker Carol Weiss har undersøgt instrumentel anvendelse af evaluering og er nået frem til, at den finder sted ret sjældent (Weiss & Bucuvalas, 1980). Hun har i stedet som den første lagt et langt bredere perspektiv i tid og rum på evalueringsanvendelse. I stedet peger hun på den oplysende brug som et positivt aspekt ved evalueringsarbejde. Her anvendes evalueringsresultater ikke direkte i forhold til evaluanden. I stedet spredes viden fra evalueringer i form af ”knowledge creep”, gradvis vidensspredning, inden for forskellige faglige felter, policy-felter og organisatoriske felter. Viden, begreber og ideer spredes langsomt og gradvist.

Evalueringer kan føje sig til hinanden ligesom forskningsresultater, der oplyser resten af samfundet uden at noget enkeltresultat nødvendigvis kan påberåbe sig særlig sandhedsværdi. På længere sigt antager man imidlertid, at de mest pålidelige evalueringsresultater vinder over de andre. De gode ideer spredes, og de dårlige taber i den oplyste debat.

Den oplysende evalueringsanvendelse er således positivt ladet. Man kunne hævde, at de samme processer, som leder til oplysende brug af evalueringer, også kan lede til formørkende brug af evalueringer. Fortalere for oplysende evalueringsanvendelse vil dog fastholde, at evalueringer i det lange løb er principielt formålstjenlige. Det formål, der tilgodeses, er imidlertid det større offentlige oplysende formål, ikke den enkelte evaluerings direkte formål. Forholdet mellem en evaluerings formål og funktion er således langt mere subtilt end tilfældet er i kontrol, læring, strategisk og taktisk evalueringsanvendelse. Den enkelte evaluering har nemlig ofte til formål at betjene en konkret opdragsgiver med en specifikt efterspurgt evalueringsviden, mens evalueringens oplysende funktion sjældent ophøjes til at være det primære evalueringsformål, medmindre der er tale om egentlig forskning med evaluering som sidefunktion. Mange evaluatore og i særdeleshed evalueringsforskere vil trods alt gerne satse så meget som muligt på evalueringernes oplysende funktion, idet man er opmærksom på, at en overdreven fokus på den enkelte opdragsgivers anvendelse af den enkelte evaluering på længere sigt kan svække både evaluators frihed til at problematisere overordnede policy-beslutninger og svække den overordnede, langsigtede samfundsmæssige gavn af evaluering (Henry, 2000).

Tilhængere af den oplysende evalueringsanvendelse hænger således en stor del af deres forventninger på en knage af håb: Den vigtigste anvendelse er ikke en funktion af evalueringens formelle sigte, snarere tværtimod. I stedet håber man på en indirekte og langsigtede oplysende funktion af eva-

luering, hjulpet på vej af en offentlighed, der medvirker til at optimere udbyttet af mange evalueringer over tid. Offentligheden antages således at rumme en ”medvirkende motor”, der trækker evalueringerne i retning af formålstjenlighed.²

Carol Weiss’ begreb om oplysende evalueringsanvendelse er udviklet i en tid, hvor medier ikke spillede samme rolle, som de gør i dag. I dag bærer nyhedsmedier, internationale organisationer og internettet evalueringsresultater hurtigt videre i tid og rum. I den forbindelse tiltager evalueringer sig nogle funktioner, som næppe let lader sig indordne under den oplysende funktion, som Weiss fandt. Her tænkes på

- evalueringer som nyheder. Nyhedsmedier bringer historier, hvor evalueringsresultater sættes ind i nyhedsmæssige sammenhænge. Nyheder er ofte sort-hvide og konfliktprægede. Nyheder kan ikke rumme en sags lange og komplekse forhistorie. Nyheder glemmes hurtigt. Sådanne kriterier spiller formodentlig med, når evalueringsresultater præsenteres i medieme.
- evalueringer som images. Et image er en forestilling, som udmærker sig ved sin billedskabende kraft. For eksempel er forestillingen om et internationalt læsekapløb et image, som har knyttet sig til de offentlige billeder af de senere års internationale læseundersøgelser (Stronach, 1999).
- evalueringer kan bruges som PR og reklame. Offentlige organisationer, der befinder sig i en konkurrencesituation, vil formodentlig i stigende grad benytte eventuelle positive evalueringsresultater i markedsføringsmæssig sammenhæng. Tolkningen af evalueringsresultater kan der med ske med markedsføringsproget som kode.
- evalueringer som forbrugerinformation. Her benyttes evalueringsresultater af en offentlig organisation som en slags forbrugertest, der kan understøtte en beslutning om at vælge den pågældende institution til eller fra.
- offentliggørelse af evalueringsresultater kan endvidere virke som et forstærket incitament for den enkelte institution til at forbedre sine resultater. Incitamentet kan formodentlig både sætte sig igennem på indholdet i ydelsen såvel som i forsøg på at påvirke målingen af ydelsens kvalitet.
- evaluering kan endvidere virke som socialt mærkat. Det sker, når bestemte institutioner og/eller bestemte sociale kategorier via evaluering stemples svarende til deres evalueringsscore (f.eks. skoler nederst på en

rangliste). Det er en kendt sag fra socialforskning, at sådanne mærkater kan skabe deres egen realitet, ikke mindst når de mærkede efterfølgerne forsøger at frigøre sig fra mærkningen eller forsøger at udnytte den.

Disse mulige anvendelser af evaluering på den offentlige arena antyder, at evalueringer kan have funktioner, der ikke let lader sig indordne under Carol Weiss' optimistiske begreb om "oplysning".

I begreber om de to sidste anvendelsesformer – som benævnes hhv. den symbolske og den konstitutive – brydes mere radikalt med forestillingen om formålstjenlighed i evalueringsanvendelsen.

Den symbolske evalueringsanvendelse

Når evalueringer anvendes symbolsk, handler det ikke om at bruge evalueringresultaterne, hverken i forhold til evaluanden eller i forhold til oplysning på den større, samfundsmæssige arena. I stedet handler det om at signalere, at man er en seriøs fagperson, ledelse eller organisation. At stå for evaluering er positivt i sig selv. Signalet sendes om, at man er parat til at lytte, at man tager velinformerede beslutninger (Feldman & March, 1981), og at man er omstillingsparat og fleksibel. Man tager kontingens på sig, når man evaluerer, og dermed signalerer man, at man er på omgangshøjde med den refleksive modernitets kulturelle betingelser (Dahler-Larsen, 1998).

I denne forstand betyder evaluering som ritual en tilslutning til den refleksive modernitets kulturelle koder. Et sådant ritual er imidlertid ikke ligegyldigt. I et demokrati er det vigtigt, at der gives beretninger om ansvarlighed (March & Olsen, 1995). Et andet vigtigt aspekt af den symbolske evalueringsanvendelse er, at den fungerer som hjælpemiddel for en eller flere af de andre evalueringsformer på listen. Man kan slå på den symbolske evalueringsanvendelse for at enten fremme eller modvirke f.eks. kontrollerende, lærende, strategisk eller taktisk evalueringsanvendelse. Det symbolske kan således ikke nødvendigvis isoleres som rent symbolsk (Tompkins, 1987; Røvik, 1998).

Konstitutiv evalueringsanvendelse

Den syvende og sidste, såkaldte konstitutive evalueringsanvendelse, opererer med, at evalueringer øver formende og skabende indflydelse på virkeligheden. Denne indflydelse rækker imidlertid hinsides det planlagte, tilsigtede og gennemskuelige. Evalueringer virker ikke mindst gennem det sprog, som de bærer frem. Evalueringssproget udnævner noget til at være "en evaluand". Mere eller mindre underforstået er der et vist problem ved

denne evaluand, for eksempel ”der er for store forskelle i effektiviteten på institution X og Y”. Evalueringskriterierne udtrykker, hvad der er vigtigt ved evaluanden.

Sådanne konstitutive virkninger af evalueringer kan sætte sig i tre aspekter af den sociale virkelighed, nemlig i det saglige aspekt, det tidlige aspekt og endelig i sociale relationer og identiteter.

Det er en klassisk iagttagelse i organisationsteorien, at måden en organisation evalueres på, kan komme til at påvirke dens opfattelse af, hvad der er indholdsmæssigt og sagligt væsentligt ved arbejdets resultater (Ridgway, 1956; Jasinsky, 1956) – inklusive, at en organisation, der evalueres dysfunktionelt, vil komme til at fungere dysfunktionelt. F.eks. blev den polske møbelindustri før jerntæppets fald vurderet på, hvor mange tons møbler, den kunne producere pr. år. Det medførte, at det polske møblement blev verdens tungeste.

I en undervisningspraksis kan effekten være den såkaldte ”teaching-to-the-test”. Det betyder, at undervisningen kommer til at handle om at score så højt som muligt på givne test- eller evalueringskriterier. Disse kriterier bliver målsætninger i sig selv.

Visse evalueringskriterier er kulturelt favoriserede i en given kulturel epoke, uanset om deres forhold til en given praksis er særlig gennemtænkt. For eksempel har brugertilfredshedskriteriet i vor tid opnået status som et kriterium, der ofte tages for givet i evalueringsarbejde. Dette kriteriums selvfølgelighedspræg giver af og til problemer. For eksempel tøver politiet med at offentliggøre antallet af brugerklager fordelt på politistationer, fordi, som det siges, politiet mener at have andre opgaver end at stille brugere tilfreds.

I undervisningssektoren giver brugertilfredshedskriteriet også problemer. Det skyldes, at der er en treleddet relation mellem den gode undervisning og brugertilfredsheden. For det første er visse forhold ved brugertilfredsheden ganske vist sammenfaldende med den gode undervisning. Det gælder f.eks. hvorvidt læreren formår at facilitere et positivt læringsmiljø, hvor deltagerne i en undervisning af egen drift arbejder videre med stoffet, simpelthen fordi det er lykkedes læreren at formidle, hvor spændende det er.

For det andet findes der aspekter af brugertilfredsheden, der er den gode undervisning uvedkommende. Om der f.eks. er forklippede huller i udleverede A4-ark, således at de direkte kan sættes ind i deltagerens arbejdsmappe, kan af og til betyde liv eller død for brugere, men er næppe en iboende egenskab ved god undervisning som sådan. For sådanne forholds ved-

kommende står brugertilfredsheden teknisk udtrykt ortogonalt på den gode undervisning – de har ingenting med hinanden at gøre. De måler noget forskelligt.

Endelig er der for det tredje nogle forhold, hvor brugertilfredsheden kan være negativt korreleret med den gode undervisning. Det gælder ikke blot, når undervisningen stiller uventet store krav til deltagerens arbejdsindsats eller tilpasning til nye læringsformer. Det gælder også, når man må aflære noget, man tror man godt ved, for at få plads til den nye faglige begrebsdannelse og indlæring. En forståelse af visse fysiske love kræver, at der midlertidigt dispenseres fra sund fornuft. Tilsvarende forudsætter en tilegnelse af f.eks. psykologiske eller sociologiske tankegange, at man kan parkere gængse indlærte selvforståelser og kulturforståelser. At aflære noget, man tror, man ved, er næppe udpræget brugervenligt.

Hvis brugertilfredshedskriteriet imidlertid fastlåses som institutionelt givet evalueringskriterium, kan den konstitutive effekt være, at underviserne forsøger at undgå de situationer, hvor den gode undervisning ikke korrelerer med brugertilfredsheden og i stedet fremme de situationer, hvor den gør. Men hvis man forsøger at styre uden om udfordringer til brugernes forforståelser og selvforståelser og uden om ubekvemme læringssituationer, kan selve undervisningens faglige og pædagogiske indholdsbestemmelse påvirkes.

De ovenstående eksempler på konstitutive virkninger viser den såkaldte indikatorfiksering, altså at evalueringskriteriet eller indikatoren så at sige ender med at blive mere vigtig end det gode arbejde som sådan. Indikatorfikseringen findes, når simple målbare evalueringskriterier kommer til at stå for hele det gode arbejde, uanset om disse kriterier er dårlige udtryk for kvaliteten i det, der skal måles. I metodisk sprog vil man her tale om, at indikatoren har *lav validitet*, den måler ikke det, den skal måle. Indikatorfikseringen findes imidlertid også i mere avanceret form. Den findes, når indikatoren begrebsligt set *bliver* definitionen på kvalitet. Det sker, hvis diskussionen af, hvad god læsning er, erstattes af en operationel (målbart) definition på læsning eller med andre ord netop en indikator. Magtperspektivet her ligger lige for: En avanceret indikatorfiksering sikrer ikke blot, at adfærden styres i retning af indikatoren, den sørger også for at sætte sig definatorisk på det begreb, som hele indsatsen handler om.

Der kan opstå konstitutive virkninger ikke blot i forhold til det saglige indhold i en offentlig indsats, men også i forhold til dens tidsorientering. En hospitalsafdeling kvalitetsvurderes f.eks. løbende på, om den inden for senest 14 dage efter undersøgelse af en patient tilsender en epikrise (beskri-

velse af undersøgelsens resultater) til patientens praktiserende læge. Den pågældende afdeling sender imidlertid ofte en blodprøve til klinisk-kemisk afdeling, som er mere end 14 dage om analysen. For at leve op til kvalitetsdefinitionens deadline fremsender afdelingen alligevel en epikrise til praktiserende læge, selv om brevet er indholdsmæssigt tomt. Når undersøgelsesresultaterne senere faktisk forelægges, fremsendes den egentlige epikrise. For at leve op til kvalitetskriteriernes tidsdimension må afdelingen således sende to epikriser, en for at leve op til kvalitetssystemet og en for at leve op til formålet med en epikrise.

I England lavede politikerne en aftale med nogle hospitaler om at reducere ventelisterne til gengæld for en ekstrabevilling. Ventelisternes omfang blev defineret som antal måneder gange antal personer fra første konsultation på hospitalet til den faktiske indlæggelse. Det viste sig hurtigt, at projektet lykkedes. Imidlertid klarede hospitalerne opgaven ved at undlade at tage patienter ind til første konsultation. I dette eksempel som i det forrige har evalueringen af noget tidligt ud over de tidlige konsekvenser tillige nogle indholdsmæssige.

Effekterne af undervisning, ikke mindst af samfundets nye medlemmer, har i princippet et vældigt langt tidsperspektiv. Evalueringer må i princippet forkorte effektbilledets tidshorizont, i og med at der skal findes udtryk her og nu for undervisningens værdi og/eller effekter i evalueringøjeblikket. Dette problem kan ikke løses metodisk; begrænsningen er rent logisk. På grund af gabet i tid mellem undervisningens principielt meget langsigtede virkninger og evalueringens bundethed til nutiden er undervisning moden for en række forestillingsmæssige og ideologiske konstruktioner af, hvad der kan tænkes at ske efter evalueringen og fremover som resultat af undervisningen. Men evalueringssystemer må alligevel operere med en eller anden vedtagne deadline. Det, som kan fanges inden deadline, tæller med.

Også på kortere sigt er en evalueringens billede af undervisningen tidsfølsomt. Menneskelige oplevelser, erkendelser og nye erfaringer vurderes forskelligt i lyset af den valør, som de tilskrives i efterfølgende fortællinger; disse fortællinger er altid fæstnede på et bestemt udsigelsestidspunkt (Horsdal, 1999: 119). Nogle undervisningsforløb giver hurtige erkendelsesmæssige gevinster, andre langsomme. Deltageres efterfølgende påskønnelse heraf kan være hurtig eller langsom. Intet evalueringensbillede fæstnet i tid er derfor nødvendigvis "det korrekte". Men hvert tidspunkt kan give sit billede.

Evalueringernes konstitutive effekt langs den tidsmæssige dimension består i, at man kan komme til at indrette en indsats efter, at (erkendelsen

af) dens evaluerbare effekter skal vise sig på det tidspunkt, hvor en given evaluering finder sted. En evalueringsmæssig deadline vil, som flere af ovenstående eksempler viser, sjældent være neutral over for bestemmelsen af indsatsens indhold. Den vil heller ikke være neutral over for konstitutionen af forskellige sociale relationer og identiteter. Hermed menes groft sagt den sproglige markering af, hvem der har noget at skulle have sagt i relation til evaluanden.

Samtidig med, at en evaluering kommunikerer om evaluandens indhold og dens tidsperspektiv, metakommunikerer den også om de involverede parter identiteter og relationer. En given evaluering antyder, hvem der har ret til at udtale sig, hvilke rettigheder og pligter man har i forhold til evaluanden, hvad der er forventeligt og normalt, og hvem der har ansvar for at iværksætte evt. ændringer som følge af evalueringsarbejdet. Evalueringen er også med til at bestemme, om de identiteter, der knytter sig til evaluanden, skal have lov til fortsat at eksistere. Hvis evalueringen udfordrer evaluanden, udfordrer den også de mennesker, der får deres identitet gennem evaluanden.

Af særlig identitetsmæssig interesse er i vor tid kategorien ”brugeren”. Ganske ofte former evalueringer i vor tid forholdet mellem de involverede parter som et forhold mellem en bruger og komplementært dertil en serviceleverandør. Brugeren er historisk en relativt ny kategori (Vedung, 1998). Konstituerende for brugeren er typisk en individualiseret valgsituation og et sæt præferencer, der konfronteres med et givet servicetilbud udmønter sig i en tilfredshedsgrad. Det bemærkes, at brugeren ikke bestemmes ud fra et ansvar for politikdannelsen eller dens konsekvenser. Det bemærkes også, at brugeren (ligesom sin nære slægtning for-brugeren) samtidig med sin tilsyneladende sociale statusfremhævelse er genstand for en lang række kontrol-, socialiserings- og normaliseringsmekanismer (Bauman, 1983).

Evalueringer vurderer ikke blot den offentlige indsats, men er også anvendelige til at kategorisere forskellige brugere. På nogle hospitaler tilordnes patienter en DRG-værdi, dvs. et regnskabsmæssigt kronebeløb, som tilfalder den behandlende afdeling ud fra patientens ”diagnose-relaterede gruppe”. Hospitalsafdelingens blik på patienten kan påvirkes af, om det vurderes, at man faktisk kan behandle patienten for det stipulerede beløb. I hvert fald evalueres afdelingen på, om det lykkes.

Gennem tests og evalueringer placeres skolens elever i forhold til, om deres faglige og personlige udvikling er foran, på niveau med, eller bagud i forhold til det forventede på deres klassetrin, ligesom et eventuelt særligt

behov for specialundervisning kan komme på tale. Gennem deltagelse i eksamener, tests og evalueringer lærer nogle elever gradvist at blive egnede subjekter for evaluering.

Også på de offentligt ansattes side kan evalueringer naturligvis være medkonstituerende. For eksempel kan lærere skildres som fagprofessionelle, som konsulenter og sparringspartnere på elevernes læreprocesser, som serviceleverandører og som ansatte. Mange har traditionelt antaget, at læreridentiteten som andre fagprofessionsidentiteter er mere eller mindre givet med uddannelsen og jobbet; nyere forskning anlægger imidlertid det perspektiv, at fagprofessionel identitet står til løbende konstruktion. I et sådant perspektiv tilbyder professionelle og organisatoriske processer som evaluering nogle byggesten, som læreren kan arbejde videre med i sit løbende identitetsarbejde (Halford & Leonard, 1999: 107). Evalueringer giver tilbagemeldinger at tolke på, og enhver bruger tilbagemeldinger fra omverdenen i sit identitetsarbejde. Forskellige slags evalueringer giver imidlertid ganske forskellige muligheder. Evalueringer baseret på standardiserede mål har en tendens til at trække det professionelle element ud af den professionelle; en professionel er nemlig i princippet en, der har et mandat til at bruge kompleks viden til at træffe varierende beslutninger med i forhold til komplekse situationer. Er der høj grad af standardisering i arbejdet (som nogle evalueringer stipulerer), så er behovet for professionel ageren så meget mindre. Andre evalueringer lægger imidlertid op til, at den fagprofessionelle netop træder i karakter. Det er også af betydning, hvilke sammenligninger, den fagprofessionelle eventuelt udsættes for. På nogle arbejdspladser evalueres den fagprofessionelle i forhold til gennemsnittet af sine kolleger. Alt andet lige vil et sådant system incitere til, at man ikke engagerer sig i kollegialt samarbejde og erfaringsudveksling.

Evalueringer har også konstitutive virkninger på hele den offentlige organisation, der evalueres.

I et vist omfang kan en organisation selv være med til at præge sin identitetsdannelse via evalueringsarbejde i og med, at den kan vælge at spejle sig i forskellige symbolske felter i omgivelserne, når den afsøger evalueringsarbejdets muligheder. I et projekt med frie kostskoler er det f.eks. set, at den enkelte skole kunne bruge evaluering til at sætte sig selv i tale som henholdsvis fællesskab, virksomhed, offentlige institution, pædagogiske miljø og politisk aktør (Dahler-Larsen, 2003b).

En organisation kan også afvise sådanne muligheder, hvis den vurderer, at dens identitet fastholdes bedst, hvis evalueringsarbejdets reflektive muligheder slet ikke afprøves. Dette kan imidlertid have betragtelige omkost-

ninger i form af tab af symbolsk anerkendelse, hvis det tidligere gennemførte ræsonnement om vor tids symbolske valorisering af evalueringsprocedurer er korrekt. Ikke at ville svare på evalueringsudfordringen er også et svar, som omverdenen kan tolke på. Dertil kommer, at evalueringskravet i dag stadig oftere formuleres som et obligatorisk, eventuelt lovgivningsmæssigt bindende krav.

Ligesom tilfældet er med evalueringers indholdsmæssige og tidsmæssige virkninger, er deres virkninger på identitet og sociale relationer næppe givet på forhånd. Lovbestemte tiltag kan selvfølgelig gøre deres. Men ganske ofte kan man formode, at den konstitutive effekt afhænger af, hvor stærkt f.eks. faglige og kollegiale værdier fastholdes i arbejdet, hvor stærk institutionel basis brugernes repræsentanter tildeles, hvilket bolværk den lokale organisationskultur eventuelt udgør, hvilke belønnings- og sanktionssystemer der i øvrigt spiller ind på en offentlig ydelses tilrettelæggelse og udførelse, samt hvorvidt forskellige kontrol- og evalueringssystemer eventuelt spiller sammen eller modvirker hinanden.

En oplagt indvending mod begrebet om konstitutive virkninger er naturligvis således, at ikke alle evalueringer når at have nogen væsentlige konstitutive virkninger. Mange glemmes formodentlig, inden de når at gøre nogen væsentlig forskel.

I figur 3.1 opsummeres de tre aspekter af en ydelse, som konstitutive virkninger kan ramme, nemlig det indholdsmæssige, det tidslige og det identitetsmæssige. Det foreslås endvidere, at de konstitutive virkninger af evaluering kan manifestere sig i tre forskellige slags ændringer, kaldet førsteordens, andenordens og tredjeordens ændringer.

I førsteordensændringer påvirkes alene det sprog, som man bruger til at beskrive problemer, kriterier og løsninger med. I andenordensændringer træffes en beslutning – en sanktioneret begivenhed, der markerer en officiel opmærksomhed, for eksempel organisatorisk og/eller politisk opmærksomhed. Det kan for eksempel være en beslutning om at tildele ekstra timer til et hidtidigt underprioriteret indsatsområde. I tredjeordensændringer påvirkes den institutionelle orden på et bestemt område. Et eksempel herpå er oprettelsen af et evalueringsinstitut.

Figur 3.1. kan bruges til at vise, at ikke alle konstitutive virkninger af evaluering når lige langt. Nogle evalueringer når kun at demonstrere en smule af deres egen evalueringstale og klinger derefter af. Andre kan imidlertid have langt mere vidtrækkende virkninger.

Konstitutive virkninger kan forplante sig både på langs og på tværs i figuren. En evaluering af, hvorvidt elever har nået fastsatte mål i tredje, sjette

og niende klassetrin er formuleret som tidslige krav, men kan have konsekvenser både for undervisningens indhold og for sociale identiteter og relationer – specielt for de børn, som viser sig ikke at nå de fastsatte mål.

Figur 3.1. Typer af konstitutive virkninger

Domæne:	Ændringer af:		
	Første orden	Anden orden	Tredje orden
Indhold	Tale	Beslutning	Institutionelt arrangement
Tid			
Sociale identiteter og relationer			

En analyse af "Fælles Sprog", et monitoreringssystem på ældreområdet kendt blandt andet for sine "stregkoder", har således vist, at systemet gennem sin løbende overvågning medvirker til at flytte forankringen af omsorgen fra den indholdsmæssige og den sociale dimension til den tidslige, idet hele Fælles Sprog er bygget op på tidsenheder afsat til varetagelsen af bestemte på forhånd kodede omsorgsproblemer. Til gengæld virker den stramme tidsstyring naturligvis tilbage på den leverede omsorgs faktiske indhold (de konkrete ydelser som varierer fra gang til gang trues) og på relationen mellem de ældre og hjemmehjælperne (dobbelts umyndiggørelse) (Højlund & la Cour, 2001: 98-99).

Med andre ord er der både en vis omsættelighed og en vis afsmitning mellem de forskellige lag i skemaet. Virkninger kan forplante sig i flere led, når aktører reagerer på evalueringer, og der reageres på reaktionerne. Hvis en skole for eksempel offentligt præsenteres som "nederst på en rangliste" kan dens lærerkreds og ledelse beslutte sig for at sætte bestemte handlinger i gang. Skolen kan forsøge at skaffe ekstra midler begrundet i dens særligt vanskelige forhold. Der kan også ske en øget uddifferentiering af elever til specialklasser for, at disse ikke skal forstyrre det evalueringsbillede af skolens "normale elever", som den først og fremmest bliver målt på. Specialklasseleverne og deres forældre kan reagere på denne uddifferentiering. Samtidig kan nogle af de fagligt stærkere elever og deres forældre samt nogle af lærerne forlade skolen for ikke at blive delagtiggjort i skolens negative image i offentligheden. Skolen kan også sætte supplerende evalueringsarbejde i gang for at fremme andre billeder af skolen end ranglistens evalueringsbillede.

De konstitutive virkninger kan således ligge i kæder efter hinanden og tillige gå i forskellige retninger. De tolkningsmæssige rammer, der gøres gældende om evalueringen, er vældig afgørende for, om en evaluering medfører stærke konstitutive virkninger. Evaluerings introduktion på et nyt område er med andre ord en kompleks proces. Sahlin-Andersson (1996) beskriver en sådan indlejningsproces som en *redigering* efter visse regler, Røvik (1998) beskriver den som en *oversættelse*. Begge metaforer er interessante og understreger, at der ikke er tale om en umiddelbar og mekanisk videreførsel af et bogstaveligt budskab i en evaluering. Men metaforen om redigeringsregler gør måske processen mere programmerbar end den reelt er, og metaforen om oversættelse antyder en mulig validitetskontrol (nemlig tilbageoversættelse til originalsproget) som ikke forekommer at være nødvendig eller relevant, når et givet konkret område af praksis lever videre med sin egen version af den indoptagne evaluering.

Under en sådan betragtning er udfaldet af et evalueringsforløb ikke gjort hverken med en bestemmelse af evalueringens officielle formål eller med en fastlæggelse af bestemte evalueringskriterier. Tværtimod kan evalueringen tillægges forskellig mening undervejs via en aktiv og reflektiv medspil- len fra involverede aktører i langvarige og komplekse evalueringsforløb.

Alt efter hvor indgribende de konstitutive virkninger er, kan man tale om en række reaktionsmønstre omfattende

- afvisning, hvor evaluering så vidt muligt holdes ud fra en bestemt praksis
- hykleri, hvor evaluering foretages, men det samtidig sikres, at virkningerne kun omfatter udtalelser og erklæringer på det retoriske plan
- tæmning, hvor evaluering foretages, men der sorteres i evalueringen, i metoder, kriterier, evaluander, resultater og anvendelser af evaluering så man søger at sikre, at der kun findes virkninger af en type, som man kan leve med i en given organisatorisk sammenhæng
- kolonialisering, hvor evalueringens konstitutive virkninger over en bred kam og i dybden kommer til at præge efterfølgende praksis inden for et givet område.

Stærke tilhængere og modstandere af evaluering som sådan har ofte en tendens til at hænge sig fast i de to yderste af disse muligheder. Snarere end at tilskrive et af disse udfald automatisk til i en given situation vil det imidlertid nok være mere fornuftigt at lade det komme an på en nærmere undersøgelse, hvilket af udfaldene, der bedst passer.

Sammenfatning

I dette kapitel er præsenteret syv forskellige begreber om evalueringsanvendelse. Fortalere for evaluering har en tendens til kun at hæfte sig ved en enkelt positiv anvendelsesform, som om den var den eneste mulige. Modstandere af evaluering har en tendens til kun at hæfte sig ved en enkelt negativt ladet anvendelsesform. I begge tilfælde antages det, at evaluering kun har én anvendelsesmæssig bestemmelse.

Mens de anvendelsesformer, som er skitseret i dette kapitel, ikke er gensidigt udelukkende, lukkes op for muligheden for, at der i praksis kan være flere anvendelser af samme evaluering. Anvendelsesproblemet er således ganske komplekst, ikke mindst i takt med at virkningerne af en evaluering kan spredes gennem tid og rum, for eksempel via elektroniske medier.

Konsekvensen heraf er, at en empirisk undersøgelse højst kan opnå at indfange nogle vigtige og illustrative eksempler på en evaluerings anvendelser, man næppe kan levere en udtømmende beskrivelse.

Noter

1. Den amerikanske evalueringsforsker Michael Scriven argumenterer imidlertid for, at sondringen mellem de to (som han betegner summativ versus formativ evaluering) udelukkende er en sondring angående anvendelsen. Den samme evaluering kan altså bruges til kontrol eller læring, alt efter hvem der modtager evalueringsresultaterne, og hvad der gøres ved dem. Synspunktet er sundt i den forstand, at sigte og design ikke behøver at sige særlig meget om den faktiske anvendelse, en pointe som uddybes i det følgende (Scriven, 1991).
2. Ifølge en dansk forsker (Rieper, 1996) er vi i Danmark imidlertid ikke særlig gode til at organisere os sådan, at evalueringernes oplysende funktion fremmes (det afhænger af netværk, tidsskrifter, konferencer m.m. Og det handler om ikke at være blufærdig over andres vurdering af ens arbejde).

Kapitel 4

Evalueringsprocessens komponenter

I det foregående kapitel er evalueringens anvendelsesmuligheder beskrevet. Det er sket uafhængigt af, hvordan evalueringen forløber, herunder hvem, der blander sig i evalueringen, og hvordan den i øvrigt er forløbet. I dette kapitel analyseres evalueringsprocessen nærmere, dels i form af de dele, processen kan opdeles i, dels i form af aktører med hver deres rationaler, som kan blande sig i processen.

Evalueringsprocessens momenter

Ofte tænker man på et evalueringsforløb som en proces med enhedspræg. Ikke mindst hvis man tænker en evaluering som resultat af én ganske bestemt magtudøvelse, vil man have en tilbøjelighed til at se hele evalueringsforløbet under en sådan sammenhængende forståelse. Den samlede proces er ganske forudsigelig – som legetøjstogets rute. Ser man derimod evalueringsprocessen som en tur i radiobilen, er der ingen klar linje i hele ruten. Ruten er i stedet et resultat af de mange sammenstød, der sker i de forskellige faser i turen.

I det følgende gives et bud på en opdeling af et evalueringsforløb. Opdelingen er en analytisk konstruktion. I praksis kan momenterne overlape tidsligt. Aktører tager formodentlig flere andre momenter i betragtning, når de tager beslutning om et enkelt moment. Der er derfor ikke nødvendigvis nogen fastlagt kronologi og nogen klar adskillelse mellem momenterne – af hvilken grund, der netop opereres med ”momenter” snarere end med faser.

1. Initiering

Selv om flere forskellige aktører kan se en fordel ved evaluering, sættes den først i gang, når nogen faktisk tager et initiativ dertil. Med et initiativ til evaluering kan en given indsats flyttes fra at være hinsides diskussion til at kunne diskuteres, altså fra at være uantastet til at være antastet.

I et magtperspektiv er det formodentlig ikke ligegyldigt, hvor initiativet kommer fra. Initiativet giver et første signal om et muligt ”ejerskab” til evalueringen, og initiativet spiller ind på, hvornår evalueringen finder sted. Initiativet er imidlertid ikke altid udtrykt af en konkret aktør, for eksempel hvis der findes et lovgivningsmæssigt krav om evaluering.

2. Dagsorden

I dagsordenfasen bestemmes evalueringens formål og emne. Evalueringsspørgsmålet bestemmes, altså hvilket spørgsmål, evalueringen skal besvare. At gøre en indsats til genstand for evaluering sender et dobbelttydigt signal, nemlig at indsatsen er væsentlig nok, men også problematisk nok til at blive evalueret. Ikke mindst afgrænsningen af evaluanden (evalueringsgenstanden) er interessant i et magtperspektiv: Hvad skal evalueres, og hvad holdes udenfor? I dagsordenen indgår endvidere implicit eller eksplicit nogle værdier og/eller standarder, som evaluanden holdes op imod.

I dagsordenen formuleres således den fortolkningsmæssige ramme for, hvad evalueringen handler om, og hvor god en offentlig ydelse skal være for at være god. I det omfang, som de efterfølgende momenter er selvstændige, kan dagsordenens beslutninger eventuelt ændres og nyfortolkes. Men af og til udstedes der i dagsordenfastsættelsen løfter og erklæringer på skrift (f.eks. i et kommissorium), som efterfølgende binder evalueringprocessen i visse henseender.

3. Vidensforvaltning og -organisering

Dette moment vedrører organiseringen af evalueringen, herunder spørgsmål som "hvem gør hvad", og "hvem har adgang til hvilke informationer undervejs". Disse spørgsmål kan være ganske signifikante i et magtperspektiv. Der kan for eksempel over evalueringens faser være ganske forskellige arbejdsdelinger mellem interne og eksterne parter, hvilket kan give ganske forskellige muligheder for indflydelse (Dahler-Larsen, 2003a). Der kan også knyttes symbolske fordele til selve det at indsamle data (Feldman & March, 1981), ligesom der kan være symbolske demokratiske fordele ved at lade andre blive spurgt og inddraget undervejs. Også spørgsmål om opbevaring, deling og offentliggørelse af evalueringens viden og i hvilken form, herunder om data skal anonymiseres, hører til vidensforvaltningen. Anonymisering medvirker til at skabe bestemte typer af data, men hindrer også at der tages personligt ansvar for de enkelte udsagn. Disse aspekter af evalueringsforløb er naturligvis magtmæssigt relevante. Man udfordrer eller beskytter de personer og enheder, hvis evalueringens resultater lægges frem. Ved ikke at lægge informationer frem kan man endvidere forhindre andre aktører i at handle på disse oplysninger.

4. Design

Til dette moment hører valg af evalueringsmodel og metode, herunder en metode til eventuel effektbestemmelse. Her ligger også valget af evalue-

ringskriterier, valg af tidspunkt og valg af svarpersoner i en eventuel interview- eller spørgeskemaundersøgelse. Alle disse valg øver indflydelse på, hvilket billede evalueringen tegner af evaluanden. Om billedet bliver favorabelt eller ej afhænger et stykke ad vejen af netop disse valg. I valget af design påvirkes også den troværdighed, som evalueringen kan have i forskellige kredse alt efter disse kredses viden om og tiltro til forskellige designs.

5. Dataindsamling

Selv om visse beslutninger i dataindsamlingsfasen er bestemt af andre beslutninger, f.eks. i evalueringens design, rummer dataindsamlingsfasen også en del muligheder for selvstændig indflydelse. Der kan være tale om decideret manipulation med data eller blot visse diskrete udeladelser. Men også når gængse regler følges, er der indflydelsesmuligheder i dataindsamlingsfasen.

At medvirke i dataindsamlingen giver muligheder for at påvirke data både på giver- og modtagersiden. Som udfylder af selv et ganske stramt formuleret skema har man stadig frihed til at svare ja eller nej. I nogle modeller for selvevaluering skal en organisation selv svare på opstillede evalueringsspørgsmål. I så tilfælde afgiver man sine egne vurderinger samtidig med, at man medvirker i dataindsamling.

Som evaluator træffer man også i dataindsamlingen beslutninger, der i videre forstand påvirker evalueringsprocessen. Støder man på praktiske forhindringer i dataindsamlingen (adgangsproblemer, praktiske problemer, ressourceproblemer, modstand), kan disse være ganske reelle barrierer, men samtidig også blive brugt som påskud for at gå denne eller hin vej i evalueringssarbejdet. Når dataindsamlingen er relativt lidt forprogrammeret, som det for eksempel typisk er tilfældet i kvalitativ metode, træffer man i tæt samspil med dataindsamlingen tillige en række beslutninger om begrebsdannelse og tematisering i også i dataindsamlingsfasen.

At være med i dataindsamlingen og observere denne giver tillige mulighed for efterfølgende at give mere eller mindre legitimitet til data – man kan eventuelt vælge at fortælle om de iagttagelser, der svækker tiltroen til data, hvis man finder dette gavnligt.

6. Analyse og sammenfatning

I analyse og sammenfatning af evalueringsresultater sker en selektion, fortolkning og udpegning af, hvad der er det vigtigste. Forskellige analytiske principper vil lægge op til, at forskellige aspekter er vigtige. Afbalancerin-

gen af forbehold versus pointer kan gøres ganske forskelligt. Selv i en strikt statistisk test kan to forskellige signifikansniveauer føre frem til, at et resultat præsenteres som enten holdbart eller tilfældigt opstået.

7. Gyldiggørelse

I gyldiggørelsen ligger en social definition af, hvad der kan gælde som valid viden. Om data er "hårde" eller "bløde", er ikke egenskaber ved selve data, men ved de reaktioner, som man inden for et givet socialt felt knytter til de pågældende data (Hastrup, 1993). Gyldighed er med andre ord et spørgsmål om social tilskrivelse af gyldighed. Et første skridt er allerede taget i analyse og sammenfatning af evalueringresultaterne, hvor evaluator ofte argumenter for evalueringresultaternes gyldighed, typisk ud fra metodiske spilleregler og således eventuelt i en vis balance med metodiske forbehold og begrænsninger. Men mange andre forhold bestemmer gyldiggørelsen af den producerede evalueringens viden i bredere sociale, organisatoriske og politiske sammenhænge. En forudsætning er, at der overhovedet er opmærksomhed på evalueringen. Men en af de knappest ressourcer i komplekse organisatoriske systemer er netop opmærksomheden (Simon, 1969). Mange andre forhold end evaluering konkurrerer om beslutningstageres opmærksomhed (Weiss, 1983). Opmærksomhed mod en bestemt problem kan ofte kun tildeles i et ganske kort tidsvindue i en større politisk proces (Kingdon, 1984).

Opmærksomhed må tillige kombineres med relevans for at evalueringen kan opleves som et gyldigt grundlag for handling. Relevans bestemmes blandt andet af den enkelte aktørs rettedhed mod diverse praktiske gøremål i den verden, som den enkelte aktør tager for givet (Schutz, 1975). Relevans og gyldighed i politiske beslutningssammenhænge er formodentlig en meget kompleks størrelse, som ikke mindst bestemmes af muligheden for at koble bestemte problemer og bestemte løsninger på måder, der er ganske unikke for bestemte tidspunkter i komplekse beslutningssystemer (March & Olsen, 1976), lige som muligheden for at danne flertal om en sag, eventuelt på grundlag af forskellige mål eller uklarhed i målformuleringen, formodentlig spiller en rolle (Baier, March & Sætren, 1986).

8. Konsekvensdragelse

I konsekvensdragelsen ligger en fortale for, en legitimering af og en iværksættelse af handlinger som følge af den indsamlede evalueringens viden. I nogle tilfælde ligger det i en evalueringens opdrag eller i evaluators funktion i

øvrigt at udlede nogle anbefalinger af evalueringsarbejdet. I så fald er første skridt taget til en konsekvensdragelse.

Grundlæggende forskellige strategier kan vælges, hvis et program eller en programkomponent viser sig at fungere dårligt: reparation eller udskiftning. Under reparation hører sådanne strategier som efteruddannelse, kompetenceudvikling, organisatorisk læring osv. Under udskiftning hører nedlæggelse kombineret med oprettelse af noget andet, flytning af ansvar for opgaver og/eller øget markedsstyring således at dårligt fungerende enheder "dør af sig selv". I nogle tilfælde ligger det i en evaluerings dagsorden, at den overvejer en eller flere af disse strategier i sine anbefalinger. Men uanset et sådant oplæg rummer konsekvensdragelsesfasen ofte et frirum for forskellige forslag fra evaluators side. I mange tilfælde kan et givet problem logisk set løses ved hjælp af forskellige funktionelt lige gode tiltag. I nogle tilfælde har evaluator selv interesser i at fremme bestemte forslag til løsninger. Internt i organisationer kan ledelseshierarkiet, interesser og organisationskulturen sætte snævre rammer for, hvilke ændringsforslag, der overhovedet kan diskuteres.

Forskning i den politiske konsekvensdragelse af evalueringer i komplekse politiske systemer viser, at sammenhængen mellem evalueringens viden og politisk handling sjældent er direkte (Weiss, 1977; 1983; Albæk, 1988), men ganske ofte præget af tidsforskydning, partielle tolkninger og et samspil mellem mange forskellige beslutningstagere. Politiske beslutninger præges tit af ganske bemærkelsesværdige koblinger mellem "problemer" og "løsninger" (March & Olsen, 1976).

Ovenstående opdeling i otte forskellige momenter er som nævnt begrebslig/analytisk. I praksis kan der være mere eller mindre stærke sammenhænge mellem faserne, og disse sammenhænge kan være åbne eller skjulte. Et givet moment kan rumme beslutninger, som aldrig træffes eksplicit, men nærmest formes som en konsekvens af andre beslutninger i løbet af evalueringen.

Opdelingen er imidlertid stadig analytisk gavnlige for at understrege, at evalueringsforløbet forskellige dele ikke behøver at være forment af de samme kræfter. Sågar kan en løs kobling mellem en evalueringens momenter udnyttes i et magtperspektiv: Måske vil man give visse momenter væk til andre interessenter, hvis man blot selv beholder kontrollen over de mest indflydelsesrige momenter. Tilsvarende kan det være nødvendigt at insistere hårdt på et ganske bestemt moment, hvis man anticiperer, at beslutninger herom kan bruges til at træffe efterfølgende beslutninger om andre momen-

ter, som eventuelt medfører bindende konsekvenser. Alt dette afhænger formentlig af, hvilke aktører der blander sig i evalueringsforløbet.

Evalueringens aktører og rationaler

Mange aktører kan blande sig i en evaluering. Hver især kan de bringe forskellige rationaler med ind i evalueringsarbejdet. Med rationale menes en måde at tænke, tale og ræsonnere på, der kan præge, hvad en evaluering handler om, og hvordan den forløber. Med den indledende kapitels ord søger hvert af disse rationaler at fremme nogle argumenter eller tænkemåder som uantastelige.

I det følgende vil der blive præsenteret seks sådanne rationaler. Der præsenteres med andre ord *hypoteser* om, hvilke aktører der ud fra hvilke hensyn kunne tænkes at øve indflydelse på evalueringsarbejde. Der mindes om, at disse ræsonnementer handler om analytiske typer af aktører, ikke konkrete personer. Det kommer an på en prøve i den efterfølgende analyse, om disse typer har genklang i casestudier.

1. Det metodiske rationale

Ifølge denne hypotese er evalueringer i høj grad præget af metodiske hensyn. Ideen er, at det er afgørende for en evaluerings pålidelighed og troværdighed, at den metodiske kvalitet i evalueringen er så høj som muligt. Afgørende nøgleord er derfor for eksempel metodisk design, validitet, reliabilitet, generalisering osv. Det metodiske rationale er først og fremmest repræsenteret af en metodisk kyndig evaluator. Men det kan også være repræsenteret af interessenter omkring evalueringen, for hvem metodisk troværdighed er afgørende.

2. Det politiske rationale

Ifølge denne hypotese udtrykker politikere ofte deres målsætninger for offentlige indsatser meget uklart. Det hænger dels sammen med, at mange af deres vedtagelser er baseret på flertal bestående af interessenter med forskellige interesser. Det hænger også sammen med, at det giver større manøvrerum i fremtiden til at mene, hvad man har lyst til, om indsatsen. Af disse grunde er politikeres opdrag til evaluatorene ofte tilsvarende uklare. Politikere har i øvrigt almindeligvis temmelig begrænset interesse for evalueringer, fordi det giver langt større signalværdi at sætte noget nyt i gang end at undersøge, om det virker (Baier, March & Sætren, 1986), og fordi mange andre faktorer end evalueringer konkurrerer om politikernes op-

mærksomhed (Weiss, 1983). Til gengæld kan politikere i glimt udvise stor interesse for evalueringer, specielt når evalueringer strategisk eller taktisk kan bruges i forhold til dagsordener, der i øvrigt har politikernes interesse. Den strategiske eller taktiske brug forudsætter ikke, at der er noget særligt nært forhold mellem evalueringens indhold og den politiske handling, der foreslås. Det afgørende er, at politikeren kan hive evalueringen ind i argumentationen for sin sag, når det i øvrigt er opportunt for politikeren. I sådanne situationer er politikere uhyre bevidst om indholdet i beslutningerne, men også om hvordan hans/hendes navn og omdømme knyttes til bestemte sager.

3. Det administrative rationale

Aktuelle organisationsteoretikere peger på, at organisationer i stigende grad præges af en abstrakt og selvstændiggjort diskurs om, hvordan moderne organisationer skal ledes og udvikles. Der er tale om almene organisationsforståelser og -opskrifter, som spredes via medier, uddannelse og konsulenter, og som sjældent tager højde for de institutionelle og faglige særtræk ved hver enkelt organisation (Røvik, 1998). En god organisationsopskrift kan tilsyneladende bruges alle vegne, møder ingen værdimæssig modstand og lader sig let indordne i abstrakt ledelsestænkning. Den fremtræder som et kendetegn på moderne og effektiv ledelse.

Det administrative rationale ser evaluering som en organisationsopskrift, der helst skal forløbe glat og desuden være ledelsesmæssig relevant. Idealet er, at evaluering skal foretages rutinemæssigt og naturligt sammenflette en måde at samle data ind på og en måde at styre organisationen på. Er organisationen bureaukratisk, kan opstilling af et sådant procedure-styre endvidere let blive til et mål i sig selv.

Det administrative rationale vil typisk være repræsenteret ved embedsmænd eller andre, der har ansvar for store organisationers drift og omdømme.

Afhængig af situationen og omgivelsesbetingelserne kan evalueringer også forvaltes som organisatoriske ceremonier, der først og fremmest har til hensigt at fremstille organisationen i et positivt lys i forhold til omgivelserne (Meyer & Rowan, 1977) – den såkaldte symbolske evalueringsanvendelse. Organisationens omgivelser består i vidt omfang af andre bureaukrater. Bureaukrater anerkender nemlig bureaukrater, der benytter sig af moderne organisationsopskrifter.

4. Det fagprofessionelle rationale

I praksis udføres mange af den offentlige sektors opgave af fagprofessionelle (læger, lærere, dommere osv.). Den professionelle er karakteriseret ved selv at have et vist handlerum i opgaveløsningen. Meningen er, at man ud fra sin uddannelse og erfaring skal foretage kvalificerede valg i forskellige komplekse situationer. Graden af professionalisme hænger snævert sammen med, om den pågældende erhvervsgruppe har autonomi til selv at bestemme de kriterier, dens arbejde vurderes ud fra (Scott et al., 1967). I bedste fald – set fra professionens synspunkt – sker evaluering af de professionelle kun via andre medlemmer af samme profession. Typisk vil en profession foretrække evalueringskriterier, der specifikt beskriver processer, der er særegne for professionen. Ud fra et fagprofessionelt rationale vil man således ofte søge at sikre, at en evaluering er så konsistent som mulig med professionens eget vidensgrundlag og normsæt.

5. Brugerens rationale

I de senere år har ”brugeren” som kategori spillet en styrket rolle i forhold til den offentlige sektor. I kategorien ”bruger” rendyrkes et ganske snævert aspekt af tilværelsen, nemlig de interesser, forventninger og bekymringer, som knytter sig til modtagelsen af en bestemt offentlig service. Det administrationspolitiske ideal er ofte, at staten skal være mere responsiv (Jørgensen & Melander, 1999), hvilket forudsætter lydhørhed i forhold til brugere. Brugerens position søges styrket gennem brugerbestyrelser, klagemuligheder samt dialogmetoder og undersøgelser såsom evaluering.

6. Den kulturelle baggrundsstøjs rationale

Formodentlig præges evalueringer ikke kun af synlige aktører, der er på scenen for at præge den enkelte evaluering. Formodentlig betyder ”scenografien” og ”bagtæppet” også en hel del. Det forsøges indfanget i begrebet om ”den kulturelle baggrundsstøjs rationale”. I dette rationale spiller almindelige forventninger og ”tidsånden” en vigtig rolle i prægningen af evalueringer.

Det vil formodentlig være lettere at få en evaluering til at fremstå som uantastelig, hvis den opleves som værende i overensstemmelse med udbredte normer og værdier, og hvis den drager konklusioner, som svarer til, hvad mange forventer. Det er formodentlig lettere at få en evaluering til at gøre en forskel, hvis dens anbefalinger er i samsvar med ændringsprocesser, der alligevel er undervejs i samfundet. En sådan kulturel ”baggrunds-

støj” må derfor formodes at være en ressource for aktører, der kan appellere hertil for at styrke deres argumentation.

Man skal ikke undervurdere styrken af den kulturelle grundtone i det moderne samfund, der anslår overordnede værdier som rationalitet og retfærdighed (Meyer, Boli & Thomas, 1994). Men naturligvis former hvert samfund sig hvert år med sin særlige udgave af ”den kulturelle baggrundsstøj”.

Hvor de ovenstående fem rationaler typisk vil blive inddraget af hver deres grupper af aktører, kan dette sjette rationale formodentlig trækkes på af flere forskellige aktører i forhold til en evaluering.

Alt andet lige er det mest sandsynligt, at den kulturelle baggrundsstøj appelleres til af aktører, der er forandrings fortalere. Formodentlig opstilles evaluering med den kulturelle baggrundsstøjs støtte som en udfordring til ”forældede” eller ”stivnede” institutioner og måder at gøre ting på.

Imidlertid skal man heller ikke for de første fem rationalers vedkommende overvurdere, i hvilket omfang rationalerne faktisk slår igennem empirisk hos de forventede grupper af aktører. For det første kan aktørerne i virkeligheden opføre sig ganske anderledes, end disse simple skitse-rationaler antyder. For det andet kan nogle aktører forsøge at udtale sig på andres vegne og dermed appellere til disse andres rationaler. Sidst, men ikke mindst, kan der være et kompliceret spil mellem flere forskellige rationaler i løbet af en evalueringsproces.

Alt i alt antyder dette kapitel, at ”en evaluering” ikke nødvendigvis har noget særskilt enhedspræg, men i stedet kan bestå af forskellige momenter, der hver især præges af forskellige aktører.

Kapitel 5

Cases: Metode

Afgrænsningen af evalueringsforløb

I dette og de følgende kapitler præsenteres en undersøgelse baseret på case-studier. Der er foretaget otte casestudier af hver sit evalueringsforløb. Afgrænsningen af evaluering er sket ud fra den definition, der blev præsenteret i indledningen. Det betyder, at arbejde inden for f.eks. monitorering og akkreditering indgår i undersøgelsen, også selv om ordet evaluering ikke direkte bruges af de involverede aktører. I de fleste tilfælde kalder aktørerne i casestudierne imidlertid også selv evalueringen for evaluering.

I nogle tilfælde sættes spot på en enkelt drøftelse af en ændring i et evalueringssystem. I andre tilfælde følges et mere omfattende evaluering fra noget, der ligner et startpunkt til noget, der ligner et slutpunkt.

Afgrænsningen af et evalueringsforløb i tid og rum er som antydnet metodisk og praktisk en vanskelig sag. Det skyldes imidlertid, at den også teoretisk er en vanskelig sag. Afgrænsningen i tid er vanskelig, blandt andet fordi der jo altid ligger noget forud for starten af et forløb. Et initiativ til evaluering har altid igen sin forhistorie. Tilsvarende kan konsekvensdragelsen af en evaluering række ganske langt ud i tid efter evalueringens egentlige afslutning, hvis politiske aktører vælger fortsat at argumentere ud fra den tidligere foretagne evaluering. Nogle evalueringer er fæstnet i vedvarende systemer til evaluering, som pr. definition ikke har en afslutning. Mange evalueringer medfører andre evalueringer. Også afgrænsningen af evalueringsforløb i rum er vanskelig. Mange forhold spiller ind på evalueringer, og de spiller tilbage igen. I det omfang, der handles om evalueringsforløb, er det ikke entydigt, om de andre "vare" i handlen også skal regnes med til evalueringsforløbet.

Trods alt giver det mening at afgrænse evalueringsforløb. Aktører i de undersøgte cases er selv med til at skabe en vis sammenhæng i et evalueringsforløb i og med, at de henviser til begivenheder, de mener, er sammenhængende.

Et stykke ad vejen er afgrænsningen af et evalueringsforløb alligevel mere eller mindre pragmatisk sat, ikke mindst i lyset af, at undersøgelsen naturligvis må have en ende på et tidspunkt. De fleste af casestudierne er empirisk afsluttet i eftersommeren 2002.

Udvalget af evalueringsforløb

Det påberåbes ikke, at de undersøgte evalueringsforløb er repræsentative for alt evalueringsarbejde i Danmark. Men de er udvalgt, således at de repræsenterer en vis variation. Der er variation med hensyn til politikområde og emne, i hvert fald er både undervisning, ældrepleje og sygehusområdet med. Der er variation med hensyn til forvaltningsniveau (kommunalt og centralt niveau). Der er variation med hensyn til graden af konflikt og mængden af offentlig omtale (fra ingen omtale til forsidestof). Der er også stor variation med hensyn til omfanget af det evalueringsarbejde, der undersøges, ligesom både enkeltstående evalueringer og organisatoriske rutine-procedurer til evaluering er med.

Dataindsamling

Hver enkelt case er typisk blevet belyst ved hjælp af dokumenter og interviews. Blandt dokumenterne tæller baggrundsmateriale, notater, tidligere rapporter, avisartikler m.m.. Der har typisk været foretaget 3-6 interviews med nøglepersoner inden for hver case, for eksempel en kvalitetsmedarbejder, en leder og to medarbejdere. Interviewpersonerne er udvalgt dels for at give mest mulig indsigt i selve evalueringsforløbet og dels for at give forskellige synspunkter på dette. Interviewene er optaget på bånd og derefter udskrevet.

I de fleste tilfælde er dokumentstudier og formelle interviews suppleret med uformelle interviews i felten (f.eks. hos hjemmehjælpere), deltagelse i møder samt mere uformelle samtaler, hvor oplysninger ”opsnuses” snarere end indsamles formelt. I almindelighed har der været velvilje til at deltage, men der har også hos flere interviewpersoner været en tilbageholdenhed med hensyn til at tale om forhold, som ikke passer med deres arbejdsorganisations almindelige måde at præsentere sig udadtil på. I andre tilfælde har medarbejdere imidlertid udtalt sig uhyre bramfrit, dog under løfte om anonymisering. Adgang til feltarbejdet har i nogle tilfælde været let, andre gange præget af adgangsproblemer, hvor det var svært at se en god begrundelse.

I enkelte tilfælde er en case udarbejdet uden interviews, men i stedet ud fra dokumenter, mødereferater m.m. I de tilfælde, hvor et evalueringsforløb kan beskrives mest indgående ved ikke-formelle data, er sådanne benyttet. I de situationer er taget mange notater. For hver case er oprettet en sagsmappe, som fylder fra 30 op til mere end 200 sider.

Repræsentation af caseforløbene

For at holde styr på det omfattende materiale er udarbejdet tilblivelseshistorier om hver enkelt evaluering eller "casestories", der sammenfatter de vigtigste begivenheder i hvert enkelt evalueringsforløb. I det følgende gengives disse casestories i så komprimeret form som muligt. Skelettet i hver enkelt af disse casestories består af de afgørende beslutninger i hver enkelt moment i evalueringsforløbet undervejs; holdninger og synspunkter fra forskellige parter undervejs gives også, dels som begrundelser for aktørernes handlinger i forhold til evalueringsforløbet, dels som vidnesbyrd om de værdier, normer og mentaliteter, der brydes i diskussionerne om evalueringer.

Der findes ingen objektiv måde, hvorpå hele evalueringsforløbet kan fortælles. Det betyder imidlertid ikke, at en casestory kan fortælles på hvilken måde, det skal være. Fortællingen kan ikke være inkonsistent med data, der foreligger om forløbet. Triangulering mellem interviewudsagn og dokumenter har flere gange været gavnlige i opbygningen af casestories om evalueringsforløbene.

Det er imidlertid en del af hele evalueringsspillet væsen, at evalueringsforløb ikke opleves på samme måde af alle, der er involveret. Hvis nogen for eksempel giver udtryk for at blive dårlig behandlet i et evalueringsforløb, så kan dette udsagn ikke blot være et udsagn *om* evalueringsforløbet, men også et udsagn *i* evalueringsforløbet, der tæller i en eventuel kamp mod de andre parter i evalueringsforløbet. Samme status har spekulationer om de andre parters motiver. Det er således logisk umuligt at "få det hele med".

Fortællingernes opbygning er i denne bog primært styret af hensynet til at illustrere eller modbevise pointer i forhold til det skitserede begrebsapparat angående magt og evaluering. Samtidig er det en del af casestudiets væsen, at en virkelighedsnær case altid er lidt rigere og lidt mere sammensat, end hvad teorien siger. Den mest korrekte afreportering af en case kan derfor godt være den, der lader visse tolkningsmuligheder stå åbne snarere end kun at fremme en eneste tolkning.

Casebeskrivelserne er af meget forskellig længde afhængig af sagens kompleksitet, evalueringsforløbets varighed og adgangen til data.

Anonymisering

En af de vanskeligste metodiske beslutninger i denne undersøgelse har været at finde den rigtige balance mellem eksakthed og anonymisering. For eksakt gengivelse af alle detaljer tæller hensynet til efterprøvning. For ano-

nymisering af data tæller hensynet til at få de mest åbne udsagn fra respondenter, som personligt og organisatorisk kan have noget i klemme i forbindelse med evaluering. For anonymisering tæller også hensynet til at beskytte enkeltpersoner, som har medvirket i god tro, mod at medier efterfølgende sætter urimeligt spot på enkelte af deres udsagn.

Interviewpersoner er blevet gjort bekendt med undersøgelsens baggrund og formål. De har fået at vide, at oplysninger, der kan henføres til enkelte personer og organisationer, så vidt muligt tildækkes. Sådanne oplysninger er derfor i enkelte tilfælde udeladt og/eller erstattet med uvæsentlige, fiktive detaljer (såsom navnet X-by). Interviewpersonerne har imidlertid også fået at vide, at det ikke i alle tilfælde er muligt at sløre identiteten af de deltagende parter. I nogle tilfælde har vi kun en enkelt institution i landet, der kan udføre den pågældende opgave, som omtales i en case.

I sådanne tilfælde og måske i flere tilfælde vil læseren have en chance for at gætte, hvem der måske kan være tale om, også selv om det ikke direkte fremgår af teksten. Imidlertid anmodes der hermed om forståelse for anonymiseringsprincippet. Anonymisering har været et vigtigt metodisk skridt for at sikre, at offentligt ansatte har kunnet omtale forhold på deres egen arbejdsplads, som de finder problematiske.

Analyse

De otte cases afrapporteres samlet og analyseres samlet. De betragtes samlet som en kilde til viden om, hvordan magt udfolder sig i evalueringsforløb. Der er opstillet et komparativt design, hvor forskelle i evalueringsforløb søges systematisk forklaret ud fra forskelle i baggrundsfaktorer. Ikke mindst inspireret af Stake (1994), som taler for at repræsentere hver case uden at drage for stærke sammenligninger med andre, har der været en ønske om at repræsentere hver case på sine egne præmisser og opsummere nogle pointer om evaluering og magt for casene tilsammen, men altså ikke overdrive en generalisering af de lovmæssigheder, der kunne uddrives af sammenligninger mellem casene.

Kapitel 6

Cases: Evaluering i praksis

Case 1. Akkreditering af hospitalsvæsenet i Y-by

Som baggrund for arbejdet med akkreditering af sygehusvæsenet i Y-by nævner en kvalitetsmedarbejder Sundhedsstyrelsens udmeldinger om kvalitet og den gradvise glidning fra ”patient” til ”bruger”, som ansporer til øget fokusering på målbar kvalitet i den enkelt behandlingsforløb. ”Der var mange impulser, der ligesom sagde, at nu er vi nødt til at kunne fortælle og dokumentere, hvor dygtige vi er”. Man mente, at en form for selvevaluering ikke ”havde samme værdi” som en udefrakommende vurdering.

Efter besøg i USA har man oplevet akkrediteringsmodellen. ”Den amerikanske model er god”, som en interviewperson siger, ”fordi den er så veletableret”. Vedkommende forklarer, at i USA er akkreditering en forudsætning for, at et hospital får kunder, fordi forsikringsselskaber kun vil dække udgifterne, når hospitalet er akkrediteret. Derfor handler en meget stor del af akkrediteringsprocessen om patientsikkerhed og organisatorisk kompetence, alt sammen for at undgå fejl, der koster dyrt for forsikringsselskaberne.

”Det startede med, at en enkelt medarbejder var meget optaget af det, og så arbejdede man med det i hospitalsdirektørkredsen. Det var ikke noget, der kom nedefra”. Den pågældende medarbejder nævner, at han via rejser i USA ”kan se hvad der rører sig”, og ”fanger det sidste nye” før andre i Danmark.

Adspurgt om, hvorfor akkreditering blev valgt frem for andre kvalitetsudviklingsværktøjer peger en kvalitetsmedarbejder på, at ”akkreditering kan bruges som overordnet paraply for alt kvalitetsarbejde”. De andre tiltag (medarbejderen nævner en række eksempler) kan benyttes som redskaber i forhold til akkrediteringen.

Selve akkrediteringen sker dels i form af akkrediteringsmyndighedens gennemgang af skriftligt materiale fra hospitalet, dels en vurdering ved et besøg af akkreditorer på stedet over 3-5 dage. Teamet vurderer ikke direkte, om den givne behandling er fagligt rigtig eller forkert, men i stedet om de igangsatte aktiviteter er i overensstemmelse med hospitalets egne mål og protokoller samt internationale standarder, hvor disse findes. Et vigtigt element er også, hvordan der på hospitalet følges op på eventuelle brist i kvaliteten. Nedskevne politikker og journaler spiller her en stor rolle, ligesom

de fysiske rammer inspiceres. I vurderingsprocessen gør teamet brug af en scoringsmanual, som omsætter kvalitative vurderinger til et pointsystem, der munder ud i en samlet vurdering af hospitalet. Der rundes af med en mundtlig og skriftlig rapport til hospitalet. Derefter opnås endelig akkreditering på et af fem niveauer alt efter, hvilke forbedringer der skal til for at sikre fortsat akkreditering. Akkrediteringen er gyldig i tre år og kan fornys ved en gentagelse af processen.

Akkrediteringen sker i forhold til ca. 360 standarder. ”De er lavet af et internationalt panel af fagpersoner”. Det er et diskussionspunkt, i hvilken grad disse standarder er fuldstændig ubøjelige i akkrediteringsprocessen, og i hvilket omfang de kan gøres til genstand for lokal forhandling. Flere gange understreger en leder, at standarderne ikke stod til forhandling – idet troværdigheden af evalueringen siges at blive styrket af, at hospitalet selv ikke kunne påvirke standarderne. Tværtimod er hele standardtænkningen en væsentlig del af filosofien i akkrediteringen: ”Hvis du mindsker variation, får du også en bedre kvalitet. Altså noget med ensartethed”.

En anden medarbejder siger: ”Vi fik nogle indrømmelser. Der var nogle ting, vi overhovedet ikke gad have noget med at gøre. Det var f.eks. sikkerhed omkring spædbørn (...) Akkreditørerne kører op og ned i den her elevator, og vi havde fået meget ros, og det var vi selvfølgelig glade for, og så pludselig stiger en af dem ud på tredje sal, og der står en baby ude foran en rygestue, og han siger: ”Der må ikke stå babyer alene!” Moderen sad indenfor og røg. Fra vores synspunkt var det glimrende, at hun ikke udsatte babyen for røg, men i Amerika må du ikke det der, for de har kæmpeproblemer med bortførelse af spædbørn (...). Heldigvis kom oversygeplejersken tøffende, og han skulle så se medicinskabet og afdelingen, og det så glimrende ud”.

Undervejs i processen er der en del modstand og kritik, ikke mindst fra lægeside. Godt nok mener man, at akkrediteringen er et ganske effektivt ledelsesværktøj, men man kritiserer, at den fokuserer på ”at medicinen bliver håndteret rigtigt fra lager til medicinskab, at brandudstyret er checket og på plads osv.” Men kernen, den lægelige og plejemæssige ydelse, bliver ikke godt nok belyst, argumenterer man. ”Det er den kritik, der gjorde det umuligt at argumentere for, at denne akkrediteringsmodel skulle være landsdækkende. Der er for gode argumenter imod, nemlig at det ikke rammer der, hvor der var allermest brug for kvalitetsforbedringer. Og så begynder det ene amt efter det andet at udvikle egne modeller”.

Fra yngre lægers side bliver det kritiseret, at den lægelige uddannelse ikke indgår i akkrediteringsarbejdet. En leder svarer igen, at ”lægerne med

nogen ret har kritiseret projektet for ikke at handle om forskning og uddannelse, men det skyldes, at akkreditering tager udgangspunkt i patientens møde med hospitalet. Uddannelse og forskning er ikke blandt det, der umiddelbart påvirker patientens oplevelse". Lederne nævner dog, at det kunne være værd at få med på længere sigt. "Men vi syntes ikke, der var tid nok til at gøre det kvalificeret, for laver vi standarder, skal de også være rigtig gode". Yngre læger udfordrer endvidere hospitalet på spørgsmålet om uddannelse ved at sige, at når hospitalet ikke inddrager uddannelse i akkrediteringsarbejdet, er grunden, at hospitalet ikke kan komme godt igennem en evaluering på dette område.

På et tidspunkt opdager man også, at man har glemt at informere en medarbejdergruppe. Man indkalder den til møde, men de møder ikke op, fordi de har misforstået ordet akkreditering.

Fra ledelsens side mener man at have gjort meget med hensyn til information om projektet. Kvalitetsafdelingen har uddannet 600 på et dagskursus i akkreditering, og der er holdt mange åbne møder og udarbejdet meget materiale. Men man erkender, at det kan have knebet med at nå ud i organisationens yderste led.

Der opstår en del problemer frem til selve akkrediteringen med, at procedurevejledninger skal ud til den enkelte medarbejder via IT, men mange kan ikke anvende computere rigtigt eller har ikke adgang hertil. Computerens procedurer kriteses for at være uoverskuelige, og nogle procedurevejledninger blev først færdige så kort for akkrediteringsbesøget, at personalet ikke havde mulighed for at sætte sig ind i dem.

Det værste set fra en kvalitetsmedarbejders synspunkt er, at der ikke er afsat penge og personale specielt til akkrediteringen. Timerne skal findes i den almindelige "pressede" hverdag, inklusive tid til møder og undervisning. Kvalitetsmedarbejderne er heller ikke i starten helt klar over, hvor omfattende et projekt, de sætter i gang.

Man gør forsøg på at udpege de allervigtigste standarder for medarbejderne, for eksempel ved farvekodning.

De forskellige afdelinger viser sig at være meget uens med hensyn til, hvor alvorligt man tager arbejdet med procedurerne. En kvalitetsmedarbejder fortæller, at det er ganske hårdt arbejde at sikre, at afdelingerne ikke indbyrdes får at vide, hvor lidt de langsomste afdelinger arbejder med sagen – det vil nemlig virke demotiverende, som det hedder. Det får blandt andet som konsekvens, at flere afdelinger sidder og bruger tid på at udvikle procedurer for det samme område, siger en medarbejder.

Under selve akkrediteringen bliver der etableret en hotline mellem kvalitetsafdelingen og de enkelte afdelinger, der akkrediteres. Derigennem formidles information om, hvad akkreditørerne specielt går efter. ”Hvis en afdeling mangler en proceduremappe, kan vi hurtigt sende en af sted med et eksemplar”. En medarbejder fortæller, at det er showpræget, hvad der foregår. For eksempel bliver der gjort ekstra rent i anledning af akkrediteringsbesøget.

Efter akkrediteringen oplever mange en stor træthed, og ledelsen er klar over, at det ikke er tiden til nye store processer på hospitalet. En medarbejder nævner, at man efter akkrediteringen er faldet tilbage til gamle rutiner, fordi de nye rutiner er blevet forceret igennem uden at bundfælde sig.

I tiden lige efter akkrediteringen fremsendes en del kritik. En del af dagspressen citerer sygeplejersker for at sige, at ”det var latterligt og spild af penge i forhold til, at vi mangler apparatur til måling af iltindholdet hos patienterne (...) Vi ville gerne fortælle, at en del af vores udstyr er forældet, og at vi virkelig trængte til noget nyt, blandt andet kan man ikke høre apparaternes alarmer ude på gangen. Men ledelsen sagde, at det havde ikke noget med det at gøre”. ”Jeg har aldrig kigget i proceduremappen. Den er virkelig tyk og vi har aldrig fået at vide, at vi kunne tage os en time til at læse i den”.

Fra ledelsesside tilbagevises kritikken. Der peges på, at kritikerne er få, oftest anonyme og at mindst en af dem ikke har været ansat under selve akkrediteringsprocessen.

Fra ledelsesside nævnes, at hvis der har været tale om fusk, må der sættes navn på, hvilke afdelinger det er sket i, således at ledelsen kan tage konsekvensen.

Der udarbejdes en evaluering af forløbet baseret på spørgeskemaer til medarbejdere. Da det viser sig, at ca. 70 pct. af medarbejderne føler sig tilfredse med forløbet, bliver det af ledelsen brugt som argument for, at kritikken ikke er berettiget. (Af rapporten fremgår det, at 71 pct. er enige eller delvis enige i, at de har været tilstrækkeligt involveret i akkrediteringsforløbet).

Ikke mindst som konsekvens af arbejdet med akkreditering forventer man til sidst at have omkring 20.-25.000 procedurevejledninger på hele hospitalet. Man har fået bevilget ansættelse af en person, der skal sidde og sikre, at der ikke lige pludselig er fem procedurevejledninger om det samme emne. Man skal også sikre sig, at hver af dem er evidensbaseret og helt automatisk opdateres årligt. Men man skal også gemme de gamle procedu-

revejledninger, sådan at når der kommer patientklager over tidligere behandling, så kan man afgøre, om den tids procedurevejledning blev fulgt.

En medarbejder vurderer efterfølgende: ”Det er en succeshistorie, fordi vi blev akkrediteret med kryds og bolle, så det politiske budskab, man ville sende ud om institutionen, det kom ud, og det er fint”. ”Men resultatet var ikke optimalt, fordi metoden ikke indeholder nogle ordentlige måder at måle kvaliteten på, fordi den mangler to meget tunge områder, altså selve den faglige behandling og udviklingsområdet for personalet og forskningen”.

Case 2. Stregkoder i ældreplejen i S-by

En vicedirektør i den kommunale forvaltning i S-by forklarer, at ønsker udmeldt fra Socialministeriet, KL og regeringen om mere styr på hele hjemmehjælpsområdet indgik i baggrunden for kommunens interesse for et stregkodesystem, som registrerer plejeydelsen automatisk ude ved dørkarmen hos den enkelte ældre. Samtidig skulle systemet være så lidt bureaukratisk og ressourcekrævende som muligt og så skal det i øvrigt kunne spille sammen med økonomisystemet. Med i baggrunden indgik også avisoverskrifter, hvor der stod, at hjemmehjælperne drak kaffe og ikke kom ud og fik lavet det, de skulle.

Den tekniske løsning, man valgte, opfylder de krav og retningslinjer, der er opstillet i KL's projekt Fælles Sprog.

I forbindelse med lanceringen af systemet er der en vis skepsis ikke mindst fra de faglige organisationers side, hvor man er bekymret for systemets brug til kontrolformål. Der bliver imidlertid aftalt et sæt regler, således at systemet ikke skal bruges til kontrol af den enkelte medarbejder. Man må ikke lave direkte tidslige sammenligninger pr. opgave mellem medarbejderne.

Ifølge vicedirektøren er der en vis forståelse blandt personalet for, at der er brug for at dokumentere, at man er på banen og laver et godt stykke arbejde, selv om ikke alle deler den udlægning.

Ledelsen har svært ved at se et alternativ til den form for elektronisk opsamling af data, som foretages. Der aflægges nemlig mange tusind besøg om ugen hos kommunens gamle. ”Når der kommer en klage, så er det her jo rent faktisk også et stærkt redskab i forhold til, at medarbejderen rent faktisk kan dokumentere, at ”jeg har været der, med de og de opgaver, med de og de funktioner”. Så skal man ikke til at rekonstruere, hvad var det lige, og hvornår var det nu, det står derinde helt præcist. Kom klokken det og gik klokken det, og jeg lavede sådan og sådan”.

Ledelsen ser i dag hjemmehjælpen som et lille driftsselskab. Tidligere skulle man bare overholde et budget. I dag sker afregningen på baggrund af den såkaldte ATA-tid (ansigt til ansigt). ”Når du åbner døren og går ind, er det ATA”. Hver gruppe kan sammenligne sin indtjening, effektivitet og ATA-tid med de andre grupper. Hvis ikke borgeren får den bevilgede tid, reduceres der i indtjeningen. Med henblik på at maksimere indtjeningen gælder det om at holde andre registrerede tider nede som for eksempel ruteplanlægning, sygefravær, tværfagligt samarbejde osv.

Stregkodesystemet bruges i praksis ikke blot af hjemmehjælpere, men af hele plejepersonalet inklusive social- og sundhedshjælpere, assistenter, plejere, plejehjemsassistenter og sygehjælpere.

To medlemmer af plejepersonalet er ærgerlige over at have mistet visitationskompetencen. Visitatoren kender ikke fru Jensen på samme måde som selve plejepersonalet. Der er også ærgrelse over, at sygeplejerskernes arbejde ikke er visiteret; dermed sættes der mere fokus på kontrol af plejepersonalet, siger de to.

En tillidsmand nævner – langt inde i et interview – at fagforeningen har været unødvendigt meget modstander af stregkodesystemet. ”Jeg mener helt sikkert, at hvis man går ud og passer de ældre på de præmisser, man skal, så kan man sagtens gå med en scanner i hånden. (...) Der var måske blevet for meget af ”sådan nogen ting” (et eksempel med en bruger, der kun skulle have lavet en snaps), og byrådet og sådan noget, de ved jo godt, hvordan tingene er, for de hører jo også ...”.

En hjemmehjælper beskriver arbejdsdagen med scanneren således: ”Vi har sådan et stregkodehæfte, hvor man har en kode på sig selv. Det første, vi gør, er at tage scanneren og strege ind på vores kode, så man kan se, det er mig, der bruger den scanner. Når vi så går ind til klokken syv til orienteringsmødet, så streger vi på uregistreret tid, og når vi går ud derfra, så sætter vi den på transport og kører ud til borgeren og starter med at scanne ind på den kode, borgeren har på sin dørkarm. Og så har vi jo borgerens opgaver liggende i de kørelister, vi har, og når vi har gjort det, vi skal, scanner vi ud på døren, og så slår den selv over på transport. Sådan kører det så dagen igennem. Til sidst kommer man ind og afleverer sine nøgler og scanner og giver besked”.

En anden hjemmehjælper understreger, hvor vigtigt det er at bruge scanneren rigtigt. Taler man med en pårørende efter at have scannet ud, står den jo på transport. Mange går ud med 4-5 timers transport, så har man jo ikke brugt scanneren rigtigt”.

Scanningen er sådan, at man kun skal udføre tre scanninger, nemlig opgaverne samlet, ind og ud. Man skal ikke scanne de enkelte opgaver særskilt, så længe man yder det, som er visiteret.

Der er mulighed for at gøre nogle nødvendige ting ud over de rekvirerede opgaver, f.eks. hvis brugeren den dag har særlige behov. ”Så kan man ændre på det ved at melde det til gruppelederen, som checker det af med visitatoren. Eventuelt kommer visitatoren ud igen”, fortæller en medarbejder. ”Hvis jeg har en time til rengøring ude hos en borger, og kun bruger tre kvarter, så er der også mulighed for, at jeg kan gøre noget, som borgeren gerne vil have lavet”.

Direkte adspurgt under et interview oplyser en medarbejder, at den enkelte bruger ikke på den enkelte dag skal kvittere for modtagelsen af hver enkelt ydelse. Der er således ikke direkte kontrol af, om de visiterede ydelser faktisk leveres, om de tager mere eller mindre end den fastsatte tid, og om hvordan kvaliteten af ydelsen er. Hvis der på grund af kurser og fravær er lidt ekstra borgere på, kan man ”cutte lidt for at få det til at hænge sammen”, men det udligner sig over en længere periode, og det holder gruppelederen meget nøje øje med – især hvordan det kører med ATA-tiden. Man kan således godt mingelere lidt med tiden i det daglige. ”Vi har det sort på hvidt, hvornår hjælperen kommer og går, man kan ikke snyde, vel, hoveddøren skal lukkes op, for at man kan scanne, den sidder lige indvendigt på dørkarmen (...), men man kan godt strege, at man har udført nogle opgaver, selv om man ikke har udført dem, men man kan jo ikke snyde med, om man har været der”, fortæller en medarbejder, ”men det er ikke noget, vi har haft fokus på eller snakket om”.

Brugeren er i princippet kun informeret om hvilke ydelser, han/hun er tildelt, ikke hvor lang tid. En hjemmehjælper nævner en strategi til at undgå stress som medarbejder under strekkode-systemet. ”Man skal ikke fortælle en bruger, hvor meget tid, de har, fordi de hænger sig i tiderne. Det er jo det, der har givet balladen til strekkodesystemet, fordi det har kollegerne også gjort.(...) Så snart du fik den her (scanneren) har man fokuseret alt for meget på det og det var det, man fik stress af. Man skal bare glemme alt det, der står ude i siden (tiderne). Man skal bare køre efter sin sunde fornuft”.

”Men medarbejderne føler sig stadig kontrolleret”, siger en tillidsmand. Der er fortsat en vis uro angående strekkodesystemet. En medarbejder har netop været til medarbejdermøde om sagen, og siger: ”Så er der nogen, der siger, at med scanneren, der kan man da se, hvad vi laver, men det er jo at foregøgle folk, at man kan se, hvad de laver, man kan se, at de har været

der”. ”Er der sygefravær og der ikke er afløsere at sætte ind, så gives der bare besked på, at man skal scanne ind og ud og være ligeglad med de tider, der er sat. Og så virker systemet ikke alligevel, fordi hvis man er der kortere tid, så får man jo stadig kun penge for den tid, man er der, så regnestykket kommer ikke til at hænge sammen”.

Aften- og nathjemhjælperne fandt det på et tidspunkt særligt vanskeligt at overholde de visiterede tider. Man holdt derfor et møde med forvaltningen, hvor det blev aftalt kun at scanne en uge om måneden i hvert fald i et halvt år, mens folk udefra lavede en særlig undersøgelse af problemet.

Ledelsen lyttede blandt andet til personalet på dette punkt, fordi ”det var tungt, en tung scanner, og om aftenen er der så mange kortvarige besøg, og vi kunne kun lige nå at få den op af posen. Det er ikke det med kontrollen i al almindelighed, det er det andet synspunkt”.

I den tid, hvor man har kørt uden scanner, mener nogen fra personalet, at det er ”nemmere at mingelere”. ”Jeg kan tage den ene og du kan tage den anden, det er nemmere, når det ikke er bundet så stramt op”.

Daghjemhjælperne har nøje fulgt udviklingen blandt aftenhjemhjælperne. Aftenpersonalet har ikke optjent ATA-tid, mens de har kørt uden scanner. I stedet har man aftalt, at beløbet har været fastlagt på forhånd. Nogen har sagt: ”Når aftenholdet kan blive fri, hvorfor kan vi så ikke ...?”

På et tidspunkt blev der taget initiativ til at gå i gang med værdibaseret ledelse på ældreområdet i S-by. Det skyldes ifølge en medarbejder, at ”ledelsen tog konsekvensen af reaktionerne fra personalet, de har oplevet deres hverdag meget firkantet og personalet følte det som en stor belastning. Der var ubalance i tingene, simpelthen. Så er vi altså nødt til at koncentrere os om den bløde del. Så det er sådan set ledelsen, i erkendelse af eller på baggrund af de reaktioner, de har mødt gennem de senere år, der har besluttet, at nu skal vi have lagt noget i den anden vægtskål”. Medarbejderen siger: ”I øjeblikket bliver vore ledere kun målt på evnen til at overholde budgettet, men jeg mener, at værdigrundlaget bør ind”.

Case 3. Monitorering af ældreplejen i Z-by kommune

Fælles Sprog er et system til vurdering af en borgers funktionsevne og et katalog over ydelser, som borgeren kan tildeles. Borgerens behov vurderes på ni områder, herunder for eksempel ”personlig pleje”, ”daglig husførelse”, ”mental og psykisk tilstand” osv. Ideen er at udtrykke forskellige borgers behov for hjælp og den hjælp, de modtager på en nogenlunde standardiseret måde. Baggrunden er ikke mindst de store og stigende udgifter

på ældreområdet. Allerede i begyndelsen af 1990'erne diskuterede KL således hjemmehjælpsrådets manglende dokumentation af ressourceforbrug og den generelt manglende fælles informationsopsamling (la Cour & Højlund, 2001: 9).

I paragraf 110 i lov om social service, der trådte i kraft i 1998, nævnes et skema angående den individuelle hjælp til borgeren, som skal rumme oplysninger om, ”hvilke opgaver hjælpen omfatter, formålet med hjælpen og for hvilken periode hjælpen gives”. Borgeren har herefter et retskrav på den visiterede ydelse.

Der har i mellemtiden været udviklet en del systemer til at opgøre behovet for hjælp samt registrere de visiterede ydelser og leveringen heraf.

Fælles Sprog er et sådant system. Der har været en ganske massiv kritik af systemet. Det vigtigste kritikpunkt har været, at Fælles Sprog standardiserer hjælpen i en sådan grad, at borgeren og hjemmehjælperen ikke længere har mulighed for i den enkelte situation selv fleksibelt at afgøre, hvilken hjælp borgeren har brug for den pågældende dag. Socialministeren fik 1.4.2001 iværksat en ændring af loven om social service, som blandt andet giver hjemmel til en mere fleksibel tilrettelæggelse af hjemmehjælpen.

I 1999 begyndte Z-by kommune at bruge Fælles Sprog. Ifølge afdelingslederen på området udgør Fælles Sprog et godt instrument til at gøre et serviceniveau juridisk holdbart, fordi den enkelte ydelse omskrives til en standard, skriftliggøres og dermed bliver et synliggjort område.

Fælles Sprog har ifølge afdelingslederen flere funktioner: Det bliver et arbejdsredskab for den, der visiterer – og man ændrer samtidig strukturen, så kun få personer kan visitere. For det andet bliver det et pædagogisk værktøj til at synliggøre over for personalet, hvorfor den ene ældre skal have hjælp, og den anden ikke skal. For det tredje kan man til enhver tid fortælle politikerne, hvilke ydelser der kommer til gengæld for de brugte penge på området. For det fjerde kan man respondere på klager fra pårørende. Man kan ud fra Fælles Sprog nøjagtigt forklare en pårørende, efter hvilke kriterier, hjælpen er tildelt.

Et af de vigtigste argumenter for at indføre systemet er ifølge afdelingslederen retssikkerheden. ”Mit personale var ikke tidligere klædt på til at foretage skriftlige afgørelser”. Og så var der hensynet til, at den enkelte borger fik den rigtige hjælp. På dette punkt nævnte afdelingslederen specielt, at undersøgelser har vist, at især de stærkeste af de svageste traditionelt set har været bedst til at sikre sig omsorg fra plejepersonalet. På den baggrund er et tildelingssystem baseret på mere objektive kriterier at foretrække.

I Z-by kommune har man ifølge afdelingslederen forsøgt at lade de ansatte komme til en forståelse af, at Fælles Sprog er en sag for alle ansatte og ikke kun for visitatoren. Personalet blev undervist i grupper, og der var temadage, og endelige tog visitator ud i grupperne for at høre, hvordan Fælles Sprog og visitationen hænger sammen.

Generelt siges der at have været stor enighed i Z-by om at indføre Fælles Sprog, men sygeplejerskerne var ængstelige for, om systemet fratog dem deres faglighed. I sygeplejerskernes blad har der blandt andet været artikler om, at det at udfylde prædefinerede vurderings- og ydelseskemaer ikke kræver særlige sygeplejefaglige kompetencer. Man har i stedet søgt at undgå en automatik i sammenhængen mellem borgerens (skematiske) funktionsniveau og ydelsestildelingen, ligesom man har argumenteret for vigtigheden af som faggruppe selv at have visitatorfunktionen.

En visitator mener ikke, at Fælles Sprog fungerer faktisk optimalt i Z-by. Hun mener for eksempel, at kriterierne er for meget præget af symptomer i forbindelse med demens, en lidelse, der aktuelt har været meget fokus på. Til gengæld mener hun, at Fælles Sprog er et godt dokumentationsredskab i forhold til politikerne.

Efter at Fælles Sprog blev indført i Z-by, har man kun én visitator i hele kommunen. ”Nu er der altså mere ens behandling uanset hvor du bor i kommunen”. ”Fordelen ved Fælles Sprog er, at der er et stykke papir, hvor det fremgår hvad visitator har vurderet. Det giver kontinuitet i plejen. Som regel vil man drøfte om visitationen er korrekt efter 3-6 måneder”, fortæller en sygeplejerske.

Et eksempel på et visitationsbesøg forløber som følger. En kvinde på 87 har fået en dårlig skulder og derfor bedt om rengøringshjælp. Det er en ”let” visitation, så sygeplejersken går relativt let henover en del af spørgsmålene i visitationsskemaet. Til sidst udarbejdes et dokument, som borgeren skal underskrive og selv får en kopi af. Det særlige ved denne visitation er, at ægtefællen på 91 er gået i seng, idet han ligger i sengen det meste af dagen. Efter end visitationsbesøg fortæller visitator, at kvinden faktisk var for god til at få tilbudt 1,5 time hver 14. dag, men hun gjorde primært på grund af ægtefællens tilstand.

I visitationsguiden er der ved de enkelte ydelsestyper afsat normeret tid til de pågældende opgaver. Men den normerede tid er kun vejledende. Visitator kan justere op eller ned. En hjemmehjælper kan eventuelt vende tilbage til visitator og forklare, hvorfor man ikke mener, at de visiterede ydelser kan nås på den afsatte tid.

Man har ikke i sig selv et system, der registrerer, om de visiterede ydelser faktisk leveres præcist som aftalt. Men hvis man på grund af sygdom eller lignende er nødt til at aflyse et hjemmehjælperbesøg, bliver en ny aftale noteret, ligesom der hos demente ligger en bog i hjemmet, hvor hjemmehjælperen kan kommunikere med pårørende, så man kan samle op på evt. ændrede aftaler.

Ændrer man på den visiterede hjælp, skal det registreres i såvel det lovpligtige kardeks som i de papirer, der relaterer sig til Fælles Sprog. Er det en ændring, som strækker sig over længere tid, skal den ifølge en hjemmesygeplejerske registreres i computeren. ”Det er utroligt vigtigt, at det er dokumenteret, at det er borgeren selv, der har fravalgt sit bad til fordel for en gåtur”.

”Medarbejderen kan også bruge Fælles Sprog til at argumentere for det, som de mener borgeren skal have, så de ikke skal føle, at de stjæler sig til at give borgeren det, som de mener, der er behov for”, siger afdelingslederen.

Afdelingslederen finder, at Fælles Sprog også er et godt redskab til at løse uenigheder mellem faggrupperne. ”Sygeplejersken er hurtig til at sige, at den og den borger skal have en tryghedsalarm installeret, fordi vedkommende er nervøs for at være alene hjemme. Hun er jo den omsorgsfulde person. Så kommer terapeuten, der skal installere det og siger: Hvorfor skal den borger have en tryghedsalarm, hun skal hellere have genoptræning, så hun ikke er bange for at være alene. Så kommer de til mig som afdelingsleder og siger, at hende den anden, hun er dum. Så tager jeg fat i Fælles Sprog og siger: Hvad er det så, der indikerer, at borgeren skal have en fællesalarm? Så snakker vi om, hvorvidt borgerens mobilitet er nedsat således, at hun er en 3'er, eller hun i forhold til psykisk funktionsevne er en 3'er, eller hun har en kronisk lidelse som 3'er, eller hun har to ud af de tre samstemmende på 2'er-niveau, så er hun berettiget til nødkald. Så kan de to medarbejdere gå frit ud i verden. Det vil sige, jeg får debatten hen fra at diskutere, om den ene er mere omsorgsfuld end den anden, fordi det refererer til dem som personer, så tvinger jeg dem til at tage stilling til, hvor er borgerens funktionsevne, og hvor meget er hendes nervøsitet invaliderende for hende. Hvis hun i stedet bliver kategoriseret som 2'er, så kan jeg kommunikere derudfra med borgeren og med mit personale og med det sociale nævn”.

Afdelingslederen nævner, at det er en helt bevidst ledelsesstil, hvor hensigten er, at de forskellige arbejdsgrupper integrerer sig i hinanden. Imidlertid er politikerne ikke i særlig grad repræsenterede. ”Jeg tror, jeg har lanceret Fælles Sprog tre gange for politikerne, og de sidder tilbagelænet og si-

ger, at det lyder godt. Deres involvering er ikke særlig stor, de tager ikke ansvar for at styrke en bestemt ting eller gå ind i en arbejdsgruppe. Jeg laver kvalitetsstandarder og så godkender de. Men de er orienteret om, hvordan det foregår”.

Case 4. De internationale læseundersøgelser

IEA (The International Association for Evaluation of Educational Achievement) opstod sidst i 1950'erne som en non-profit-sammenslutning af forskningsinstitutioner fra hele verden inden for uddannelsesforskning. Man kan således forestille sig, at Sputnik-chokket og 1960'ernes uddannelseseksploration har indgået i denne institutions omverdenshistorie i de første år.

I 1985 bliver der nedsat en styringskomite til at se på en international læseundersøgelse; den første af disse for Danmarks vedkommende. Undervisningsminister Bertel Haarder skriver senere, at en af de beslutninger, han efterfølgende har været gladest for, var beslutningen om som minister at bevilge et tocifret millionbeløb, således at Danmark for første gang kunne tilmeldes undersøgelsen (Haarder, 1997: 9). Haarders motivation er ikke mindst en formodning om, at den socialdemokratiske vision om lighedsskolen har ført til store problemer med de grundlæggende faglige kundskaber (pp. 9-10).

Efter grundigt forarbejde foregik undersøgelsen i 1990-91, den publiceres internationalt i 1992, og i Danmark bliver de danske resultater i bearbejdet form og med supplerende undersøgelser offentliggjort i 1994 (Mejding, 1994).

Hovedresultaterne er, at danske ni-årige læser dårligere end ni-årige i mange andre lande, og selv om resultatet er relativt bedre for 14-årige, ligger også denne gruppe lavere end elever fra lande som Sverige og Finland.

Undersøgelserne viser blandt andet, at lærernes forventninger til elevernes læsevner ligger endnu lavere end den faktiske score hos børnene i tredje klasse (Mejding, 1995).

Resultaterne vækker heftig debat. Avisernes overskrifter er af typen ”Danmark i bund”, ”Bundrekord”, ”Fallit for læsesystem”, ”Ukvalificerede lærere giver ukvalificeret undervisning” osv.

Der sker meget kritiske drøftelser af undersøgelsens metode. Man drøfter, i hvilket omfang skolesystemerne overhovedet er sammenlignelige. Efterfølgende argumenterer ikke mindst DPI's repræsentant for en sammenligning især med de nordiske lande, hvor undervisningssystemerne er relativt sammenlignelige.

Man drøfter også forholdet mellem læseforståelse og læsehastighed. Disse to er blandet sammen i den første udgave af den internationale undersøgelse. Læseforståelse og læsehastighed skilles imidlertid ad i de efterfølgende publikationer. Man drøfter også, om det har gjort en forskel for undersøgelsens resultater, om danske børn måske er mindre vant til den type tests, som undersøgelsen bygger på.

En forsker indvender, at udvælgelsen af items (enkeltspørgsmål) i den internationale læseundersøgelse er bestemt mere af statistiske og matematiske overvejelser end af et veldefineret begreb om læsning. Læseprøverne måler, hævdes det, først og fremmest et operationelt læsebegreb, som er skabt gennem selve undersøgelsens statistiske metode (Mylov, 1997).

Alligevel kan undersøgelsens hovedbudskab ikke manes i jorden. Der drøftes skyld, ansvar og forbedringsmuligheder. Seminariernes indsats på læsningsområdet kritiseres. Antallet af timer og ressourcer til skolen drøftes. I de følgende år sættes en række politiske initiativer i værk i forhold til skolen og læsningen. Der er ny folkeskolelov med styrkelse af danskfagets timetal. På læreruddannelsessiden udbredes ansvaret for læsning til at være en integreret del af arbejdet i flere fag. Der sættes penge af til læseforskning. I Ringkøbing Amt kommer dagplejemødre på kursus i at fange børnenes opmærksomhed med højt læsning.

Efter en indledningsvis modstand mod undersøgelsens metode og konklusioner virker en del lærere efterhånden som indstillet på at arbejde med en forbedring af læseindsatsen.

De følgende år er præget af flere undersøgelser af skolen i Danmark. I Nordlæs-projektet, hvis resultater udkommer i 1996, bekræftes hovedresultaterne fra tidligere vedrørende Danmarks placering i forhold til de andre nordiske lande. Også IEA laver flere undersøgelser, herunder de såkaldte Timss-undersøgelser. OECD sætter de såkaldte PISA-undersøgelser i værk (Rektor ved DPU, Lars-Henrik Schmidt forklarer i øvrigt i *Folkeskolen* (1999, nr. 18), at IEA foretager forskningsbaserede pædagogiske undersøgelser, mens OECD er interesseret i at udvikle indikatorer, der kan sige noget om forholdet mellem investeringer og præstationer. OECD er således et internationalt regeringsorgan uden forskningsmæssig baggrund).

I projektet Danlæs kortlægger Danmarks Pædagogiske Institut (DPI) læseniveauet i slutningen af 1. klasse i 75 kommuner. Flere dagblade kræver aktindsigt hos de kommuner, der har deltaget i undersøgelsen. Dermed berøres et tema, der efterfølgende kommer til at præge debatten, nemlig i hvilken form og på hvilket detaljeringsniveau (evt. kommuneniveau og skoleniveau), resultater af læseundersøgelser skal offentliggøres.

I en artikel i *Jyllands-Posten* 7.6.1997 opregner den daværende undervisningsminister Ole Vig Jensen under overskriften "En af verdens bedste folkeskoler" en række initiativer, som er sat i gang for at opruste inden for læseområdet. Til gengæld nævner han, at de elever, der deltog i den "den forældede syv år gamle læseundersøgelse", nu har forladt skolen, og at "alle eksperter siger samstemmende, at folkeskoleeleverne staver og læser bedre, end man nogensinde har gjort før i dette land".

Et evalueringsinstitut grundlægges, en lov om Danmarks Evalueringsinstitut vedtages 12.5.1999, og instituttet får til opgave at evaluere stort set hele uddannelsessektoren. En politiker ser i den forbindelse følgende sammenhæng: "Danmarks Evalueringsinstitut er et svar på de internationale evalueringer: Vi må selv gøre et hjemmearbejde og sørge for, at vi har nogenlunde hånd i hanke med, hvordan udviklingen er" (Elisabeth Arnold, tale 9, onsdag, 2.12.1998). I forbindelse med evalueringsinstituttets dannelse er der betragtelig debat om, hvilke informationer instituttet skal indsamle, og hvordan de skal offentliggøres.

En politiker, Marianne Jelved, spørger: "Er det virkelig meningen, at vi her i jeg havde nær sagt Grundtvigs fædreland skal have et Danmarks Evalueringsinstitut, der har som et af sine formål at rangordne institutionerne i forhold til hinanden?" Til gengæld taler hun for et vist arbejde med "kvalitetssystematik". "Det ord dækker jo bare over, hvordan vi får et fælles sprog til at tale om kvalitet" (Jelved, tale 13, 2.12.1998).

Det kommer til at stå i loven om Danmarks Evalueringsinstitut (LBK nr. 905 af 12.9.2000), at evalueringsrapporter ikke må indeholde oplysninger om karaktergennemsnit for fag, hold eller samtlige eksaminer eller prøver på en institution eller grupper af institutioner, bortset fra landsgennemsnit (paragraf, 8, stk. 3). Det kommer også til at fremgå af bemærkningerne til loven, at evalueringsinstituttets formål ikke er at rangordne specifikke udførende enheder i uddannelsessystemet.

Debatten om, hvad der skal offentliggøres og i hvilken form, fortsætter. I en artikel i *Politiken* 5.11.1999 kræver Venstres næstformand, Lars Løkke Rasmussen således, at "alle læse-, stave- og regneundersøgelser, der betales med skatteborgernes penge, skal offentliggøres".

Den daværende undervisningsminister, Margrethe Vestager, afviser offentliggørelse af testresultater fordelt på enkeltskoler i *Politiken*, 26.10.1999. Hun er specielt betænkelig ved, hvad formålet med en sådan offentliggørelse skulle være.

Direktør for Danmarks Pædagogiske Institut, Lars Henrik Schmidt udtrykker i *Folkeskolen* (2000, nr. 13) ærgrelse over, at metoden i Danlæs-

undersøgelsen fremover spoles af, at skoler må bringe sig i position til at score godt på læsetests, når resultaterne skal bruges til sammenligning af skoler. Han frygter også, at der i fremtiden bliver fokus på at vælge mellem skoler snarere end at forbedre undervisningen.

Statsminister Andres Fogh Rasmussen refererer i sin nytårstale 2003 til undersøgelser, der har skildret den danske skole som dårligt placeret i internationale sammenligninger: ”Det er trist at læse internationale undersøgelser, som viser, at danske børn og unge lærer for lidt på en række vigtige områder. Vi skal have højnet det faglige niveau. Det er betingelsen for at klare sig i fremtiden.” Under undervisningsminister Ulla Tørnæs træder i sommeren 2003 en lov i kraft om, at hver skole skal offentliggøre sine karaktergennemsnit.

Case 5. Evaluering af historie med samfundskundskab i det almene gymnasium

Evalueringer foretaget af Danmarks Evalueringsinstitut kommer typisk i værk ved, at evalueringsinstituttets bestyrelse laver en handlingsplan over planlagte evalueringer, som godkendes i Undervisningsministeriet.

I Danmarks Evalueringsinstituts evalueringsrapport *Historie med samfundskundskab i det almene gymnasium* (2001: 9ff.) står, at formålet med denne evaluering er at evaluere kvaliteten af faget historie med samfundskundskab i det almene gymnasium i dag med særligt henblik på undervisningens mål, progression over de tre år samt de forudsætninger, eleverne udvikler i forhold til de videregående uddannelser”. I øvrigt nævnes Undervisningsministeriets og skoleejernes roller samt en række mere specifikke aspekter af undervisningen og dens rammebetingelser.

Evalueringen er en af to evalueringer af specifikke fag (det anden handler om fysik). Alligevel er en del af evalueringens problemstillinger af mere bred relevans for gymnasiet som sådan, f.eks. i forbindelse med spørgsmål om pædagogikken, mundtlighed/skriftlighed osv.

En deltager i evalueringsprocessen hævder, at kommissoriet for evalueringen har lagt op til en ”ganske omfattende, nærmest systematisk evaluering, hvor alle aspekter af faget stort set er relevante. Det har givet evalueringsgruppen en meget stor frihed til at fokusere. På den måde afspejler indholdet og konklusionerne tankerne i evalueringsgruppen og på skolerne mere end i den politiske sfære, om man vil”.

Evalueringen har været organiseret med en evalueringsgruppe med ansvar for rapportens konklusion og anbefalinger. Evalueringsgruppen var sammensat af nordiske eksperter inden for historie og didaktik samt danske

eksperter med indsigt i historiefaget. Evalueringsinstituttets projektgruppe har skrevet rapporten i samarbejde med evalueringsgruppen. Evalueringsinstituttet har haft det overordnede praktiske og metodiske ansvar for evalueringen og for, at der er anvendt hensigtsmæssige og pålidelige metoder. 15 udvalgte gymnasier har medvirket i evalueringen og har ikke formelt haft mulighed for at undslå sig. På hver af disse er foretaget selvevaluering og besøg af evalueringsgruppen. Endvidere er ved et konsulentfirma foretaget en interview- og en spørgeskemaundersøgelse blandt historielærerne.

Af gymnasielærere bliver selvevalueringsprocessen bliver beskrevet både positivt og negativt. Man er utilfreds med, at der ikke er sat et vist antal timer af til arbejdet, men nogle steder fortæller man, at der har været gode diskussioner og en god intern proces.

Gymnasielærere fra nogle skoler har rejst kritik af spørgeskemaundersøgelsen – det var vanskeligt at finde ud af, hvor meget eller lidt man skulle skrive til de åbne spørgsmål.

Besøgsrunden blev af en deltager set som en positiv oplevelse, idet ”der var en meget høj grad af faglig dialog”. Vedkommende kunne mærke, at evalueringsgruppen bestod af personer med indsigt i fagets problematik og arbejdsmetoder.

Inden offentliggørelsen af evalueringsrapporten blev der holdt en høring. Nogle deltagere nævner, at ”mødet var præget af en saglig og nøgtern debat, og det var meget glædeligt at opleve, at lærergruppens og elevgruppens protester mod at flytte historieopgaven til slutningen af 2.g blev hørt, således at denne anbefaling blev taget ud af den endelige rapport”. ”Saglige synspunkter blev hørt, og det synes jeg var fint”.

Et af evalueringsgruppens medlemmer nævner i et interview, at der er en interessant debat internt i gruppen mellem forskellige faglige synspunkter, ikke mindst hvad angår samfundskundskab (evalueringen handler om historie med samfundskundskab). Man drøfter for eksempel muligheden for at anbefale at gøre samfundskundskab til et særligt fag. Men den opfattelse kommer til at vinde tilslutning i gruppen, at skolevirkeligheden i dag ikke ser sådan ud, at man kan foreslå flere et-årsfag med lavt timetal. Personer med kendskab til skolen siger: ”Nej, det går ikke den vej”.

Et af evalueringsgruppens medlemmer bemærker sig et interessant spil mellem gruppen og Evalueringsinstituttets folk. ”Man kan nemt komme i en situation, hvor evalueringsgruppen bliver styret af sekretariatsfunktionen, og næste version foreligger lige med det samme bagefter. De arbejder hele tiden, men vi har andre jobs”. Personen forklarer, at vil man have indflydelse, får man det ved at arbejde i de indledende faser. ”Jeg sætter så

meget tid af, som det tager". På et tidspunkt var der ifølge et medlem af evalueringsgruppen et decideret forhandlingssammenbrud, og der var ikke altid god stemning mellem parterne. "Men evalueringsgruppen har det endelige ansvar".

Selve rapporten tager mange temaer op. I resumeet omtales behovet for både at udvikle mundtlighed og især skriftlighed i faget, temaer angående historisk bevidsthed og identitet, en styrkelse af fagets progression gennem gymnasiets tre år, en opprioritering af ikke-europæiske temaer, arbejdsformer, efteruddannelse, elevmedbestemmelse, øget IKT i undervisningen m.m.

Der er en vis pressedækning på rapporten. Aviser taler om "alvorlig kritik" og "manglende historiebevidsthed" og "for lidt udviklingsarbejde".

Der er i øvrigt i samme periode også artikler, hvor forskellige personer fra uddannelsesverdenen udtaler sig om svingende eller direkte svigtende kvalitet i evalueringsinstituttets rapporter.

Nogle historielærere synes, at de kom til at fremstå i dårligere lys i pressen, end evalueringsrapporten egentlig begrundede. Fra Evaluering sinstituttets side fandt man, at man gerne ville gøre opmærksom på det positive, men en pressemeddelelse fokuserer på nogle få problemer, og det giver alt i alt et mere negativt billede end sammenfatningen i rapporten, som igen giver et lidt mere negativt billede end hele rapporten. Man vil gerne øve sig i at flette mere af det positive ind i resumeer og pressemeddelelser. Man er imidlertid opmærksom på, at hvad angår pressen, er man "i periferien af vores professionalisme" og "vi er ude i nogle sammenhænge, hvor vi har meget lidt indflydelse på, hvordan tingene sker".

En anden mener, at pressemeddelelsen i en vis forstand skulle være med til at retfærdiggøre instituttets eksistensberettigelse. Et medlem af evalueringsgruppen var ked af den negative presseomtale, men samtidig også ærgerlig over, at mange lærere ikke gjorde sig den "fundamentale akademiske umage at læse rapporten".

Evaluering sinstituttet arrangerede efterfølgende en konference om opfølgning af evalueringen. Ifølge en deltager kom konferencen i højere grad til at handle om rapporten end om fremtiden og perspektiverne. Det blev nævnt af en anden deltager, at ministeriet var "sure", fordi man ikke mente, at en sådan konference burde iværksættes af evalueringsinstituttet.

Efterfølgende udtaler Historielærerforeningens formand sig meget negativt: "Instituttet manipulerer med realiteterne i øst og vest – vi kan ikke bruge deres evaluering til noget som helst".

Det hævdes efterfølgende, at en stor del af den skarpeste kritik kommer fra en af de deltagende skoler, som ikke fik så positiv evaluering, som den havde håbet.

I Danmarks Evalueringsinstitut reflekterer man over forløbet og overvejer mere fokuserede kommissorier, således at der bliver færre anbefalinger, og de enkelte anbefalinger måske bliver sværere at komme udenom.

Efterfølgende mener et medlem af evalueringsgruppen, at man alt i alt har lavet en ganske god rapport, som har været debatskabende. En lærer mener, at det i et vist omfang er banalt, hvad der står i rapporten. ”Det kunne hvem som helst i faget have sagt sig selv. Det havde man ikke behøvet at sætte så stort et apparat i gang for at finde ud af (...) Hvis gymnasiet havde et frafald på 40 pct., og eleverne ikke var i stand til at klare sig på de videregående uddannelser, så kunne der måske være grund til at tage problemerne op, men det er jo ikke det, der er situationen i dag”.

Case 6. Evaluering af håndarbejds læreruddannelsen

En rapport har i 1998 dokumenteret store beskæftigelsesproblemer for uddannede håndarbejds lærere. Ledigheden var mere end 20 pct., og næsten halvdelen af de arbejdstimer, som de beskæftigede leverede, fandt sted i jobs ”uden brug af håndarbejds læreruddannelsen” (Bennekou & Jørgensen, 1998: 27). Der har i de seneste år frem til rapportens udarbejdelse været en stigning i produktionen af håndarbejds lærere, som ikke er modsvaret af en tilsvarende udvikling i efterspørgslen.

Evaluering af håndarbejds læreruddannelsen foretages i regi af Danmarks Evalueringsinstitut og sættes i gang i marts-april 2001. I sommeren og efteråret kører et forløb med selvevaluering på håndarbejds seminarierne, arbejdsmarkedsundersøgelse og institutionsbesøg. Rapporten afsluttes og fremsendes til Undervisningsministeriet januar 2002.

På seminariet i T-by har man arbejdet indgående med at involvere alle medarbejdergrupper i selvevalueringen. Der er udarbejdet en detaljeret tidsplan, hvor evalueringsgruppen laver midtvejsrapporteringer til medarbejderne, og hvor f.eks. pedel, køkken og rengøring involveres. Også de studerende involveres, først til at godkende en plan for selvevaluering, senere til at kommentere på selvevalueringen og drøfte den på et fællesmøde. I selvevalueringen skal man besvare en række spørgsmål, som på forhånd er opstillet, herunder f.eks. spørgsmål om pædagogisk efteruddannelse af underviserne, meritoverførsler, planer for internationalisering af uddannelsen og målsætninger for efteruddannelse af medarbejderstaben fremover.

Den 25.4.2002 afholder ministeriet et møde med håndarbejdsseminarierne, hvor det meldes ud, at kun ét uddannelsessted skal videreføres, og at udbudet af pladser skal reduceres dramatisk.

Det indvendes fra seminarieside via et brev til ministeren, at ”den gængse fremgangsmåde med, at seminarierne arbejder videre ud fra evalueringsresultatet, følges ikke”. Anmodet om at kommentere på dette svarer ministeren (Uddannelsesudvalget, 2. samling, bilag 387): ”Jeg kan oplyse, at ministeriet ikke har en standardprocedure for opfølgning på en evaluering. Proceduren afhænger af evalueringens resultat. Håndarbejdsevalueringen indeholder en række anbefalinger, der henvender sig til Undervisningsministeriet, og en række, der henvender sig til seminarierne. Ministeriet har valgt at følge op på de anbefalinger, der er rettet mod Undervisningsministeriet, dels ved at reducere optaget i 2002 til den eksisterende håndarbejds-læreruddannelse, dels ved at lægge op til ganske væsentlige ændringer på området med virkning allerede fra optaget 2003, hvor der arbejdes hen imod udbud med meget begrænset optag af en ny professionsbacheloruddannelse til faglærer i håndarbejde (...).” Men seminarierne rådes derudover til ”at følge op på de af evalueringsrapportens anbefalinger, der retter sig mod dem, i det omfang, det i lyset af, at uddannelsen i den nuværende form er under afvikling, er meningsfuldt”.

Den 28.5.2002 får flere håndarbejdsseminarier foretræde for Folketingets Uddannelsesudvalg. De argumenterer blandt andet for, at deres evalueringresultater ikke er så ringe endda, at det er urimeligt at vurdere en eksisterende uddannelses i lyset af en kommende lov (professionsbachelor), og at evalueringer ikke tidligere er blevet brugt til slagtning af uddannelser, men i stedet til vurdering, justering og kvalitetssikring. Der peges endvidere på, at man føler, at Undervisningsministeriet i sin konsekvensdragelse går videre, end hvad evalueringspanelet har anbefalet. Det siges for eksempel at være ministeriets egen ”tillægsgdagsorden”, at den nye uddannelse kun skal udbydes et sted, og at reduktionen i udbuddet skal være så drastisk, som det nu foreslås.

Undervisningsministeren (Ulla Tørnæs) er den 19. juni i samråd med Uddannelsesudvalget, idet udvalget har bedt hende svare på, hvilke planer regeringen har for håndarbejds-læreruddannelsen, herunder antallet af uddannelsessteder og -pladser.

Af ministerens talepapir fremgår, at hun finder evalueringen ”usædvanlig og meget alvorlig”. Hun lægger her både vægt på beskæftigelsessituationen og på, at de studerende udtrykker utilfredshed med uddannelsen, og at aftagere i produktionsvirksomheder er kritiske over for håndarbejds-lærer-

nes teknisk-faglige kvalifikationer. Ministeren nævner, at evalueringspanelet ” anbefaler, at en ny uddannelse inden for området i givet fald bør etableres på professionsbachelorniveau. Uddannelsen bør målrettes en læreruddannelse med et håndværksmæssigt indhold. Formål, professionsområde og fagudbud skal strammes op i forhold til den eksisterende. Optaget til uddannelsen skal tilpasses beskæftigelsesmulighederne inden for området” (Ministerens talepapir af 19.6.2002, Uddannelsesudvalget, bilag 438: 3).

Undervisningsministeren har på den baggrund besluttet, at den eksisterende håndarbejds læreruddannelse afvikles og afløses af en ny fagligt og dimensioneringsmæssigt smal professionsbacheloruddannelse i håndarbejde. Den nye uddannelse skal have snævert fokus på det tekstile område. Optaget reduceres fra sommeren 2002. Uddannelsen skal fremover placeres i CVU-regi, idet der dog tages hensyn til de enkelte hidtidige håndarbejdsseminariers ønsker om den nærmere karakter af den ny institutionsmæssige tilknytning.

Ministeren lukker i talepapiret op for den mulighed, at to institutioner (hhv. øst og vest for Lillebælt) kan udbyde hver 25 pladser på den nye uddannelse, idet beskæftigelsesudviklingen dog skal følges nøje, idet ”regeringen ikke vil uddanne til arbejdsløshed, (...) navnlig når der på mange områder er mangel på uddannede” (Ministerens talepapir af 19.6.2002, Uddannelsesudvalget, 2. samling, bilag 438: 5)

I ministerens overvejelser om placering af udbudet på to institutioner indgår evalueringsresultaterne samt regionale overvejelser, konkret at ”yderkantområderne ikke bliver snydt” (Ministerens talepapir af 19.6.2002, Uddannelsesudvalget, 2. samling, bilag 438: 7).

Da ministeren ud fra en samlet overvejelser foreslår CVU Midt-Vest samt CVU København og Nordsjælland, gør ministeren sig i sit talepapir en del overvejelser over, hvorledes der kan ydes kompensation til de områder i landet, der mister uddannelsespladser.

Der forekommer en del uro efterfølgende, ikke mindst foranlediget af håndarbejdsseminarier, der er usikre på deres fremtid. Undervisningsministeren præciserer, at der ikke er tale om at ”nedlægge uddannelsen i evalueringsinstituttets navn”, men at omstille den ”i lyset af de problemer, evalueringsrapporten har afdækket” (Ministerens svar på spørgsmål 64, Uddannelsesudvalget, 2. samling, bilag 419: 2)

Nogle af deltagerne oplever det som en lidt ”kafkask” situation, at de har medvirket indgående i, hvad de selv mente, var en åben og konstruktiv selvevaluering, når de nu ser, at deres uddannelsesinstitution i praksis ender med at få et dødsstød.

En lokalpolitiker, der forsvarer de lokale interesser knyttet til uddannelsespladser i sin egn, skriver i et læserbrev, at den manglende dialog og formen for evaluering ”vil få følger i fremtiden, idet en stor del af evalueringerne er af frivillig art, og man må forvente, at de uddannelser, der i fremtiden skal medvirke til sådanne evalueringer, vil skæve til, hvordan det er gået med håndarbejdsuddannelsen”.

I et andet læserbrev hævdes det, at det er en kulturarv, der er sat på spil. Den dårlige beskæftigelsessituation blandt håndarbejds lærere hævdes at være et resultat af mange års forringelser på folkeoplysningen. Håndarbejde og håndværk udgør en betydelig del af kulturarven, men nedprioriteres i en tid, hvor industriens interesser lyder, at ”bedst er billigst” og hvor betjening af tastaturet på en computer efterhånden er den eneste anerkendte form for håndarbejde.

Case 7. Evaluering af undervisning på handelshøjskolen i Q-by

På en handelshøjskole har man en praksis med at lade studerende udfylde evalueringsskemaer i alle undervisningsforløb. Dekanen nedsætter på et tidspunkt et udvalg, der skal se nærmere på, om denne praksis skal fortsætte, og om man eventuelt kunne tænke sig andre evalueringsformer og et andet skema. Man kunne f.eks. tænke sig stikprøveevalueringer og evt. mere dybtgående evalueringer i særlige tilfælde. Spørgsmålet om ressourcer til evaluering indgår ikke eksplicit i udvalgets kommissorium, ligesom det heller ikke indgår, hvorfor dekanen ser et behov for at tage evalueringssystemet op til revision.

Der gisnes om, at dekanen har forpligtet sig via kontraktstyring til at rapportere noget om arbejdet med evaluering og kvalitetssikring. Der gisnes også om, at dekanen har problemer med de ressourcer, der afsættes til at indkode alle studerendes skemaer fra alle fag.

Udvalget drøfter en lang række spørgsmål og laver herefter en indstilling, som lægges frem på et åbent debatmøde. Udvalget indstiller, at det eksisterende spørgeskema omarbejdes, så der lægges lidt mindre vægt på kontrolanvendelsen og mere vægt på information til den enkelte lærer om, hvad der kan gøres bedre. Udvalget indstiller endvidere, at hvert enkelt studienævn kan lave sin egen evalueringsmodel, eventuelt inspireret af det omarbejdede evalueringsskema. Udvalget tager stilling til, om nogle undervisningsaktiviteter kan udelades af evalueringssystemet, og indstiller at vejledning udelades. Denne omfatter nemlig kun ganske få studerende ad gangen. Udvalget tager også stilling til, om skemaet skal være heldækken-

de eller stikprøvebaseret. Efter en del overvejelser foreslås ”heldækkende”, fordi det vil være vanskeligt at argumentere for, hvilke lærere der skal udvælges til særlig evaluering, hvis ikke alle er med.

På debatmødet opnås der en god tilslutning til udvalgets tanker om evaluering. Der rejses også spørgsmål om, hvorfor det eksisterende evalueringssystem ikke først er blevet grundigt evalueret, og udvalget begrundet det med ressourceknaphed. De studerende efterlyser i øvrigt mere synlige konsekvenser af evalueringssystemet. Det kommer frem på mødet, at selv om kontrolfunktionen af og til har et dårligt ry, er den faktisk nødvendig ikke mindst for studienævn, der er forpligtede på at overvåge kvaliteten af f.eks. eksterne underviseres arbejde.

Der drøftes anonymisering og offentliggørelse. Mødet lander på, at den enkelte studerende skal have lov at svare anonymt, da der foreligger en skæv magtrelation mellem lærer og studerende. Man enes også om, at evalueringer af lærere ikke skal offentliggøres.

På et enkelt institut skaber udvalgets arbejde stor ravage. En adjunkt søger at skabe samlet modstand mod det foreslåede evalueringssystem. Hans argumenter er, at det udgør et angreb på den enkelte lærers eget råderum over undervisningen samt på lærerens faglige rettigheder. Studenterbaseret evaluering vurderes som usaglig og en folkedomstol. Adjunkten argumenterer også, at en for stor fokus på evalueringer vil give de studerende et falsk indtryk af deres egen rolle. I stedet for at samle sig om deres egen indsats vil de føle sig støttet i, at det blot handler om at vurdere læreren. Argumenterne flettes sammen med et opgør mod moderne kvantificerede og mistillidsprægede ledelsesmetoder.

Adjunkten ønsker de studerendes anonymitet brudt. Adjunkten ønsker en styrkelse af lærerens retsstilling over for evalueringresultater, herunder ankeinstanser osv.

Adjunkten er desuden oprørt over, at evalueringssarbejdet er nået så langt, uden at man har hørt lærerkollegiet på det enkelte institut. Adjunkten starter en underskriftsindsamling mod evalueringssystemet. Der er imidlertid meget begrænset tilslutning hertil. Mange er uenige med adjunkten, og kulturen det pågældende sted er ikke til underskriftsindsamlinger.

Det diskuteres, hvilken rolle de studerendes tilfredshed med undervisningen skal spille på institutionen som helhed. Adjunkten argumenterer, at det vil være usagligt at indrette sig efter dette kriterium. Flere andre lærere argumenterer, at studentertilfredsheden er vigtig, ikke mindst fordi tiltrækning og fastholdelse af studerende er afgørende faktorer i spillet om institutionens økonomi. Atter andre nævner i uformelle sammenhænge, at adjunk-

tens modstand mod evaluering måske skyldes, at han selv har dårlige evalueringresultater. Det blev drøftet, hvad solidaritet er i den forbindelse. Er solidaritet, at andre lærere slæber med de store hold og de vanskelige undervisningsopgaver for at give plads til en, som måske ikke kan indfri evaluering forventningerne?

Evalueringkonceptet implementeres som foreslået af udvalget. Senere opnår adjunkten ikke fast ansættelse.

De kanen vender på et tidspunkt tilbage og melder, at man på grund af en klemt ressourcesituation ikke ser sig i stand til administrativt at understøtte det evalueringkoncept, som udvalget foreslog. Det lokale studienævn finder selv efterfølgende de nødvendige ressourcer. Evalueringkonceptet fastholdes på dette studienævns område. Der er efterfølgende løbende diskussioner om, hvilke temaer studienævnet skal tage op som følge af evalueringerne. Enkeltproblemer i forhold til de enkelte lærere drøfter studielederen med hver af disse.

Case 8. Evaluering af daginstitutionerne i X-by

Som større by har X-by flere hundrede daginstitutioner, fritidsklubber m.m. Pædagogerne i X-by er kendt for at have en stærk fagforeningstradition. Som en af dem selv siger i spøg: ”Vi er pædagogerne i X-by, og vi er imod alting”.

På et tidspunkt rumler i forvaltningen nogle tanker om at komme i gang med kvalitetsmåling og -udvikling, ligesom byrådet på et tidspunkt vedtager en principbeslutning herom. En afdelingsleder i forvaltningen har været med til at udvikle et koncept for kvalitetsmåling af daginstitutioner og fritidsordningen baseret på spørgeskemaer til forældre. Konceptet indbefatter en pakke, hvor resultaterne kan computerbehandles nærmest automatisk, og hvor den enkelte institution kan sammenligne sig med andre uden dog at forvaltningen kan sætte navn på hver enkelt institution.

Konceptet benyttes af nogle af institutionerne. Men der er også en del debat om konceptet, dels gående på de definitioner af kvalitet, som spørgeskemaerne bygger på, dels på om forældre er pålidelige kilder til udsagn om de pågældende forhold som blandt andet handler om børnenes hverdag i institutionen.

På et tidspunkt er afdelingslederen med til et ledelsesmøde, hvor det beslutes at tage fat på implementeringen af byrådets beslutning om kvalitetsmåling. Søndagen efter ringer en journalist til ham for at høre, om der er nogen nyheder til lokalsiderne. Afdelingslederen nævner – efter eget udsagn nærmest i flæng – at der skal til at sættes gang i arbejdet med kvali-

tetsmåling. Da dette nævnes i mandagsavisen, bryder en stærk bølge af modstand løs blandt mange af kommunens institutioner. Man er oprørt over ikke at være blevet informeret om kvalitetsinitiativet gennem gængse kanaler såsom møder, samarbejdsudvalg og interne nyhedsbreve, ligesom man er bekymret over indholdet i det fremtidige kvalitetsarbejde.

Fra forvaltningens side prøver man at dæmpe gemytterne. Der vedtages et koncept, hvor hver enkelt institution selv må bestemme, hvordan den vil evaluere sin virksomhed. Kravet er blot, at evalueringresultaterne fremlægges på institutionens hjemmeside. Forvaltningen udvikler et koncept for at hjælpe institutionerne hermed, således at offentliggørelsen af evalueringerne ikke skal byde på unødige tekniske forhindringer. Et par år efter har omkring halvdelen af institutioner lagt en evaluering ud på deres hjemmeside. Der er blandt mange pædagoger fortsat betragtelig modstand mod hele ideen. Det understreges ikke mindst, at pædagogerne er uddannede til at tage sig af børn, ikke at lave målinger og beskrivelser.

Der er ingen sanktioner mod de institutioner, der er længe om at leve op til evalueringsbestemmelserne. Forvaltningen ser sin rolle som en varsom formidler af en mellemvej mellem forskellige interesser:

”Vi har fået sådan en management-bølge ind over os her i X-by lidt sent. Institutionerne vil ikke lave virksomhedsplaner, derfor har byrådet været så klogt at vedtage noget andet noget ... Jeg synes ikke, de har fået den (managementbølgen) så meget ude i institutionerne. Men vi er nødt til at gøre opmærksom på, at vi er en del af et politisk system. Det kan være sort, det kan være hvidt ... Vi bliver nødt til at lande et sted midt i, for at vi alle sammen vil have med det at gøre ...”.

Kapitel 7

Magtens hvem: Aktører og rationaler i evaluering

I kapitel 4 blev præsenteret hypoteser om, hvilke aktører med hvilke rationaler, der kunne blande sig i evalueringsforløbene. Her samles de vigtigste iagttagelser fra casene om dette emne.

Det metodiske rationale

I flere af casene spiller det metodiske rationale en temmelig tilbagetrukket rolle. Det er ikke mindst tilfældet i de situationer, hvor et evalueringsdesign ”købes” udefra. Det gælder for eksempel akkrediteringscasen og casene fra ældreplejen. I et vist omfang er også evalueringsinstituttets metode så institutionaliseret, at den ikke gøres til genstand for ny forhandling i de forløb, der er undersøgt her (historieevalueringen og håndarbejdsevalueringen).

I nogle tilfælde ønsker man slet ikke en drøftelse af den metodiske side. Da der i akkrediteringscasen kommer røster frem om, at nogle af akkrediteringens billeder af hospitalet er fabrikeret til lejligheden, er der for så vidt tale om et bidrag til en debat af pålideligheden i evalueringens metode. Ledelsen svarer med, at hvis der har været fusk, må der drages ledelsesmæssige konsekvenser ude på de enkelte afdelinger. Hierarkiet må træde i stedet for metodede-batten.

I casen med evaluering af daginstitutioner spiller et metodisk rationale heller ingen særlig rolle, men det skyldes ikke, at metoden er bundet forinden, snarere at der i casen gives frihed til, at hver institution kan vælge sin metode.

I casen med evaluering af undervisning på en videregående uddannelsesinstitution er der en vis debat om metodevalget, men ikke særskilt med udgangspunkt i et metodisk rationale. Kun i casen med de internationale læseundersøgelser er der en indgående debat om undersøgelsesmetode, validitet osv. Dette evalueringsforløb er som det eneste af de undersøgte en del af en forskningsinstitutions kerneopgaver, hvorfor det er forventeligt, at det metodiske rationale spiller en stor rolle for de ansvarlige forskere, ligesom de gerne drager i brechen for at forsvare deres metode offentligt. At der ikke har været fuldstændig enighed mellem alle internationale og danske forskere i sagen trækkes blot ind som en yderligere nuancering af det metodiske rationale.

Ellers må man i hovedtræk sige, at det metodiske rationale spiller en meget tilbagetrukket rolle i de fleste cases, især netop i dem, hvor et helt evalueringssystem købes, implementeres eller gendannes.

Det har en implikation i relation til magtanalysen. Det betyder nemlig, at metodisk kompetence i modsætning til, hvad man måske ville tro, *ikke* i sig selv udgør en meget væsentlig ressource i kampen om at præge evalueringssystemerne. I mange tilfælde er et bestemt design nemlig forudbestemt eller leveret udefra. Den mindsker den forskel, som et metodisk rationale kan spille i selve casen. Ikke mindst i akkrediteringscasen fremstilles det sågar som en styrke ved evalueringen, at metoden (og altså dermed kriterierne) ikke kan påvirkes ved lokal forhandling.

I de cases, hvor evalueringen er en del af et større system (akkreditering, stregkodesystem, institutionaliseret evalueringssystem) drejer den afgørende magt sig ikke om den metodiske kompetence til selv at snedkerere en evaluering; den handler i stedet om lovgrundlag, ressourcer, organisation til at rekvirere et evalueringssystem og styre dets implementering og anvendelse. Dette er konsistent med nogle af de nyere iagttagelser i videnssociologien: Viden skabes ikke længere af enkeltpersoner; vidensteknologier, ressourcer, samarbejde og organisation er afgørende.

Det politiske rationale

Det politiske rationale spiller en tilbagetrukket rolle i flere af casene, om end naturligvis evalueringssystemer oprindeligt er politisk godkendt. Det gælder både Evalueringsinstitutets arbejde og casene fra ældreplejen. I et tilfælde prøver folk fra forvaltningen sågar at forpligte politikere på ny på prioriteringer og udvikling af systemet, men politikerne nikker blot, og har det formodentlig fint med, at monitoreringssystemet løser de opgaver, det løser.

I et andet tilfælde (evaluering af daginstitutioner) findes der blot et politisk opdrag om evaluering, men ikke nogen nærmere specificering af formål, metode eller anvendelse.

I tilfældet med akkreditering spiller det politiske rationale også en tilbagetrukket rolle i selve evalueringssystemet; men af den grund skal man ikke overse, hvilke politiske motiver der kan være bag ønsket om at markere hospitalsvæsenet i Y-by gennem netop denne blåstemplende evaluering. Der foregår nemlig i landet flere samtidige evalueringssystemer, der alle søger at bringe sig i position til at præge et eventuelt fremtidigt fælles nationalt system.

I to tilfælde træder det politiske rationale imidlertid ekstraordinært tydeligt i karakter. Den ene er ændringen af uddannelsen af håndarbejds lærer til professionsbachelor. Her ser ministeren den "usædvanlige og alvorlige" evaluering som en anledning til at foretage større ændringer både i uddannelsens indhold, dimensionering og institutionelle placering. Hendes konsekvensdragelse rækker videre, end hvad evalueringsrapporten eksplicit foreslår. Hun gør sig imidlertid umage med at finde kompensation til de områder af landet, der mister uddannelsespladser. Hendes opfølgning på rapporten virker overraskende og krænkende på de ramte institutioner. Men lovgivningen (ikke mindst lov om Danmarks Evalueringsinstitution §§ 10 og 11) giver hende udmærket hjemmel til en sådan opfølgning. Andre parters skuffelse er således en skuffelse i forhold til deres egne forventninger, ikke en skuffelse i forhold til evalueringsforløbet formalia. At den reformerede uddannelse lægges i CVU-regi, er formodentlig ikke til ulempe for ministeren, hun tænker måske på denne konstruktion som hendes og som et system, hun gerne vil fremme.

Også i casen med de internationale læseundersøgelser træder den politiske logik meget eksplicit frem. Evalueringsresultaterne bliver åbenlyst brugt til i mange år derefter at begrunde forskellige indsatser i forhold til skolen. Et sporadisk forsøg på at kalde undersøgelsen "forældet" fra en undervisningsministers side falder til jorden; reformbølgen fortsætter i flere år derefter. Den politiske logik manifesterer sig ikke blot som en kamp om, hvilke reforminitiativer, der skal sættes i værk i skolepolitikken. Den viser sig i høj grad også som en videnspolitisk kamp om skolen. Et centralt spørgsmål er, hvilke data om den enkelte skole, som offentligheden og forældrene har krav på at kende. Efter mange års kamp på dette område tages et afgørende skridt med den seneste lov om, at skolerne skal lægge de gennemsnitlige karakterer ud på internettet.

Fælles for de to tilfælde, hvor politikere er stærkt på banen, er også det faktum, at evalueringsresultaterne i begge tilfælde er ganske alarmerende. Det inviterer alt andet lige til politisk initiativ. I historieevalueringer er kriterierne mange, vurderingerne blandede og forslagene talrige.

Det administrative rationale

Ikke mindst i casene fra ældreplejen synes implementeringen af overordnede monitoreringssystemer at være en ganske velkommen ting set fra et administrativt synspunkt. Systemerne hjælper med økonomisk overblik og dokumentation samt med at håndtere klager. Endvidere er det fra et administrativt synspunkt udmærket, at systemet hjælper med til at give forskellige

faggrupper en fælles måde at håndtere visiteringsmæssige prioriteringer på. Konflikter med personalet og især konflikter om strikt kontrol versus fleksibilitet i opgaveudførelsen løses ved små forhandlinger frem og tilbage samt diverse tilpasninger af systemet. Det overordnede administrative hensyn til overblik, styring og dokumentation lades aldrig af syne. Og det lader sig heller ikke forstyrre af, at statsministerkandidater på et givet tidspunkt forlanger mere fleksibilitet i hjemmehjælperens møde med den enkelte ældre.

Den administrative logik står også ganske stærkt i billedet i forbindelse med evaluering af daginstitutionerne. I denne case kan man som forvaltning ikke komme særlig langt med eksplicite krav om evaluering. Og opfølgningen på det vedtagne metodefrie koncept sker kun mildt og langsomt. Forvaltningen er formentlig afhængig af, at der ikke bliver for meget uro på området. Daginstitutionernes evne til at skabe ballade er således en ressource i magtspillet. Alligevel er det administrative rationale på dette område ganske bevidst og ganske tålmodigt. Det bevæger sig bare, givet omstændighederne, små skridt ad gangen.

Det administrative rationale står også ganske stærkt i akkrediteringscasen. Akkreditering tænkes som et overordnet administrationspolitisk koncept, der forener mange af hospitalets øvrige evalueringsmæssige, udviklingsmæssige, styringsmæssige og politiske tiltag. Imidlertid ender akkrediteringsforløbet med at spænde den administrative bue så hårdt, at flere af forberedelsestiltagene får et vist hektisk præg. Tillige binder forløbet administrative ressourcer både før, under og efter forløbet.

Det fagprofessionelle rationale

I ganske mange af tilfældene føler fagprofessionelle sig overset eller undermineret af evalueringer. I akkrediteringscasen beklager lægerne, at forskning og uddannelse ikke tælles med i akkrediteringen. I ældreplejen føler nogle faggrupper deres faglighed undermineret, når de mister visitationsretten. I daginstitutionscasen opponerer man mod at skulle beskrive sin pædagogiske faglighed i skrift og tal. I historieevalueringen føler mange lærere sig trådt på af hele evalueringsforløbet, selv om der dog også er eksempler på, at man har haft en god faglig dialog. Sidstnævnte skyldes ikke mindst, at der sidder anerkendte fagfolk i evalueringsgruppen. I håndarbejdscazen føler nogle fagpersoner, at deres mange kundskaber ikke anerkendes, og at den kulturarv, de repræsenterer, manes til jorden. I evaluering af undervisning på en videregående uddannelsesinstitution appellerer en lærer til sine

kolleger, sine faglige rettigheder og sin faglige organisation i kampen mod evalueringssystemet (uden dog at få opbakning i sin sag).

I ingen af tilfældene er evalueringsforløbet helt igennem initieret og bestemt af fagprofessionelle selv. Hvilke strategier vælger de fagprofessionelle over for evalueringer og evalueringssystemer?

- a. De yder modstand, råber op og forhæler evaluering. Det sker tydeligt i daginstitutionscasen. Det sker også i akkrediteringscasen, men dog kun med den konsekvens, at hele forløbet bliver stresset umiddelbart før deadline. I et særligt giftigt tilfælde i akkrediteringscasen forsøger enkelte fagprofessionelle i medierne at underminere processens troværdighed ved at nævne, at det akkrediterede billede af hospitalet er fabrikeret til lejligheden.
- b. Fagprofessionelle bruger evalueringen som anledning til at bringe egne dagsordener op. Det sker ikke mindst i akkrediteringscasen, hvor dagsordenen er forskning og udvikling og hvor man i øvrigt peger på manglende tid, ressourcer, instrumenter og rengøring, som overses i den foreliggende evaluering.
- c. Fagprofessionelle finder kompromisser mellem overordnede evalueringsmæssige og organisatoriske hensyn på den ene side og egne hensyn på den anden. Det sker i ældreplejen, hvor man finder måder at forhandle om de visiterede ydelser på eller ”mingelerer” lidt med de tider, som det tager at udføre opgaverne. Man forhandler sig til løsninger om ikke at kode enkeltopgaver eller ikke informere brugeren om, hvor meget tid, der er sat af til vedkommende, såvel som til i perioder slet ikke at scanne sit arbejde ind. Derved mørklægges nogle data. I evaluering af undervisning på en videregående uddannelsesinstitution lader man fortsat de studerendes brugertilfredshed spille en vis rolle i obligatoriske evalueringsskemaer, men man tilpasser skemaet, så det er mere gavnligt for den enkelte lærers pædagogiske udvikling. I historieevalueringen overtaler man evalueringsgruppen til enkelte småændringer inden den endelige rapport.
- d. Fagprofessionelle indordner sig under evalueringer, men bruger dem også i deres egen argumentation. Fagprofessionelle tilegner sig evalueringer, både i arbejdets indhold og i måden at ræsonnere på. Et langt stykke ad vejen leverer fagprofessionelle det input, som evalueringssystemer beder om. Det gælder ældreplejen, akkrediteringscasen og undervisningsevalueringen, hvor det daglige arbejde lader sig måle og beskrive ved hjælp af evalueringssystemets termer, og hvor

fagprofessionelt personale medvirker hertil. Men i nogle tilfælde bruges evalueringssproget så efterfølgende til at argumentere ud fra. Det gælder for eksempel, hvis der er tvister om visiteringen af ydelser til en ældre borger. Man finder et argument for at tage et særligt hensyn, hvis der er behov for det.

- e. Fagprofessionelle søger at modvirke en dramatisk konsekvensdrageelse af evalueringer. Det sker i akkrediteringscasen, hvor læger er inde på, at dette system ikke ufordøjet kan blive model for et efterfølgende nationalt system. Det sker i håndarbejdscasen, hvor seminariefolkene opponerer mod nedlæggelser – dog uden større held. Det sker i historiecassen, hvor ”drøftelser” og ”diskussion” er det måske mest centrale resultat af forløbet, så vidt det er beskrevet i casen.

I mange tilfælde virker fagprofessionelle oprørte eller krænkede over den behandling, de får i evalueringer. De fremstiller i flere tilfælde evalueringer som overgreb. I nogle tilfælde bliver en sådan følelse af noget urimeligt et input i forhandlingen om, hvordan evalueringer skal udvikle sig, fordi andre parter, for eksempel en forvaltning, er afhængig af en vis ”ro” i systemet, ligesom man ofte ikke ønsker at fremstå som krænkende. Imidlertid er fagprofessionelles følelse af krænkelthed ofte kun en krænkelthed i forhold til egne forventninger, ikke i forhold til andres forventninger eller i forhold til formalia omkring evalueringforløbet. Det ser man tydeligst i håndarbejdsmiljøets krænkelthed over, at ministeren brugte loven til at følge op på evalueringresultatet. Blandt andet derfor har den fagprofessionelle krænkelthed ikke altid nogen praktisk virkning på evalueringens udfald. Men ganske ofte er der tale om en kompliceret og uafgjort, løbende forhandling, hvor fagprofessionelles eventuelle krænkelthed og løbende stillingtagen ganske givet spiller en rolle.

Brugerens rationale

I betragtning af de senere års megen omtale af brugeren af offentlige ydelser er det bemærkelsesværdigt, så lidt brugere blander sig i de nævnte caseforløb. I akkrediteringsforløbet handler evalueringen efter sigende om brugerens møde med hospitalet, men spørgsmålet belyses ud fra et perspektiv på sagen, der minder om akkrediteringssystemets oprindelige tilknytning til forsikringselskaber i udlandet. Der fokuseres ikke mindst på sikkerhed i hospitalets procedurer. De konkrete danske brugere indgår ikke.

I ældreplejen kommer brugeren først og fremmest ind i billedet som en, der potentielt kan klage over den leverede ydelse. Systemerne indretter sig

derfor på at kunne imødegå klager. Den konkrete brugers stemme indgår ellers ikke direkte i monitoreringsarbejdet.

I lyset af de senere års fokus på gennemsækelighed og transparens kan man undre sig over, at den enkelte bruger i S-by ikke oplyses om, hvor meget tid, der er sat af til hjemmehjælperens løsning af de visiterede opgaver. Brugeren har dermed ikke adgang til oplysninger om den helt centrale faktor, som systemet er orienteret imod, nemlig tidsfaktoren. Selv om hele systemet er gearet til at måle, styre og honorere ATA-tid, så kan brugeren ikke vide, om vedkommende får den ATA-tid, som er tildelt vedkommende. Denne nedprioritering af brugerens perspektiv på sagen er ikke konsistent med den erklærede ideologi på området. Den må skyldes, at andre rationaler i dette tilfælde har vundet over brugerrationalet.

I daginstitutionscasen skal brugere indgå i evalueringsarbejdet, men det er op til den enkelte institution at finde ud af hvordan. I undervisningsevalueringen er nogle få studerende med til at drøfte ændringer i evaluerings-systemet. I evalueringsinstituttets evalueringskoncept for uddannelser indgår studerende som medspillere i selvevalueringen. Men generelt må man sige, at omtalen af brugere og især tale *på vegne af brugere* står noget stærkere i billederne af evaluering end brugernes egne udsagn.

Den kulturelle baggrundsstøjs rationale

Til denne kategori hører temaer, som er ”oppe i tiden”, men som ikke nødvendigvis på forhånd kan tilskrives nogen konkret aktør.

Denne faktor spiller en ganske væsentlig rolle i evalueringsforløbene. Flere af evalueringsforløbene starter med en konstatering af ”behovet for dokumentation”, som i abstrakt forstand tages mere eller mindre for givet. Brugerperspektivet på offentlige organisationer appelleres der ligeledes til som et nærmest obligatorisk kulturelt mønster. Det samme gælder retsliggørelsen af ydelserne og dermed klagemulighederne. Da en respons på en klage afhænger af den offentlige organisations dokumentation af sin egen indsats, er disse strømninger således flettet sammen. Pointen er imidlertid, at disse hensyn ofte indgår som ”brede begrundelser” snarere end specifikke analyserede faktorer.

Den aktuelle ”tidsånd” angriber f.eks. håndarbejds læreruddannelsen hårdt. Med til at forklare slagkraften i evalueringen af håndarbejds læreruddannelsen er formodentlig, at kulturarven og traditionen i denne uddannelse har svært ved at markere sig i forhold til den aktuelle ”tidsånd”. Da der samtidig er store beskæftigelsesproblemer, giver den almindelige kulturelle baggrundsstøj snarere et godt argument for at lade evalueringen slå til.

De internationale læseundersøgelser får formodentlig kulturel medvind i sejlene i forbindelse med en stadig stærkere forestillingsverden knyttet til global konkurrence. Tidligere værdimæssige billeder af skolen som fællesskab og som lighedsorienteret institution er antagelig stærkt svækkede i dag. I stedet kan brugerne (forældrene) i dag bedre tillade sig at stille krav til skolen, og lærernes almene omdømme er svækket. Sådanne forhold i "tidsånden" er med til at forklare de internationale læseundersøgelsers gennemslagskraft såvel som deres efterfølgende bølgevirkninger med hensyn til skolepolitik og -evaluering. (Dog tæmmes den overordnede evalueringsudvikling også af en mere traditionelt begrundet dæmper på offentliggørelse af sammenligninger og karaktergennemsnit.)

Temaer i den kulturelle baggrundsstøj er formodentlig også med til at forklare, hvorfor adjunkten i casen med undervisningsevaluering ikke får opbakning til sin protest mod evalueringsskemaerne. Hele institutionen er i dag baseret på indtægter ved at tiltrække og fastholde studerende. Går han ikke ind for at undervise med henblik på en vis studentertilfredshed og formår han ikke at leve op til dette kriterium, føler flere af de andre kolleger, at de skal trække et større læs. Med denne tolkning af tidsånden har hans modstand svært ved at få opbakning.

I historieevalueringen forlader evalueringsgruppen en ide om at foreslå samfundskundskab som et lille særskilt etårigt fag med argumentet: "det går ikke den vej". Med andre ord vil en kyndig evaluator, der søger indflydelse, tage udviklingen i tidens stemninger og fagpolitiske trends i betragtning, når der nævnes ændringsforslag og anbefalinger. *Derved bliver evaluering en slags accelerator for tidsånden. Evaluering fremmer hastigheden i tidsåndens forslag om at ændre praksisser, der ikke "følger med".*

Ud over de ovennævnte, forventede rationaler byder casene på yderligere to sæt af aktører og rationaler. De knytter sig til medierne og til evaluator.

Medier og medielogik

Casene viser, at medierne og den mediemæssige logik spiller en rolle i flere af evalueringsforløbene. I forhold til de internationale læseundersøgelser er medierne vedvarende inde i billedet, mest ofte med overskrifter af typen "Danmark i bund". I de fleste af de andre evalueringsforløb er medierne inde mere kortvarigt. Det gælder i akkrediteringscasen, daginstitutionscasen, historiecassen, håndarbejdscasen og strekcodecasen. I de fleste af disse tilfælde er det med negative overskrifter om offentlige organisationers indsats, i akkrediteringscasen dog med negative overskrifter om evalueringsforløbet snarere end om selve serviceydelsen. Medierne medvirker med

deres rationale til at forenkle og tilspidse evalueringsproblemstillingerne. Den typiske beretning er ”dette er skidt, nu må der handles”. Medieomtalen fungerer derfor ikke mindst som springbræt til politisk handling. Efterfølgende rummer mediernes forskellige stridigheder med synspunkter for og imod. Selv i de tilfælde, hvor en mediestorm er overstået, spiller bevidstheden hos de involverede aktører om tidligere medieomtale eller potentiel medieomtale formentlig en stor rolle. De få øjeblikke, hvor mediernes retter søgelyset mod en bestemt evalueret indsats, er nok mere afgørende end ganske mange andre momenter.

Evaluators eksistensberettigelse

Den evalueringsansvarlige – hvad enten der er tale om en kvalitetsmedarbejder eller et særskilt evalueringsinstitut – har tilsyneladende et særskilt rationale. Hertil knytter der sig et behov for at vise og begrunde sin eksistensberettigelse. Det kommer blandt andet til udtryk ved, at man fra evalueringsinstituttet må udsende pressemeddelelser, der kan vise, at evalueringsinstituttet når frem til vigtige og markante resultater. I akkrediteringscasen gør kvalitetsafdelingen sig relevant ved at uddanne ganske mange og gøre kvalitetsarbejdet så synligt som muligt overalt på hospitalet.

Til positionen som evaluator er der knyttet særlige ressourcer. For eksempel er der den fundamentale ressource, at evaluator ofte kan arbejde heltids med evaluering. Det giver en tids- og ressourcefordel i forhold til andre aktører, som typisk har evaluering som en aktivitet mellem hverdagens andre gøremål. En anden ressource – for en forskningsforankret evaluator – kan være den forskningsmæssige og metodiske kompetence, som evalueringsarbejdet tilføres (jf. de internationale læseundersøgelser).

For en evaluator forankret i en serviceleverende institution kan en anden ressource være muligheden for at formulere en vision på organisationens vegne, som indbefatter evalueringens betydning for hele organisationen. I det omfang, som denne vision indgår i organisationens strategi og ledelsesmæssige ”hovedstrøm”, kan det give en ganske magtfuld position. I nogle tilfælde formuleres dette mere eller mindre af en forvaltning i almindelighed (som i monitorering af ældreplejen). Hvis det evalueringsmæssige perspektiv mere eller mindre nedspilles, kan det imidlertid føre til forsinkelser og uskarphed i evalueringsbilledet (som daginstitutionscasen). Derfor står evaluator stærkest, når vedkommende er forankret i en særskilt rolle eller en særskilt institution.

Kapitel 8

Sprog og virkelighed i evaluering: Magtens hvordan

Plusmagt og minusmagt: Belysninger og mørklægning

Hvad bruges evalueringerne til at begrunde? Dette aspekt kan man kalde evalueringernes plusmagt. Der kastes lys over noget, og handlinger begrundes. Men casene viser også en ganske udpræget minusmagt: Magten til at sikre, at evalueringer mørklægger andre aspekter af sagen. Dette er ikke nødvendigvis fordækt. Ingen evaluering kan kaste lys over alt samtidig. Men reguleringen af grænsen mellem belysningen og mørklægningen er et vigtigt led i magtspillet.

Evalueringer bruges til at belyse en lang række procedurer på et hospital og begrunde øget fokus på sikre, forudsigelige procedurer. De bruges til at kaste lys over forældres vurdering af daginstitutioner. De bruges til at give luft for studerendes vurdering af undervisning på en videregående uddannelsesinstitution. De bruges til at holde styr på leveringen af visiterede ydelser til plejekrævende gamle. De bruges til at give vurderinger af uddannelser og fag fra forskellige vinkler – både selvevalueringer, arbejdsmarkedsundersøgelser og fagfolks vurderinger kommer ind. De bruges til at give internationale perspektiver på læseundervisningen i den danske folkeskole. Hver af disse evalueringer har efterfølgende en lang række virkninger, rækkende fra den daglige arbejdsplanlægning til diskussioner blandt fagfolk til symbolsk styrkelse af bestemte institutioner til en markant politisk beslutning om titelskifte, neddimensionering og ændring af institutionelt tilhørsforhold til en hel uddannelse, for ikke at nævne de store skolepolitiske implikationer af de internationale læseundersøgelser.

Samtidig med, at disse forhold belyses og handlinger begrundes, foregår der i hver enkelt case også en ganske betragtelig mørklægning af nogle forhold. Dette er ikke i sig selv opsigtsvækkende, snarere logisk indlysende, fordi ingen evaluering kan dække alt. Imidlertid er det værd at bemærke, hvilke mørklægninger, der finder sted.

I strekcodecasen mørklægges faktisk leverancen af de enkelte ydelser. Selv om systemer som dette strekcodesystem kritiseres for at overfokusere på kontrol af de enkelte ydelser helt ned til det enkelte minut, så sker en sådan registrering ikke i det pågældende system. Selv om systemet på mange måder er overvågningspræget, så giver det faktisk ikke mulighed for med sikkerhed at sige, at en bestemt ydelse er leveret på en given adresse.

Det giver kun mulighed for at sige med sikkerhed, at hjemmehjælperen har været der. Ligeledes er det mørklagt for den enkelte bruger, hvilken tid, der af systemet er sat af til pleje af vedkommende. Som led i systemets implementering er tillige sikret, at der ikke foretages sammenligninger mellem enkelte medarbejdere. Disse mørklægninger virker formodentlig tilsammen i retning af at sikre, at hele registreringsystemet ikke har en overmåde stor kontrol med hver enkelt hjemmehjælper brug af det enkelte minut.

I de internationale læseundersøgelser kølvand mørklægges sammenligninger af enkelte skoler og kommuner, i hvert fald for en tid. I en forskningspræget undersøgelse forsvares mørklægningen med hensynet til, at de enkelte skoler ikke skal igangsætte et kapløb om de bedste resultater, der vil underminere gyldigheden af undersøgelsen.

En direkte institutionel mørklægning har man i evalueringsinstituttets bestemmelser om ikke at udføre sammenligninger og fremlægge karaktergennemsnit for enkelte skoler. Denne mørklægning brydes til sidst via ny lovgivning af en ny regering.

I daginstitutionscasen sker formodentlig også nogle mørklægninger, men det er svært at identificere dem, fordi de er delegeret til hver enkelt institution. Som en konsekvens heraf mørklægges også samtidig en mulighed for at lave direkte sammenligninger mellem institutioner.

Disse iagttagelser tyder på, at magten til at mørklægge bestemte spørgsmål i et evalueringsarbejde kan være en ganske vigtig brik i magtspelet. At mørklægge et aspekt af evalueringsdata kan være en vej til at beskytte samarbejdspartnere eller institutioner, styrke sin forskningsmæssige troværdighed, undgå konkurrence og sikre sig selv en vis handlefrihed.

I flere tilfælde ofte opnås en mørklægning ved forhandling. Der er imidlertid ingen garanti for, at virkningen er stabil. Men det ser ud til, at mørklægning er et vigtigt slag i spillet om evalueringsforløbene.

Evalueringssprogets betydning: det antastelige og det uantastelige

Et andet vigtigt slag består i sprogets mulighed for at fastslå noget som uantasteligt eller antasteligt i forbindelse med evaluering. Som ramme om mange af evalueringsforløbene etableres en sprogbrug, hvor behovet for dokumentation fremstår som nærmest naturgivent og derfor uantasteligt.

En meget stor del af evalueringssproget lægger en overordnet fortolkningsmæssig ramme ned over en bestemt institution eller aktivitet.

I akkrediteringsforløbet bliver akkreditering set som en overordnet tolkningsmæssig ramme for mange andre af hospitalets aktiviteter (som der-

med underordnes denne ramme). I flere evalueringssammenhænge, ikke mindst evalueringsinstituttets, bliver begrebet kvalitet, eventuelt udmøntet i nogle underpunkter i et kommissorium, til en tilsvarende overordnet ramme, som efterfølgende kriterier og argumenter i sagen må forholde sig til.

I casene med ældrepleje bliver det ganske tydeligt, at prioriteringer og diskussioner af tværfaglig og prioriteringsmæssig art efterfølgende må forholde sig til de koder og kategorier, som monitoreringssystemet tilbyder. Ikke mindst forvaltningen har det fint med, at evalueringssystemet har denne funktion.

Visitorer bruger også de opstillede kategorier til at finjustere en ydelsestildeling med, også ved at indtænke hensyn, der ikke lige indgår i kategorierne.

I de mere subtile varianter af evalueringssproget bliver evalueringskriteriet som operationel indikator til selve begrebet om det, der er til diskussion. Idet der kan være forskellige opfattelser af, hvad god læsning er og skal være, eventuelt et mangedimensionelt begreb, kommer den måde læsesikkerhed og læsehastighed bliver målt på i særskilt grad til at udgøre selve bestemmelsen af, hvad der menes, når man taler om læsning. At også indikatoren faktisk er en konstruktion illustreres af, at de danske forskere vælger at sondre mere eksplicit mellem hastighed og sikkerhed end i hvert fald en enkelt af de udenlandske forskere. Tilsvarende diskuterer en enkelt sygeplejerske om hele behovs- og ydelseskategoriseringen bygger for meget på en enkelt diagnose, der er oppe i tiden. Men når først kriterier og kategorier er på plads, har de som nævnt en bemærkelsesværdig evne til at overtage, hvad en drøftelse af en indsats overhovedet handler om.

Forskellige kriterier i evaluering finder mere eller mindre genklang som noget uantasteligt. For eksempel i håndarbejdscasen fremstår den høje ledighedsgrad som et uantasteligt faktum ud fra et uantasteligt kriterium. Et enkelt læserbrev antaster imidlertid dette ved at hævde, at håndarbejds lærernes høje ledighed skyldes mange års nedskæringer på folkeoplysningsområdet. Pointen er altså, at ledigheden ikke er en objektiv realitet, men i stedet kontingent. Derfor er ledighedskriteriet antasteligt. Dette synspunkt ser imidlertid ikke ud til at vinde noget særligt indpas. Bekymringen for den faktiske ledighed (ikke den kontingente) har formodentlig både en tilstrækkelig politisk og mere almen opbakning til, at dette kriterium tages alvorligt i sig selv.

I akkrediteringscasen sker en interessant antastelse af evalueringens metode ved, at nogle medarbejdere giver anledning til en avisoverskrift om "fusk". Akkrediteringen forsvarer sig heroverfor på en meget interessant

måde: man laver en evaluering af akkrediteringen. Her viser det sig, at en vis procentdel af personalet udviser en vis tilfredshed med forskellige aspekter af akkrediteringen, hvorefter kritikken afvises. Når en evaluering antastes, bruger man en fornyet evaluering til at afvise antastelsen.

Evalueringsforløbene

De færreste evalueringsforløb sker uantastet. Det er forventeligt, at de må blive antastede, når de selv pr. definition antaster. I alle de nævnte evalueringsscases har der været indvendinger og diskussioner. Her ses lidt nærmere på de enkelte momenter i forløbene.

Initiativet til evaluering ligger i de nævnte cases hos politikere, i evalueringsinstitutioner, hos ledelser og forvaltninger.

I flere af dem har der været løbende forhandlinger om dagsordenfastsættelsen: hvad skal med, og hvad skal holdes ude. Det ser man ikke mindst i stregkodecasen, men også i daginstitutionscasen. Akkrediteringscasen er ligeledes præget af løbende forsøg på at påvirke evalueringens dagsorden. Dagsordenfastsættelsen i håndarbejdscasen søges genforhandlet meget sent i forløbet af de institutioner, som den rammer, dog uden held. Stregkodecasen er nok den bedste illustration af, at der er løbende og vedvarende forhandlinger om, hvad der skal ind og ud af evalueringssystemet. At systemet skal implementeres, forekommer dog at være et ufravigeligt krav. Dagsordenproblematikken er i denne case vedvarende og aldrig helt afgjort.

Vidensforvaltningen udgør ligeledes en vedvarende kontrovers i flere tilfælde. Der er åbenlys videnspolitisk strid om, hvilke institutioner, der skal belyses med hvilke data i kølvandet på de internationale læseundersøgelser. I forbindelse med vidensforvaltningen er tidligere peget på den store betydning, som mørklægning spiller under evalueringsforløb. Til gengæld for implementeringen af et evalueringsmæssigt styresystem aftales i flere tilfælde ganske distinkte mørklægningsforløb.

Designmomentet skaber måske mindre kontrovers end forventet. I mange tilfælde fremstår et design som forhåndsbestemt enten i en købt evalueringsteknologi eller i et institutionaliseret koncept for evaluering. I stedet ser det ud til at designspørgsmål omplantes som dagsordenspørgsmål, vidensforvaltningsspørgsmål, gyldiggørelsesspørgsmål og konsekvensdragelsspørgsmål.

Kun i mindre lokale ”skrædderier” af evaluering (undervisningsevalueringsscasen) samt i relation til forskning (de internationale læseundersøgelser) udgør designspørgsmål et vigtigt omdrejningspunkt.

Analyse og sammenfatning er i forskningsmæssig sammenhæng afgørende og tidskrævende (som f.eks. i de internationale læseundersøgelser). I evalueringsinstituttets koncept spiller evalueringsgruppen en stor rolle i analyse og sammenfatning, og det ser ud til, at den spilles bevidst. I hvert fald har man både diskussioner internt i gruppen og med evalueringsinstituttet om, hvor evalueringen skal "lande". En vigtig ressource er den sekretariatsmæssige arbejdskraft til faktisk at få noget skrevet. Derimod sker den i monitoreringssystemer automatiseret. Kontroverser herom bliver derfor ligesom designkontroverser omplantet til andre kontroverser.

I gyldiggørelsesfasen ser man i flere cases en del debat. Enkelte medarbejdere prøver at ugyldiggøre akkreditering ved at påpege "fusk". Hospitalet svarer igen med en ny evaluering. I undervisningscasen søger en adjunkt at ugyldiggøre studenterevalueringer med henvisning til, at det giver en forkert rollefordeling mellem elev og lærer og ved at slå på lærerens faglige rettigheder. Det vinder ikke opbakning. I håndarbejdscasen ses et interessant forsøg på at ugyldiggøre evalueringens dramatiske virkninger ved at pege på, at man selv har gjort et stort arbejde i selvevalueringen, og man havde forventet, at resultaterne ville blive brugt til udvikling. Dette forsøg på ugyldiggørelse har ikke den helt store virkning.

Konsekvensdragelsesfasen tiltrækker sig også særskilt opmærksomhed. Her er der ofte kontrovers, mest tydeligt i håndarbejdscasen og de internationale læseundersøgelser. Men der er også mere subtile konsekvensdragelser, ikke mindst i de vedvarende evalueringssystemer til monitorering af ældreplejen. Her er adfærdspåvirkningen og holdningspåvirkningen ganske svær at kortlægge eksakt, men systemerne virker, trods modificerende afværgningsforsøg, ganske indgribende i den enkelte medarbejders dagligdag og i medarbejdernes måde at diskutere og prioritere på. Der findes således en del konsekvensdragelser, som ikke påkalder sig nogen særlig opmærksomhed.

Karakteristisk for de nævnte momenter i evalueringsforløbene er, at de er meget spredt ud over de enkelte forløb. De lader sig ikke let pakke ned som afsluttede faser. Således er f.eks. stregkodecasen fortsat præget af dagsordenforhandlinger. En interessant udløber af evalueringsarbejdet her er et projekt om at arbejde mere med værdier på området. Evaluering kan føre til mange ting.

Blandt de mest produktive evalueringsforløb må regnes de internationale læseundersøgelser. I deres kølvand nærmest myldrer det med flere forskellige opfølgninger, både i forhold til læseundervisningen, lærernes uddannelse, efterfølgende undersøgelser, skolepolitik, og etablering af et helt

institut til evaluering, der, når det først er etableret, selv afstedkommer en række yderligere initiativer. Her ser man således igennem flere år en gradvis udfyldning af de forskellige mulige konstitutive virkninger både i sprog, beslutninger og nye institutioner.

Interessant er det også, at evalueringer i flere tilfælde bruges til at følge op på andre evalueringer med. Internationale undersøgelser følges op af evalueringer af den enkelte skole. En kritik af et akkrediteringsforløb følges op af en medarbejderevaluering af forløbet. En af evalueringens mest bemærkelsesværdige egenskaber er således dens evne til at skabe mere evaluering.

Flere af casene er således meget produktive i flere henseender, de skaber følgevirkninger af mange slags. Af og til byder de også på overraskelser. De akkrediteringsansvarlige blev formodentlig overraskede over, hvor mange ressourcer, en forberedelse til akkrediterings viste sig at kræve, ligesom påstandene om fusk nok ikke har været ventede. Håndarbejds lærere, der loyalt deltog i selvevaluering, har formodentlig fået sig et chok over håndarbejdevalueringens faktiske virkninger. Få har formodentlig kunnet forudse de internationale læseundersøgelses faktiske rækkevidde i dansk skolepolitik, selv om der givetvis har været politisk motivation for at tage kampe og få noget til at ske på området.

I flere tilfælde fører indvendinger og kontroverser til, at evalueringer justeres, dagsordener modificeres, anbefalinger blødgøres, og visse virkninger afbødes. Mange evalueringer har ikke lineære forløb.

Der skabes også i lange sags- og evalueringerforløb bemærkelsesværdige, nærmest paradoksale virkninger. Øgede retskrav og klagemuligheder på brugerside har ført til, at man i den kommunale verden har opfundet visitations- og registreringssystemer, der standardiserer og dokumenterer den enkeltes ydelse. Dette system kritiseres så for at være umyndiggørende og tayloristisk. Efterfølgende må man tilføje systemet lidt mere formel og specielt noget uformel elasticitet. At en af de vigtigste begrundelser for overhovedet at have monitoreringssystemer på ældreområdet er at kunne svare kvalificeret på klager fra brugere indgår ikke altid i forståelsen af sagens sammenhæng.

Evalueringssystemer kan være svar på problemer, der samtidig giver nye problemer. Mange evalueringer er ikke lineære. Deres forløb går ingen lige vej fra startpunkt til et veldefineret slutpunkt.

Kapitel 9

Konklusion: Evaluering og magt

Den almindelige tendens til at efterspørge mere dokumentation i og omkring den offentlige sektor synes at vinde flere sejre end nederlag. Flere tilfælde viser også, at det i dag er blevet almindeligt at foreslå evaluering uden nærmere begrundelse. Ofte sikres tilslutning til evaluering, uden at man har særlig klarhed på, hvad evalueringsarbejdet fører til.

Evalueringsbølgen har blandt andet ført med sig, at medarbejdere i offentlige organisationer i dag taler ganske åbent om, hvor vigtigt det er at få legitimitet og politisk kredit gennem evaluering. At evalueringer således har en åbenlys symbolsk funktion – inden resultaterne overhovedet behøver at blive brugt – er ikke længere noget, som ses som fordækt eller problematisk. Det er blot en blandt flere strategiske variable under givne omstændigheder. Allerede på dette punkt har evalueringsbølgen formodentlig indvarslet et bemærkelsesværdigt værdiskred. Ganske ofte behøver evaluering ikke begrundes nærmere. Offentligt ansatte taler påfaldende direkte om, hvor vigtigt evaluering er som symbolsk og strategisk aktivitet i sig selv.

Analysen har vist, at mange interessenter spiller ind på evaluering, herunder administratorer, politikere, fagfolk og medier. Brugere spiller en mere tilbages trukket rolle, end man måske skulle tro. Kun i få tilfælde spiller metodisk kompetence en nøglerolle. I det moderne videnssamfund er evalueringer sjældent snedkerarbejde, men kan rekvireres og bestilles og gøres til organisatoriske koncepter. Kompetence er derfor ikke en flaskehals, som ikke kan overvindes ved hjælp af ressourcer og organisation.

Men når de skal afvikles, er evalueringsforløb sammensatte og ofte genstand for kontrovers. Af og til sker der løbende forhandlinger og justeringer om deres indhold, retning og konsekvenser. Der er stor kompleksitet i mange evalueringsforløb og i deres vidtrækkende virkninger. Selv om disse virkninger kan være svære at udpege og afgrænse helt præcist, giver analysen en bedre forståelse af evaluering, end gængse forståelser ofte giver anledning til.

Evalueringsforskningen har traditionelt set været ganske bekymret om, at evalueringer ikke blev brugt særlig meget. Caseanalyserne viser, at ikke-brug måske kan forekomme, men brug finder sted, og talrige aspekter af evaluering bruges på mange måder. Der er ikke støtte til en antagelse om, at evaluering er rent symbolsk udenomsværk og ingen forskel gør. At evaluering ikke fungerer som den siger den gør, betyder ikke, at den slet ikke fun-

gerer. Der er således ingen modsætning imellem, at evaluering har en vigtig symbolsk funktion og at den har talrige indholdsmæssige, tidslige og sociale virkninger.

Der er ikke støtte til en – i øvrigt gængs – antagelse om, at evaluering er entydig magtudøvelse, der kun betjener en enkelt styringslogik eller en enkelt ideologi. Evalueringer byder på så mange kontroverser og så mange store og små forhandlinger og forskydninger i forløbene, at det er svært at tro på, at en enkelt dominerende magt skulle stå bag det hele. Hvis det var tilfældet, ville evalueringsforløbene formodentlig være mere lineære, end tilfældet er. Der ville være mindre uro i og mellem momenterne i et evalueringsforløb, end analysen faktisk har vist.

Modstand og indvendinger undervejs i et evalueringsforløb har forskellige konsekvenser. Nogle kampe er tabt, hvis de knyttes sammen med ufor- svarlige evalueringresultater eller modvind i tidsånden eller formel lovgivningsmæssig og forvaltningsmæssig modvind. Den formelle styringskæde mener det, når den mener det. Men andre indvendinger fører til forhandlinger og justeringer af dagsordener, ændret vidensforvaltning og modificerede konsekvensdragelser.

I flere evalueringsforløb har en væsentlig magt bestået i at mørklægge udvalgte data eller fordelinger af data undervejs i et evalueringsforløb. Denne magt bruges hyppigt. Selv om den kommer til udtryk negativt, som en forhindring eller forhaling, peger den på en væsentlig videnspolitisk problematik. Udbredelsen af evaluering og medier til formidling af evalueringresultater såsom computere, internettet og nyhedsmedier gør det til et uhyre vigtigt videnspolitisk spørgsmål, hvilke data om hvilke institutioner der skal offentliggøres i hvilken form.

Analysen har også peget på, at evalueringer blandt andet virker ved at tilbyde et sprog, som virker strukturerende på, hvordan man kan håndtere en institution eller en indsats. Dette sprog bliver fortolkningsramme på flere niveauer, både som administrationspolitik og organisatorisk strategi, som pejlemærker og kriterier for, hvad den gode indsats går ud på, og som pejlemærker for den enkelte medarbejders prioriteringer og tidsforbrug i det daglige. Denne konstitutive funktion af evaluering er væsentlig og rækker formodentlig videre og dybere end enkeltstående evalueringresultater. En væsentlig magt er således magten til at forme evalueringssproget. Denne iagttagelse viser, at hvis man alene efterspører evalueringers virkninger på formelle beslutninger, er fokus alt for snævert.

Forskere har vist, at i tilfældet med monitorering af ældreplejen, flyttes der ”definitoriske beføjelser fra sags- og socialdimensionen til tidsdimen-

sionen; forankringen af omsorgen bliver tidsmæssig og ikke substantiel” (La Cour & Højlund, 2001: 10). En sådan konstitutiv virkning er vigtig og ganske indgribende i forhold til hele begrebet om pleje. En tilsvarende konstitutiv virkning er fundet f.eks. i casen i Z-by, hvor plejepersonalet efter indførelsen af fælles sprog nu må argumentere ud fra monitoreringssystemets kriterier, når de drøfter tildeling af pleje til den enkelte ældre borger.

Men i casestudierne foregår samtidig fortsatte forhandlinger på forvaltnings- og medarbejderniveau om præcis, hvordan man skal gøre hvad op i tid og hvem der skal vide, hvordan hvad er gjort op. Tidsforbruget til den enkelte opgave og brugerens adgang til viden om det forventede tidsforbrug sløres. Virkningen heraf er formodentlig at tilføre de enkelte medarbejdere en smule frihedsgrader og fleksibilitet i det daglige, således at de konstitutive virkninger af styresystemet bliver modificerede og mindre entydige, selv om systemet fortsat har stærke konstitutive virkninger.

Magten til at forme evalueringssproget ligger ikke altid hos en enkelt aktør. Den er ofte afhængig af ”kulturel baggrundsstøj” i tidsånden. Evalueringer, der vil være nyttige, udfordrer institutioner og praksisser under hensyntagen til eller med henvisning til temaer og værdier, der er oppe i tiden – på jævnt dansk: ”hvor det bærer henad”. Evaluering kan i den forstand have en selvopfyldende funktion. Man viser noget er galt ved at fortælle, at det ikke lever op til det, som er på vej frem. Evaluering virker dermed som en accelerator for tidsånden. Men da tidsånden, ikke mindst fremtidens ditto, er et spørgsmål om fortolkning, består en væsentlig magt i magten til netop at pege på ”hvor det bærer hen”. Denne udpegning er af og til præget af, hvor andre forventer, det bærer hen. Et interessant fortolkningsspil kan hermed udspilles. Det gælder om at være hurtigst til at pege på, hvor andre senere godt kan se, at det bærer henad. Intet under, at dette spil af og til må rumme en del vildskud og uforudsigelige virkninger. I det lys udgør det ingen væsentlig trussel mod evalueringssbølgen, at evaluering af og til viser sig ikke at medføre de tilsigtede virkninger. Tværtimod skabes der i dag tilsyneladende blot mere efterspørgsel efter evaluering.

Intet under, at det efterfølgende kan være vanskeligt – trods evalueringssmentalitetens klare ide om ansvarsplacering – at bestemme med sikkerhed, hvem der har ansvaret for en evaluering og dens virkninger.

Bilag 1

En note om magtbegrebet

Når magten udfolder sig, knæsættes nogle principper, beslutninger eller betydninger som socialt gældende.

Blandt magtens spor tæller i denne bogs sammenhæng, hvordan evalueringer forløber, hvad de handler om, og hvordan de anvendes. Udtrykket ”magtens spor” er valgt, fordi man netop ikke kan se magten direkte; man kan kun se dens spor. Ved at betragte magtens spor slutter man tilbage til magten selv. Dette er en vanskelig proces teoretisk og begrebsligt.

Med til magtbegrebet hører alt det, som A kan trække på, når det lykkes for A at få B til at gøre noget, som B ellers ikke ville have gjort. Magtens spor sætter sig i B's adfærd. Endvidere eksisterer magten i dette perspektiv kun i og med, at B's ønsker er anderledes end A's. Der er en (potentiel) konflikt. A kan for eksempel iværksætte en evaluering af B, som B skal medvirke i, men ikke ønsker. I den henseende er magt nogle ressourcer, som er under A's kontrol.

Ganske ofte er disse ressourcer begrænsede til kun at være anvendelige på bestemte områder i bestemte situationer. Magten er situationel.

I de fleste tilfælde har B imidlertid også en vis mængde ressourcer at trække på, som er egnede til at reducere A's bestemmelse over evalueringen, herunder en eller anden grad af modstand, uvilje eller modmagt. Magten er således relativ.

Magten er også relationel, idet den måske ikke er en specifik egenskab eller ressource ved A, men snarere ”bor” i relationen mellem A og B. B spiller med i relationen for at få magten til at lykkes.

Magtmidlerne behøver ikke at blive vist eller omtalt for at virke. Alene B's forestillinger om A's magtudøvelse såvel som A's forestillinger om B's modmagt påvirker deres handlinger. Aktører, der ønsker at præge en evaluering, tager implicit eller eksplicit højde for andre aktørers indflydelsesmuligheder. Magten kan således udfolde sig gennem forestillinger. Disse forestillinger behøver ikke at blive konfronteret. De virker allerbedst, når de ikke bliver udfordret.

En vigtig magt ligger i overhovedet at udelukke en aktørs forestillinger om de handlemuligheder, der havde været til stede, hvis magten ikke havde fungeret. Når magten er avanceret, efterlader den ingen brutale spor. Her virker magten ikke nødvendigvis på adfærden, men allerede på forestillingerne om mulig adfærd. Magten virker så at sige ved at fratage bestemte

ønsker. Derved kan der også findes magt uden konflikt. Det er diskutabelt, om et magtbegreb på disse præmisser er interessant. Thomsen (2000: 54) argumenterer for, at konfliktkriteriet må med i magtbegrebet, idet ”frivillig underkastelse næppe kan siges at udgøre en synderlig interessant magtform”. Det er jeg ikke enig i. Når magten udøves i form af diskret konsensus – ved at fælles uantastede præmisser lægges til grund – er magten sublim og meget interessant, selv om dette magtbegreb byder på store vanskeligheder i en empirisk undersøgelse.

Vi kan her tale om en diskret magt, som ikke er en egenskab ved A i forhold til B, men er snarere indbygget i konstruktionen af en bestemt social orden. I den henseende kan man sige, at magten kan være strukturel og kulturel, og altså ikke længere blot en egenskab ved A eller B.

Når magten er strukturelt og kulturelt indbygget i en bestemt social orden – når magten er det, at noget gøres socialt gældende – bliver den nærmest allestedsnærværende. Det bliver svært at skille den ud, dels i rum, idet den ikke ”bor” noget bestemt sted hos nogen særlig aktør, dels i tid, fordi magten er med til at udfolde en social orden, som løbende fornyes og forandres.

I en sådan analyse bliver ”aktøren” ikke til oprindelsepunkt og ”ejer” af magten, men tværtimod selv i lige så høj grad et *resultat* af magten. En social orden – en diskursiv orden, om man vil – skaber som en af sine vigtigste effekter de subjektpositioner, der overhovedet kan siges noget og gøres noget ud fra. Som et eksempel kan det være en effekt af evaluering ikke (blot) at være instrument til at sikre brugerindflydelse, men også overhovedet at definere, hvad det vil sige at være ”bruger”. Tilsvarende gælder for andre grupper af aktører og deres ”interesser”.

Tænkes en sådan tanke igennem – at magten er, at noget gøres socialt gældende – kan man ikke identificere ”magten” i form af ”interesser”, ”ressourcer” osv., der ligger bag den producerede orden, og som ”forklarer” et givet socialt udfald (f.eks. et evalueringsresultat). Man kan blot beskrive, hvordan det går til, at noget gøres gældende. Man kan undersøge, hvordan det afgøres, hvad evalueringer skal handle om. Man kan undersøge, hvem der gør sig gældende (og gøres gældende) i et evalueringsforløb. Og man kan undersøge, hvordan det går til, at evalueringer anvendes.

Man vil altid i en given analyse kunne pege på mange andre betingelser ud over de netop udpegede, som skal til for at give et udtømmende billede af magten. Magtanalysen bliver altid partiel og perspektivisk. Nogle få forudsætninger for nogle få magtspor udvælges til nærmere undersøgelse,

og det bliver umuligt at foretage en fuldstændig entydig afkodning af disse spor.

Hvordan kan man så kende og udpege magten? I nogle magtforståelser opereres med en skelnen mellem legitimt og illegitimt, og kun den illegitime ses som værd at studere. Dette er ikke perspektivet i denne bog. Et absolut skel mellem legitimt og illegitimt kan ikke foretages på nogen indlysende, objektiv og stabil måde. Ofte handler en strid netop om, hvad der kan tages som givet som legitimt.

Fremgangsmåden er i stedet indledningsvist at skitsere nogle gængse forestillinger om magtens udfoldelse i evaluering og derefter i analysen at kontrastere sporene i de faktiske evalueringsforløb med disse gængse forestillinger. I det lys er det interessant, hvis evalueringer forløber anderledes end forventet. Det er også interessant, hvis der er strid om evalueringers tilrettelæggelse og udfald. Og det er interessant, hvis striden om evaluering handler om noget andet, end det man taler om, når evalueringer drøftes.

En sådan analyse må – indrømmet – blive partiel og perspektivisk, men forhåbentlig kan den udfordre og berige ellers gængse forestillinger om magt og evaluering.

Litteratur

- Albæk, Erik (1988). *Fra Sandhed til Information*. København: Akademisk Forlag.
- Albæk, Erik (2001). "Vidensinteresser og de mange betydninger af evaluering", pp. 21-31 i Peter Dahler-Larsen & Hanne Kathrine Krogstrup (red.). *Tendenser i evaluering*. Odense: Odense Universitetsforlag.
- Baier, Vicky E., James G. March & Harald Sætren (1986). "Implementation and Ambiguity". *Scandinavian Journal of Management Studies*, 2:197-212.
- Bauman, Zygmunt (1983). "Industrialism, Consumerism and Power". *Theory, Culture and Society*, 3:32-43.
- Bauman, Zygmunt (2002). *Fællesskab – en søgen efter tryghed i en usikker verden*. København: Hans Reitzels Forlag.
- Beck, Ulrich (1994). "The Reinvention of Politics: Towards a Theory of Reflexive Modernization", pp. 1-55 i Ulrich Beck, Anthony Giddens & Scott Lash (eds.). *Reflexive Modernization*. Stanford: Stanford University Press.
- Beck, Ulrich (1997). *Risikosamfundet – på vej mod en ny modernitet*. København: Hans Reitzels Forlag.
- Bennekou, Helle & Klaus Jørgensen (1998). *Håndarbejds lærernes beskæftigelses-situation*. København: Udviklingscentret for Folkeoplysning og Voksenundervisning.
- Castoriadis, Cornelius (1982). "The Crisis of Western Societies". *Telos*, 53:17-28.
- Christiansen, Peter Munk & Lise Togeby (2003). *På sporet af magten*. Århus: Aarhus Universitetsforlag.
- Christie, Christina A. & Marvin C. Alkin. (1999). "Further Reflections on Evaluation Misutilization". *Studies in Educational Evaluation*, 25:251-210.
- Dahler-Larsen, Peter (1998). *Den Rituelle Refleksion – om evaluering i organisationer*. Odense: Odense Universitetsforlag.
- Dahler-Larsen, Peter (2003a). "Hvad er selvevaluering", 29-40 i Peter Dahler-Larsen, Vibeke Normann Andersen, Kasper Møller Hansen & Carsten Strømbæk Pedersen (2003). *Selvevalueringens Hvide Sejl*. Odense: Syddansk Universitetsforlag.
- Dahler-Larsen, Peter (2003b). "Selvevalueringens vanskelige sejlads: Perspektiver fremover", 201-211 i Peter Dahler-Larsen, Vibeke Normann Andersen, Kasper Møller Hansen & Carsten Strømbæk Pedersen (2003). *Selvevalueringens Hvide Sejl*. Odense: Syddansk Universitetsforlag.
- Dahler-Larsen, Peter & Flemming Larsen (2001). "Anvendelse af evaluering – Historien om et begreb, der udvider sig", pp. 211-220 i Peter Dahler-Larsen & Hanne Kathrine Krogstrup (red.). *Tendenser i evaluering*. Odense: Odense Universitetsforlag.
- Danmarks Evalueringsinstitut (2001). *Historie med samfundskundskab i det almene gymnasium*. København: Danmarks Evalueringsinstitut.

- Danmarks Evalueringsinstitut (2002). *Selvevaluering i praksis*. København: Danmarks Evalueringsinstitut.
- Feldman, Martha S. & James G. March (1981). "Information in Organizations as Signal and Symbol". *Administrative Science Quarterly*, 26:171-184.
- Folkeskolen (diverse numre).
- Foss Hansen, Hanne (2001). "Evaluering: Teori og praksis i designprocessen". *Nordiske Organisasjonsstudier*, 3:43-62.
- Foucault, Michel (1980). "Power/Knowledge", i Colin Gordon (ed.). *Selected Interviews and Other Writings 1972-77 by Michel Foucault*. New York: Harvester.
- Gibbons, Michael, Camilla Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott & Martin Trow (1994). *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: Sage.
- Halford, Susan & Pauline Leonard (1999). "New Identities? Professionalism, Managerialism and the Construction of the Self", pp. 102-120 i Mark Exworthy & Susan Halford (eds.). *Professionals and the Managerialism in the Public Sector*. Buckingham: Open University Press.
- Hastrup, Kirsten (1993). "Self-Defining Worlds. The Anthropological Notion of Reality". *Danish Yearbook of Philosophy*, 28:61-79.
- Hellstern, G. M. (1986). "Assessing Evaluation Research", i Franz X. Kaufmann, Giandomenico Majone & Vincent Ostrom (eds.). *Guidance, Control and Evaluation in the Public Sector*. Berlin: de Gruyter.
- Henry, Gary T. (2000). "Why Not Use?" *New Directions For Evaluation*, 88:85-98.
- Horsdal, Marianne (1999). *Livets Fortællinger*. København: Borgen.
- Højlund, Holger & Anders La Cour (2001). "Standardiseret omsorg og lovbestemt fleksibilitet – organisationsændringer på et kærneområde i velfærdsstaten". *Nordiske Organisasjonsstudier*, 2, 3:391-117.
- Haarder, Bertel (1997). "Det vigtigste fag – en national tragedie", pp. 9-19 i *Modersmåls-Selskabets Årbog 1997*. København: C.A. Reitzels Forlag.
- Jasinsky, F.J. (1956). "Use and Misuse of Efficiency Controls". *Harvard Business Review*, XXXIV:105-112.
- Jørgensen, Torben Beck & Preben Melander (1999). *Livet i offentlige organisationer: Institutionsdrift i spændingsfeltet mellem stat, profession og marked*. København: Jurist- og Økonomforbundets Forlag.
- Kingdon, Jon (1984). *Agendas, Alternatives and Public Policies*. Boston: Little Brown & Company.
- Kirkhart, Karen (2000). "Reconceptualizing Evaluation Use: An Integrated Theory of Influence". *New Directions for Evaluation*, 88:5-23.
- La Cour, Anders & Holger Højlund. (2001). "Den fleksible hjemmehjælper". *Social Kritik*, 76:6-19.
- Lomborg, Bjørn (1998). "Metodologisk vurdering af borger/bruger-undersøgelser i Århus Amts kommuner 1995-96". *Nordisk Administrativ Tidsskrift*, 1:93-120.

- Longino, Helen E. (2002). *The Fate of Knowledge*. Princeton: Princeton University Press.
- Lovbekendtgørelse nr. 905. (2000).
- March, James G. & Johan P. Olsen (1976). *Ambiguity and Choice in Organizations*. Bergen-Oslo-Tromsø: Universitetsforlaget.
- March, James G. & Johan P. Olsen (1995). *Democratic Governance*. New York: The Free Press.
- Mark, Melvin M. & Gary T. Henry (2002). "The Multiple Methods and Pathways through which Evaluation can Influence Attitudes and Action". Paper præsenteret på European Evaluation Societys konference i Sevilla.
- Mejding, Jan (1994). *Den grimme ælling og svanerne?* København: Danmarks Pædagogiske Institut.
- Mejding, Jan (1995). "Den grimme ælling og svanerne? – En international undersøgelse". *Kvan*, 41:68-77.
- Meyer, John & Brian Rowan (1977). "Formal Structure as Myth and Ceremony". *American Journal of Sociology*, 83:340-363.
- Meyer, John W., John Boli & George M. Thomas (1994). "Ontology and Rationalization in the Western Cultural Account", pp. 9-27 i Richard Scott & John W. Meyer (eds.). *Institutional Environments and Organizations*. Thousand Oaks: Sage.
- Mylov, Peer (1997). "Den grimme danske ælling og de nordiske svaner". *Dansk Pædagogisk Tidsskrift*, 3:30-48.
- Patton, Michael Quinn (ed.) (1997). *Utilization-Focused Evaluation*. Thousand Oaks: Sage.
- Ridgway, V.F. (1956). "Dysfunctional Consequences of Performance Measurements". *Administrative Science Quarterly*, 2:240-247.
- Rieper, Olaf (red.) (1996). *Bør Danmark have en evalueringspolitik?* København: Institut for Statskundskab, Københavns Universitet.
- Røvik, Kjell Arne (1998). *Moderne Organisasjoner. Trender i organisasjonstenkingen ved tusenårsskiftet*. Bergen-Sandviken: Fagbokforlaget.
- Sahlin-Andersson, Kerstin (1996). "Imitating by Editing Success: The Construction of Organizational Fields", pp. 69-92 i Barbara Czarniawska & Guje Sevón (eds.). *Translating Organizational Change*. Berlin: de Gruyter.
- Schutz, Alfred (1975). *Hverdagslivets Sociologi*. København: Hans Reitzel.
- Schwandt, Thomas A. (2003). "Back to the Rough Ground! Beyond Theory to Practice in Evaluation". *Evaluation*, 3:353-364.
- Scott, W.R., S.M. Dornbusch, B.C. Busching & J.D. Laing (1967). "Organizational Evaluation and Authority". *Administrative Science Quarterly*, 12:93-117.
- Scriven, Michael (1991). "Beyond Formative and Summative Evaluation", pp. 19-64 i Milbrey W. McLaughlin & D.C. Phillips (ed.). *Evaluation and Education: At quarter century*. Chicago: University of Chicago Press.
- Shadish, William R., Thomas D. Cook & Laura C. Leviton (1991). *Foundations of Program Evaluation*. Newbury Park: Sage Publications.

- Simon, Herbert A. (1969). *The Sciences of the Artificial*. Massachusetts: Massachusetts Institute of Technology.
- Stake, Robert E. (1994). "Case Studies", pp. 435-454 i Norman Denzin & Yvonna Lincoln (eds.). *Handbook of Qualitative Research*. 2nd ed. Thousand Oaks: Sage.
- Stronach, Ian (1999). "Shouting Theatre in a Crowded Fire: 'Educational Effectiveness' as Cultural Performance". *Evaluation*, 2:173-193.
- Thomsen, Jens Peter Frølund (2000). *Magt og indflydelse*. Århus: Magtudredningen.
- Tompkins, Phillip (1987). "Translating Organizational Theory: Symbolism over Substance", pp. 70-96 i Fredrick M. Jablin et al. (eds.). *Handbook of Organizational Communication. An Interdisciplinary Perspective*. Newbury Park: Sage.
- Vedung, Evert (1998). *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur.
- Weiss, Carol H. (1977). "Research for Policy's Sake: The Enlightenment Function of Social Research". *Policy Analysis*, 35:31-45.
- Weiss, Carol H. (1983). "Evaluation Research in the Political Context", pp. 31-45 i Elmer L. Struering & Marilyn B. Brewer (eds.). *Handbook of Evaluation Research*. Beverly Hills: Sage.
- Weiss, Carol H. & Michael J. Bucuvalas (1980). *Social Science Research and Decision Making*. New York: Columbia University Press.
- Wirth, Wolfgang (1986). "Control in Public Administration: Plurality, Selectivity and Redundancy", pp. 569-624 i Franz X. Kaufmann, Giandomenico Majone & Vincent Ostrom (eds.). *Guidance, Control and Evaluation in the Public Sector*. Berlin: de Gruyter.
- Ziehe, Thomas (2003). "Læring i det kontrafaktiske rum". Oplæg på konference om læring på Bornholm, 10. september.

Om forfatteren

Peter Dahler-Larsen er professor ved Institut for Statskundskab, Syddansk Universitet i Odense. Forskningsmæssigt fokuserer han på evaluering som samfundsmæssigt og organisatorisk fænomen. Han er forfatter til *Den Rituelle Refleksion – om evaluering i organisationer* (Odense Universitetsforlag, 1998) samt en række andre publikationer inden for evaluering og kvalitativ metode.

Udgivelser fra Magtudredningen pr. 1. januar 2004

Bøger

Jørgen Goul Andersen, Peter Munk Christiansen, Torben Beck Jørgensen, Lise Togeby & Signild Vallgård (red.) (1999). *Den demokratiske udfordring*. København: Hans Reitzels Forlag.

Peter Munk Christiansen, Birgit Møller & Lise Togeby (2001). *Den danske elite*. København: Hans Reitzels Forlag.

Anette Borchorst (red.) (2002). *Kønsmagt under forandring*. København: Hans Reitzels Forlag.

Martin Marcussen (2002). *OECD og idespillet – Game Over?* København: Hans Reitzels Forlag.

Lise Togeby (2002). *Grønlandere i Danmark. En overset minoritet*. Århus: Aarhus Universitetsforlag.

Torben Beck Jørgensen & Kurt Klaudi Klausen (red.) (2002). *Territorial dynamik – streger på landkort, billeder i vore hoveder*. Århus: Aarhus Universitetsforlag.

Flemming Mikkelsen (red.) (2002). *Bevægelser i demokrati. Foreninger og kollektive aktioner i Danmark*. Århus: Aarhus Universitetsforlag.

Jens Blom-Hansen (2002). *Den fjerde statsmagt? Kommunernes Landsforening i dansk politik*. Århus: Aarhus Universitetsforlag.

Margaretha Järvinen, Jørgen Elm Larsen & Nils Mortensen (red.) (2002). *Det magtfulde møde mellem system og klient*. Århus: Aarhus Universitetsforlag.

Anker Brink Lund (2002). *Den redigerende magt – nyhedsinstitutionens politiske indflydelse*. Århus: Aarhus Universitetsforlag.

Finn Sivert Nielsen & Inger Sjørlev (red.) (2002). *Folkets repræsentanter. Et antropologisk blik på Folketinget*. Århus: Aarhus Universitetsforlag.

Thomas Pedersen (red.) (2002). *Europa for folket? EU og det danske demokrati*. Århus: Aarhus Universitetsforlag.

Peter Munk Christiansen & Asbjørn Sonne Nørgaard (2003). *Faste forhold – flygtige forbindelser. Stat og interesseorganisationer i Danmark i det 20. århundrede*. Århus: Aarhus Universitetsforlag.

Martin Marcussen & Karsten Ronit (red.) (2003). *Internationaliseringen af den offentlige forvaltning i Danmark – forandring og kontinuitet*. Århus: Aarhus Universitetsforlag.

Gorm Winther (red.) (2003). *Demokrati og magt i Grønland*. Århus: Aarhus Universitetsforlag.

Lise Togeby (2003). *Fra fremmedarbejdere til etniske minoriteter*. Århus: Aarhus Universitetsforlag.

Torben Beck Jørgensen (red.) (2003). *På sporet af en offentlig identitet – værdier i stat, amter og kommuner*. Århus: Aarhus Universitetsforlag.

Erik Damgaard (2003). *Folkets styre. Magt og ansvar i dansk politik*. Århus: Aarhus Universitetsforlag.

Hans Mouritzen (red.) (2003). *Er vi så forbeholdne? Danmark over for globaliseringen, EU og det nære*. Århus: Aarhus Universitetsforlag.

Hans Sode-Madsen (2003). *Farlig ungdom. Samfundet, ungdommen og ungdomskommissionen 1945-1970*. Århus: Aarhus Universitetsforlag.

Lars Bille & Jørgen Elklit (red.) (2003). *Partiernes medlemmer*. Århus: Aarhus Universitetsforlag.

Peter Munk Christiansen & Lise Togeby (red.) (2003). *På sporet af magten*. Århus: Aarhus Universitetsforlag.

Tim Knudsen (2003). *Offentlighed i det offentlige. Om historiens magt*. Århus: Aarhus Universitetsforlag.

Niels Nørgaard Kristensen (2003). *Billeder af magten. Portrætter til forståelse af magt og demokrati*. Århus: Aarhus Universitetsforlag.

Lise Togeby, Jørgen Goul Andersen, Peter Munk Christiansen, Torben Beck Jørgensen & Signild Vallgård (2003). *Magt og demokrati i Danmark. Hovedresultater fra Magtudredningen*. Århus: Aarhus University Press.

Jørgen Goul Andersen (2003). *Over-Danmark og Under-Danmark? Ulighed, velfærdsstat og politisk medborgerskab*. Århus: Aarhus University Press.

Signild Vallgård (2003). *Folkesundhed som politik. Danmark og Sverige fra 1930 til i dag*. Århus: Aarhus University Press.

Jørgen Grønnegård Christensen (2003). *Velfærdsstatens institutioner*. Århus: Aarhus Universitetsforlag.

Palle Svensson (2003). *Folkets røst. Demokrati og folkeafstemninger i Danmark og andre europæiske lande*. Århus: Aarhus Universitetsforlag.

Thomas Pallesen (2003). *Den vellykkede kommunalreform og decentraliseringen af den politiske magt i Danmark*. Århus: Aarhus Universitetsforlag.

Mogens Rüdiger (2003). *Statens synlige hånd. Om lovgivning, stat og individ i det 20. århundrede*. Århus: Aarhus Universitetsforlag.

Henrik Jensen (2003). *Europaudvalget – et udvalg i Folketinget*. Århus: Aarhus Universitetsforlag.

Skrifter

Erik Oddvar Eriksen (1999). *Is Democracy Possible Today?* Århus: Magtudredningen.

Ole Hammer & Inger Bruun (2000). *Etniske minoriteters indflydelseskanaler*. Århus: Magtudredningen.

- Jens Peter Frølund Thomsen (2000). *Magt og indflydelse*. Århus: Magtudredningen.
- Jørgen Elklit, Birgit Møller, Palle Svensson & Lise Tøgeby (2000). *Hvem stemmer – og hvem stemmer ikke?* Århus: Magtudredningen.
- Jacob Gaarde Madsen (2000). *Mediernes konstruktion af flygtninge- og indvandrerspørgsmålet*. Århus: Magtudredningen.
- Karsten Vrangbæk (2001). *Ingeniørarbejde, hundeslagsmål eller hovedløs høne? Ventetidsgarantier til sygehusbehandling*. Århus: Magtudredningen.
- Søren Laursen (2001). *Vold på dagsordenen. Medierne og den politiske proces*. Århus: Magtudredningen.
- Jørgen Goul Andersen & Mette Tobiasen (2001). *Politisk forbrug og politiske forbrugere. Globalisering og politik i hverdagslivet*. Århus: Magtudredningen.
- Erik Albæk, Peter Munk Christiansen & Lise Tøgeby (2002). *Eksperter i medierne. Dagspressens brug af forskere 1961-2001*. Århus: Magtudredningen.
- Helle Porsdam (2002). *Fra pax americana til lex americana? En diskussion af dansk retliggørelse som en påvirkning fra USA*. Århus: Magtudredningen.
- Eva Østergaard-Nielsen (2002). *Politik over grænser: Tyrkeres og kurderes engagement i det politiske liv i hjemlandet*. Århus: Magtudredningen.
- Jonathan Schwartz (red.) (2002). *Medborgerskabets mange stemmer*. Århus: Magtudredningen.
- Walter Korpi (2002). *Velfærdsstat og socialt medborgerskab. Danmark i et komparativt perspektiv, 1930-1995*. Århus: Magtudredningen.
- Steen Thomsen, Torben Pedersen & Jesper Strandskov (2002). *Ejerskab og indflydelse i dansk erhvervsliv*. Århus: Magtudredningen.
- Frank Rasmussen & Peder Andersen (2002). *Globaliseringens økonomiske konsekvenser for Danmark*. Århus: Magtudredningen.
- Carsten Greve (2002). *Privatisering, regulering og demokrati. Telestyrelsens funktion som uafhængig reguleringsmyndighed*. Århus: Magtudredningen.
- Ann-Dorte Christensen (2003). *Fortællinger om identitet og magt. Unge kvinder i senmoderiteten*. Århus: Magtudredningen.
- Thomas Schøtt (2003). *Den økonomiske elites netværk*. Århus: Magtudredningen.
- Peter Dahler-Larsen & Niels Ejersbo (2003). *Djøficering – myte eller realitet?* Århus: Magtudredningen.
- Jan H. Hermansen, Lars Bille, Roger Buch, Jørgen Elklit, Bernhard Hansen, Hans Jørgen Nielsen & Karina Pedersen (2003). *Undersøgelsen af medlemmerne af de danske partiorganisationer. Dokumentation*. Århus: Magtudredningen.
- Lars Torpe & Torben K. Kjeldgaard (2003). *Foreningssamfundets sociale kapital. Danske foreninger i et europæisk perspektiv*. Århus: Magtudredningen.
- Jens Blom-Hansen (2003). *Subsidiaritetsprincippet vendt på hovedet? EU's strukturpolitik og Danmark*. Århus: Magtudredningen.

- Jens Peter Christensen (2003). *Domstolene – den tredje statsmagt*. Århus: Magtudredningen.
- Camilla Palmhøj Nielsen (2003). *Til glæde for hvem? – om intern regulering i staten*. Århus: Magtudredningen.
- Peter Munk Christiansen & Asbjørn Sonne Nørgaard (2003). *De som meget har ... Store danske virksomheder som politiske aktører*. Århus: Magtudredningen.
- Lise Togeby, Jørgen Goul Andersen, Peter Munk Christiansen, Torben Beck Jørgensen & Signild Vallgård (2003). *Demokratiske udfordringer. Kort udgave af Magtudredningens hovedresultater*. Århus: Magtudredningen.
- Lise Togeby, Jørgen Goul Andersen, Peter Munk Christiansen, Torben Beck Jørgensen & Signild Vallgård (2003). *Power and Democracy in Denmark. Conclusions*. Århus: Magtudredningen.
- Birte Siim (2003). *Medborgerskabets udfordringer – etniske minoritetskvindes politiske myndiggørelse*. Århus: Magtudredningen.
- Anette Borchorst (2003). *Køn, magt og beslutninger. Politiske forhandlinger om barselsorlov 1901-2002*. Århus: Magtudredningen.
- Peter Dahler-Larsen (2003). *Evaluering og magt*. Århus: Magtudredningen.

www.magtudredningen.dk