



# 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](mailto:Magtudredningen@ps.au.dk)

[www.magtudredningen.dk](http://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 evalueringsvirksomhed (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*

	<b>et instrument</b>	<b>en arena</b>
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
Evalueringens 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 reproducere 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 kørernes 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.*

	<b>Monologisk</b>	<b>Pluralistisk</b>
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.

## **Evalueringssideologi**

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 evalueringssbølge præger dem, der i vor tids evalueringssbølge 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. Evalueringssbølgens 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 evalueringens anvendelse. 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 evalueringarbejde 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.<sup>1</sup>

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 spekulationen 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.<sup>2</sup>

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. Deltagerens 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 behandelende 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 evalueringssudfordringen 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 evalueringseresultaterne, 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 evalueringarbejdet. 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-









































































































































