

Håndbog i evaluering
Metoder til at dokumentere og vurdere proces og
effekt af offentlige indsatser

Redaktion
Olaf Rieper

AKF Forlaget
Juni 2004



Foto: Stine Christensen

Udsnit af bogens forfattere til seminar på Magleås Kursuscenter den 19. februar 2004.

Set fra venstre:

Trine Bille, Anders Larsen, Mikael Togeby, Leena Eskelinen, Olaf Rieper, Michael Rosholm, Eigil Boll Hansen, Leif Olsen, Emil Husted Erichsen, Lene Holm Pedersen, Kjeld Høgsbro.

Forord

Baggrunden for den foreliggende håndbog i evalueringsmetoder er et ønske om at forny AKF's tidligere håndbog på området fra 1993. Der er siden da sket ganske meget på evalueringsområdet i Danmark og internationalt. Herhjemme har vi fået en faglig evalueringsforening (Dansk Evalueringsselskab), og mange offentligt støttede evalueringsprogrammer og -projekter har set dagens lys – på statsligt niveau og på kommunalt/amtsligt niveau. Evaluering er næsten blevet et mantra og en industri med et eget marked for udbud og efterspørgsel og med mange forskellige købere (bestillere) og sælgere (evaluatorer).

Selv om der efter år 2000 er udgivet en række bøger om evaluering på dansk, er der en påfaldende mangel på en allround metodebog på evalueringsområdet. Denne håndbog sigter mod at udfylde dette behov og vil forhåbentlig kunne bruges af den nytilkomne såvel som af den mere erfarne i evalueringsfeltet.

Bogen er fortrinsvis skrevet af nuværende og tidligere medarbejder i AKF. Jeg vil gerne benytte lejligheden til at takke mine venner og kolleger for deres indsats. Kapitlerne præsenterer en bredde over sektorer og en variation i metoder, som findes i få andre værker.

Bogen er finansieret over AKF's grundbevilling.

Hvis læserne i denne håndbog finder inspiration og grundlag for at foretage (endnu) bedre evalueringer, er målet nået.

God fornøjelse!

Olaf Rieper
Juni 2004

Indhold

Indledning	9
Temadelen	13
1 Evalueringsbegrebet, modeller og paradigmer	
<i>Leif Olsen og Olaf Rieper</i>	15
1.1 Evalueringsbegrebet – et rummeligt begreb og en semantisk magnet .	15
1.2 Evalueringsmodeller.....	18
1.3 Paradigmer i evaluering	24
2 Evaluering i praksis – hvilke valg skal foretages?	
<i>Olaf Rieper</i>	34
2.1 Oversigt over spørgsmål og faseinddelingen.....	34
2.2 Præsentation af de enkelte faser	35
2.3 God evalueringsskik.....	49
3 Intern eller ekstern evaluering	
<i>Olaf Rieper</i>	51
3.1 Et kontinuum mellem ekstern og intern evaluering	51
3.2 Fordele og ulemper ved henholdsvis intern- og ekstern evaluering ...	53
3.3 Hvornår vælge hvilken slags evaluering?.....	55
4 Effektevaluering	
<i>Mikael Togeby</i>	57
4.1 Indledning.....	57

4.2	Det kontrollerede eksperiment.....	58
4.3	Brug af sammenligningsgruppe.....	59
4.4	Statistiske analyser af naturlige eksperimenter.....	61
4.5	Effektevalueringer for få midler.....	63
4.6	Effektevaluering med fokus på teori.....	65
5	Procesevaluering	
	<i>Kjeld Høgsbro</i>	66
5.1	Hvorfor vælge procesevaluering?.....	66
5.2	Forskellige former for procesevaluering.....	68
5.3	Videninteresser og valg af evalueringsform.....	75
5.4	Evaluators rolle.....	78
	Sektordelen	81
6	Vejleder til sektordelen	83
7	Evaluerings på skoleområdet	
	<i>Eskil Heinesen, Torben Pilegaard Jensen og Jill Mehlbye</i>	95
7.1	Særlige udfordringer ved evaluering af effekter af skoleressourcer ...	97
7.2	Eksempel: Effekter på elevernes uddannelsesforløb af forskelle i skoleudgifter pr. elev.....	100
7.3	Eksempel: PISA-Programme for International Student Assessment ..	106
7.4	Eksempel: Kommunernes vejledning af unge om uddannelse og erhverv.....	112
8	Evaluerings på børn og unge-området	
	<i>Jill Mehlbye</i>	119
8.1	Temaer inden for børn og unge-området.....	119
8.2	Særlige problemstillinger omkring evalueringer.....	121
8.3	Evalueringemetoder med henblik på at få en vurdering af udbyttet og effekten af indsatsen.....	122
8.4	Eksempel: Evaluering af fritidstilbuddene til børn og unge i en kommune.....	125
8.5	Eksempel: Evaluering af børns udvikling og trivsel i en kommune ...	128

8.6	Eksempel: En kommunes indsats over for særlig socialt udsatte børn og unge	132
8.7	Afrunding	136
9	Evaluering af social- og sundhedsmæssig indsats over for ældre	
	<i>Eigil Boll Hansen</i>	138
9.1	Særlige udfordringer	139
9.2	Eksempel: Akutstue i Rødning Kommune.....	142
9.3	Eksempel: Hjemmehjælpernes arbejdsprincipper og ældres velbefindende	146
10	Evaluering af integration af etniske minoriteter	
	<i>Jill Mehlbye og Eskil Heinesen</i>	154
10.1	Evalueringer på integrationsområdet	154
10.2	Særlige udfordringer ved evalueringer på integrationsområdet	155
10.3	Særlige udfordringer vedr. effektforskning	156
10.4	Eksempel: Evaluering af kommunernes integrationsindsats – en registerundersøgelse.....	157
10.5	Eksempel: Unges overgang fra grundskole til ungdomsuddannelse – en kvalitativ og kvantitativ evaluering.....	163
10.6	Eksempel: Integration af tosprogede børn og unge i daginstitution og skole – en kvalitativ og kvantitativ evaluering.....	167
11	Evaluering på handicapområdet	
	<i>Leif Olsen</i>	173
11.1	Er evaluering på handicapområdet noget særligt?	173
11.2	To eksempler: En formativ evaluering og en brugerevaluering	176
11.3	Eksempel: Rehabilitering af mennesker med traumatiske hjerneskader	177
11.4	Eksempel: Evaluering af amternes rådgivning på handicapområdet ..	183
11.5	Afrunding	191
12	Evaluering af socialt arbejde for udsatte grupper	
	<i>Kjeld Høgsbro</i>	192
12.1	Hvad kendetegner de udsatte grupper?	193
12.2	Målene for den professionelle indsats	195

12.3	Hvad er det vi kan evaluere?	196
12.4	Valget mellem forskellige evalueringsformer.....	202
12.5	Eksempel: Undersøgelser af det samlede tilbud for hjemløse, misbrugere og sindslidende.....	205
13	Evaluering af beskæftigelsespolitikken	
	<i>Torben Pilegaard Jensen og Peter Plougmann</i>	210
13.1	Introduktion og specifikke forhold	210
13.2	Eksempel 1: Anvendelsen af andre aktører i beskæftigelsespolitikken	214
13.3	Eksempel 2: En evaluering af »Lokale koordinationsudvalg«	219
14	Evaluering af aktivering ved hjælp af kvantitative metoder	
	<i>Michael Rosholm</i>	224
14.1	Indledning	224
14.2	Eksempel 1: Et simpelt mål på effekt af aktivering	227
14.3	Eksempel 2: Et kompleks mål for aktivering	231
14.4	Formidling og anvendelse	239
15	Evaluering på arbejdsmiljøområdet	
	<i>Leena Eskelinen</i>	240
15.1	Indledning	240
15.2	Eksempel 1: Udviklingstendenser i arbejdsmiljøet (på syv visionsområder)	243
15.3	Eksempel 2: Kvalitetsudvikling i forflytning	247
15.4	Afrunding	251
16	Evaluering inden for energisektoren	
	<i>Anders Larsen og Lene Holm Pedersen</i>	252
16.1	Indledning	252
16.2	Eksempel 1: Frivillige aftaler.....	255
16.3	Eksempel 2: Evaluering af vindkraftudbygningen	260
16.4	Afrunding	266

17 Cost-benefit-analyser på miljøområdet	
<i>Emil Husted Erichsen og Camilla Damgaard</i>	268
17.1 Hvad er en cost-benefit-analyse?	268
17.2 Eksempel 1: Cost-benefit-analyse af et skovrejsningsprojekt ved Odense	273
17.3 Eksempel 2: Organisk affald	279
18 Evalueringer på kulturområdet	
<i>Trine Bille</i>	288
18.1 Særlige udfordringer for evalueringer på kulturområdet	289
18.2 Evalueringsmodeller på kulturområdet	290
18.3 Economic impact-analyse af Brandts Klædefabrik i Odense	396
18.4 Værdisætning af Det Kgl. Teater	300
Forfatterne	306
Bilag: Vejledninger i evalueringer på nettet	311
Litteratur	315
Emneregister	334
Noter	345

Indledning

Målgruppe

Denne håndbog er beregnet på læsere, der står over for at skulle gennemføre en evaluering, eller som skal udarbejde en bestilling til en evaluering. Bogen er tænkt som en hjælp til ansatte i stat, amter og kommuner, men ikke kun. Også personer i private virksomheder og frivillige organisationer, der som udfører eller bestiller står over for en evalueringssopgave, vil kunne hente inspiration her. Bogen henvender sig til mest til læsere, som har praktiske erfaringer med eller teoretisk baggrund i evalueringsmetoder. Men også nye i evalueringfeltet vil kunne have glæde af håndbogen, herunder studerende.

Bogens opbygning

Bogen består af to dele, en temadel og en sektordel. Temadelen består af kapitler, som tager evalueringstemaer op på tværs af sektorer. Efter et begrebsafklarende kapitel følger et kapitel om, hvordan man griber evaluering an i praksis, hvilke valg man står overfor, og hvilke faser en evalueringssproces normalt gennemgår. Herefter følger en omtale af intern og ekstern evaluering og deres respektive fordele og ulemper. Temadelen afsluttes med to kapitler om to meget udbredte evalueringssmodeller, henholdsvis effektevaluering og procesevaluering.

Temadelen giver således en oversigt over forskellige slags evalueringer og viden om, hvordan man griber evaluering an i praksis.

Den efterfølgende sektordel indeholder en række kapitler om evalueringer på fx arbejdsmarkedsområdet, skoler, udsatte grupper mv. Hvert af disse sektorkapitler giver konkrete eksempler på, hvordan evalueringer på

de pågældende sektorområder er blevet foretaget, og hvilke generelle udfordringer der er til evalueringer på det pågældende område. Der gives eksempler på evalueringer, der kræver mere beskedne ressourcer, og evalueringer, der kræver større ressourcer. Sektordelen er den mest omfattende, og eksemplerne heri er bogens rygrad. Tankegangen er, at vi lærer gennem eksempler, og at vi i høj grad er i stand til at overføre andres erfaringer til egen praksis.

Udvalget af sektorer og evalueringseksempler afspejler de områder, hvor AKF de senere år har gennemført evalueringer. Det betyder, at sektorområder som fx sygdomsbehandling, udviklingsbistand og de videregående uddannelser ikke er medtaget.

I starten af sektordelen er indsat et vejledningsafsnit til sektorkapitlerne, som læseren kan bruge som »nøgle«, idet den muliggør forskellige indgange til bogen.

Bogens ambition

Ambitionen er at give læseren kendskab til en vifte af evalueringsmetoder og -modeller som inspiration til selv at vælge evalueringsdesign. Men det er ikke bogens ambition at give vejledning i, hvordan dataindsamling og -analyse foretages i praksis. Her er henvist til speciallitteraturen, idet en sådan ambition ville sprænge alle praktiske rammer.

Forskellige måder at anvende bogen på

Læseren kan bruge bogen på forskellige måder:

- Der kan tages udgangspunkt i, at man ønsker at gennemføre fx en effektevaluering. Man vil typisk starte med at læse kapitlet om effektevalueringer i bogens temadel, for derefter at se i »Vejleder til sektordelen«, hvilke eksempler der er på netop effektevaluering på forskellige sektorområder.
- Der kan tages udgangspunkt i, at man ønsker at gennemføre en evaluering, men ikke ved, hvilke særlige hensyn der gør sig gældende på det pågældende sektorområde. Her vil det relevante sektorkapitel tilbyde oplysninger om udfordringer og problematikker for evaluering på dette sektorområde og pege på evalueringsvejledninger specifikt for området.

- Der kan tages udgangspunkt i, at man ønsker at kunne vælge mellem flere forskellige evalueringsdesign og -metoder, men savner oversigt over relevante muligheder. Her vil »Vejleder til sektordelen« give en oversigt over de forskellige fremgangsmåder, der er anvendt i eksemplerne, som man kan lade sig inspirere af.

Bogens fokus og afgræsning

Bogens fokus er som nævnt præsentationen af metoder, der faktisk er blevet anvendt til evalueringer på forskellige sektorområder. Hertil kommer som nævnt udvalgte temakapitler om begreber og praktisk fremgangsmåde.

Bogen indeholder derimod ikke sociologisk/politologiske perspektiver på fænomenet evaluering, eller forklaringer på evalueringsaktiviteternes udvikling. Derimod sigter bogen mod at give læseren en forståelse af den variationsrigdom i metoder og fremgangsmåder, som evalueringsfeltet rummer, og som eksemplerne illustrerer. De specifikke metoder er belyst gennem eksemplerne og dermed sat i en evalueringssammenhæng.

Temadelen

1 **Evalueringensbegrebet, modeller og paradigmer**

Leif Olsen og Olaf Rieper

Dette kapitel redegør først for indholdet af begrebet »evaluering«, som det bruges i denne håndbog. Herefter følger en præsentation af en bred vifte af forskellige modeller for evalueringer. Kapitlet afsluttes med en introduktion til en række grundlæggende anskuelser (paradigmer), der bevidst eller ubevidst præger evalueringspraksis, både når det gælder opstart, gennemførelse og formidling af evalueringernes resultater.

1.1 **Evalueringensbegrebet – et rummeligt begreb og en semantisk magnet**

Evalueringensbegrebet er, som andre levedygtige begreber, særdeles rummeligt, hvilket giver anledning til diskussioner om generelle afgrænsninger og afgrænsninger i forhold til konkrete evalueringer. Det er således væsentligt, når man står foran at skulle iværksætte evalueringer at gøre sig overvejelser over og formulerer, hvad der afgrænser evalueringensbegrebet i den konkrete situation. I evalueringenslitteraturen har der især været kontroverser om evalueringensbegrebets afgrænsning i forhold til monitorering og i forhold til (samfunds)videnskabelig forskning. Monitorering er en speciel form for evaluering, hvor der foretages regelmæssige registreringer af den givne indsats (fx støtte til indførelse af ny teknologi, forsøg med handleplaner), så der fremkommer målinger af såvel implementering som output (se nedenfor). Formålet med monitorering er at kontrollere indsatsen undervejs i forløbet og eventuelt foretage justeringer på baggrund af de gennemførte målinger. Vurderingerne af målingernes resultater og fortolkninger heraf

ligger imidlertid ofte uden for selve opgaven med at monitorerer, hvorfor det kan diskuteres, om der egentligt er tale om evaluering, der har til formål at vurdere givne indsatser (jf. definitionen nedenfor).

Evalueringer kan også være forskning, når de udføres på forskningsmæssige præmisser, hvilket blandt andet indebærer inddragelse af teorier, brug af forskningsbaserede metoder og bemanding med forskere, der har deres forskningsmæssige frihed til at gennemføre evalueringen og formidle resultaterne. Forskningsbaserede evalueringer kan i nogle tilfælde nærme sig grundforskning og i mindre grad foretage konkrete vurderinger af givne indsatser, hvilket giver anledning til diskussion af grænsen mellem forskning og evaluering. I praksis er der flydende grænser mellem evalueringer som monitorering (udredning og konsulentvirksomhed) og forskning. Mange evalueringer gennemføres fx i administrative miljøer og af konsulenter, hvor der også kræves, at evalueringerne lever op til forskningsmæssige krav til metodeanvendelse og dokumentation.

Definitionen af evaluering i denne håndbog er:

Evaluering er systematisk indsamling af informationer og data, der giver grundlag for at skabe viden om gennemførelse, organisering og virkninger af konkrete indsatser, der er sat i værk for at påvirke forhold i samfundet. Evalueringer gennemføres med henblik på praktisk anvendelse, hvilket kan ske gennem formidling af resultater undervejs eller efter at evalueringen er gennemført.

For at præcisere og afgrænse denne definition af evaluering gennemgår vi kort definitionens nøglebegreber:

Indsatser: Udgangspunktet for enhver evaluering er et konkret fænomen, som af bestemte grunde og af bestemte parter ønskes vurderet. Det er vigtigt at klarlægge, hvilket fænomen der er tale om, idet det har betydning for alle øvrige trin i gennemførelsen af evalueringer. Indsatser kan være aktiviteter i både offentligt og privat regi, fx offentlige serviceydelser, regulering af virksomheder, informationskampagner, projekter, programmer, lovgivning og politikområder.

Skabe viden om gennemførelse og virkninger: Evalueringer har til formål at foretage en vurdering af den evaluerede indsats. Vurderingerne kan foretages på baggrund af udvalgte kriterier, der belyses og analyseres gennem informationer og data, der er indsamlet til formålet eller allerede

findes fx som administrative eller registerbaserede data. Vurderingskriterier kan fx udledes af indsatsens mål, interessenters opfattelse eller de kan være baseret på fagligt begrundet indsigt, fx om sindslidendes behov. I nogle tilfælde fremlægger evaluators imidlertid sine resultater uden vurdering, som herved overlades til evalueringens brugere, fx beslutningstagere.

Gennemførelse: Her er tale om den proces, hvorigennem indsatsen fx en service leveres, eller et program gennemføres. Gennemførelsen kan være interessant i sig selv, men kan også være en forudsætning for de resultater, der opnås gennem indsatsen og resulterer i output, fx »antal sager afsluttet« eller »antal kursustimer afholdt«.

Virksomheder eller effekter: Her er tale om indsatsens umiddelbare virkninger, som fx læsefærdighed, der er resultat af indsatsen »danskundervisning«, og indsatsens virkninger på længere sigt, fx kvalifikationer på arbejdsmarkedet (kaldet henholdsvis outcome og impact i den engelsksprogede litteratur). På dansk anvendes bl.a. begreberne præstationer og resultater, men der er ingen klar sprogbrug på området.

Systematisk indsamling af informationer og data: Grundlaget for at kunne vurdere virkninger og effekter af indsatser er, at evaluator tilvejebringer de relevante informationer og data, som evalueringens analyser kan bygge på, og viden kan udvikles ud fra. Der stilles således krav om, at evaluator arbejder systematisk og mestrer de relevante metoder til at fremskaffe informationer og data. Herudover skal evaluator være i stand til at vurdere eller sørge for at fremskaffe faglige vurderinger af, hvor gyldige og robuste de indsamlede informationer og data er.

Med henblik på praktisk anvendelse: Evalueringer iværksættes på grund af et ønske om viden blandt en eller flere parter med interesse i den pågældende indsats, samt økonomiske midler til at iværksætte evalueringer fx et ministerium, et amt, en kommune, institutioner, interesseorganisationer mv. Der kan også være tale om, at evalueringer iværksættes på baggrund af et påbud om, at indsatsen skal evalueres fx som betingelse for at opnå støtte til den pågældende indsats, hvilket er ret udbredt. Formålene med den praktiske anvendelse kan placeres i et kontinuum mellem udvikling og læring på den ene side og kontrol og sanktioner på den anden side. Hvis en evaluering fx viser, at en given indsats er god, så kan den praktiske anvendelse bestå i, at det beslutes, at der ikke skal foretages ændringer.

Men ellers kan ændringer spænde fra små justeringer, til radikal omlægning eller ligefrem nedlæggelse af en indsats. Det ligger dog sjældent i evaluators magt at sikre anvendelse af evalueringer, idet mange andre faktorer påvirker anvendelsen af evalueringer. Der kan fx skelnes mellem instrumentel, oplysende, legitimerende og taktisk anvendelse af evalueringer, se nærmere herom i det følgende kapitel.

Definitionen af evaluering stiller betydelige krav til evaluator, men ikke krav til, hvor og af hvem evalueringer skal og kan gennemføres. Evalueringer kan såvel udføres i forskningsmiljøer, konsulentmiljøer som i andre miljøer, hvor evaluatorene (interne som eksterne) har de nødvendige rammer og kompetencer til at leve op til de faglige krav, som gennemførelse af evalueringer stiller. Evalueringer kræver under alle omstændigheder, at evaluator kan leve op til de almindeligt anerkendte faglige krav til dataindsamling, analyse, dokumentation og formidling.

Spørgsmålet om, hvordan man kan og bør evaluere konkrete indsatser, må nødvendigvis besvares konkret, idet evalueringer må »skræddersyes« for at være relevante for de parter, der ønsker vurderinger af givne indsatser. I det følgende vil vi se på forskellige modeller, der kan give inspiration og danne ramme for diskussionen af, hvordan man kan »skræddersy« konkrete evalueringer, hvilket kan være et broderi over en model eller et patchwork, der integrerer elementer fra flere forskellige modeller.

1.2 **Evalueringsmodeller**

Betegnelsen evalueringsmodel bruges her som et udtryk for forskellige typer af evalueringer, der har en række lighedspunkter og adskiller sig fra andre typer af evalueringer. Modellerne er en måde at skabe en vis orden på i en ellers uoverskuelig mængde af forskellige fremgangsmåder på evalueringsområdet. Faglitteraturen indeholder mange forslag til typologier over evalueringsmodeller. Dahler Larsen og Krogstrup 2003 beskriver tre modeller. Vi har valgt at præsentere en mere udfoldet typologi, der indeholder ti evalueringsmodeller, hvoraf nogle kan deles op i yderligere en række modeller. Den valgte typologi, der er inspireret af Vedung 1997 og Foss Hansen 2003, illustrerer den meget brede vifte af modeller, der benyttes til evalueringer.

Skema Oversigt over evalueringsmodeller – hovedspørgsmål, styrker og svagheder

1.1 svagheder

Evalueringsmodel (beslægtede betegnelser)	Hovedspørgsmål	Styrker	Svagheder
Målopfyldelses-evaluering (resultatevaluering)	I hvilken grad er indsatsens mål opfyldt? Har indsatsen bidraget til målopfyldelsen?	Simpel logik. Demokratisk legitimitet Evaluators neutral, værdier er givet med målet.	Oftest uklare og flere mål for en indsats. Der ses bort fra utilsigtede konsekvenser, skjulte dagsordener og gennemførelsen af indsatsen.
Effektevaluering (målfri evaluering, sideeffektevaluering)	Hvilke effekter af indsatsen er opnået og så uden for indsatsens målområde?	Fokus på sideeffekter, som ofte forekommer.	Der ses bort fra gennemførelsen af indsatsen.
Procesevaluering (evaluering af processen fra indsatsens start til modtagelse hos målgruppe)	Hvordan gennemføres indsatsen? Hvordan forklares mangelfuld proces?	Identifikation af elementer i indsatsen som hæmmer henholdsvis fremmer modtagelse hos målgruppen og muliggør forbedring af indsatsen. Potentiale for teori-baseret evaluering, som muliggør teori-udvikling.	Effekter er ikke omfattet.
Formativ evaluering (udviklingsorienteret evaluering) Evaluering foregår sideløbende med, at indsatsen udvikles og i interaktion mellem indsats og evaluator.	Som under proces-evaluering, samt: Hvordan forbedres indsatsen sideløbende med evalueringen?	Indsatsen får støtte til gennemførelse via evaluator og begge kan opnå læring.	Evaluators rolle er vanskelig grundet krydspres fra indsatsen og fra rekvirent.
Systemmodeller (sammensat evaluering)	Er indsatsens faktiske forudsætninger som specificeret? Er indsatsen blevet gennemført som planlagt? Er de opnåede effekter i overensstemmelse med de tilsigtede?	Kombinerer fordele ved målopfyldelses-evaluering og procesevaluering.	Kombinerer svagheder ved målopfyldelses-evaluering og procesevaluering.

Evalueringsmodel (beslægtede betegnelser)	Hovedspørgsmål	Styrker	Svagheder
Økonomiske modeller: a. cost-efficiency b. cost-effectivness c. cost-benefit	a. Er produktiviteten tilfredsstillende? b. Er effektiviteten tilfredsstillende? c. Er nytten tilfredsstillende?	Indsatsens omkostninger medgår, og opnået effekt ses i forhold til medgåede ressourcer.	Resultaterne er afhængige af gyldige effektindikatorer, som kan være vanskeligt målbare.
Brugerevaluering (brugerne er målgruppen for indsatsen)	Hvordan erfarer brugerne indsatsen?	Da brugerne er indsatsens målgruppe, er deres vurdering vigtig.	Tendens til at foretage for overfladiske tilfredshedsundersøgelser, der ikke gyldigt måler brugernes erfaringer. Svært at anvende på indsatser, der består af tvang, samt over for svage brugere.
Interessentevaluering (interessenter er de aktører, der har en interesse i indsatsen)	Hvordan erfarer interessenterne indsatsen?	Interessenternes værdikriterier bliver afdækket. Interessenternes kendskab til indsatsen udnyttes. Interessenterne vil mere sandsynligt anvende evalueringens resultater.	Interessenterne kan være svære at identificere og afgrænse. Interessenterne kan være tilbøjelige til at ytre sig. Svage grupper blandt interessenterne kan blive nedtonet.
Kollegial evaluering (peer review)	Hvordan vurderes professionelles indsatser af eksperter inden for samme profession?	Ved komplekse indsatser kan kollegial evaluering være eneste mulighed.	Kan virke unødigt selvbeskyttende for professionen. Resultater af kollegiale vurderinger kan være usikre.

Det er vigtigt at understrege, at modellerne både kan benyttes hver for sig og i kombination. Skema 1.1 giver en stikordsoversigt over de ti evalueringsmodeller, hvorefter vi giver enkelte eksempler på brug af »rene« og »kombinerede« modeller.

Et eksempel på en målopfyldelsesevaluering finder vi fx i AKF's evaluering af Energikonsulentordningen, hvor konsulenter skulle gennemføre varmesyn i enfamiliehuse og foreslå energibesparelser. Evalueringen havde bl.a. til formål at evaluere, om de foreslåede foranstaltninger eller justeringer blev gennemført. Og hvor meget varme og brændsel der blev sparet? Gennem evalueringen fandt man bl.a. frem til, at der ved ca. 5.000 gennemførte varmesyn i 1990 blev foreslået ca. 15.500 energibesparende for-

anstaltninger. Tre år efter var 25% af de foreslåede foranstaltninger gennemført, og de gennemførte projekter havde reduceret husholdningernes energiforbrug med 6%.

Et eksempel på en kombineret resultat- og procesevaluering finder vi fx i AKF's evaluering af Socialministeriets forsøgsprojekt »Sygemeldt – og i gang« (Eskelinen og Koch 1997). Forsøget havde til formål at udvikle socialarbejdernes arbejdspraksis. Evalueringen fandt sted sideløbende med udviklingsprojekter i seks kommuner og omfattede både en proces- og en resultatevaluering. Procesevalueringen førte til en erfaringsopsamling af de nye metoder og arbejdsprocedurer, som blev udviklet i projektet, samt en række gode råd for sagsbehandlingen på området, fx at det er vigtigt, at kommunerne hurtigt indkalder til opfølgningssamtaler og foretager disse efter konkrete vurdering af den enkelte klients situation. Resultatevalueringen bidrog med en række målinger af, hvad klienterne havde fået ud af kommunernes indsatser, fx viste det sig, at 38% af klienterne endte sygemeldingsperioden med, at de vendte tilbage til deres tidligere arbejde eller andet arbejde, mens 16% af klienterne ophørte med at arbejde.

Et eksempel på en *effektevaluering* er givet i kapitlet om evaluering på skoleområdet, ved den såkaldte PISA-undersøgelse. PISA-programmet – Programme for International Student Assessment – er etableret i et samarbejde blandt OECD-medlemslande. Formålet er at belyse, hvor godt unge mennesker er forberedt til at møde udfordringerne i dagens informations-samfund. PISA er karakteristisk ved ikke at vurdere kompetencerne ud fra specifikke læseplaners indhold, men ved i stedet at tage udgangspunkt i de livsområder, som anses for at være vigtige for unge menneskers fremtidige livssituation. Effektevaluering er uddybet i kapitel 4.

Der er ikke tale om en målopfyldelsesevaluering i snæver forstand, idet effekterne ikke er afgrænset til bestemte nationale skolepolitiske mål, men det forventes alligevel, at evalueringen kan forsyne uddannelsespolitikere, uddannelsesadministratorer og praktikere med en omfattende vurdering af læringsresultater målt ved slutningen af den undervisningspligtige periode, og derved bidrage til uddannelsespolitiske beslutninger.

Et eksempel på en *formativ evaluering* er givet i kapitlet om evaluering på handicapområdet. Procesevalueringer er uddybet i kapitel 5. Evalueringen rettede sig mod en holistisk og tværfaglig rehabiliteringsmodel for

voksne med traumatisk hjerneskade. Der var tale om en ny behandlingsmodel, der blev indført og var under udvikling ved et genoptræningscenter på kolonien Filadelfia i Dianalund. Der er tale om en bred og dyb evaluering, for så vidt angår den indholdsmæssige kompleksitet, idet hjerneskader og den kombinerede behandling heraf gennem medicinske, psykologiske og sociale indsatser indeholder meget sammensatte problemstillinger.

Evalueringen havde fire mål: 1) at give rehabiliteringscentret tilbagemeldinger på deres arbejde med en helhedsorienteret rehabiliteringsmodel undervejs i en 2-årig evalueringsproces, 2) at pege på justeringer og udviklingsmuligheder i rehabiliteringscentrets organisation, 3) at delagtiggøre en bredere offentlighed i de problemstillinger, der knytter sig til implementering af den holistiske og tværfaglige rehabiliteringsmodel og 4) at sætte modellen ind i den internationale diskussion omkring behandling af traumatiske hjerneskader.

Det blev valgt at gennemføre en procesevaluering, der blev designet som en formativ evaluering. Det var således de processuelle aspekter af genoptræningscentrets arbejde med at implementerer genoptræningsmodellen, der blev evalueret. Herudover var det sigtet med evalueringen at give rehabiliteringscentret tilbagemeldinger såvel undervejs i evalueringen som efterfølgende, hvilket netop er kendetegnende for og styrken ved formative evalueringer.

Et eksempel på en *systemmodel* er undersøgelsen af indsatser over for hjemløse, misbrugere og sindslidende, som er beskrevet i kapitlet om evaluering af udsatte grupper. Inden for Socialministeriets evalueringsprogram fra 2001 var der således et ønske om en mere omfattende evaluering af tilbuddene til hjemløse, misbrugere og sindslidende i Danmark. Her drejede det sig om sammenhængen i det samlede tilbud, og undersøgelsen skulle blandt andet besvare spørgsmålet om, hvilke modeller man havde for denne indsats, om tilbuddet var dækkende, og om der var forskelle i den måde forskellige grupper var dækket ind af tilbuddet på.

Evalueringen omfattede forskellige delundersøgelser. For det første en etnografisk undersøgelse, hvor en etnograf opholdt sig i området i nogle perioder og gennemførte a) en kortlægning af de forskellige tilbud til de forskellige målgrupper og disse tilbuds principper og målsætning, b) en undersøgelse af brugernes livsverden, deres gruppemæssige tilhørsforhold,

deres hverdag, brug af forskellige tilbud og personlige karriere, c) en undersøgelse af samspillet mellem brugere og professionelle inden for udvalgte tilbud, og d) en spørgeskemaundersøgelse blandt brugere af udvalgte tilbud (en brugerevaluering).

For det andet en »auditundersøgelse«, hvor man samlede udvalgte professionelle i de forskellige regioner til en diskussion af nogle af de cases, som etnograferne havde sat fokus på. Diskussionerne, der blev båndoptaget, berørte spørgsmål som professionel etik, begrænsningerne i tilbuddet og dilemmaerne i de professionelle krav (en kollegial evalueringsmodel).

For det tredje en organisationssociologisk undersøgelse af den koordinering af tilbuddet, som foregik på forvaltningsniveau i et samspil mellem forskellige sektorer og mellem amter og kommuner.

Et eksempel på anvendelse af en *økonomisk evalueringsmodel* (cost-effectiveness) er givet i kapitel 16 i eksemplet om vindkraft. Her sammenlignes de samfundsøkonomiske elproduktionsomkostninger, som bl.a. inkluderer produktionsomkostningerne, eksterne miljøeffekter og beskæftigelseseffekter for vindkraft og to referencealternativer: et naturgasfyret kraftværk og et kulfyret kraftværk.

Af metoder blev bl.a. anvendt omkostningsanalyse, værdisætning af tekniske emissionsloefficienter og spørgeskemaundersøgelse af betalingsvillighed for miljøgoder.

Et eksempel på *brugerevaluering* er beskrevet i kapitel 11 med et eksempel på handicapområdet, hvor to forskellige brugergrupper af amternes rådgivning på området blev gennemført. Der blev dels gennemført en spørgeskemaundersøgelse af kommunalt ansatte som brugere, og dels en interviewundersøgelse af mennesker med handicap eller deres pårørende.

En *interessentevaluering* er fx anvendt i eksemplet med »andre aktører i beskæftigelsespolitikken«, kapitel 13. Her interviewes en række interessenter på området: i tre AF-regioner, hos »andre aktører« ledige i udvalgte aktiveringsforløb og deltagere på kurser.

Et eksempel på *kollegial evaluering* er en evaluering, som blev foretaget af en forskningsfinansieringsordning i det norske Kommunenes Sentralforbund (Rieper og Kristiansen 1994). Her blev sammensat fire vurderingsgrupper efter de temaer, som forskningen omhandlede: kommu-

nal økonomi, organisation, helse- og socialområdet og national udvikling. Vurderingsgrupperne blev sammensat af seniorforskere fra Sverige og Danmark og suppleret med en nøgleperson med norsk, administrativ baggrund. Hver vurderingsgruppe fik et udvalg af forskningsrapporter til vurdering efter på forhånd opstillede vurderingskriterier. Vurderingen blev foretaget i grupperne ved dels en skriftlig vurdering af hver rapport, dels pointgivning for hvert kriterium.

Der er i indledningsafsnittet til håndbogens sektordel henvist til de eksempler, der er anvendt i hvert sektorkapitel, og hvilken evalueringsmodel der kendetegner hvert eksempel.

Valget af en eller en kombination af forskellige evalueringsmodeller bør som sagt »skræddersyes« til den konkrete situation, hvilket kræver, at evaluator tager stilling til en lang række praktiske, metodiske og teoretiske spørgsmål, hvoraf nogle er meget grundlæggende. En måde at tale om de mest grundlæggende spørgsmål for evalueringer og forskning på generelt er ved brugen af begrebet paradigme, som sammenfatter de grundlæggende anskuelser, som evaluator og forsker tager udgangspunkt i. Som afslutning på dette kapitel har vi valgt at præsentere fire grundlæggende paradigmer, som vi på mange punkter mener kan bidrage til at afklare evalueringens grundlæggende udgangspunkt og dermed være en hjælp i arbejdet med valg af evalueringsmodeller og udarbejdelse af det samlede design for konkrete evalueringer.

1.3 Paradigmer i evaluering

Et videnskabeligt paradigme er et samlet begreb for det sæt af grundlæggende og accepterede anskuelser, der er styrende for forskning og evaluering, såvel i teori som i praksis. Der er ikke kun et paradigme for forskning og evaluering, men flere forskellige, der alt efter videnskabsteoretisk perspektiv kan supplere hinanden eller er uforenelige som ild og vand. Et paradigmes grundlæggende anskuelser kan være sprogligt ubeskrevet og virke gennem normer og traditioner, der følges af forskere og evaluører inden for bestemte forsknings- og evalueringsmiljøer. Paradigmer kan også i væsentligt omfang være beskrevet og dermed gøres til genstand for refleksion, hvilket gør det muligt at foretage et bevidst valg af det grundlæggende

de paradigme for såvel forskning som evaluering, hvor det er muligt at holde sig inden for et paradigme eller bevidst at kombinere flere paradigmer.

Når vi kort indfører begrebet paradigme i denne sammenhæng, er det, fordi vi mener, at refleksion og stillingtagen til spørgsmålet om valg af paradigme vil styrke både evaluators og rekvirenters grundlag for at foretage de øvrige valg i forbindelse med planlægning og gennemførelse af konkrete evalueringer.

Evalueringslitteraturen indeholder flere forskellige typologier over paradigmer. Nogle opererer med 2-3 paradigmer (Patton 1997), andre med 4-5 paradigmer (Launsø og Rieper 2000, Foss Hansen 2003). Vi har valgt at præsentere fire forskellige paradigmer, der illustrerer variationen og kan bruges i forbindelse med diskussioner om valg af paradigme eller kombinationer heraf i konkrete evalueringer.

Skema 1.2 giver en oversigt over de fire paradigmer, som vi har valgt at præsentere i denne sammenhæng. Der er tale om henholdsvis: 1) det positivistiske, 2) det fortolkningsvidenskabelige, 3) det kritiske og 4) det handlingsrettede paradigme. Disse fire paradigmer adskiller sig fra hinanden på en række afgørende punkter. Efter stikordsbeskrivelsen af hvert paradigme i skemaet følger en uddybende beskrivelse af hvert paradigme med henvisninger til evalueringslitteraturen.

Skema 1.2 Eksempler på fire paradigmer og deres kendetegn

Paradigme Kendetegn	Empirisk-analytisk	Fortolkningsvidenskabeligt (hermeneutisk)	Kritisk	Handlingsrettet
Evalueringens sigte	Beskrive Forklare	Forstå Fortolke	Afsløre Afdække	Udvikle indsats
Evalueringens generelle spørgsmål	Hvilke Y er konsekvenser af X?	Hvilken mening tillægger X et bestemt fænomen Y i hvilken kontekst?	Hvilke kræfter virker bag aktørernes rygge, og hvordan kan de synliggøres?	Hvordan udvikler aktører indsatsen ud fra viden tilegnet under evaluering?
Fremherskende design	Eksperiment Statistisk variansanalyser	Casestudier	Vælges efter behov	Aktionsforskning Formativ evaluering
Fremherskende metode	Kvantitativ, fx Registeranalyser Survey	Kvalitativ, fx Dybdeinterview Feltarbejde	Vælges efter behov	Flere, vælges efter behov
Evaluators rolle	Neutral Distanceret	Deltagende observatør	Distanceret samfundskritisk blik	Proceskonsulent Rådgiver
Evalueringens anvendelsesorientering	Instrumentel Kontrol af målnåelse og utilsigtede effekter	Konceptuel Indsigtsgivende	Ændrer værdier Reformere Omvælte	Forbedre indsats i samspil med aktører i indsatsen

Det empirisk-analytiske paradigme

Den grundlæggende anskuelse inden for dette paradigme er, at den sociale virkelighed kan beskrives og forklares som ting, og det derfor er muligt at overføre de naturvidenskabelige idealer om positivisering og objektivitet på det sociale område. Antagelsen er, at man ved hjælp af naturvidenskabelige metoder kan opnå objektiv og sikker viden. Forestillingen er endvidere, at resultaterne af forskning og evaluering umiddelbart kan lægges til grund for politiske og administrative beslutninger. Anskuelserne bygger således på den såkaldte instrumentelle anvendelse af evalueringens resultater.

Inden for evalueringens litteratur er dette paradigme tydeligt hos Campbell og Stanley 1963 og kommer til udtryk i titlen på en kendt artikel: *Reforms as experiments* (Campbell 1969). I dansk sammenhæng er denne tradition opsummeret af Kjeld Møller Pedersen 1979.

Historisk har dette paradigme været fremherskende i de tidlige evalueringsmiljøer i 1960'erne og 1970'erne, og det præger stadig i betydelig grad evaluatorers selvforståelse og arbejde. Blandt årsagerne hertil er, at det er et magtfuldt paradigme, der er udbredt inden for store dele af sam-

fundsvideenskaberne og i særdeleshed inden for lægevidenskaben. Det er desuden en udbredt videnskabsanskuelse, som mange spontant associerer til. Det er også magtfuldt, fordi det tilsyneladende i særlig grad knytter an til den vigtige afdækning af forholdet mellem årsager og virkninger, som er et centralt spørgsmål i evalueringer, hvor det undersøges, om indsatsen (og dens elementer) er årsag til ønskede eller uønskede effekter af indsatsen. Inden for de øvrige paradigmer kan man også arbejde med forholdet mellem årsager og virkninger, men på andre måder, som historisk ikke har haft den samme gennemslagskraft.

Inden for det positivistiske paradigme håndteres det kontrafaktiske spørgsmål: Hvad var der sket, hvis indsatsen ikke havde fundet sted? Spørgsmålet håndteres typisk ved at anvende en eller anden variation af det eksperimentelle design, hvor man ud over en gruppe, der påvirkes af den givne indsats, også undersøger en kontrolgruppe, der ikke udsættes for indsatsen, se kapitlet om effektevaluering.

Det fortolkningsvidenskabelige paradigme

Det fortolkningsvidenskabelige paradigme bygger på den grundlæggende anskuelse, at viden og indsigt i sociale forhold nødvendigvis må bygge på forståelse og indlevelse i de sociale fænomener, der undersøges. Undersøgelsen af sociale forhold kan således ikke sidestilles med undersøgelse af naturen, som ontologisk adskiller sig fra sociale fænomener. Inden for evalueringsslitteraturen er dette paradigme tydeligt præsenteret af Guba og Lincoln (1990), og i dansk sammenhæng kan nævnes Krogstrup 1997 og 2003.

Det følger heraf, at evalueringer inden for det fortolkningsvidenskabelige paradigme må bygge på teorier, metoder, informationer og data, der bidrager til at skabe forståelse af forudsætninger for og gennemførelse af de indsatser, der skal evalueres. Inden for dette paradigme gælder det om at skabe grundlag for forståelse af de processer, der leder frem til, at indsatsen gør en forskel eller ej i bestemte sammenhænge (kontekster).

Forskellige aktørers meninger om processer og resultater skal identificeres og fortolkes, og evaluators forforståelse skal tages i betragtning og udvikles. Evaluator er ikke en neutral og distanceret observatør i forhold til den evaluerede indsats, men evaluators forforståelse indgår reflekteret i

fortolkninger og evaluators forforståelse kan ændre sig i løbet af selve evalueringen.

De anvendte metoder er ofte kvalitative, idet de skal give indsigt i de meninger og betydninger, som aktører tillægger processer og resultater.

En evalueringsrapport anskues som en fremlæggelse af forskellige fortolkninger af proces og resultater, og ikke som en objektiv beskrivelse. Anvendelsen bliver derfor ofte af konceptuel karakter, fx kan brugere ændre synspunkt, få en anden vinkel på indsatsen og en ny indsigt.

Det kritiske paradigme

Inden for det kritiske paradigme er det en grundlæggende antagelse, at den sociale verden og sammenhængene heri ikke altid er som det, de umiddelbart giver sig ud for at være. For at forstå og forklare fænomener i den sociale verden er det derfor nødvendigt med kritiske teorier om og analyser af, hvad der eventuelt gennem sig under »overfladen« og udgør den »virkelige verden«. Den såkaldte Frankfurterskole har fostret de fremmeste repræsentanter for udviklingen af kritisk teori. Jürgen Habermas (Eriksen 2003) er den måske mest kendte repræsentant, der bl.a. har udviklet teorien om kommunikativ interaktion. I evalueringsslitteraturen er Fischer (1999) et eksempel på en evaluator, der arbejder inden for det kritiske paradigme. Det er vanskeligt at finde eksempler på konkrete evalueringer, men man kan tænke på evalueringer af samfundsmæssige konfliktfelter, fx inden for miljøpolitikken og inden for sundhedspolitikken.

Inden for det kritiske paradigme er det ambitionen at blottlægge skjulte sammenhænge og afsløre illigitime påvirkninger af indsatser og deres virkninger. Evaluator får her rollen som kritiker af fremherskende værdier og institutioner og vil hente sine værdier fra samfundsvidenskabernes filosofi og teorier.

Det handlingsrettede paradigme

Den grundlæggende antagelse inden for det handlingsrettede paradigme er, at den socialt relevante viden og indsigt bedst opnås, når forskning og evaluering indgår i direkte sammenhæng og interaktion med de sociale fænomener, der undersøges. Herudover er det et ideal inden for dette paradigme, at det undersøgte fænomen påvirkes positivt i den proces, hvori under-

søgelse og evalueringer gennemføres. Der arbejdes således med en normativ tilgang, hvor målet er at forbedre indsatsen *samtidig med*, at evalueringen pågår.

Inden for samfundsforskningen er det handlingsrettede paradigme bl.a. repræsenteret ved aktionsforskning (Borchmann 2003, Nielsen 1996, Hegland 1981). Når det gælder evaluering, ligger paradigmet til grund for den tidlige formative evaluering og den nyere »empowerment«-evaluering. I dansk sammenhæng er denne evalueringspraksis beskrevet af Høgsbro og Rieper (2001) og i Dahler-Larsen og Krogstrup (2003, ss. 80-95).

Evaluering er per definition handlingsrettet og sigter mod praktisk anvendelse. Arbejder man inden for det handlingsrettede paradigme, bruges begrebet imidlertid til at sige, at evaluator tillægges en udviklingsopgave i samarbejde med indsatsens medarbejdere og interessenter. Dette udviklingsarbejde finder sted i løbet af evalueringsprocessen gennem interaktion mellem evaluator og indsatsens medarbejdere, jf. kapitlet om procesevaluering.

Appendiks 1

Eksempler på typologier over paradigmer

The Evaluation Methods Paradigms Debate Summary of Emphases: Thesis, Antithesis, Synthesis

	<i>Thesis: Originally Dominant »Scientific« Paradigm</i>	<i>Antithesis: Originally Competing »Alternative« Paradigm</i>	<i>Synthesis: Utilization-Focused Evaluation, Paradigm of Choices</i>
Purpose	Summative	Formative	Intended use for intended users
Measurement	Quantitative data	Qualitative data	Appropriate, credible, useful data
Design	Experimental designs	Naturalistic inquiry	Creative, practical, situationally responsive designs
Researcher stance	Objectivity	Subjectivity	Fairness and balance
Inquiry mode	Deduction	Induction	Either or both
Conceptualization	Independent and dependent variables	Holistic interdependent system	Stakeholder questions and issues
Relationships	Distance, detachment	Closeness, involvement	Collaboration, consultative
Approach to study of change	Pre-post measures, time series, static portrayals at discrete points in time	Process-oriented, evolving, capturing on-going dynamism	Developmental, action-oriented. What needs to be known to get program from where it is to where it wants to be?
Relationship to prior knowledge	Confirmatory, hypothesis testing	Exploratory, hypothesis generating	Either or both
Sampling	Random, probalistic	Purposeful, key informants	Combinations, depending on what information is needed
Primary approach to variations	Quantitative differences on uniform, standardized variables	Qualitative differences, uniquenesses	Flexible: Focus on comparisons most relevant to intended users and evaluation questions
Analysis	Descriptive and inferential statistics	Case studies, content and pattern analysis	Answers to stakeholders' questions
Types of statements	Generalizations	Context-bound	Extrapolations
Contribution to theory	Validating theoretical propositions from scientific literature	Grounded theory	Describing, exploring, and testing stakeholders' and program's theory of action
Goals	Truth, scientific acceptance	Understanding, perspective	Utility, relevance: Acceptance by intended users

Kilde: Patton 1997, s. 270.

Dominant Paradigm: Operational Definition of Evaluation Quality in the 1970s

<i>Dimension of Evaluation Quality</i>	<i>Coding Scheme (higher number = higher quality)</i>
Sampling	1 = Systematic random 0 = Nonrandom, cluster, or nonsystematic
Data analysis	2 = Quantitative 1 = Qualitative and quantitative 0 = Qualitative
Statistical procedures	4 = Multivariate 3 = Descriptive 2 = Ratins from qualitative data 1 = Narrative data only 0 = No systematic material
Impact Procedures Design	3 = Experimental or quase-experimental randomization and control groups 2 = Experimental or quasi-experimental without both randomization and control groups 1 = Londitudinal or cross-sectional without control or comparison groups 0 = Descriptive, narrative

Kilde: Patton 1997, s. 273.

Dimensions of Competing Methodological Paradigms

<i>Qualitative/Naturalistic Paradigm</i>	<i>Quantitative/Experimental Paradigm</i>
Qualitative data (narratives, descriptions)	Quantitative data (numbers, statistics)
Naturalistic inquiry	Experimental designs
Case studies	Treatment and control groups
Inductive analysis	Deductive hypothesis testing
Subjective perspective	Objective perspective
Close to the program	Aloof from the program
Holistic contextual portrayal	Independent and dependent variables
Systems perspective focused on interdependencies	Linear, sequential modeling
Dynamic, ongoing view of change	Pre-post focus on change
Purposeful sampling of relevant cases	Probabilistic, random sampling
Focus on uniqueness and diversity	Standardized, uniform procedures
Emergent, flexible designs	Fixed, controlled designs
Thematic content analysis	Statistical analysis
Extrapolations	Generalizations

Kilde: Patton 1997, s. 299.

Evalueringsparadigmer

Anskuelser om: Paradigme:	Evalueringsopgavens indhold	Evalueringens organisatoriske og samfundsmæssige rolle	Det analytiske fokus i evalueringen	Metode	Evaluators rolle
Positivistisk	Beskrive, måle, vurdere og forklare	Social ingeniørkunst med afsæt i kontrol af hidtidig indsats	Adfærd og effekt	Kvantitativ	Neutral, distanceret observatør
Hermeneutisk	Forstå og fortolke processer, situationer og kontekster	Bidrage til indsigt og oplysning	Mening	Kvalitativ	Deltagende observatør
Kritisk	Problematisere	Provokere, omvælte, ændre værdier	Sociale, politiske mønstre	Multiple, vælges efter behov	Aktivist
Deltagelsesbaseret	Udvikle praksis og problemløsning	Empowerment	Interessenters opfattelse af problemer og løsninger	Selvevaluering, dialog, konfrontation og forhandling	Proceskonsulent
Pragmatisk	Besvare spørgsmål, tilgodese informationsbehov	Gøre en forskel, nytte afhænger af kontekst	Problemstillinger, svar	Multiple, vælges efter behov	Professionel

Kilde: Foss Hansen 2003, s. 24.

Appendiks 2

Eksempler på definitioner af »Evaluering«

The application of social science techniques to the appraisal of social-action programs has come to be called evaluation research. (International Encyclopedia of the Social Sciences, 1968).

Careful retrospective assessment of the merit, worth and value of administration, output, and outcome of government interventions, which is intended to play a role in future, practical action situations. (Vedung 1997, p. 3).

Policy evaluation uses a range of research methods to systematically investigate the effectiveness of policy interventions, implementation and processes, and to determine their merit, worth, or value in terms of improving the social and economic conditions of different stakeholders. (Cabinet office, UK: The Magenta Book, chap. 1, p. 7, 2003).

2 **Evaluering i praksis – hvilke valg skal foretages?**

Olaf Rieper

Evalueringsprocessen kan opdeles i en række valg illustreret ved spørgsmål, hvis besvarelse kan anvendes som retningsgivende for, hvordan man planlægger en evaluering, og som også kan anvendes som støtte ved læsning og vurdering af en evalueringsrapport.

De her angivne spørgsmål bygger på en pragmatisk grundindstilling til evaluering, og de vil være nyttige i de fleste former for evalueringer. De går i lidt forskellige variationer igen i flere lærebøger i evaluering (Mehlbye, Rieper og Togeby 1993, Sverdrup 2002, Rossi, Lipsey og Freeman 2004). Spørgsmålene repræsenterer forskellige faser i planlægningen af en evaluering, og der anvendes derfor i det følgende både begreberne faser og spørgsmål.

2.1 **Oversigt over spørgsmål og faseinddelingen**

Til forskel fra megen anden samfundsforskning er en evaluering karakteriseret ved, at nogle aktører forventes at skulle anvende evalueringens resultater i praksis. Derfor er især de første spørgsmål (1., 2., 3. og 4. spørgsmål) specielle for evaluering, mens de efterfølgende spørgsmål 7-10 (fra problemformulering til formidling) er gældende for samfundsforskning i almindelighed.

1. Hvad er evalueringens genstand?
2. Hvad er evalueringens kontekst?
3. Hvem skal anvende evalueringen? Og til hvad? (evalueringens formål)

4. Hvilke værdikriterier (om nogle) skal evalueringens resultater bedømmes ud fra?
5. Hvad er evalueringens teoretiske ramme (hvis eksplicit)?
6. Hvilke spørgsmål skal besvares gennem evalueringen (evalueringsspørgsmål, problemstillinger)?
7. Hvilken evalueringsmodel (evt. kombinationer) skal vælges?
8. Hvilke metoder skal bruges til dataindsamling, og hvordan skal de gennemføres?
9. Hvordan skal databearbejdning og -analyse/fortolkning foretages?
10. Hvordan skal evalueringens resultater formidles?
11. Sammensætning af evalueringsteam, tidsplan og budget.

2.2 Præsentation af de enkelte faser

I det følgende vil de enkelte faser kort blive omtalt. Der vil som gennemgående eksempel blive anvendt en evaluering af et lederudviklingsprogram på et sygehus (Hatting og Rieper 1991).

Hvad er evalueringens genstand?

Selve det, der gøres til genstand for evaluering skal i udgangspunktet beskrives med en detaljeringsgrad, så evaluator/bestiller har en god forståelse af, hvad evalueringen retter sig imod, og hvordan genstanden afgrænses. En mangelfuld afgrænsning og forståelse af evalueringens genstand/indsats kan føre til misforståelser senere i evalueringsprocessen og i værste fald gøre evalueringen ubrugelig.

Beskrivelsen kan indeholde følgende punkter:

Indhold: hvilke ressourcer og aktiviteter består indsatsen af? Hvordan er den organiseret?

Klarhed i formål og rationale: hvad er indsatsens formål, er de klart beskrevet, er der flere formål og mistanke om skjulte formål? Det er vigtigt for evaluator at have et indtryk heraf, især hvis der er tale om en målopfyldelsevaluering, men også af hensyn til interesser for at anvende evalueringens resultater. Rationalet er indsatsens egentlige begrundelse.

Grad af konfliktende interesser: er indsatsen omgærdet af stærk politisk opmærksomhed, af modstridende interesser, eller er det en aktivitet,

der ønskes justeret på operationelt, teknisk niveau? Stærke konflikter kan betyde forsøg på at påvirke evaluatør og forsøg på at bortforklare uønskede resultater. Ved at være opmærksom på risikoen for noget sådant fra evalueringens start kan evaluatør tage skridt i retning af fx at inddrage stærke interessenter i evalueringens tilrettelæggelse og vælge design og metoder, som er robust over for kritik.

Grad af nyhed: er det en indsats, man har drevet gennem længere tid, eller en ny aktivitet, der er igangsat, fx som forsøg? En veletableret indsats, som er i normal drift, er ofte præget af institutionel forankring, faglige traditioner og faste rutiner. Evaluatør må erhverve en vis indsigt i disse forhold for at kunne vælge relevante fremgangsmåder og vurdere ambitionsniveauet for anvendelse af evalueringen.

Indsatsens kompleksitet: er indsatsen præget af overskuelige og forudsigelige processer og virkningsmekanismer, eller er der tale om, at vejene fra indsats til virkning er svære at gennemskue og virkninger vanskeligt forudsigelige? Mere komplekse indsatser stiller store krav til valg af evalueringsmodeller og -metoder, og til at begrebsliggøre virkningsmekanismer og effekter.

Eksempel: evaluering af lederudvikling på et sygehus

Indhold og nyhed: Lederudviklingsprogrammet var en nystartet aktivitet i amtet, og blev betragtet som et forsøg. Det blev gennemført af en privat konsulent. Der var nedsat en styregruppe for et lederudviklingsprogram bestående af sygehusets topledelse, uddannelseslederen og konsulenten. Det forløb over 2,5 år og bestod af 10 samlinger af to dages varighed med deltagerne opdelt i fire hold med hver ca. 30 deltagere, der deltog i alt ca. 120 af sygehusets ledere, dvs. overlæger, sygeplejeledelse, laborantchefer m.fl. Motivationen fra lederne var blandet.

Klart formål og rationale: lederudviklingsprogrammets formål var nøje beskrevet og gik i korthed ud på at udvikle den enkelte og give deltagerne redskaber og færdigheder til at lede. Rationalet var at få lederne væk fra en faglig orienteret ledelse til en professionel ledelse. Rationalet var ikke klart nedskrevet.

Lav grad af konflik: Der var interesse for lederudviklingsprogrammet fra amtets sundhedsdirektør, men ikke fra politisk side. Sygehusets topledelse var meget opmærksom og interesseret.

Medium kompleksitet: Der var ikke tidligere erfaringer med et sådant lederudviklingsprogram i sygehusregi, så der var ikke en færdig model for lederudviklingsprogrammet. På den anden side forelå megen erfaring med lederudvikling i andre offentlige organisationer, og konsulenten var erfareren på området.

Hvad er evalueringens kontekst?

Kontekst vil sige de omgivelser, der præger indsatsen. Man kan også kalde konteksten for indsatsens rammebetingelser. Disse kan indeholde institutionelle forhold som lovgivning, rutineprægede handlemønstre, magtfulde interessenter og økonomiske forhold, som indvirker på indsatsen og dens muligheder for at virke. Konteksten vil være forskellig, alt efter om evalueringens genstand er et projekt på lokalt, nationalt eller internationalt niveau. Fx kan evalueringens genstand være en lokal institution, en bestemt velfærdsydelse, fx hjemmehjælp, et virkemiddel i regeringens politik, fx aktiveringsprogrammer, eller hele politikområder, fx handlingsplan for et bedre arbejdsmiljø. Men genstanden kan også være EU's strukturfondsprogrammer, som gennemføres i flere europæiske lande. Det er oplagt, at evaluator må have en vis viden om indsatsens kontekst for at kunne tilrettelægge evalueringen hensigtsmæssigt.

Hvem skal anvende evalueringen? Og til hvad?

Dette er et grundlæggende spørgsmål, som bør besvares tidligt i planlægningen af en evaluering, idet det vedrører selve evalueringens formål og dermed udstikker retningslinjer for det videre evalueringsarbejde, dvs. de følgende faser.

I nogle tilfælde er det enkelt at identificere de forventede anvendere af en evalueringens resultater, som kan være bestiller, bestemte beslutningstagere eller -fora, et råd, udvalg eller en komite. Men der kan også være tale om en mere diffus skare, som man ikke kan identificere eller afgrænse helt entydigt på forhånd, fx ledere og medarbejdere i amternes specialråd-

givning, eller beslutningstagere i de nye EU-medlemslandes ministerier, der har til opgave at forvalte midler fra EU's strukturfonde.

Selve begrebet »forventede anvendere« er heller ikke entydigt. I snæver betydning henviser det til de personer, som må forventes at handle på grundlag af evaluering (*i eksemplet med lederudvikling: styregruppen for programmet og sundhedsdirektøren i amtet*), i en bredere betydning er det personer, som må forventes at sætte sig ind i evalueringens resultater, fordi de antages at have en kundskabsinteresse i evalueringen, selv om de ikke direkte vil handle ud fra den. *I eksemplet med lederudviklingsprogrammet: andre sygehusledelser, som overvejer at gennemføre lederudvikling, uddannelseskonsulenter i diverse fagforeninger for sygehuspersonale.*

Det næste er, hvad de forventes at anvende resultaterne til. Her skelner litteraturen om anvendelse af evaluering mellem instrumentel anvendelse, konceptuel eller oplysende anvendelse, anvendelse til legitimering og taktisk anvendelse. Den instrumentelle anvendelse er den »klassiske«, hvor informationer fra evalueringer indgår som beslutningsgrundlag for at foretage forbedringer af eller afvikle en indsats. Den oplysende anvendelse indebærer, at nogle bliver klogere, ser anderledes på indsatsen, ændrer holdninger e.l. Ved anvendelse som legitimering søger bestiller at retfærdiggøre allerede truffne beslutninger eller handlinger ud fra den senere foretagne evaluering. Ved taktisk anvendelse søger bestiller at udsætte beslutninger eller undgå at blive draget til ansvar eller at bruge evalueringen til at vise et image udadtil som en rationel organisation. *I eksemplet med lederudvikling: styregruppen ønsker at anvende evalueringens resultater til at justere afslutningen af programmet (instrumentelt), men også at vise over for omverdenen og indadtil, at man tager programmet alvorligt og vil forbedre det (taktisk anvendelse).*

En afklaring af den forventede anvendelse vil hjælpe evaluator/bestiller til at stille relevante spørgsmål i evalueringen og til at vælge i de resterende faser af evalueringsprocessen.

Hvilke vurderingskriterier skal evalueringens resultater bedømmes ud fra?

Begrebet »evaluering« stammer fra latin og betyder »vurdere«. Spørgsmålet er, hvad kilden til vurderingskriterierne er, og hvem der fastlægger dem.

Lad os først slå fast, at der kan gennemføres evalueringer, uden at evaluator nødvendigvis skal foretage vurderinger. I de tilfælde, som ikke er få, overlades vurderingen til bestiller eller andre interessenter. Det kan ske i mål-frie evalueringer – »just give them the facts«. I mange tilfælde foretager evaluator vurderinger uden at tydeliggøre kriterierne herfor. Det er kritisabelt, men vil forekomme på områder, hvor der er så stærk enighed om det ønskelige, at kriterierne ligger i kulturen.

Vurderingskriterier kan opstilles på grundlag af henholdsvis beskrivende værdisætning og foreskrivende værdisætning. Tages udgangspunkt i *beskrivende værdisætning*, udledes værdikriterier af indsatsens mål, eller af andre interessenters udtalte værdier, fx værdier hos indsatsens brugere. Værdisætning med udgangspunkt i politisk vedtagne mål tillægges en demokratisk værdi og har en høj grad af legitimitet, fordi det i princippet er folkevalgte politikere, der har besluttet, hvilke værdier der skal lægges til grund for offentlig service og regulering.

Ved *foreskrivende værdisætning* er det derimod evaluator, der opstiller kriterierne. Det kan være på grundlag af fagligt begrundet viden om indsatsens bidrag til problemløsning (fx undgå klimaopvarmning ved energispareprogrammer) eller på grundlag af faglig viden om brugeres behov, fx i forbindelse med beskyttelse af unge ved tvangsfjernelse fra hjemmet, eller ved pleje af demente. Det er indlysende, at der kan være forskellige opfattelser af de foreskrivende værdier, og evaluator må være beredt på at forsvare og begrunde sine kriterier.

Ved evalueringer, hvor der ønskes et mål for omkostningseffektivitet, måles indsatsens effekt i fysiske enheder eller pengeværdi og sættes i forhold til indsatsens omkostninger. Fx måles energispareaktiviteters effekt som: kWh sparet/omkostninger i kr. ved energispareprogrammet, eller ved regionale udviklingsprogrammer: antal job skabt ved indsatsen/indsatsens omkostninger. Værdikriteriet er indbygget i begrebet »omkostningseffektivitet«, men vurderingen af, om en funden omkostningseffektivitet bør vurderes som tilfredsstillende eller ej, afhænger af sammenligninger med standarder, som fx sammenligninger med fortiden, tilsvarende indsatser i ind- og udland, »bedste praksis«, (benchmarking), brugeres og interessenters forventninger og professionelle standarder.

Ud over de formulerede mål med den givne indsats kan der være tale om generelle værdier i den pågældende kultur, fx proceduremæssige værdier (såsom lighed for loven) og demokratiske (deltagelses) værdier.

I eksemplet med lederudvikling kan man sige, at der i rationalet for programmet ligger værdier om omkostningseffektivitet (professionel ledelse sikrer et veldrevet sygehus), men evalueringen havde ikke til formål at måle omkostningseffektivitet, da denne først evt. ville blive indhøstet på længere sigt. Men da målene for programmet var detaljeret beskrevet, kunne evalueringens resultater holdes op mod disse delmål, dvs på grundlag af beskrivende værdisætning: fx var det et delmål, at deltagerne skulle anvende værktøjer til medarbejderudvikling. Derfor var ét kriterium, i hvilket omfang deltagerne havde gennemført medarbejdersamtaler med personalet på deres afdeling.

Hvad er evalueringens teoretiske begrebsramme (hvis eksplicit)?

Evalueringer uden eksplicit formulerede begrebsrammer kan få et meget lokalt præg og være svære at overføre til andre, lignende indsatser. Ved at formulere en teoretisk ramme kan man se evalueringen i forhold til et bredere kundskabsfelt. Rammen kan påvirke den måde evalueringsspørgsmålene stilles på, og den kan afgrænse og præcisere analysen af de indhentede data.

Der skelnes mellem to sæt af teorirammer: Teori om målområdet og teori om indsatsen.

Teori om målområdet handler om de kræfter, der påvirker det område, som den aktuelle indsats retter sig imod, men uden indsatsen så at sige. Teori om målområdet bliver sjældent formuleret i evalueringer, men viden herom indgår i et eller andet omfang, ofte afhængig af foreliggende faglige traditioner på området og evaluators faglige baggrund.

Evaluator vil kunne tilegne sig teori om indsatsens målområde ved at læse oversigtsværker eller reviewartikler på området eller ved at tale med fagligt indsigtfulde personer på området.

I lederudviklingsprogrammet havde vi i baghovedet organisationsteori om ledelse af faglige bureaukratier, hvor stærke professioner har magt og er afgørende for driften. Og teori om de kræfter, der former lederroller

i en sådan organisation, fx kollegial kontrol og accept, professionelles uvillighed til at agere i et hierarki, faglige afværgemekanismer osv. Den ene evaluator have en faglig baggrund i organisationsteori, den anden havde ansættelse som økonom på et sygehus.

Begrebsramme om indsatsen (interventionsteori, programteori eller theory of action) er antagelser om, hvad der bringer indsatsen til at fungere. Teori om indsatsen er således forestillinger om, hvilke elementer i indsatsen der påvirker målgruppen gennem hvilke mekanismer. Teori om målområdet er derimod antagelser om, hvilke andre forhold end indsatsen, der påvirker målgruppen.

Teori om indsatsen foreligger sjældent i færdig form, men skal udledes og sammensættes fra forskellige kilder, fx indsatsens medarbejdere og ledelse, faglig teori og viden på området.

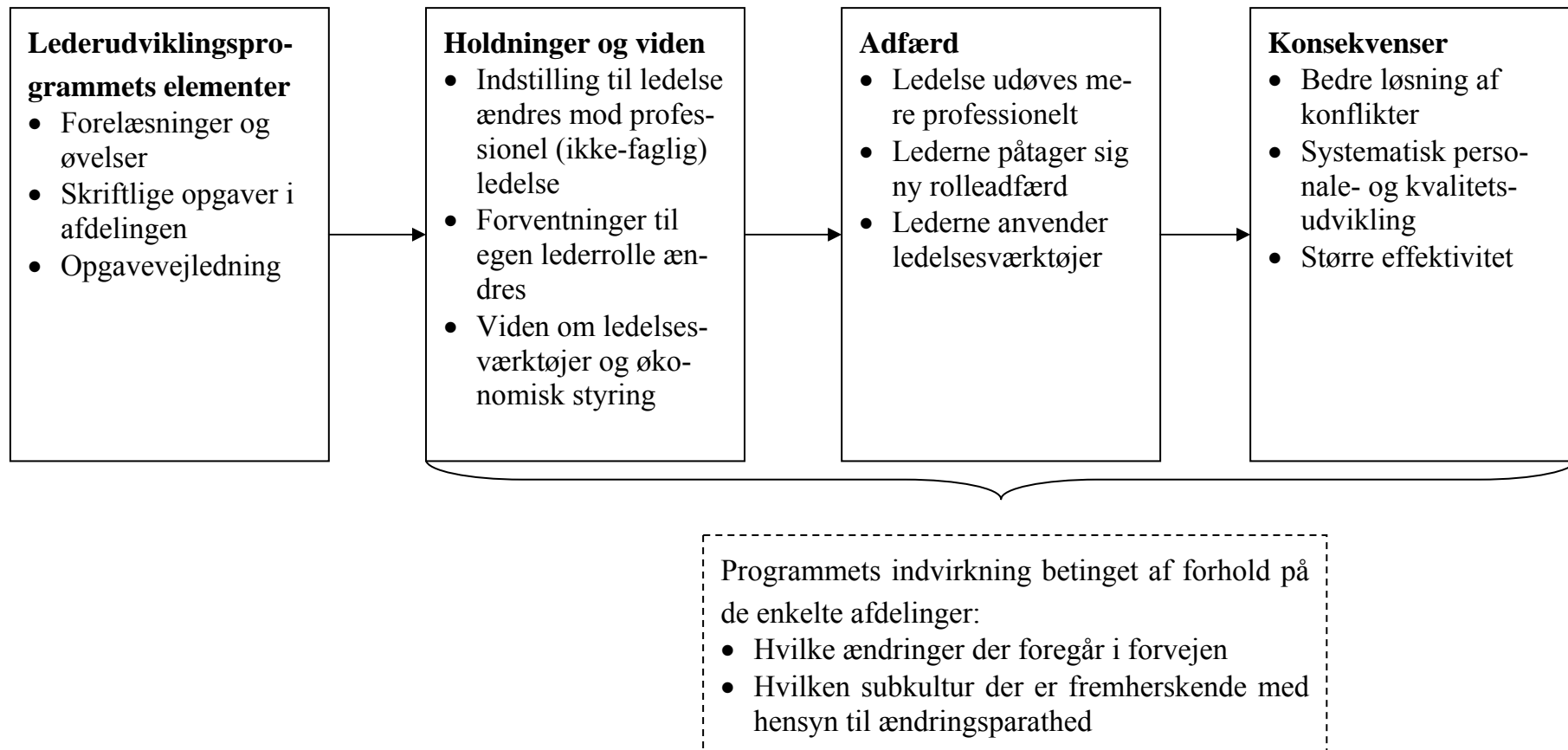
Begrebsrammen vil ofte være formuleret i prosatekst, men kan også afbildes i et kassepilediagram, der viser de formodede årsag-virkningssammenhænge. Sådanne diagrammer kan opstilles i færdig form, inden evalueringen påbegyndes, men kan også udvikles undervejs i evalueringsprocessen. Diagrammet tjener til at udvælge årsag-virkningssammenhænge, der er særligt vigtige at fokusere evalueringen på, og viser samtidig, indirekte, hvad der er udeladt.

Teori om indsatsen kan også udvikles i analysefasen

I evalueringen af lederudviklingsprogrammet gav den foreløbige analyse af datamaterialet anledning til at opstille antagelser om, at de enkelte sygehusafdelinger i forvejen var ulige langt i ændringsprocesser, og at det bevirkede, at udbyttet af programmet var forskellig fra afdeling til afdeling. En efterfølgende analyse viste, at ledere i de enkelte afdelinger, der havde fået størst udbytte af programmet, også var de afdelinger, hvor man var længst fremme i en ændringsproces, før lederudviklingen startede.

Diagram:

Eksempel på begrebsramme (teori om indsats) fra evaluering af et lederudviklingsprogram



Hvilke spørgsmål skal besvares gennem evalueringen (evalueringsspørgsmål, problemstillinger)?

På grundlag af evalueringens formål og viden om genstandsområdet fastlægges de evalueringsspørgsmål eller problemstillinger, som evalueringen skal besvare. Såfremt evalueringens forventede brugere udgør et mindre antal, kan det være hensigtsmæssigt at drøfte med dem, hvad deres informationsbehov er, og hvilke evalueringsspørgsmål de er interesseret i at få belyst.

I eksemplet med lederudviklingsprogrammet startede evaluatoren med at interviewe hver af de forventede brugere, dvs. medlemmer af sygehusets direktion, uddannelseslederen, konsulenten og amtets sundhedsdirektør. Herefter indkaldte evaluatoren samtlige parter til et møde, hvor evalueringens spørgsmål blev drøftet på grundlag af en skitse fra evaluatoren. Herefter blev evalueringsspørgsmålene endeligt fastlagt.

Evalueringsspørgsmålene vil ofte blive opstillet i et hierarki med hovedspørgsmål efterfulgt af mere detaljerede underspørgsmål. Disse underspørgsmål er således retningsgivende for konstruktionen af de dataindsamlingsinstrumenter (spørgeskemaer osv.), der vælges anvendt, og den efterfølgende analyse.

Arbejdet med at formulere evalueringsspørgsmål kan ikke gøres grundigt nok, da de er knastakselen for hele evalueringen. Selv om det er ønskeligt at formulere evalueringsspørgsmålene på et ganske detaljeret niveau, bør der også være muligheder for at opdage nye problemstillinger undervejs i evalueringsprocessen.

Eksempel på evalueringsspørgsmål i evalueringen af lederudviklingsprogrammet

Hvordan indvirker lederudviklingsprogrammet på ændringer i deltagernes holdninger, indstillinger, synspunkter:

- *vedrørende målformulering og målstyring?*
- *vedrørende opfattelse af egen lederrolle og andres lederrolle?*
- *med hensyn til egen personlige udvikling?*
- *med hensyn til at bruge et fælles »sprog« som bærer af en fælles forståelse?*

- *med hensyn til synliggørelse/klargøring af spændinger og konfliktstof?*

Hvordan indvirker lederudviklingsprogrammet på deltagernes adfærd?

Sker der ændringer i:

- *kommunikation med patienter, herunder inddragelse af patienters ønsker og synspunkter?*
- *faktisk udført personaleudvikling/ledelse herunder anvendelse af bestemte værktøjer i personaleledelse?*
- *anvendelse af teknikker til ressourcestyring?*
- *gennemførelse af målformulering og målstyring?*
- *lederstil?*
- *det faktiske samarbejde mellem afdelinger?*
- *faktiske måder at gennemføre konfliktregulering på?*

Hvordan indvirker videregående ændringer på lederudviklingsprogrammet?

- *Hvilke ændringer er sket i kvaliteten af patientrelationerne?*
- *Hvilke ændringer er der sket i personalets arbejdsstrivsel/engagement?*
- *Hvilke ændringer er der sket i produktiviteten?*

Hvilken evalueringsmodel (evt. kombinationer) skal vælges?

Her skal omtales på, hvilket grundlag man kan vælge evalueringsmodeller, jf. Foss Hansen 2001. Valg af evalueringsmodeller kan i praksis være præget af 1) forhandlinger mellem evaluator, bestiller og evt. andre interessenter, 2) et bestemt koncept eller design, som anvendes som standard på et givet sektorområde og 3) evaluators tidligere erfaringer med bestemte evalueringsmodeller, 4) at bestiller egenhændigt bestemmer, hvilken evalueringsmodel der skal anvendes.

Ad 1) Forhandlinger mellem evaluator, bestiller og evt. andre interessenter: Forskellige interessenter kan have forskellige ønsker om, hvilken evalueringsmodel der skal anvendes – og evaluator kan have sine ønsker. Udfaldet kan blive, at den stærkeste vinder, at man opnår et kompromis, eller at forskellige evalueringsmodeller kombineres. *I eksemplet med lederudviklingsprogrammet accepterede sygehusets direktion evaluators forslag om*

at gennemføre en survey samt gruppeinterview med alle de ledere, som deltog i programmet.

Ad 2) Et bestemt koncept eller design, som anvendes som standard på et givet sektorområde:

På bestemte sektorområder kan der være en udbredt opfattelse af, at bestemte evalueringsmodeller bør benyttes. Anvendelse af en bestemt eller nogle få evalueringsmodeller kan hidrøre fra, at bestemte modeller internationalt er blevet standard på området, fx inden for udviklingsbistand, eller at et bestemt koncept er blevet anvendt af en dominerende evaluator, fx på uddannelsesområdet i Danmark. En bestemt model kan få forrang, fordi den er velbeskrevet og afprøvet i mange tilfælde, fx BIKVA-modellen på det sociale område.

Ad 3) Evaluators tidligere erfaringer med bestemte evalueringsmodeller:

Evalueringstiljører kan udvikle kompetencer med bestemte evalueringsmodeller, som de derfor foretrækker. Kompetencerne kan bunde i bestemte faglige uddannelser (fx er økonomer uddannet i at gennemføre effektmålinger ved hjælp af statistiske metoder), eller at et evalueringstiljører har udviklet en særlig ekspertise med bestemte evalueringsmodeller eller bestemte metoder (fx at kombinere kvantitative og kvalitative metoder).

Ad 4) Bestiller bestemmer egenhændigt, hvilken evalueringsmodel der skal anvendes.

Det ses også ofte, at bestilleren af en evaluering i sit udbudsmateriale forudsætter, at bestemte evalueringsmodeller bliver anvendt.

Kriterier for valg af model

Men uanset, hvordan valg af model foregår, hvad skal evaluator eller bestiller lægge vægt på i valget af model(ler)? Det abstrakte svar er, at man bør vælge den model, der bedst kan give svar på evalueringsspørgsmålene. Men i praksis kan et givet sæt evalueringsspørgsmål give rum for flere modeller, og evalueringsspørgsmål kan udformes efter, at modellen er valgt.

Der kan peges på to udgangspunkter for at vælge model: evaluering-

gens funktion, evalueringens genstand.

Med udgangspunkt i en evalueringsfunktion som kontrol fra overordnet instans kan nærliggende valg være målopfyldelses- eller effektmodel, men også monitorering. Er evalueringsfunktionen derimod læring for indsatsens medarbejdere, vil valg af formativ model være relevant.

Evalueringens genstand (indsatsen) kan være kendetegnet ved forskellige grader af kompleksitet i forholdet mellem indsats og virkninger, og med forskellig målbarhed for virkningerne. Programmer for at opnå energibesparelser har klart målbare effekter (målt i energienheder), og aktiviteter, der giver effekter (virkningsmekanismer), kan identificeres med rimelig sikkerhed. Her vil typisk en målopfyldelsesmodel være egnet. Er der derimod tale om indsatser over for mennesker, der både har stofmisbrug, og som har sindslidelser, er effektmålet mindre klart og indsatsernes virkningsmekanismer svære at afdække. Her skal typisk vælges en systemmodel eller flere modeller i kombination, jf. kapitlet om evaluering af socialt arbejde for udsatte grupper.

Hvilke metoder skal bruges til dataindsamling, og hvordan skal de gennemføres?

Det grundlæggende princip er igen, at evaluator skal vælge den eller de metoder, der bedst kan besvare de stillede evalueringsspørgsmål.

Metodemulighederne er mange og adskiller sig ikke fra metoderne i samfundsvidenskaberne som sådan. Her skal derfor blot gives en oversigt med henvisninger til speciallitteraturen. Højt strukturerede metoder er indsamlingsinstrumenter med forhåndsklassifikation af de ønskede oplysninger, over for de lavt strukturerede metoder med lav grad af forhåndsklassifikation.

Oversigt over metoder til dataindsamling (eksempler)

	Spørgemetoder	Observationsmetoder	Dokumentarmetoder
Højt strukturrede (Olsen 2004)	Postspørgeskemaer. Telefoninterview. (Olsen 2004)	Observation af smågrupper baseret på afkrydsningsskema.	Registeroplysninger. Databanker.
Lavt strukturrede	Dybdeinterview. (Kvale 1997) Fokusgruppeinterview. (Hatting og Rieper 1991, Halkier 2002)	Etnografisk feltarbejde. Deltagerobservation. (Kristiansen og Krogstrup 1999)	Dagbøger. Mødereferater.

Hvordan skal databearbejdning og -analyse/fortolkning foretages?

Her gives også en oversigt, idet databearbejdning og -analyse er generelt for samfundsvidenskaberne og ikke specielt for evaluering. Igen er det helt centralt, at analysemetoderne vælges, så de bedst muligt med de givne ressourcer til evalueringen vil kunne give svar på evalueringsspørgsmålene.

Oversigt over metoder til analyse og fortolkning af indsamlede oplysninger (eksempler)

	Oplysninger indsamlet ved spørgemetoder	Oplysninger indsamlet ved observation	Oplysninger fra foreliggende materialer
Kvantitative analysemetoder	Analyser ved hjælp af statistiske programpakker (Nielsen og Kreiner 1999)		Bearbejdning og Statistisk analyse af registerdata (Kluve og Schmidt 2002).
Kvalitative fortolkningsmetoder	Meningskondensering. Grounded theory (Kvale 1997, Fog 1997). Narrative analyser (Hansen og Sehested 2003)	Fortolkning af oplysninger indsamlet ved etnografisk feltarbejde (Andersen m.fl. 1992).	Socialkonstruktivisme (Hansen og Sehested 2003)

Hvordan skal evalueringens resultater dokumenteres og formidles?

Formidlingen afhænger primært af evalueringens formål: hvem skal bruge evalueringen (målgruppe) og til hvad (anvendelsestype og -tid)? Men for al formidling gælder et generelt dokumentationskrav til evalueringer, nemlig at man skal kunne gennemskue de anvendte metoder, så læseren selv kan danne sig en mening om resultaternes troværdighed. Uanset om man væl-

ger enkle eller avancerede metoder, bør læseren af evalueringsrapporten kunne se, hvordan metoderne konkret er anvendt. Ofte vil bilag, der beskriver de anvendte indsamlings- og analysemetoder, kunne indløse dette krav. Beskrivelsen bør være så detaljeret, at læseren i princippet vil kunne anvende de samme metoder.

Målgruppe og formidlingsform

Et første skridt i at fastlægge, hvordan en evaluering skal formidles, er at anvende den tidligere bestemmelse af evalueringens målgruppe (som er gjort i fasen om evalueringens formål) og afdække, hvilken formidlingsform der er mest velegnet til målgruppen. Ved en afgrænset og veldefineret målgruppe kan evaluator simpelthen spørge målgruppen eller repræsentanter for denne. Og målgruppens brug af forskellige formidlingskanaler kan identificeres.

På denne baggrund vælger evaluator så, evt. i samråd med bestiller, formidlingsform: mundtlig, skriftlig, og/eller billeder. Den mundtlige form er velegnet til en begrænset målgruppe og er normalt en effektiv formidlingsform, men begrænset til relativt få personer. Den skriftlige form er den mest udbredte i form af rapporter og artikler, der når relativt mange og kan gøres let tilgængelig ved offentliggørelse på internettet. Billeder i form af fotos og levende billeder anvendes som supplement til de to andre former, men står sjældent alene. Billeder er imidlertid meget indtryksgivende, fx kan fotos og videooptagelser fra etnografisk feltarbejde og fra fokusgruppeinterview formidle forhold, som vanskeligt lader sig beskrive på skrift.

Formidlingskanal og -tidspunkt

Internettet er en effektiv formidlingskanal og vil med bredbåndsforbindelse kunne formidle på tekst, billeder og tale.

Fagtidsskrifter og magasiner, som er rettet mod den pågældende målgruppe, er et velegnet medie til at nå mange om en evalueringens hovedresultater.

Massemedier (radio og tv) når de store befolkningsgrupper og er mest egnet til at gøre opmærksom på, at evalueringen eksisterer, så særligt interesserede kan rekvirere yderligere oplysninger.

Konferencer, seminarer o.l. er velegnede kanaler til at nå et mindre antal personer.

Tidspunktet for formidlingen kan både ligge under selve evalueringsprocessen og efterfølgende. Såfremt formidlingen sker, inden evalueringen er afsluttet, vil det i sagens natur dreje sig om formidling af foreløbige resultater, begreber og perspektiver. Dette kan finde sted som led i en formativ evaluering eller som led i, at en følgegruppe følger evalueringen og drøfter foreløbige resultater. Den såkaldte procesanvendelse er ganske udbredt (Claus Rebien, Kim Forss).

Sammensætning af evalueringsteam, tidsplan og budget

Sammensætningen af den gruppe, der gennemfører en evaluering, sker normalt tidligt i planlægningen, men bør ikke ske, før man har et overblik over, hvilke kompetencer evalueringen kræver. Evalueringsteamet skal indeholder de relevante metodekompetencer samt den nødvendige viden om evalueringens genstand (indsatsen og dens kontekst). Den ansvarlige evaluator kan bemande teamet fra egen organisation, eller danne et konsortium med andre evaluatore, såfremt der mangler kapacitet eller kompetencer i egen organisation. Samarbejdet kan finde sted i form af ligestillet partnerskab eller ved anvendelse af underleverandører til velafgrænsede opgaver, fx ses ofte, at en hovedleverandør anvender underleverandører til telefoninterview eller til at stå for tekniske specialopgaver.

Tidsplan og budget opstilles inden for de givne rammer for evalueringen. Budgettet vil kunne opstilles i et regneark med evalueringsopgaver specificeret i forspalten (lodret kolonne) og med tidsintervaller opstillet i vandret række. De enkelte celler kan så indeholde det anslåede tidsforbrug for de enkelte opgaver fordelt på fx uger.

2.3 God evalueringsskik

Enhver professionsgruppe med respekt for sig selv har udarbejdet standarder for god professionel adfærd, etik mv.

Således også evaluatore. De professionelle faglige evalueringssforeninger i forskellige lande har udarbejdet sådanne standarder, se fx guidelines for god practice in evaluation på www.evaluation.org.uk

I Danmark har Dansk EvalueringsSelskab haft en arbejdsgruppe om emnet, se Kjærsgaard (2002), hvorfra disse bud stammer:

»10 bud« på god evalueringsskik

1. Brug altid rigelig med tid til at planlægge og designe et evalueringsprojekt.
2. Gennemfør en »foranalyse«, hvis der er brug for at få konkretiseret den problemstilling, der skal evalueres.
3. Ingen må spille med »lukkede kort«, når en evaluering skal opstartes og gennemføres.
4. Afklar mest muligt i forbindelse med udarbejdelse af kontrakt og projektbeskrivelse.
5. Alle interessenter bør inddrages/informerer på relevante tidspunkter af en evaluering.
6. Test analysernes bæredygtighed og rækkevidde gennem midtvejsmøde eller ved præsentation af de foreløbige resultater/løsningsforslag.
7. Aftal altid på forhånd, hvem der har rettighederne til rapporterne, og hvem der må offentliggøre resultaterne.
8. Når rapportudkastet ligger der, må rekvirenten ikke overspille sin ejerrolle, og evaluator må garantere høj faglig standard.
9. Formidling må indtænkes i projektet og målrettes de relevante målgrupper.
10. Evaluator og rekvirent mødes efterfølgende for at drøfte og kritisk gennemgå hele evalueringsforløbet.

3 Intern eller ekstern evaluering

Olaf Rieper

3.1 Et kontinuum mellem ekstern og intern evaluering

Det er almindeligt at skelne mellem forskellige måder at organisere gennemførelsen af evalueringer på. En vigtig, gennemgående skelnen i evalueringsslitteraturen er mellem intern og ekstern evaluering. En intern evaluering gennemføres af medarbejdere i den samme organisation, som er ansvarlig for den indsats, der evalueres. Ekstern evaluering varetages af medarbejdere fra en anden organisation end den, der er ansvarlig for indsatsen. Skellet er i princippet klart nok, men i praksis er der snarere tale om et kontinuum mellem ekstern og intern evaluering end om to helt forskellige måder at organisere evaluering på.

Intern evaluering foregår antagelig i relativt stort omfang i kommuner og institutioner (Albæk og Rieper 2001, Dahler-Larsen 2002). Der er knyttet såvel ulemper som fordele til både intern og ekstern evaluering. Det behandles i andet afsnit i dette kapitel, som afsluttes med overvejelser over, hvornår man skal vælge det ene eller det andet. Men først om kontinuummet mellem intern og ekstern evaluering.

I tabel 3.1 er en udfører defineret som den, der er ansvarlig for at foretage evalueringen, mens underleverandøren er en, som gennemfører bestemte dele af en evaluering, fx dataindsamling eller analysearbejde. Bestiller er den, der ønsker evalueringen gennemført og finansierer den. Anvender er defineret som aktører, der formodes at anvende resultaterne af evalueringen.

Tabellen illustrerer kontinuummet ved at vise nogle kombinationer af disse aktører, og det angives, hvor man normalt vil anvende begreberne intern og ekstern evaluering.

Tabel 3.1 Eksempler på kontinuum mellem ekstern og intern evaluering¹

	Intern	Ekstern
Udfører		
Underleverandør		
Bestiller		
Anvender		

Bem.: For overskuelighedens skyld er streger undladt mellem nogle af cirklerne:

1 Opstillingen er inspireret af Vedung 1997, s 116.

Et eksempel på kombination 3 er en sygehusledelse, som bestiller en evaluering af deres lederudviklingsprogram fra et sektorforskningsinstitut, som står for hele evalueringen og er ansvarlig for evalueringsrapporten. Det er sygehusledelsen, som anvender evalueringens resultater til at tilrettelægge en fremtidig lederudvikling.

Et eksempel på kombination 6 er en kommune, som har kørt et forsøg med internetbaseret (elektronisk) kommunikation mellem borgere og forvaltning/politikere. Kommunalbestyrelsen ønsker, at projektet evalueres, og en medarbejder ansat i kommunens udviklingsafdeling er ansvarlig for opgaven. Denne medarbejder anvender et konsulentfirma til at stå for telefoninterview og fokusgruppeinterview, mens hun selv foretager analyse og skriver rapporten til kommunens ledelse. Såfremt den, der skulle anvende evalueringens resultater, var en ekstern organisation, fx KMD (Kommunedata), er det et eksempel på kombination 5.

Den interne evaluator kan være placeret forskellige steder i den organisation, der er ansvarlig for den evaluerede indsats. I eksemplet ovenfor er evaluator placeret i en udviklingsafdeling (stabsafdeling under kommunaldirektøren), og hvis vedkommende også står for projektet med internetbaseret kommunikation, er der tale om en selvevaluering. Den interne evaluering kan således være organisatorisk forankret sammesteds eller på afstand af indsatsens ledelse og medarbejdere. I et amt har man eksempelvis en kvalitets- og udviklingsafdeling, som varetager brugerevalueringer i form af patientsurvey på sygehuse, dvs. en intern evaluering af kombination 4, såfremt amtet og dets institutioner betragtes som én organisation, men en ekstern evaluering af kombination 3, såfremt sygehuset anskues som en organisation for sig, og som både er bestiller og anvender.

3.2 **Fordele og ulemper ved henholdsvis intern- og ekstern evaluering**

Evalueringsslitteraturen indeholder overvejelser om fordele og ulemper ved interne henholdsvis eksterne evalueringer. En bog, som udelukkende handler om intern evaluering og interne evalueringerfunktioner, er Sonnichsen 2000, hvis forfatter har en baggrund som leder af FBI's evalueringssenhed i USA.

Som det fremgår af nedenstående skema, er kriterierne for fordele og ulemper set fra såvel evaluators som indsatsens position. Centralt står spørgsmålet om evaluators grad af (hierarkisk) afhængighed i forhold til dem, der har ansvaret for den indsats, der evalueres. Antagelsen er den rimelige, at afhængighed giver risiko for, at evaluator vil tage hensyn til indsatsens ledelse og dermed afgrænse evalueringen fra at tage kontroversielle emner op, nedtone negative resultater eller i værste fald fikle med evalueringens resultater. Under alle omstændigheder vil en intern evaluering ofte fremstå som mindre troværdig over for omgivelserne. På den anden side er en ekstern evaluator, som outsider, mindre forpligtiget over for opfølgning af evalueringen og forbedringer af indsatsen.

Der kan endvidere være en risiko for, at den interne evaluator ikke har tilstrækkelig faglig erfaring med og ekspertise i evaluering. Dette er en al-

vorlig svaghed ved intern evaluering og kan forværres af den nævnte afhængighed mellem evaluator og organisationen, som også står for indsatsen. På den anden side kan en kompetent intern evaluator afværge evt. forsøg på påvirkning fra indsatsens side.

Den interne evaluators detaljerede viden om indsatsen skal afbalanceres med det forhold, at interne evaluators kan være for »indfødte« og dermed mangle blik for vigtige forhold.

Det må regnes for en vigtig styrke ved intern evaluering, at erfaringerne fra evalueringen beholdes personbåret i organisationen. Der indhøstes normalt mere viden i en evaluering end evaluator formår at formidle i skriftlig form. Tilmed kan viden fra flere evalueringer akkumuleres over tid som læring i organisationen.

Fordele henholdsvis ulemper ved intern evaluering

Fordele Evaluator...	Ulemper Evaluator...
Er forpligtiget over for organisationen	Er afhængig af organisationens ledelse
Har viden i detaljer om indsatsens struktur, funktion og brugere	Er for indforstået med forhold vedr. indsatsen og kan dermed være blind for nogle aspekter
Kan reagere hurtigt på ønske om at igangsætte evaluering	Kan mangle faglig evalueringskompetence
Kan fungere som organisationens hukkommelse	
Kan løbende støtte beslutninger i organisationen med evalueringsresultater	Kan være blind for, hvad organisationen har behov for
Har lettere adgang til data i organisationen	
Kan fleksibelt varetager andre opgaver end evaluering i organisationen	Kan blive belastet af andre opgaver
Kan let følge op på, om evalueringens anbefalinger bliver fulgt	Kan komme i etiske dilemmaer
Kan sikre læring og kontinuitet med hensyn til evalueringer i organisationen	

De vigtige fordele ved den eksterne evaluator kan samles i begreberne uafhængighed og faglig evalueringskompetence. Uafhængighed af indsatsens ledelse og organisation kan være en vigtig faktor for at legitimere evalueringens neutralitet og troværdighed. Men begrebet kan gradbøjes. Mange evaluators opererer på et marked og vil være interesseret i at blive anvendt

en anden gang af bestilleren. Dette kan også udgøre et (markedsmessigt) afhængighedsforhold, som kan være nok så stærkt.

En ekstern evalueringssenhed vil normalt kunne mønstre en større faglig evalueringskompetence end en intern. Dette skyldes, at eksterne evalueringer ofte vil have større erfaring med og være mere specialiseret i bestemte evalueringssopgaver, end interne har mulighed for. Eksterne evalueringer vil også lettere kunne sammensætte et evalueringsteam med differentieret ekspertise. Fx kan en evaluering, der både kræver statistisk analysekompetence og erfaring med kvalitative analyser, sammensættes af evalueringer med netop sådanne specialkompetencer.

Skal en organisations interne evalueringssystemer underkastes evaluering, ligger det næsten i sagens natur, at en ekstern evaluering er at foretrække. Men ikke ubetinget, fx kan man fra centralforvaltningen i en kommune evaluere institutioners brugerevalueringer, altså en intern evaluering af kommunens institutioners evalueringssystemer.

Fordele henholdsvis ulemper ved ekstern evaluering

Fordele	Ulemper
Er uafhængig af den organisation, der står for indsatsen	
Kan anlægge nye, udefrakommende perspektiver på indsatsen	Mangler insideviden om indsats og organisation
Kan evaluere en organisation, inkl. interne evalueringssenheder	
	Har vanskeligere adgang til data i organisationen om indsatsen
Kan have mere specialiseret evalueringskompetence	

3.3 Hvornår vælge hvilken slags evaluering?

Fordele respektive ulemperne antages imidlertid at være betinget af, hvilket hovedformål evalueringen har. Evalueringer, som har til formål at kontrollere og evt. drastisk ændre eller nedlægge en indsats, vil normalt blive organiseret som en ekstern evaluering. Det samme vil evalueringer af politisk kontroversielle indsatser, fx en evaluering af de offentlige myndigheders indsats over for Christiania (Jæger m.fl. 1993).

Evalueringer, der derimod har et forbedrende, justerende sigte over for mindre kontroversielle indsatser, vil ofte med fordel kunne tilrettelægges som intern evaluering, da evaluator dermed kan udnytte sin position som insider. Evaluering, der skal indgå i læreprocesser, vil ofte med fordel kunne organiseres som intern evaluering, da organisationen i så fald vil kunne trække på det dybtgående kendskab og den kontinuitet, som en intern evaluator kan tilbyde.

Kontinuummet mellem intern og ekstern evaluering muliggør, at bestiller kan kompensere for svagheder ved henholdsvis intern og ekstern evaluering ved at inddrage samarbejdspartnere og underleverandører til bestemte evalueringsopgaver. Fx kan manglende faglig kompetence hos intern evaluator kompenseres ved at inddrage ekstern rådgiver eller underleverandører til bestemte opgaver.

Oversigt over forhold, der peger på henholdsvis intern eller ekstern evaluering

Evalueringsfunktion:	Intern evaluering	Ekstern evaluering
Kontrol	–	+
Læring	+	–
Politisk kontroversiel indsats	–	+
Store ændringer af indsats forventes	–	+
Mindre justering af indsats forventes	+	–

4 Effektevaluering

Mikael Togeby

4.1 Indledning

I dette og det følgende kapitel omtales nærmere to af de evalueringsmodeller, der blev nævnt i kapitel 1. Det er valgt at uddybe henholdsvis effektevalueringer og procesevalueringer, idet de udgør de vigtigste og mest gennemprøvede evalueringsmodeller.

Effektevalueringer finder effekten af en indsats. Procesevalueringer finder fremmende og hæmmende forhold for at opnå en effekt fra en indsats. I praksis er de to typer evalueringer som yin og yang. Det er almindeligvis umuligt at gennemføre en god effektevaluering, hvis man ikke ved, hvilke forhold som har betydning for at opnå en effekt, dvs. har en god teori, og det er netop, hvad procesevalueringen kan bidrage med. Tilsvarende er det nødvendigt for procesevalueringen at have et fokus på effekten, fx hvilke barrierer der findes for at opnå en effekt. Selv om metoderne til dataindsamling og analyse ofte er forskellige, så kan det være hensigtsmæssigt at gennemføre de to typer evalueringer efter en koordineret plan.

Hvor effektevalueringer med statistiske analyser kan have fokus på tal, så kan procesevalueringer supplere med ord og begreber.

Effektevalueringer kan anvendes til at dokumentere, hvilken effekt som kan tilskrives en given indsats. Det er ofte offentligt finansierede aktiviteter, som evalueres. Og ofte ønsker opdragsgiverne at vide, om pengene er givet godt ud. Et par eksempler:

- Fører en informationskampagne til færre trafikuheld?
- Fører praktik på arbejdspladser til en hurtigere integration af flygtninge?

- Fører rådgivning om energibesparelser til, at der bliver realiseret flere energibesparelser?

I disse eksempler er fokuseret på aktivitetens slutmål. Hvis opgaven er at undersøge, om visse forudsætninger for at opnå et resultat er opfyldt, fx om bilisterne forstår og husker informationskampagnens budskaber, så vil man almindeligvis kalde det for en procesevaluering (se kapitel 5). Der er gode argumenter for, at effekt- og procesevalueringer gennemføres samtidig eller koordineret.

Effekterne kan både omfatte forudsete og uforudsete, positive og negative og inden for såvel som uden for det tilsigtede område. Hvis målet fx er integration af flygtninge, så kan positive eller negative påvirkninger af familiemønstret være eksempler på effekter uden for det tilsigtede område. Effekten kan endvidere opgøres på kort eller længere sigt. Det er som regel praktisk umuligt at dække alle disse typer af effekter, men det kan være hensigtsmæssigt at være opmærksom på, hvilke typer effekter som evalueringen tænkes at omfatte.

Den store udfordring i forbindelse med effektevalueringer er at fastlægge årsag-virknings-forholdet. Ét er at dokumentere, at der sker en ændring af et problems omfang (færre trafikuheld) – en anden og langt sværere opgave er at dokumentere, hvilken del af ændringen (virkning) som ske på grund af kampagnen (årsag). Grunden til, at det er svært, er, at der sker andet end netop indsatsen, som kan have betydning for ændringen i problemet. Er der kommet flere biler på vejene? Har der været debat om hastighedsgrænserne, eller har vejret været anderledes i før- end i efterperioden?

4.2 Det kontrollerede eksperiment

Det kontrollerede eksperiment er et stærkt design, som imidlertid ofte er vanskelig at anvende. Når dette design præsenteres først, så er det for at illustrere de forhold, som bør have opmærksomhed, når man afviger fra betingelserne for det kontrollerede eksperiment, fx den tilfældige udvælgelse.

I et kontrolleret eksperiment udvælges indsatsgruppe og kontrolgruppe tilfældigt, og derved sikres det, at de to grupper er helt ens, statistik set.

Ved at alle modtager »en behandling« (en pille), sikres mod en forsøgseffekt. Den efterfølgende analyse er meget enkel, idet forskellen i eftersituationen alene kan tilskrives påvirkningen fra den rigtige pille i forhold til den virkningsløse placebo-pille. En tilfældig udvælgelse betyder, at der fra samme totalpopulation skal vælges et passende antal tilfældigt til de to grupper.

En styrke ved det kontrollerede eksperiment er, at man ikke behøver at have en teori for indsatsen. Evalueringen kan gennemføres som en black-box-evaluering.

I praksis er det imidlertid sjældent, at det kontrollerede eksperiment kan anvendes i forbindelse med evalueringer. Det skyldes, for det første, at det kan være vanskeligt at fortage den helt tilfældige udvælgelse af deltagere. Hvis alle i fx en kommune tilbydes en bestemt aktivitet, så kan der ikke fra samme totalpopulation udvælges personer til kontrolgruppen, da der ikke er nogen, som ikke er blevet udsat for indsatsen. Der kan også være etiske problemer med at anvende en kontrolgruppe. Kan man tilbyde personer med et alvorligt behov »en placebo-pille«?

Der er også ulemper ved at anvende en black-box-metode. Hvis man ikke ved, hvilke elementer som har været vigtig for at opnå en effekt, så kan man risikere, at man ved en gentagelse ikke bevarer de vigtige ingredienser i aktiviteten.

4.3 Brug af sammenligningsgruppe

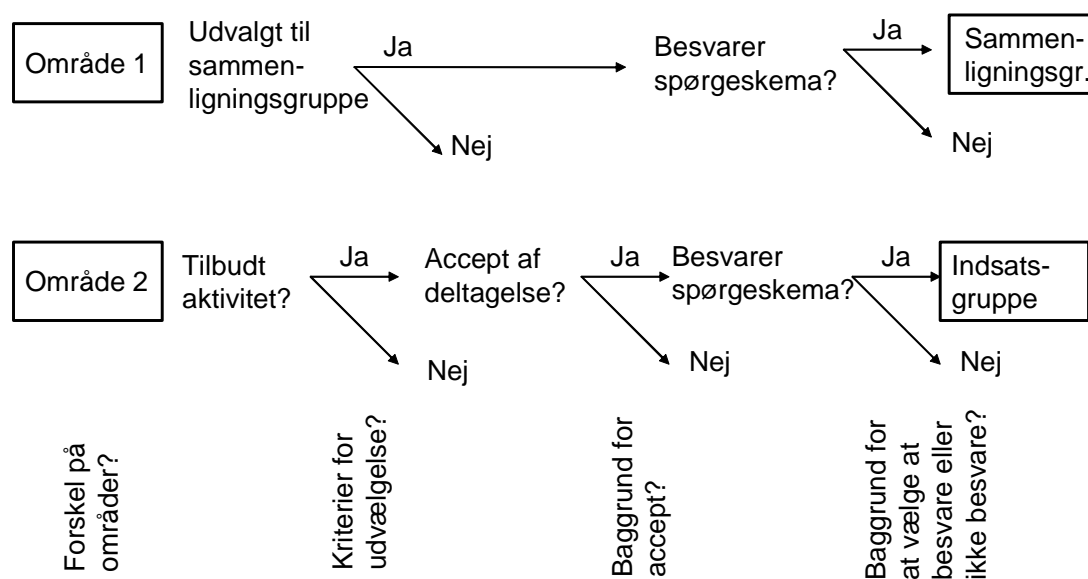
Et alternativ til det kontrollerede eksperiment er at slække på kravet om, at indsatsgruppe og kontrolgruppe skal vælges tilfældigt fra samme totalpopulation. I stedet kan man vælge en sammenligningsgruppe, som i så høj grad som muligt ligner indsatsgruppen, men som ikke er udvalgt fra samme totalpopulation. Det kan fx være en lignende befolkningsgruppe fra en nabokommune. Om det er en god ide at anvende nabokommunen, afhænger af, hvilke forhold som er vigtige for indsatsen.

I praksis vil sammenligningsgruppen ofte afvige fra indsatsgruppen på væsentlige punkter, og det er derfor nødvendigt at korrigere for dette. Forskellen kan gå på den variabel, som indsatsen handler om (fx arbejdsløshed, helbred, antal trafikuheld, energiforbrug) og/eller på baggrundsvariab-

ler (køn, uddannelse, indkomst, alder). Hvis sammenligningsgruppen fx har en højere indkomst end indsatsgruppen, og denne variabel har betydning for det resultat, der fremkommer, og dermed indsatsens succes, så må man korrigere for indkomstens betydning.

Figuren herunder viser med et tænkt eksempel, at der kan være mange forhold, som kan give forskelle i en sammenlignings- og en indsatsgruppe. Ud over indsats- og sammenligningsgrupper er der i figuren i realiteten fem andre grupper (som har svaret nej til et af udvælgelsesspørgsmålene). Generelt kan det forventes, at alle disse syv grupper vil være forskellige. Forskellene skyldes ikke alene indsatsen, men en række valg taget af den enkelte eller af ledelsen af indsatsen eller evalueringen. Hvis fx dem, som siger ja til at deltage, er markant anderledes end dem, som takker nej, så kan det »forvride« grupperne. I princippet skal betydningen og vægtingen af alle valg kendes, før man med sikkerhed kan beskrive effekten af indsatsen.

Et eksempel på udvælgelse af indsats- og sammenligningsgruppe



Når målgruppen selv kan vælge at deltage i indsatsen og vælge, om de vil besvare et spørgeskema, så vil der optræde en form for selv-selektion. Dette kan føre til en skævvridning af resultaterne, som kan gå begge veje. Det kan føre til, at resultaterne enten er dårligere eller bedre, end de i virkelig-

heden er. En rygestopkampagne kan tiltrække dem, som er særlig motive- ret – eller der kan være en overvægt af rygere, som har prøvet at stoppe med at ryge mange gange uden held. Tilsvarende kan valget om at bruge tid på at udfylde et spørgeskema være begrundet af særlig interesse for em- net eller en særlig stærk holdning imod det. Hvis man har adgang til nogle grundoplysninger om totalpopulationen (fx familieform, titel, boform), kan disse anvendes til at beskrive de ændringer, som sker i hvert trin.

4.4 **Statistiske analyser af naturlige eksperimenter**

I stedet for at anvende data fra et arrangeret eksperiment (fx et kontrolleret eksperiment eller med udvalgt indsatsgruppe og sammenligningsgruppe) kan anvendes data fra en virkelig indsats. Dette kaldes til tider for det na- turlige eksperiment. Her er ingen fra centralt hold blevet udvalgt. Indsatsen er måske blevet annonceret bredt, og nogle har reageret, og andre har ikke reageret.

En generel måde at tackle udfordringerne med forskelle i indsats- og sammenligningsgruppe på er at anvende en statistisk model. Modellen kan anvendes til at korrigere for forskelle i baggrundsvariabler og til at be- stemme betydningen af de forskellige valg. Metoden kan anvendes til at beskrive betydningen af indsatsen, når der er korrigeret for en række for- hold.

Ofte anvendes en regressionsmodel, hvor fx den variabel, indsatsen handler om (y), beskrives som en funktion af et antal variabler (x_i). y kan således fx være antal trafikuheld, energiforbruget, indkomsten eller eksa- menskarakteren. Med lidt andre typer modeller kan y også være sandsyn- ligheden for at komme i arbejde eller varigheden af arbejdsløshed.

Den generelle form er:

$$y = a_0 + a_1 * x_1 + a_2 * x_2 \dots$$

Computeren finder ud fra de tilgængelige data koefficienterne (a_i), som gi- ver den bedste sammenhæng. Typisk vil en af variablerne være en dum- myvariabel (enten 0 eller 1), som beskriver, om individet har deltaget i ak-

tiviteten. Koefficienten til denne variabel beskriver indsatsens effekt. I praksis kan man fx finde, at indsatsen (beskrevet ved variabelen x_1) har haft den og den effekt (a_1), når der er korrigeret for en række variabler (x_2 til x_n , hvor n er sidste baggrundsvariabel), dvs. en betinget effekt. Fx kan det dokumenteres, at energiforbruget faldt med a_i for de familier, som tog imod rådgivningen, når der er taget hensyn til udviklingen i indkomsten (x_2) og antallet af hjemmeboende børn (x_3).

I eksemplet er det, der skal forklares (y), en almindelig variabel, som kan have mange værdier. Hvis det, der skal forklares, er en logisk variabel (ja/nej, 0/1), kan dette håndteres i en logistisk regression. Særlige metoder er velegnet til at analysere data, som handler om varighed (fx af ledighed – se kapitel 14). Ligeledes er der metoder til at analysere paneldata, dvs. gentagne observationer af de samme individer.

Der er stor fleksibilitet i modelopbygningen. Der kan indgå kontinuerte variabler (fx indkomst eller alder) – og disse kan indgå lineært (dvs. at effekten af øget indkomst er den samme over hele spektret: For hver 10.000 kr. indtægten øges, øges variabelen y med a_2) eller ikke-lineært (fx: effekten varierer U-formet med indtægten, dvs. størst effekt for små og store indtægter, eller der kan være trinværdier: Der er en ekstra effekt, når alderen overstiger 50 år). Også forhold, som ikke kan måles på en enkel skala, fx uddannelse, arbejdsfunktion eller helbredsoplysninger, kan indgå.

Data består typisk af individdata fra både før og efter indsatsen og fra indsats- og sammenligningsgruppen. Ved at anvende individdata, i stedet for aggregerede data, bevares mest mulig information, og det er muligt at opbygge mere virkelighedsnære modeller. Data kan indsamles til formålet via survey-metoder (fx telefoninterview eller spørgeskemaer) eller bestå af eksisterende registerdata – eller en kombination. Ved at anvende registerdata er det overkommeligt at komme op på et stort antal observationer, hvilket er en fordel i den statistiske analyse.

Der er flere udfordringer ved denne metode. For det første er det vigtigt, at der haves oplysninger om alle de vigtigste forhold, som påvirker resultatet af indsatsen. Dette betyder, dels at man skal have en forståelse af, hvilke forhold der er vigtige (en god teori), og dels skal det være muligt at indsamle disse oplysninger. Udeladte variabler kan give misvisende sammenhænge. Det skal også bemærkes, at fundne statistiske sammenhænge

ikke nødvendigvis er kausale sammenhænge. Forsigtighed og et kritisk blik er vigtige elementer i analysen.

Når alle relevante data foreligger, er det en udfordring at udforme den bedste model. Dette er en balance mellem at inddrage relevante parametre og ikke at gøre modellen unødvendig detaljeret. Der findes metoder og traditioner for, hvordan denne afvejning bør ske.

En vigtig del af modelarbejdet handler om at få formuleret effekten af indsatsen rigtigt. Det kan være fristende at søge at dokumentere en gennemsnitlig effekt, fordi det er enkelt og ikke kræver en detaljeret teori. Imidlertid kan det være meget givtigt at søge fx at dokumentere, at visse grupper af personer, i visse specifikke sammenhænge, har en særlig markant glæde af indsatsen. Det betyder, modellen skal detaljeres således, at der estimeres effekter for forskellige grupper. Hvis den gennemsnitlige effekt af en indsats er begrænset, kan det have stor værdi at vide, at der for en delpopulation findes positive resultater. Det kan være anledningen til at udvikle indsatsen, fx ved at fokusere på disse grupper.

Eksempler på evalueringer baseret på statistiske analyser er fx Effekter af aktiveringsindsatsen (Arbejdsministeriet, 2000), Evaluering af aktivering ved hjælp af kvantitative metoder (kapitel 14 i denne håndbog), Evaluering af integration af etniske minoriteter (kapitel 10 i denne håndbog), Effekter af elevens uddannelsesforløb af forskelle i skoleudgifter pr. elev (kapitel 7 i denne håndbog). Analyse af energiafgifter, frivillige aftaler og investeringstilskud (fx Bjørner og Jensen 2002). Mohr (1995) beskriver en række evalueringsdesign med forskellige former for udvælgelse af deltagere og efterfølgende modelbaseret analyse.

4.5 **Effektevalueringer for få midler**

Hvis det er vigtigt at dokumentere årsag-virknings-sammenhængen, så er der ofte ingen vej uden om at analysere data i en form for statistisk model eller anvende det kontrollerede eksperiment. Dette kan imidlertid være omkostningsfuldt. Dels kræves mange data, og dels kræver det erfaring at gennemføre en statistisk analyse.

Hvad gør man, når det ikke kan forsvares at anvende så mange ressourcer?

En mulighed er at slække på kravet om en absolut beskrivelse af effekten, dvs. ændre spørgsmålet om, »hvilken effekt kan alene tilskrives denne aktivitet til et spørgsmål« til »kan det sandsynliggøres, at denne aktivitet har større effekt end tidligere eller lignende aktiviteter?« Ved at gøre antagelser om centrale forhold kan undersøgelsen billiggøres. Fx kan det antages, at den samfundsmæssige påvirkning har været den samme for to beslægtede aktiviteter, eller det kan antages, at deltagerne er sammenlignelige. Ulempen er, at troværdigheden og sikkerheden reduceres, når sådanne forhold ikke undersøges, men alene antages. Det bliver lettere for kritikere at stille spørgsmål til evalueringen.

Årsag-virknings-sammenhængen kan søges beskrevet ved hjælp af såkaldt skyggekontrol: I stedet for sammenligningsgruppe og en før-måling kan man spørge indsatsgruppen om deres vurdering af tidsperspektivet og betydningen af indsatsen. *Har kurset givet dig relevante kompetencer? Hvilke? Har de været medvirkende til, at du fik arbejde?* Det er oplagt, at dette introducerer en række usikkerhedselementer, men erfaringen viser, at selv indledende undersøgelser med betydelig usikkerhed kan give overraskende og relevante indsigter.

Mange evalueringer fokuserer på holdning og vurderinger i stedet for konkrete resultater eller adfærd. Hvis målet med aktiviteten ikke er holdning, men fx beskæftigelse, energibesparelser eller at undgå trafikulykker, så kan det være en særdeles usikker tilgang at beskrive viden og holdninger. Først og fremmest er det veldokumenteret, at der ikke er nogen entydig sammenhæng mellem holdninger og konkret adfærd. En positiv holdning til energibesparelser behøver ikke at hænge sammen med et lavt energiforbrug. Selv om det er umiddelbart logisk og forståeligt, at det ud fra et trafiksikkerhedssynspunkt må være positivt med en positiv holdning til at overholde fartgrænser, så kan den positive holdning let være det eneste, der er kommet ud af en kampagne.

Uanset hvilke evalueringsdesign der anvendes til effektevalueringer, så er det vigtigt i formidlingen at gøre opmærksom på de antagelser og begrænsninger, som ligger i valget af fremgangsmåde.

4.6 Effektevaluering med fokus på teori

I tilfælde, hvor det ikke er muligt at indsamle store datamængder og anvende statistiske metoder, kan en anden tilgang være at fokusere på at opbygge en teori, som beskriver væsentlige aspekter i forbindelse med indsatsen og dens effekter. En teori kan omfatte selve aktiviteten, deltageres forudsætninger og væsentlige aspekter af omgivelserne, men også en beskrivelse af, hvornår og i hvilke situationer aspekterne er vigtige. Teorien (som altså ikke behøver at være meget abstrakt eller kompliceret) kan underbygges af informationer fra den aktuelle aktivitet. Det kan være informationer fra interview, casestudier eller spørgeskemadata, som kan anvendes til at underbygge troværdigheden af teorien.

Pawson og Tilley (1997) giver i bogen *Realistic Evaluation* flere eksempler på teorier, som er ganske jordnære. De lægger vægt på, at der er mennesker indblandet, og at deres måde at forstå og agere på er vigtig i mange offentlige indsatser. Teorien bør omfatte sådanne forhold. De anbefaler, at man tager udgangspunkt i deltagerne og dem, som leverer indsatsen, når man skal beskrive en relevant teori. Man kan til en start interviewe et begrænset antal deltagere om, hvad der er vigtigt for at opnå en effekt – og om, hvad der kan hindre en positiv effekt. Efter disse indledende dataindsamlinger kan resultaterne kombineres med tidligere erfaringer, fx fra tidligere undersøgelser og forskning, og den indledende teori kan opstilles. Derefter kan der indsamles yderligere information på de områder, som vurderes at være mest kritiske i forhold til at opnå en troværdig beskrivelse. Der kan fx gennemføres en spørgeskemaundersøgelse med et større antal deltagere eller et casestudie, hvor et begrænset antal sager belyses fra flere sider. Casestudier er særlig velegnede til at bidrage med nye ideer og vinkler til en beskrivelse af relevant teori.

En sådan tilgang til effektevaluering ligger i grænseområdet op til procesevalueringer og kan evt. gennemføres med begge formål for øje.

Spørgsmålet om, hvorvidt evalueringen bidrager troværdigt til dokumentation af effekten, kan bedømmes af indsigtfulde, uafhængige personer. Hvis en sådan kritiker kan påpege forhold, som er overset, og som kan forventes at have væsentlig betydning – ja så er arbejdet ikke gjort godt nok.

5 Procesevaluering

Kjeld Høgsbro

I dette kapitel vil jeg se på de forskellige former for procesevaluering. Formålet med kapitlet er, at man så vidt muligt ved, hvad man kan vælge imellem, og har gjort sig klart, hvad man har brug for, når man skal evaluere eller rekvirere en evaluering. Den ene form for evaluering er ikke bedre end den anden, men de tilfredsstillende forskellige videninteresser.

Jeg vil endvidere komme ind på samspillet mellem forskere (konsulenter) og praktikere i forbindelse med procesevalueringer, herunder også de forskellige, indimellem modstridende forventninger til evaluator og spørgsmålet om evaluators loyalitet over for de involverede og evaluators relation til interne konflikter i organisationen. Denne del af afsnittet fokuserer muligvis for meget på de konflikter, der kan være, men så kan man jo glæde sig over de konfliktfrie forløb samtidig med, at man ikke bliver overrumplet af konflikterne, når de opstår.

5.1 Hvorfor vælge procesevaluering?

Effektevalueringer, som de er beskrevet i det foregående, hvor man undersøger effekten af en bestemt indsats, er i mange tilfælde fuldt tilstrækkelige. Men som det altid er tilfældet, drejer det sig om den videninteresse, evalueringen udspringer af. Eller for at stille det op som et konkret spørgsmål: *Hvad har man brug for at vide mere om?* Hvis man vil vide, om en institution lever op til de målsætninger, man har fastlagt for den, er effektevalueringer for eksempel helt relevante. Effektevalueringer kan også

bruges som en tilbagevendende kontrol med, at institutionerne overholder givne krav til standard og kvalitet.

Når vi taler om nye tilbud, nye institutioner og virksomheder, som stadig befinder sig i en udvikling, kan en effektevaluering imidlertid give anledning til en fejlagtig vurdering af en ny organisations *muligheder*.

Figur 5.1 giver en enkel illustration af dette problem. Hvis A og B er to organisationer, som vurderes ud fra nogle givne målsætninger, kan vi forestille os, at A fra starten var født med stærke ressourcer og en god evne til at løfte opgaven. Derefter opstod der imidlertid konflikter i personalegruppen. De, i øvrigt kompetente medarbejdere, arbejdede i hver sin retning og kæmpede indbyrdes om deres indflydelse på organisationen. Dette tappede organisationens ressourcer og stjal tiden og opmærksomheden fra de egentlige opgaver. Vigtige medarbejdere sagde op, og nye kom til. Ledelsen mistede kontrollen med udviklingen.

Anderledes gik det i B. Her var medarbejderne fra starten ret nyuddannede, og selv om ledelsen var fagligt kompetent, tog det noget tid, før de unge medarbejdere vænnede sig til arbejdsmarkedet og fik rutinen i udførelsen af opgaverne. Til gengæld var de fleksible og åbne over for hinanden og meget opmærksomme over for den fælles opgave og de tværfaglige udfordringer, den rummede.

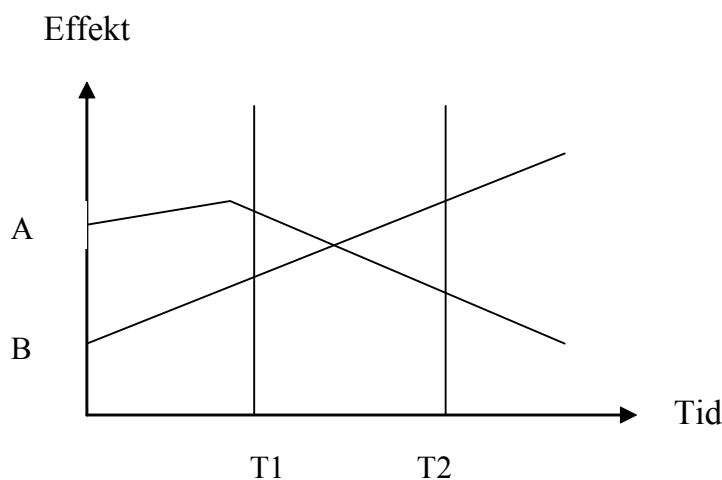
I årene efter deres oprettelse udviklede organisationerne sig vidt forskelligt. I A bredte konflikterne sig, hvilket efterhånden påvirkede resultatet i negativ retning. I B udviklede den individuelle kompetence sig hos hver enkelt medarbejder, samtidig med at samarbejdet mellem medarbejderne styrkedes. Dette påvirkede effekten i positiv retning.

En effektevaluering foretaget på tidspunkterne T1 og T2 vil give to forskellige resultater. I værste tilfælde kunne det føre til lukningen af B efter den første evaluering. Hvis vi ikke kender den udvikling, der foregår i A og B, vil en effektevaluering således indebære en risiko for en uhenigtsmæssig beslutning. Vi kan selvfølgelig vedtage, at vi nu har givet institutionen en tilstrækkelig tid til at udvikle sig, men vi kan også have et ønske om at få en større indsigt i organisationens *potentialer* og *muligheder*.

Her kommer procesevalueringerne ind i billedet. Disse evalueringsformer følger institutionen/virksomheden i en bestemt afgrænset periode

og vurderer udviklingen i organisationen og resultater. Effektevalueringen vil i ringe grad have data, som tillader, at man udtaler sig om, *hvorfor* resultaterne var gode eller dårlige. Procesevalueringen følger derimod organisationens udvikling indefra og kan pege på mulige forklaringer. Samtidig kan den komme i kontakt med en indre dynamik i organisationen, som på sigt vil kunne svække eller understøtte effekten af organisationens arbejde. Vi vil altså blive gjort opmærksomme på udviklingstendenser, der på sigt kan forbedre eller reducere effekten af organisationens arbejde. Man kan således forestille sig, at en god procesevaluering af A og B kunne pege på accelererende konflikter i A, allerede inden de fik indflydelse på effekten af A's arbejde, og at den kunne pege på »naturlige« begyndervanskeligheder i B, som ville løse sig hen ad vejen. Procesevalueringen ville kunne afsløre, at B's problemer skyldtes de nye medarbejders usikkerhed, og samtidig se mulighederne i deres engagement og ledelsens understøttelse af dette engagement.

Figur 5.1 Tænkt udviklingsforløb i to institutioner, A og B



5.2 Forskellige former for procesevaluering

Flere af de amerikanske introduktioner til evalueringsforskningen opdeler evalueringer i to typer: *Summative* og *formative*. Summative evalueringer tilstræber en indsamling af data, der kan belyse funktionen af organisationens indsats, så eventuelle eksterne bevillingsgivere kan beslutte, om det er umagen værd at give støtte til dette arbejde. Formative evalueringer har derimod til formål at forsyne aktørerne i organisationen med feed-back un-

dervejs i processen, så processen kan reguleres i positiv retning (Fitz-Gibbon og Morris 1987: 11, Herman, Morris og Fitz-Gibbon 1987:13). Der ligger ikke i denne skelnen en præmis om, at summative evalueringer ikke kan samle data over tid eller følge processen nøje, selv om summative evalueringer indimellem omtales synonymt med effektevalueringer (Erik Albæk 1988: 94). Derimod ligger der en langt mere alvorlig distinktion mellem begreberne summativ og formativ. Summative evalueringer søger at være objektive, neutrale og uafhængige af aktørerne. Formative evalueringer har derimod via deres feedback indflydelse på udviklingen og kan blive delagtiggjort i udviklingen. Bevillingsgiveren kan derefter stille sig det spørgsmål, om evaluator kan give en neutral vurdering af det produkt, vedkommende selv har været med til at skabe. Omvendt vil deltagerne inden for en summativt evalueret organisation berettiget kunne spørge, hvorfor evaluators kritiske betragtninger ikke med det samme blev forelagt organisationen, så man i tide kunne rette op på fejlene. Der opstår således et dilemma mellem på den ene side en videninteresse i objektive, neutrale beskrivelser af indsatsens organisering og resultater, og på den anden side det grundlæggende ønske om at bruge evalueringen direkte til en udvikling af organisationen. Dette dilemma fører til et valg mellem en summativ og en formativ evaluering¹.

Der rejser sig herefter en problemstilling, som inden for amerikansk evalueringforskning har ført til en skelnen mellem forskellige generationer af evalueringstraditioner. Inden for de første to generationer af evalueringer (i USA indtil 1960) lavede man evalueringer, som ukritisk tog udgangspunkt i de målsætninger, der var stillet op for organisationen. Inden for den tredje generation af evalueringer besluttede man, at dette var utilstrækkeligt. Man måtte ofte gå ind i en fundamental diskussion af organisationens målsætninger, fordi de indimellem var uklart og uprofessionelt formuleret, og indimellem var misforståede set i forhold til den overordnede funktion, indsatsen skulle have i et samfund under udvikling (Guba og Lincoln 1989: 29-31).

Men hermed var slagsmålet om evalueringforskningens dilemmaer ikke slut. Både summative og formative evalueringer havde inden for de første tre generationer af evalueringforskning det særkende, at de prøvede at operere på et værdineutralt grundlag for organisationens målsætninger

(Guba og Lincoln idid.). Men i praksis eksisterer der ikke en sådan værdineutral samfundsmæssig interesse på baggrund af hvilken, man kan formulere målsætningerne. Hvis vi fx taler om behandlingsmæssige tilbud til kræftpatienter, så kan de på en og samme tid øge patientens overlevelseschance og mindske vedkommendes livskvalitet (Høgsbro 1998). Patienten kan have en interesse i at få sagt ordentligt farvel, når det ikke kan være anderledes. Andre patienter kan blokere følelsesmæssigt over for et budskab om, at der ikke er mere at gøre, og den forskningsaktive læge kan have en interesse i at få afprøvet en ny behandling. Sygeplejerskerne kan modsat se deres professionelle faglige identitet knyttet til den livskvalitet, de kan sikre patient og pårørende. Afdelingen, eller den danske sundhedssektor som sådan, kan have en interesse i at øge den gennemsnitlige overlevelsestid, så tallene bliver pænere for Danmarks vedkommende. Medicinalindustrien kan have en interesse i at få et passende patientgrundlag for at kunne komme først med den medicin, alle håber på vil blive opfundet en dag. Livskvalitet og livsforlængelse fremstår således som modstridende præmisser for indsatsens målsætning. Disse præmisser skal afklares, før man kan tale om »god« og »dårlig« behandling.

Guba og Lincoln (1989:32-38) peger således på tre aspekter af den værdineutrale evalueringsforskning, som påkalder sig en kritisk opmærksomhed.

1. En tendens til ukritisk at identificere sig med de stærke aktørers, eventuelt bevillingsgiveres værdier og interesser. På denne måde bliver evalueringen et styringsinstrument i hånden på de stærke aktører. De øvrige involverede får ikke mulighed for at få vurderet indsatsen på deres eget grundlag.
2. En manglende evne til at forholde sig til den værdipluralisme, som hersker inden for det felt, man beskæftiger sig med. En manglende evne til at forholde sig til de forskellige aktørers forskellige krav til indsatsen og deres relation til forskellige værdigrundlag.
3. En ureflekteret binding til et naturvidenskabeligt paradigme, hvor det er forskerens opgave at finde frem til »sandheden« med al den autoritet, dette paradigme giver forskernes konklusioner.

I stedet foreslår de en evalueringstradition, som i langt højere grad lægger op til en kritisk vurdering af de forskellige aktørers krav til indsatsen. En tilgang, som går i dialog med de forskellige aktører og forstår sig selv som en del af en proces, hvor værdier, målsætninger og effektvurderinger brydes samtidig med, at organisation og indsats konstrueres i en uafsluttet bevægelse, hvor ingen, heller ikke evaluator, har det sidste ord og dermed retten til den endegyldige sandhed om indsatsen. De kalder denne tilgang for *responsiv konstruktivistisk evaluering* (ibid.: 38ff).

Denne kritik går igen i en lang række nye evalueringsformer i 90'erne, der således er udtryk for den fjerde generation af evalueringer. Disse kan alle kaldes videreudviklinger af den formative evaluering, idet evaluator går ind og påvirker processen undervejs, men de forholder sig mere bevidst til de forskellige aktørers og interessenters værdigrundlag og mål. Og de forsøger at fremme en dialogisk proces omkring værdier og målsætninger, som gør, at de forskellige aktører bliver konfronteret med disse forskelle og tvunget til at forholde sig til dem som en del af organisationens og indsatsens udvikling.

I det efterfølgende vil jeg omtale tre forskellige former for videreudvikling af den formative evaluering, som det er relevante at vælge imellem i en given situation. Det drejer sig om *BIKVA*, *Empowermentevaluering* og *Støttet selvevaluering*. En sådan omtale kan give anledning til uenigheder mellem forskellige forskere omkring disse definitioner, idet hver tradition har sine fortalere, som sikkert vil mene, at definitionen er for snæver. Desuden er der selvfølgelig altid tale om en forenkling, når evalueringsformer med et kompliceret design skal omtales med få sætninger. Men for overblikkets skyld kan det være gavnligt med en sådan kortfattet omtale af deres særlige karakteristika.

BIKVA (*Brugerinddragelse i kvalitetsvurdering*) er en evalueringsmetode, der navnlig retter sig mod organisationer, der dels indgår i et samspil med en bestemt brugergruppe, dels har en målsætning om at udvikle en større forståelse for denne gruppes behov og problemer. I BIKVA-modellen indgår evaluator i en *mediation*, hvor de to parter får bearbejdet deres forestillinger om hinanden. Kommunikationen forbedres, og man sikrer, at organisationens mål og midler befordrer et bedre samarbejde mel-

lem de to grupper. BIKVA-evalueringen begynder med at interviewe de forskellige grupper, hvorefter vurderinger, krav og værdigrundlag fremlægges for de andre grupper i en opadstigende spiral, hvor organisationens forskellige aktører bliver sig deres forskelligheder bevidst. Samtidig påbegyndes en proces, hvor grupperne må tage stilling til hinandens oplevelse af situationen og begynde at finde ud af, hvorvidt det er muligt at konstruere et fælles værdigrundlag for organisationen, og hvorvidt det er muligt at nå til enighed om en fælles vurdering af organisationens problemer og potentialer (Krogstrup 1996 og 1997).

Empowermentevaluering er ligeledes en specialiseret form for formativ evaluering, idet overvejelser omkring nødvendige feedback-mekanismer for organisationens udvikling er inde i billedet fra starten af. I den tid, evaluator er i organisationen, eksperimenteres der bevidst med den positive indflydelse, evaluator kan have på organisationens udvikling, og der fokuseres på spørgsmålet om, hvordan organisationen selv kan overtage disse funktioner når evalueringsperioden er slut. Empowermentevaluering foretager et entydigt valg af værdigrundlag for evalueringen. Målet er »simpelt hen« at udvikle organisationen ved at forbedre deltagernes evne til selvevaluering og refleksion. Evaluator fungerer som træner i denne proces, og træningen foregår på deltagernes egne præmisser (Fetterman 1996: 5).

Støttet selvevaluering er en discountudgave af empowermentevalueringen. Der er også her tale om, at man fokuserer på de evalueringsmetoder, organisationen selv kan håndtere, men evaluator indgår kun som rådgivende konsulent for denne evaluering. Der er derfor også tale om færdige koncepter for evaluering, idet evaluator ikke får nogen større indsigt i den pågældende organisations specielle forhold. Samarbejdet med evaluator kan dog i høj grad udvikle organisationens bevidsthed om sammenhængen mellem mål og midler og mulighederne for en større grad af kontrol med organisationens udvikling. En særlig udvikling af denne form for evaluering er foregået på Center for Evaluering i Århus. Konceptet er blevet kaldt »Praksisforskning i videnskabende netværk« (Ramian 2004, 1999 og 1997). Man starter med at efterspørge erfaringer vedrørende nogle bestemte problemstillinger inden for praktisk socialt arbejde. De praktikere, der har sådanne erfaringer, tilbyder man en optagelse i et netværk, hvor man lærer bestemte evalueringsredskaber og regler for videnskabelig dokumen-

tation. I den efterfølgende periode skal praktikerne så undersøge og dokumentere resultaterne af deres arbejde. Den fremadskridende selvevaluering bliver løbende diskuteret på seminarer, som samler praktikerne inden for netværket. Til sidst afrapporteres erfaringerne fra de forskellige forløb til offentligheden.

Fetterman omtaler det som en entydig opgave at fastlægge et værdigrundlag for empowermentevalueringen, og Guba og Lincoln kritiserede de første generationer af evalueringer for at ignorere værdikonflikter og identificere sig ukritisk med bevillingsgiverens værdigrundlag. Men ingen af disse bidrag formår at ophæve den vanskelige problemstilling omkring værdigrundlaget for evalueringen (House 1993: xvi). Internt er organisationerne nemlig ofte opdelt i grupper med forskellig værdigrundlag. Nogle grupper vil dele værdigrundlag med bevillingsgiverne, mens andre vil være i opposition til dette værdigrundlag. Grupperne kan også have forskellige fortolkninger af det fælles værdigrundlag.

På en lokal skole, der havde fået støtte fra et større nationalt udviklingsprogram, var målsætningen at komme ungdomskriminaliteten til livs ved at etablere »fornuftige« fritidsaktiviteter for de unge (Høgsbro et al. 1990: 67ff.). Men efterhånden som evalueringen skred frem, viste det sig, at denne målsætning var irrelevant for mange af deltagerne. Målsætningen var blevet til i et samspil mellem projektledere og lokale beboere, hvor ideen om forebyggelse af ungdomskriminalitet var en idé, der kunne skabe opbakning omkring projektet. For mange af deltagerne i projektet var målsætningen dog først og fremmest at få skabt nogle aktiviteter for de 10-14-årige i lokalsamfundet, som i en årrække havde truet med at udvikle sig til en soveby. Hvis evaluator havde taget den officielle målsætning alvorligt, ville vedkommende være kommet på en svær opgave, for der var stort set ingen ungdomskriminalitet i området.

I andre tilfælde kan evaluator komme i tvivl om, hvorvidt vedkommende skal identificere sig med bevillingsgiverens interesser, det professionelle personales eller brugernes. Muligvis kan evaluator opstille et fjerde værdigrundlag, som er mindst lige så relevant, og som afspejler en almen samfundsmæssig interesse. I eksemplet ovenfor skal vedkommende fx tage stilling til relevansen af at støtte fritidstilbud til de 10-14-årige under på-

skud af at ville forebygge kriminalitet. Uanset valget af evalueringsform må man således foretage et valg mellem forskellige værdigrundlag, hvilket rejser spørgsmålet, om man skal tage udgangspunkt i bevillingsgiverens værdigrundlag og målsætninger, i en af de involverede grupper eller i et tredje standpunkt, som forskeren anser for at repræsentere en almen samfundsmæssig interesse.

Indimellem kan det være nærmest omsonst at skulle presse den anvendte forskning ind i en form, hvor man på den ene eller den anden måde skal forholde sig til forholdet mellem effekter og målsætninger. I forbindelse med empowermentevalueringer kan man nå til en grænse, hvor evaluator i sin tilknytning til svagt stillede grupper får indkredset de kræfter, disse grupper er op imod, og begynder at arbejde som deres »advokat« (Fetterman 1996: 12-13). Hermed drejes fokus væk fra gruppernes egen aktivitet til et fokus på aktivitetens kontekst, dvs. de samfundsmæssige rammer for disse grupperes aktiviteter. Det kan blive centralt for gruppens »empowerment«, at den får en større forståelse for de kræfter, aktiviteten er op imod. Dermed har empowermentevalueringen bevæget sig over i den tradition, der er blevet kaldt *aktionsforskning*. En forskningstradition, hvor forskeren stiller sin forskning til rådighed for bestemte grupper for at forbedre deres evne til at handle strategisk korrekt i forholdet til stærke samfundsmæssige aktører. I et eksempel fra 90'ernes danske samfundsforskning drejede det sig om at udvikle et bæredygtigt produktionskoncept i fiskeriindustrien i et samarbejde mellem forskere og medarbejdere. Et sådant koncept skulle tage højde for både internt formulerede behov hos medarbejderne og eksterne forhold i form af markedet for produkterne og de organisatoriske sammenhænge, organisationen måtte respektere (Aagaard Nielsen 1996).

En anden udvikling væk fra evalueringens begreb foregår inden for det koncept, som kaldes *forskningsbaseret erfaringsopsamling* (Høgsbro et al. 1990, Høgsbro 1995, 1996 og Høgsbro et al. 1999). Denne tradition repræsenterer en udvikling af den oprindelige værdineutrale formative evaluering til et koncept, hvor man bevidst tager udgangspunkt i et alment samfundsmæssigt værdigrundlag, og fokuserer på aktiviteten som en praksis, der eksponerer bestemte sociale dynamikker og betingelser. I eksemplet med skolen ovenfor var det således svært og kunstigt at opfinde en ny mål-

sætning, som udviklingsprocessen kunne evalueres i forhold til. Det var straks mere interessant at se eksemplet som en case, der viser en særlig dynamik i samspillet mellem lokalsamfund og »ildsjæle«, hvor kriminalitetsfrygten i et lokalsamfund udnyttes til at få opbakning til nogle nødvendige fritidstilbud for de unge. Det interessante er her den kulturelle adskillelse mellem unge og ældre og spørgsmålet om, hvorvidt lokalsamfundet kan takle en sådan dobbelt begrundelse for aktiviteten, uden at tilbuddet bliver stigmatiseret i de unges øjne som et »tilbud til de kriminelle«, eller de unge i lokalsamfundet bliver stigmatiseret som »potentielt kriminelle«. Ved at definere evalueringdesignet som en forskningsbaseret erfaringsopsamling havde forskerne frihed til selv at definere, hvilke erfaringer fra projektet der havde en almen samfundsmæssig interesse. Selve evalueringsperspektivet blev dermed forladt til fordel for et fokus på aktivitetens kontekst, og hele aktivitetens udvikling blev set som et »socialt eksperiment«, der eksponerede vilkårene for kulturelle aktiviteter i danske lokalsamfund. Spørgsmålet var dermed ikke mere, om bevillingsgiveren havde fået noget for pengene, eller om indsatsen havde opnået den effekt, initiativtagerne forventede.

I forbindelse med forsøgs- og udviklingsprogrammer kan begge de to sidstnævnte traditioner være relevante alternativer til egentlige evalueringer. Aktionsforskningen, især når det drejer sig om initiativer, hvor det vigtigste er at få svage grupper til at organisere sig og få formuleret en fælles strategi for løsningen af deres situation, og den forskningsbaserede erfaringsopsamling, især når der er tale om et behov for at sikre sig, at erfaringerne med betingelserne for dette udviklingsarbejde bliver fastholdt.

5.3 **Videninteresser og valg af evalueringsform**

Som det skulle være fremgået af det foregående, er der en del valgmuligheder, man skal tage stilling til, før man vælger design for en evaluering. Valget mellem forskellige former for procesevaluering afhænger i første omgang af, hvilken viden man er interesseret i at få, hvem der skal have denne viden, og hvad den skal bruges til.

Hvilken viden er man interesseret i? Er man alene interesseret i, om indsatsen har været god eller dårlig, eller er man også interesseret i, hvad

der har gjort indsatsen god eller dårlig? Dette kan være et valg mellem effektvurdering og procesvurdering. Er man interesseret i at få afprøvet alternative strategier undervejs samtidig med evalueringen, skal man vælge de formative vurderingsmetoder. Er man primært interesseret i at udvikle organisationen optimalt på dens egne præmisser, skal man nok vælge empowermentvurderinger eller måske gå længere og vælge en form for aktionsforskning. Hvis samspillet mellem brugere og tilbud skal udvikles, bør man overveje BIKVA-metoden. Og hvis der er tale om forsøgs- og udviklingsprogrammer, som under alle omstændigheder slutter på et eller andet tidspunkt, og det mest interessante derfor er den erfaring, man har fået med den slags initiativer, skal man vælge en forskningsbaseret erfaringsopsamling.

Hvem har brug for denne viden? Hvem er »man« i spørgsmålene ovenfor? Er det en ekstern bevilgende myndighed, der vil vide, om de har fået noget for pengene, eller er det organisationen selv, som skal finde ud af sin egen placering i forhold til markedets behov og rekvirenternes krav? Dette er afgørende for det værdigrundlag og de målsætninger, der lægges til grund for evalueringen. Skal man eventuelt overlade det til forskeren at beslutte, hvad der er relevante mål for succes? Eller skal man lægge flere forskellige værdigrundlag og målsætninger til grund for evalueringen og diskutere de forskellige resultater, det giver?

Hvad skal denne viden bruges til? Skal den bruges til at forbedre den pågældende organisation, eller skal den alene levere grundlaget for en beslutning om institutionens fortsættelse? Skal den levere et input til et koncept for andre, eventuelt fremtidige, organisationer og tiltag på området, eller skal den pågældende viden alene bruges til at udvikle netop denne organisation?

Svarene på alle disse spørgsmål bestemmer både vurderingsform og måden den afrapporteres på. Hvis man vil prioritere ressourcerne til evalueringen rigtigt, er det derfor vigtigt at få afklaret disse spørgsmål. Den viden, som kan interessere andre, er måske banal og triviell for medlemmerne af organisationen, og omvendt kan interne konflikter, som medlemmerne har stor interesse for at få belyst, være ligegyldige og trivielle for andre. Effektmålinger, der kan være væsentlige for en bevilgende myndighed, kan være uden værdi for fagfolk, der er mere interesseret i at høre, hvad kon-

ceptet egentlig går ud på, og hvorvidt praksis har været i overensstemmelse med dette koncept.

Figur 5.2 giver et ultrakort og forenklet billede af de omtalte evalueringsformer med vægt på to af de ovennævnte dimensioner. I den vandrette dimension er spørgsmålet, om man primært er interesseret i viden om sammenhænge mellem mål og midler i den pågældende organisation, eller man tillige er interesseret i at vide noget om de ydre betingelser (konteksten) for arbejdet. I den lodrette dimension er spørgsmålet, om undersøgelsen primært skal tage udgangspunkt i organisationens eget værdigrundlag, eller om det i højere grad er andre aktører, eventuelle bevillingsgivere, som fastlægger, hvad der er interessant/uinteressant, og hvad der er præmisserne for vurderingen af indsatsens kvalitet.

Figur 5.2 De syv former for procesevaluering

Primært fokus	Mål-middel-sammenhænge	Kontekst sammenhænge
Primær målgruppe		
Organisationen selv	Empowermentevaluering Støttet selvevaluering	Aktionsforskning BIKVA
Esterne aktører, bevillingsgivere	Formativ evaluering Summativ procesevaluering	Forskningsbaseret erfaringsopsamling

I det første felt har vi de evalueringer, som tager udgangspunkt i organisationens eget værdigrundlag og fokuserer på udvikling af en evalueringspraksis inden for organisationen, som styrker organisationen internt og dens omdømme eksternt.

I andet felt har vi de evalueringer, som tager udgangspunkt i organisationens eget værdigrundlag, men eventuelt fokuserer lige så meget på omgivelsernes (kontekstens) betydning for organisationens succes. Visse faser i empowermentevalueringen kan bevæge sig over mod dette felt. BIKVA-metoden er placeret i dette felt, fordi den forudsætter, at organisationen tillægger brugerindflydelsen en positiv betydning og aktivt arbejder for at fremme dialogen mellem personale og brugere. Metoden er samtidig ikke entydigt rettet mod evalueringsaspektet, men først og fremmest rettet mod en udvikling af kommunikation og samspil inden for den kontekst, organisation og brugere er placeret i.

I det tredje felt har vi de evalueringer, som tillader sig en distanceret holdning til organisationens eget værdigrundlag og målsætning, enten fordi værdigrundlag og målsætning bestemmes af en ekstern bevillingsgiver, eller fordi forskeren sætter resultaterne op imod en almen samfundsmæssig funktion, som indsatsen antages at skulle have. I dette felt har jeg placeret det, som ovenfor er blevet kaldt 1. til 3. generation af summative og formative evalueringer.

I fjerde felt har jeg placeret den forskningsbaserede erfaringsopsamling som en type evalueringsforskning, der dels tager udgangspunkt i den almene samfundsmæssige funktion og betydning, bevillingsgiveren eller forskeren selv tillægger initiativet/indsatsen/tilbuddet, dels sætter fokus på de forhold i omverdenen, som modarbejder eller understøtter de pågældende initiativer/indsatser/tilbud.

5.4 Evaluators rolle

Evaluators rolle i forbindelse med procesevalueringer er alt andet end ukompliceret. Følgende spørgsmål, som let bliver rettet mod evaluator, kan give et billede på de konflikter, som indimellem kan opstå:

Hvad er det egentlig, han vil?

Synes hun egentlig, det er godt nok, det vi laver?

Har han overhovedet forstået vores specielle arbejdsmetoder?

Hvilke forudsætninger har hun egentlig for at forstå det, vi arbejder med?

Hvad mon han vil skrive om os i den endelige rapport?

Vil vi overhovedet selv få noget ud af al den tid, vi bruger på evalueringen?

Disse spørgsmål bliver selvfølgelig særligt påtrængende, når evaluator på forhånd har til opgave at distancere sig fra organisationens eget værdigrundlag og anlægge en udefra kommende kritisk vurdering af organisationens målsætninger og evne til at leve op til dem. Evaluator vil her ofte skulle kæmpe for at få plads til evalueringen. Vedkommende skal indimellem fastholde aktørerne på det opdrag, de har fået sammen, og insistere på adgang til aftalte data. Det kan være, at man oprindeligt har lagt en plan for evalueringen, som har været meget ambitiøs, og som ikke er tilstrækkelig

grad har taget højde for den belastning, dataindsamlingen vil påføre de stakkels medarbejdere. Det kan også være, at en ekstern rekvirent har bevilget penge til et ganske særligt formål, som vedkommende gerne vil have evalueret. Men i løbet af evalueringsperioden glemmes det fra tid til anden, hvad bevillingen egentlig var givet til, og evaluator må så fastholde organisationen omkring den fælles forpligtelse.

Som eksempel kan nævnes et forsøg med supervision af personalet på en kræftafdeling (Høgsbro 1998). Supervisionen, der blev varetaget af en ekstraordinært ansat psykolog for eksterne midler, skulle forbedre personalets evne til at tale med patienter og pårørende. Den eksterne bevillingsgiver ville i denne forbindelse gerne have evalueret effekten af en sådan ekstraordinær foranstaltning. Men da først psykologen var blevet en fast del af dagligdagen, opdagede personalet, at det var meget rart for arbejdsmiljøet, at personalet kunne snakke med en psykolog om deres problemer i almindelighed, og andre afdelinger syntes, det var uretfærdigt, at de ikke også fik deres del af psykologens tid. Her måtte evaluator gå ind og fastholde det egentlige formål med ansættelsen.

For alle typer evalueringer gælder det imidlertid, at organisationer er fyldt med mennesker, om man så må sige. Og disse mennesker er inddelt i grupper, hvis målsætninger for organisationen er mere eller mindre forskellige. Ledelsen bygger ofte en pragmatisk bro mellem disse forskelle, forsøger at gennemsætte sine egne eller nedtone forskellene, så organisationen fremstår mere homogen, end den er. Så snart evaluator er til stede, oplever medlemmerne, at bestemte målsætninger bliver officielt blåstemplet, og der kan opstå en kamp om, hvem der kan påvirke evaluators opfattelse af målsætningen.

Mange målsætningsdiskussioner kan skjules under intetsigende højprestige-formuleringer. Og betydningsfulde medarbejdergrupper kan have den indstilling, at de ikke gider målsætningsdiskussioner, fordi de »alligevel ikke betyder noget i den praktiske dagligdag«. Evaluator har imidlertid til opgave at evaluere organisationen og må derfor insistere på, at målsætningerne bliver så konkrete, at de kan stilles op mod målbare effekter. Dette kan medføre irritation, beskyldning for pedanteri og forstyrrelse af den konsensus, man oplevede, der var omkring opgaveløsningen.

En organisation indeholder også ofte et hierarki af professioner og

grupper, som tildeles forskellig betydning for organisationens opgaveløsning. Når evaluator begynder at spørge ind til denne forståelse, opfattes det allerede, som om evaluator stiller spørgsmål ved berettigelsen af denne rangordning, og nogle grupper begynder at ane konsekvensen af en evaluering, som stiller spørgsmål ved deres plads i hierarkiet.

Når grupper føler sig truet, vil et meget belejligt trick være at mistænkeliggøre evaluators hensigt, indsigt og metoder. Dermed har man taget en væsentlig del af brodden af evaluators kritik og den forskydning af gruppernes indflydelse, som kan blive resultatet. Og da det er svært at komme med en sådan kritik efter, at evalueringens resultat er gået én imod, må man hellere, for en sikkerheds skyld, sprede en generel mistænkeliggørelse allerede i processen.

Selv om evaluator således forsøger at holde sig neutral i processen, tilstræber en ligelig inddragelse af alle grupper, omtaler alle med respekt uden hverken at rose eller dadle nogen bestemt gruppe, kan vedkommende blive både syndebug og fælles fjende i processen. Generelt kan man sige, at jo større indre konflikter og modsætninger der eksisterer i organisationen, jo hårdere er det at være evaluator. Selv om man har de bedste intentioner om at støtte organisationens udvikling og grundlæggende sympatiserer med dens målsætninger.

Alt andet lige vil de evalueringsformer, som tager udgangspunkt i organisationens eget værdigrundlag og målsætninger, være mindst konfliktskabende. Når deltagerne føler sig overbevist om, at synliggørelsen af dens resultater ikke er noget, som udenforstående har påtvunget dem, og når de oplever, at den mest barske kritik kun har til formål at forbedre organisationen ud fra deres egne præmisser, vil de være mere åbne over for at stille deres viden og erfaring åbent til skue.

Det samme kan til en vis grad siges om den forskningsbaserede erfaringsopsamling, idet denne ikke primært fokuserer på, om deltagerne har været gode eller dårlige til at løse opgaven, men primært fokuserer på de erfaringer, man har fået med denne type tilbud/indsats og de udfordringer, sådanne initiativer er konfronteret med. Den forskningsbaserede erfaringsopsamling kan herefter udvikle sig til et berigende samarbejde mellem forskere og praktikere, hvor man afsøger muligheder og grænser for indsatsen på det pågældende område.

Sektordelen

6 Vejleder til sektordelen

Formålet med nedenstående oversigt er at give læseren en mulighed for hurtigt at få en oversigt over eksemplerne i de enkelte sektorkapitler. Hvis man fx er interesseret i en bestemt metode eller evalueringsmodel, kan man se, i hvilke eksempler de er brugt.

Oversigten tilbyder således læseren en nøgle til at finde de sektorkapitler og eksempler, der er relevante.

I nedenstående skema er i første kolonne angivet for hvert **sektorkapitel** (skrevet med fed skrift) de *eksempler* på konkrete evalueringer, som kapitlerne indeholder (skrevet med kursiv).

For hvert eksempel kan læseren i den vandrette række se en kort angivelse af den pågældende evaluering 1) formål, 2) de temaer eller hovedspørgsmål, evalueringen belyser, 3) det paradigme, der ligger bag de anvendte metoder, 4) de evalueringsmodeller, der er anvendt, og endelig 5) de anvendte dataindsamlingsmetoder.

De forskellige evalueringsmodeller og -paradigmer er nærmere beskrevet i kapitel 1.

Oversigt over eksemplerne i hvert sektorkapitel

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Skoleområdet <i>Effekter på elevernes uddannelsesforløb af forskelle i skoleudgifter pr. elev</i>	Kan udgøre en del af det politiske beslutningsgrundlag vedr. tildeling af ressourcer til folkeskolen.	Har forskelle mellem kommunerne i ressourceanvendelse pr. elev en effekt på, hvor langt eleverne når i uddannelsessystemet efter folkeskolen?	Empirisk-analytisk	Effektevaluering	Datagrundlaget er 1) AKF's forløbsregister, for en 10% tilfældig stikprøve af befolkningen, hvortil der er koblet oplysninger om forældre, 2) ekstra registerdata for elevernes skoleforløb i 8. og 9. klasse.
Skoleområdet <i>PISA-undersøgelsen</i>	Bidrager med viden om, hvorfor danske unges kompetencer er, som de er, sammenlignet med unges i andre lande.	Undersøgelsen vedrører tre faglige færdigheder: læsning, matematik og naturfag med fokus på læsning. Målinger af elevernes personlige og sociale kompetencer. I alt 32 lande har deltaget i undersøgelsen.	Empirisk-analytisk	Effektevaluering	Test inden for læsning, matematik, naturvidenskabelige fag og personlige kompetencer.

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Skoleområdet Kommunernes vejledning af unge om uddannelse og erhverv	At sikre, at flest mulig unge kommer videre i ungdomsuddannelsessystemet efter afsluttet grundskole.	Hvordan er kommunens vejledning af unge om uddannelse og erhverv organiseret? Hvordan søger kommunerne at sikre, at unge får den nødvendige vejledning? Hvilket indhold, metoder og målsætning har vejledningen? Hvordan er de unges fremtidsforventninger?	Empirisk-analytisk	Procesevaluering	Spørgeskemaundersøgelsen til kommunernes koordinator/leder af ungdomsvejledningen. Casekommuneundersøgelser ved interview med vejledere og ledere samt elever fra 9. eller 10. klasse.
Børn og unge <i>Evaluering af fritidstilbudene til børn og unge i en kommune</i>	At pege på typer af aktiviteter, som børn og unge ønsker, samt foreslå nye typer af aktiviteter.	Hvordan anvender børn og unge deres fritid? Er der særlig udsatte grupper af børn og unge uden faste fritidsaktiviteter? Savner børn og unge aktivitetsmuligheder i deres fritid? Hvad synes børn og unge er vigtigt i et fritidstilbud? Hvilke fritidsaktiviteter efterlyses af børn og unge?	Empirisk analytisk	Brugerevaluering.	Interview med nøglepersoner på området. Spørgeskemaundersøgelser, blandt skolernes 4- til 10-klasseelever, og blandt de unge mellem 16-18 år uden for folkeskolen. Interviewundersøgelse blandt grupper af unge, som ikke forventedes at deltage i spørgeskemaundersøgelsen.

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Børn og unge <i>Børns udvikling og trivsel i en kommune</i>	Danne baggrund for en drøftelse blandt kommunens børn og unge, skolens lærere, forældre og forvaltninger om miljøet og trivslen blandt børn og unge i skolerne.	Hvordan oplever eleverne det faglige undervisningsmiljø i skolen? Hvordan oplever eleverne det sociale liv i klassen? Hvordan er elevernes faglige niveau set i forhold til deres oplevelse af læringsmiljøet og det sociale liv i klassen? Hvilke sammenhænge er der mellem hjemmemiljø, fritidsanvendelse og elevernes faglige miljø og trivsel i skolen? Hvilke børn og unge befinder sig i en særlig »truet« position?	Empirisk-analytisk	Brugerevaluering	Spørgeskemaundersøgelse stilet til alle børn i folkeskolernes 5.-9. klasse. Faglige færdighedsprøver i matematik, stavning og læsning til de samme børn.

Emne: <i>Sektor: Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Børn og unge <i>En kommunes indsats over for særlig socialt udsatte børn og unge.</i>	At danne grundlag for vurderinger af, hvilken støtte og hjælp der bør ydes børn og deres forældre for at forebygge eventuelle senere eller yderligere vanskeligheder hos børnene.	Hvordan oplever forældre og børn selv familiens situation? Hvilken betydning har forældrenes brud for børnenes dagligdag? Hvilken støtte og hjælp får børn og forældre til afhjælpning af deres problemer? Hvilken yderligere eller alternativ støtte bør tilbydes?	Empirisk-analytisk. Fortolkningsparadigme	Brugerevaluering. Interessentevaluering.	Tre casestudier, som omfattede: En journalundersøgelse af et udsnit af familiesager. En kvalitativ interviewundersøgelse med mødre, fædre og børn fra opløste familier. Interview med sagsbehandlere og andre medarbejdere.
Indsats over for ældre <i>Akutstue i kommune</i>	Inspiration for aktører i primær og sekundær sundhedssektor vedr. alternativer til indlæggelser	Forebygges indlæggelse på sygehus? Dækkes behov?		Effektevaluering Procesevaluering	Interview med hhv. ledere og medarbejdere. Dokumentaranalyse. Sygehusstatistik.
Indsats over for ældre <i>Hjemmehjælpernes arbejdsprincipper</i>	Bidrag til at udvikle arbejdet i den kommunale hjemmepleje	Hvad betyder hjemmehjælpernes arbejdsmåde for modtagernes velbefindende?	Empirisk-analytisk Fortolkningsparadigme	Procesevaluering Effektevaluering	Spørgeskema til ledelse og adm. medarbejdere. Spørgeskema til hjemmehjælpere. Interview med hjemmehjælpere.
Integration af etniske minoriteter <i>kommunernes integrationsindsats</i>	Benchmarking af kommunernes succes med integration for at forbedre indsatsen	Hvad er effekten af de enkelte kommuners samlede integrationsindsats?	Empirisk-analytisk	Effektevaluering	Registeranalyse baseret på administrative registre.

Emne: <i>Sektor: Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Integration af etniske minoriteter <i>Unge overgang fra grundskole til ungdomsuddannelse</i>	At forbedre indsatsen for at styrke unges uddannelsesvalg og forblive på påbegyndt uddannelse.	Hvad er afgørende for, hvordan unge fra etniske minoriteter klarer sig i uddannelsessystemet?	Empirisk-analytisk.	Procesevaluering. Brugerevaluering.	Spørgeskemaundersøgelse. Kvalitative interview.
Integration af etniske minoriteter <i>Integration af tosprogede børn</i>	Er ressourcerne anvendt hensigtsmæssigt – evt. anden anvendelse?	Hvordan klarer børnene sig i daginstitutioner? Hvad forventer interessenterne?	Empirisk-analytisk.	Interessentevaluering.	Semi-strukturerede interview. Gruppeinterview. Spørgeskemaundersøgelser.
Handicapområdet <i>Rehabilitering af mennesker med hjerneskader</i>	Grundlag for at forbedre samt offentliggøre erfaringer med en holistisk genoptræningsmodel	Implementering af ny behandlingsmodel.	Fortolkningsparadigme	Formativ evaluering	Spørgeskemaer til professionelle klienter og pårørende over en 4-års periode. Fokusgruppeinterview med repræsentanter for faggrupper.
Handicapområdet <i>Amtslig rådgivning</i>	At skabe overblik, opmærksomhed og give ideer til forbedring af rådgivningen	Hvad er brugernes udbytte af rådgivningen? Hvilke ydelser leveres til brugerne? Hvordan er samarbejdet mellem rådgiver og brugere?	Empirisk-analytisk Fortolkningsvidenskabelige	Effektevaluering Procesevaluering	Spørgeskemaer til kommunale medarbejdere. Gruppeinterview med kommunale medarbejdere.

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Socialt arbejde for udsatte grupper <i>Tilbud til hjemløse, misbrugere og sindslidende</i>	Som led i Socialministeriets evalueringsprogram 2001 af danne grundlag for politiske beslutninger på området.	Hvilke modeller havde man for indsatserne? Er tilbuddene dækkende? Er der forskelle i den måde forskellige grupper er dækket af tilbud?	Empirisk-analytisk Fortolkningsvidenskabelige	Procesevaluering Effekt på implementering af lovgivning	Etnografisk feltundersøgelse om brugernes livsverden mv. Auditundersøgelse. Interview med nøglepersoner i kommuner, amter og frivillige organisationer. Spørgeskema blandt brugerne.
Arbejdsmarked <i>Anvendelsen af andre aktører i beskæftigelsespolitikken (kontaktforløbet).</i>	At forberede beslutningsgrundlaget for arbejdsmarkedsrådene i tre AF-regioner. At fremme AF's tilrettelæggelse af en hensigtsmæssig administration i samspillet med andre aktører.	Evalueringen har tre overordnede temaer: 1. Processen/samspillet mellem AF og aktørerne. 2. Deltagernes opfattelse af opholdet hos anden aktør. 3. Effekten af andre aktørers indsats.	Empirisk-analytisk	Proces- og effekt-evaluering. Interessentevaluering.	Interview med nøglepersoner i tre AF-regioner samt hos andre aktører. Interview med ledige, der har deltaget i forløbene. Fokusgruppeinterview med deltagerne. Spørgeskemaanalyse blandt alle ledige, der har deltaget på de relevante hold. Registeranalyse.

Emne: <i>Sektor: Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Arbejdsmarked <i>Lokale koordinationsudvalg</i>	Give baggrund for, at udvalgene kan forbedre deres indsats.	På hvilken måde har udvalgenes indsats medvirket til at fremme det rummelige arbejdsmarked på det lokale plan og forbedret samarbejdet mellem de berørte aktører?	Empirisk-analytisk	Proces- og effekt-evaluering.	Spørgeskemaundersøgelser til medlemmer af lokale koordinationsudvalg og sekretariater. Casestudier af udvalg.
Aktivering af ledige <i>Simpel effektmål for aktivering</i>	At give en kommune grundlag for selv at måle effekter af aktivering.	Hvorledes påvirker aktivering sandsynligheden for at finde et job?	Empirisk-analytisk	Effektevaluering	Registeranalyse på grundlag af kommunale registre.
Aktivering af ledige <i>kompleks effektmål for aktivering</i>	At give grundlag for tilpasning af kommunernes aktiveringspolitik.	Hvorledes påvirker aktivering sandsynligheden for at finde et job under hensyn til andre faktorer?	Empirisk-analytisk	Effektevaluering	Registeranalyse på grundlag af kommunale registre.

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Arbejds miljø <i>Udviklingstendenser i arbejdsmiljøet på syv visionsområder</i>	Overvågning af regeringens handlingsprogram for et rent arbejdsmiljø 2005.	Sker der en udvikling på arbejdsmiljøområdet inden for handlingsprogrammets ti-årige periode? Metodisk sigte: at udvikle redskaber til at følge ændringer i arbejdsmiljøarbejdet.	Empirisk-analytisk	Monitorering af arbejdsmiljøets tilstand i relation til regeringens arbejdsmiljøpolitik. En slags målopfyldelsevaluering på policy-niveau.	Registeroplysninger om helbred og arbejdsforhold fra flere registre og fra eksisterende spørgeskemaundersøgelser. Virksomhedsundersøgelse om forebyggende aktiviteter, telefoninterview samt observationer og besøgsinterview. Uddrag fra journalsystemer om gennemførte indsatser.

Emne: <i>Sektor: Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Arbejds miljø <i>Kvalitetsudvikling i forflytning (bevægeapparatet)</i>	Udviklingsprojekt for at mindske nedslidningen i plejearbejdet gennem nye arbejdsmetoder på plejehjem.	Bliver det fysiske arbejdsmiljø bedre? Kan der skabes kvalitetsudvikling og forankring af viden på arbejdspladsen?	Empirisk-analytisk	E kombination af proces- og effekt-evaluering	Registeroplysninger om sygefravær, arbejdsskader og personaleomsætning. Spørgeskemaundersøgelse vedr. besvær i bevægeapparatet. Spørgeskemaundersøgelse vedr. besvær i bevægeapparatet, ressourcepersonordning, forflytningsteknik, proces og forankring. Fokusgruppeinterview med ressourcepersoner og den øverste ledelse.
Miljøområdet <i>Skovrejsning</i>	Dialog med andre myndigheder om del af regionplan	Er det ud fra en samfundsøkonomisk betragtning mest hensigtsmæssigt at drive landbrug eller rejse skov på givne områder?	Empirisk-analytisk	Effektevaluering. Cost-benefit-analyse	Registeranalyse Foreliggende analyser
Miljøområdet <i>Organisk affald</i>	Grundlag for overvejelser om at øge genanvendelse af organisk affald	Hvordan fås mest miljø for pengene i håndtering af organisk dagrenovation?	Empirisk-analytisk	Effektevaluering. Cost-benefit-analyse.	Registeranalyse Foreliggende analyser

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Energi <i>Frivillige energiaftaler – Implementering og effektivitet</i>	Sigtet var at komme med anbefalinger til politiske beslutningstagere – særligt EU-kommissionen – om, hvordan reguleringen på CO ₂ -området kunne forbedres.	Hvad er elementerne i det danske aftalesystem? Hvordan iværksættes de konkret på virksomhederne?	Empirisk-analytisk, Fortolkningsvidenskabelige	Kombineret proces- og effektevaluering.	Casestudier af energintensive virksomheder inden for papirindustrien og mælkekondenseringsindustrien. 22 semistrukturerede interview, som blev suppleret med andet dokumentarisk materiale.
Energi <i>Evaluering af vindkraftudbygningen</i>	Udgangspunktet var en politisk debat om vindkraftens fordelagtighed. Det var hensigten at belyse de samfundsmæssige hensyn.	Vindkraftudbygning på 1000 MW sammenlignes med to referencealternativer: Et kulfyret henholdsvis et naturgasfyret kraftværk. Vindkraften er interventionen, og det kul/naturgasfyrede kraftværk er det kontrafaktiske tilfælde.	Empirisk-analytisk	Økonomisk model: Cost-effectiveness.	De »rene« produktionsomkostninger for vindkraft, dvs. eksklusive moms og afgifter. Værdisætning af fordele og ulemper, som vindmøllerne har i forhold til traditionelle kul- eller naturgasfyrede kraftværker.

Emne: Sektor: <i>Eksempler</i>	Formål	Temaer eller hovedspørgsmål	Paradigmer	Evalueringsmodeller	Metoder og datakilder
Kultur <i>Economic impact analyse af Brandts Klædefabrik i Odense</i>	Kulturministeriet ønskede at få økonomisk impact-metoden og de positive resultater fra udenlandske undersøgelser afprøvet i dansk sammenhæng.	Brands Klædefabriks betydning for forbrug, indkomst og beskæftigelse skulle belyses.	Empirisk-analytisk	Effektevaluering. Økonomisk model: Cost-effectiveness.	Publikumsinterview: Personlige interview med publikum på stedet. Regionale og nationale økonomiske effekter beregnet ved hjælp af samfundsøkonomiske modeller.
Kultur <i>Værdisætning af Det Kgl. Teater</i>	Ikke-brugerne finansierer størstedelen af teatrets drift over deres skattebillet. Det er derfor interessant at undersøge, om borgerne generelt synes, at Det Kgl. Teater »er pengene værd«. Dette kan have en umiddelbar politisk interesse.	Hvordan er borgernes betalingsvillighed for det Kgl. Teater? og hvad motiverer deres betalingsvillighed?	Empirisk-analytisk	Økonomisk model: betalingsvillighed.	En tilfældigt udvalgt stikprøve af den danske befolkning over 16 år. I alt 1.843 personer er blevet interviewet pr. telefon.

7 Evaluering på skoleområdet

Eskil Heinesen, Torben Pilegaard Jensen og Jill Mehlbye

Der har i mange år været en særlig interesse for at foretage evalueringer på uddannelsesområdet og ikke mindst på skoleområdet. En af årsagerne hertil er, at udgifterne på dette område er betydelige og er vokset kraftigt i de sidste mange tiår i takt med, at uddannelse har fået en stadig større betydning i det moderne samfund. Der er derfor et betydeligt og voksende behov for at sikre, at ressourcerne på området udnyttes bedst muligt. Der har i Danmark og i andre lande været foretaget et meget stort antal meget forskellige evalueringer relateret til grundskolen og andre dele af uddannelsessystemet. Evalueringerne har en meget forskelligartet natur, idet de fokuserer på forskellige dele af skoleområdet, på forskellige tiltag, pædagogiske principper, undervisningsformer eller ressourcer, og på forskellige processer, resultater eller outcomes. Det er derfor ikke enkelt at opstille generelle kriterier eller retningslinjer for, hvad der kendetegner gode evalueringer på uddannelsesområdet. På hjemmesiden for American Evaluation Association (www.eval.org) findes (under *publications*) forskellige vejledninger i og standarder for evaluering, hvoraf nogle er specielt målrettet uddannelsesområdet, herunder skoleområdet.

I Danmark er etableret et særligt institut, som ifølge loven har til opgave, at »samle national og international erfaring med uddannelsesevaluering og kvalitetsudvikling samt løbende at udvikle og forny metoder til evaluering og kvalitetsudvikling«, nemlig Danmarks Evalueringsinstitut. På instituttets hjemmeside (www.eva.dk) findes (under *metode*) en generel beskrivelse af de elementer, der typisk indgår i instituttets evalueringer.

Den betydelige interesse for evalueringer på skoleområdet har været en af drivkræfterne bag meget omfattende internationale undersøgelser,

hvor elevernes færdigheder på en række områder måles på sammenlignelig måde i forskellige lande, ligesom væsentlige aspekter af de forskellige landes skolesystemer søges sammenlignet. Således har International Association for the Evaluation of Educational Achievement (IEA) foretaget sådanne undersøgelser for en række lande (dog ikke Danmark) i 1964, 1970-71, 1980-84 og 1994-95. Data fra ikke mindst den sidste undersøgelse, kaldet TIMSS (Third International Mathematics and Science Study), har været anvendt i en række analyser, herunder evalueringer. En anden omfattende international undersøgelse kaldes PISA (Programme for International Student Assessment). Den dækker OECD-medlemslande (herunder Danmark), samt enkelte andre lande. AKF indgår i det konsortium, der har stået for den danske del af PISA-undersøgelsen. Undersøgelsen beskrives nærmere i afsnit 3 nedenfor. Datamaterialet fra både TIMSS og PISA kan benyttes til komplekse evalueringer af forskellige landes skolesystemer. I Danmark har PISA-undersøgelsen vakt betydelig opsigt, ikke mindst fordi den viser, at danske elevers færdigheder på en række centrale områder i gennemsnit er ringere end i andre nordiske lande, samtidig med at Danmark har et højt udgiftsniveau på skoleområdet.

I dette kapitel fokuseres specielt på forskellige former for effektevaluering ved hjælp af kvantitative undersøgelser, og der gives et eksempel på en kvalitativ evaluering.

Først diskuteres generelt de særlige udfordringer, der er i forbindelse med effektevalueringer på skoleområdet. Dernæst beskrives i afsnit 2 et eksempel på evaluering af effekten af kommunale forskelle i skoleudgifter pr. elev på elevernes uddannelsesforløb efter grundskolen. I afsnit 3 beskrives et eksempel på en evaluering baseret på den omtalte PISA-undersøgelse. Desuden gives et eksempel på en overvejende kvalitativ evaluering af vejledningen af de unge i overgangen fra grundskolen og videre til ungdomsuddannelse og erhverv.

7.1 Særlige udfordringer ved evaluering af effekter af skoleressourcer

Der er en meget omfattende international litteratur om statistisk evaluering af effekter af skoleressourcer. De mål for skoleressourcer (eller input), der anvendes, kan fx være udgifter pr. elev, klassestørrelse, elevlærer-ratio eller lærernes uddannelsesniveau. Effekterne måles normalt på output eller outcomes som elevernes karakterer eller testresultater i særligt tilrettelagte prøver. Der er også studier, som måler effekterne på den løn, eleverne opnår på arbejdsmarkedet efter afslutning af deres uddannelse, og enkelte undersøgelser belyser effekter på elevernes uddannelsesvalg efter grundskolen, fx antal års uddannelse, eller om de gennemfører en uddannelse på et givet niveau (fx svarende til en dansk ungdomsuddannelse). Der er i litteraturen ikke konsensus om størrelsen af effekterne, eller om der overhovedet er statistisk signifikante effekter. Forskellige undersøgelser kommer til meget forskellige resultater. En årsag hertil er, at input og output (outcomes) i »produktionsprocessen« kan måles på mange måder.

En anden årsag er, at det metodisk er meget vanskeligt at måle egentlige *kausale* effekter (jf. kapitlet om effektvurdering) af skoleressourcer (dvs. at bestemme årsagssammenhænge), fordi det niveau af skoleressourcer, en given elev modtager, kan være korreleret med elevens grundlæggende evner eller forudsætninger (dvs. at det er endogent). Det er med andre ord ikke tilfældigt, om en given elev modtager et højt eller et lavt niveau af skoleressourcer. Det afhænger af en række valg, som tages af forældre, lærere, skoler, kommuner og stat. Hvis fx forældre, som investerer meget tid i deres børns uddannelse, også vælger en skole til deres børn, som har mange ressourcer pr. elev (enten via valg af bopæl eller via valg af privatskole), vil der være en tendens til, at effekten af skoleressourcer overvurderes i empiriske analyser. Hvis omvendt forældrene reducerer lektiehjælp mv., hvis børnene går på en god skole med mange ressourcer, så vil der være en tendens til, at effekten af skoleressourcer undervurderes i empiriske undersøgelser. Hvis kommunerne allokere (fordeler) forholdsvis mange ressourcer til skoler, hvor en stor andel af eleverne har indlæringsvanskeligheder eller kommer fra socialt svage hjem, vil der også være

en tendens til at undervurdere effekten af skoleressourcer i empiriske undersøgelser. Tilsvarende hvis sådanne kriterier benyttes på den enkelte skole i fordelingen af ressourcer mellem forskellige klassetrin eller mellem forskellige klasser på samme klassetrin. Der kan fx være en tendens til, at en tilgang af elever til en skole primært placeres i klasser, der fungerer godt, på trods af at de i forvejen er større end klasser, der fungerer mindre godt. Og hvis en klasse fungerer dårligt, kan der være en tendens til, at eleverne i højere grad forlader skolen til fordel for andre skoler, hvorved klassen bliver mindre relativt til andre klasser, således at de målte skoleressourcer (i form af fx lærertimer pr. elev) øges.

Hvor alvorlige sådanne potentielle problemer med endogenitet er for en given empirisk undersøgelse, kommer blandt andet an på, hvordan skoleressourcer måles. Hvis der er tale om ressourcer pr. elev (fx udgifter pr. elev) på kommuneniveau, vil allokeringsbeslutninger inden for kommunen ikke have betydning (bortset fra, at de kan påvirke sandsynligheden for, at bestemte grupper af forældre vælger privatskole), ligesom forældrenes valg af bopæl (skoledistrikt) inden for kommunen heller ikke vil have betydning. Hvis aggregeringsniveauet for den variabel, der måler skoleressourcer, er højt (fx lærer-elev-ratio på kommuneniveau i stedet for lærer-elev-ratio på skoleniveau), så vil endogenitetsproblemerne alt andet lige være mindre. Det kan være en forklaring på, at man ofte finder de største effekter af skoleressourcer i undersøgelser, der måler skoleressourcer på et højt aggregeringsniveau (se fx Burtless, 1996).

I statistiske undersøgelser er den mest almindelige måde at tage højde for endogenitetsproblemer på at inddrage så mange relevante baggrundsvARIABLER som muligt i den statistiske model, der estimeres. Det er typisk variable vedr. elevernes forældrebaggrund og de socioøkonomiske forhold i lokalområdet. Det har også været strategien i nogle af AKFs undersøgelser, fx den, der beskrives nærmere i afsnit 2 nedenfor. I denne undersøgelse anvendes endvidere et mål for skoleressourcer på et højt aggregeringsniveau (kommunale skoleudgifter pr. elev), hvilket mindsker endogenitetsproblemerne, jf. ovenfor.

Selv om man har gode data for forældrebaggrund mv., kan der imidlertid være endogenitetsproblemer. Det kan således ikke udelukkes, at der er væsentlige forhold, som der ikke er data for, og som spiller en rolle for

både elevernes »resultater« (målt ved fx karakterer, uddannelsesforløb eller arbejdsmarkedskarriere) og for, hvilket niveau af skoleressourcer de modtager. For eksempel er der sjældent mål for forældrenes engagement i deres børns uddannelse eller lignende, og selv om der er sådanne mål via en spørgeskemaundersøgelse, kan de være endogene i den forstand, at forældre til fagligt svage elever alt andet lige vil føle et større behov for at engagere sig i deres børns uddannelse. Et andet eksempel: selv om der er mål for klasse- eller skolekammeraters sociale baggrund, kan der være uobserverede forhold vedr. den måde, klassen fungerer på, som har betydning for klassestørrelse og andre skoleressourcer, og for elevernes resultater.

For at komme ud over sådanne problemer har flere undersøgelser i de senere år anvendt *eksogene* kilder til variation i skoleressourcer som »instrumenter« for skoleressourcer. Et instrument for skoleressourcer er en variabel, som er korreleret med skoleressourcer, men som *ikke* er korreleret med elevernes grundlæggende evner eller forudsætninger. Når man estimerer effekten af skoleressourcer ved brug af et instrument, udnytter man altså kun en del af den samlede variation i skoleressourcer. Hvis instrumentet er gyldigt, har denne del af variationen ikke noget at gøre med de forskellige beslutninger og valg, som har betydning for fordelingen af skoleressourcer på bestemte elever. Et klassisk eksempel på eksogen variation er egentlige eksperimenter, hvor elever og lærere tilfældigt fordeles på henholdsvis store og små klasser; resultaterne af et sådant eksperiment i USA er beskrevet i Krueger (1999). Andre eksempler er »kvasi-eksperimenter« baseret på administrative regler (Angrist og Lavy 1999 og Hoxby 1999). Som instrument for klassestørrelse benyttes en administrativ regel for maksimal klassestørrelse. Ideen er, at når antallet af elever på et givet klassetrin på en skole kommer over et bestemt niveau, så er der stor sandsynlighed for, at der kommer en ekstra klasse, dvs. at der sker et stort fald i den forventede klassestørrelse. Hvis det kan betragtes som tilfældigt, om en given elev går på en skole, hvor elevtallet er lige under eller lige over det kritiske punkt, så kan variationen i gennemsnitlig klassestørrelse omkring dette punkt bruges som udgangspunkt for at måle effekten af ændringer i klassestørrelsen på elevernes resultater (fx testresultater eller uddannelseslængde). Man taler om et kvasi-eksperiment, fordi denne tilfældige variation minder om variationen i et egentligt eksperiment. De tre nævnte undersø-

gelsler har alle fokuseret på klassestørrelsens effekt på elevernes testresultater (i fx læsning og matematik). De to førstnævnte undersøgelser finder signifikant positive effekter af en reduktion af klassestørrelsen, mens den sidste finder, at der ikke er nogen signifikante effekter (inden for det variationsområde for klassestørrelsen, som afspejles i data). Det nævnte kvasi-eksperiment har også været anvendt i AKFs forskning; resultaterne herfra tyder på, at en reduktion af klassestørrelsen har positive effekter på elevernes uddannelsesforløb efter grundskolen (Browning og Heinesen 2003).

7.2 **Eksempel: Effekter på elevernes uddannelsesforløb af forskelle i skoleudgifter pr. elev**

Baggrund

AKF tog i 1998 initiativ til drøftelser med Kommunernes Landsforening, Undervisningsministeriet, Finansministeriet og Danmarks Lærerforening om en række konkrete forslag til projekter om ressourceanvendelse og kvalitet i folkeskolen. Det er et emne, der var meget fokus på sidst i 1990erne, da folkeskolen stod over for en meget stor stigning i elevtallet på grund af, at de årgange af børn, der nærmede sig skolealderen, var væsentlig større end de årgange, der forlod skolen. Der var derfor fra flere sider en bekymring for, dels at de samlede udgifter til folkeskolen ville vokse betydeligt, hvis ikke det lykkedes at reducere (stigningen i) udgifter pr. elev, dels at et reduceret ressourceforbrug pr. elev kunne føre til en forringelse af folkeskolens kvalitet.

To af de nævnte projekter blev senere finansieret af Det kommunale Momsfond, Finansministeriet og AKF, og der blev nedsat en følgegruppe med repræsentanter for KL (Kommunernes Landsforening), Finans- og Undervisningsministeriet, Danmarks Lærerforening og AKF. Det ene projekt fokuserer på forskelle mellem kommuner (og over tid) i udgifter og lærerressourcer pr. elev og årsager til disse forskelle, mens det andet fokuserer på effekten af disse forskelle på elevernes uddannelsesforløb efter grundskolen. Det er dette sidste projekt, som beskrives nærmere nedenfor.

Emne og formål

I dette projekt blev det undersøgt, om de betydelige forskelle, der er mellem kommunerne i ressourceanvendelse pr. elev, har en effekt på, hvor langt eleverne når i uddannelsessystemet efter folkeskolen. Der blev især fokuseret på, om elever, der har gået i skole i kommuner med høje folkeskoleudgifter pr. elev, har en større sandsynlighed for at gennemføre en ungdomsuddannelse end elever, som har gået i skole i kommuner med lavere udgifter – og i givet fald hvor stor effekten er. Med en ungdomsuddannelse menes en gymnasial eller erhvervsfaglig uddannelse. Hypotesen er, at større udgifter pr. elev alt andet lige vil føre til højere kvalitet af skolen, hvilket bl.a. vil give sig udslag i, at flere elever bliver i stand til at gennemføre (og faktisk gennemfører) en uddannelse efter grundskolen. Grundlæggende må ressourceforbruget naturligvis have betydning på et eller andet niveau. Men det spørgsmål, der undersøges, er, om en øgning af ressourcerne pr. elev givet det niveau, de har i Danmark, har signifikante positive effekter på, hvordan eleverne klarer sig. Eller udtrykt på en anden måde, om de eksisterende kommunale forskelle i ressourceforbrug har betydning for elevernes uddannelsesforløb. Som beskrevet i Graversen og Heinesen (1999) er der store kommunale forskelle (om end de er reduceret i de seneste år). I 1970'erne og 1980'erne var skoleudgifterne pr. elev således ca. dobbelt så høje i kommunen med de største skoleudgifter i forhold til niveauet i kommunen med de laveste udgifter, og ca. 35% af kommunerne havde et udgiftsniveau, der lå mindst 10% under eller over landsgennemsnittet.

Sandsynligheden for at gennemføre en ungdomsuddannelse blev således valgt som den grundlæggende indikator for folkeskolens kvalitet. Der er flere årsager hertil. For det første, at det i mange år har haft høj politisk prioritet at sikre, at en større andel af de unge gennemfører en ungdomsuddannelse, bl.a. fordi det har afgørende betydning for chancerne for senere at finde varig beskæftigelse på arbejdsmarkedet. For det andet, at der på det tidspunkt, hvor undersøgelsen blev gennemført, ikke var tilgængelige oplysninger om en anden oplagt kvalitetsindikator, nemlig karakterer ved folkeskolens afgangsprøve. For det tredje, at sandsynligheden for at gennemføre en ungdomsuddannelse afhænger af et forholdsvis bredt spektrum af færdigheder, som er centrale i folkeskolen, nemlig foruden rent boglige

færdigheder (som kan måles ved karakterer) også sociale færdigheder, selvstændighed, motivation for at lære mv.

Resultaterne fra projektet kan benyttes som en del af det politiske beslutningsgrundlag vedr. tildeling af ressourcer til folkeskolen, en problemstilling, der (jf. diskussionen ovenfor) er særlig aktuel i en situation med voksende elevtal, hvor der skal tages beslutninger om, i hvilket omfang de samlede udgifter til folkeskolen skal øges i takt med elevtallet, og i hvilket omfang man skal søge at begrænse udgiftsvæksten, således at udgifter pr. elev kan reduceres.

Evalueringsmodel, data og analysemetoder

Der er tale om en effektevaluering, idet effekten af mængden af skoleressourcer pr. elev i folkeskolen på elevernes efterfølgende uddannelsesforløb søges estimeret.

Datagrundlaget for evalueringen er dels AKF's forløbsregister, baseret på administrative registre i Danmarks Statistisk, for en tilfældig stikprøve på 10% af befolkningen, hvortil der (for årgange født efter 1960) er koblet oplysninger om forældre, dels ekstra registerdata for elevernes skoleforløb i 8. og 9. klasse, herunder hvilken skole de gik på, og dels data for ressourceforbrug pr. elev i folkeskolen opgjort for hver kommune over en periode på 20 år samt en række variabler for socioøkonomiske forhold på kommuneniveau. Ud fra registrene har vi detaljerede oplysninger for de samme personer over flere år. Vi kan således følge de unges vej igennem uddannelsessystemet, og vi har oplysninger om deres familiemæssige baggrund, herunder deres forældres socioøkonomiske forhold (målt ved uddannelse og indkomst mv., se beskrivelsen nedenfor).

Da undersøgelsen blev gennemført, var registrene opdateret til og med 1996. I analyserne anvendes data (fra 10%-stikprøven) for de seks årgange, der blev født i perioden 1965-70. For disse årgange har vi nemlig registeroplysninger om deres forældre i det år, barnet var 15 år, og samtidig kan vi følge de unge, til de er 25 år. Alene personer, der gik i folkeskolen i 8. klasse, er med i analyserne, da der ikke findes data for ressourceanvendelse pr. elev i privatskoler. Efter denne afgrænsning indgår der 37.933 personer i undersøgelsen.

Mængden af skoleressourcer pr. elev måles ved de kommunale netto-driftsudgifter pr. elev i den kommune, hvor eleven har gået i skole. Der findes udgiftsdata tilbage til 1977. For hver elev beregnes de gennemsnitlige udgifter pr. elev over hele skoleforløbet fra 1. til 9. klasse (eller for så mange år som muligt). Som nævnt er der data for, hvilken skole eleven gik på i 8. og 9. klasse, men ikke for tidligere klassetrin. De gennemsnitlige udgifter pr. elev er derfor for hver elev beregnet ud fra oplysninger om forældrenes bopælskommune hvert år i skoleforløbet (ud fra en antagelse om, at hvis eleven gik i folkeskolen i 8. klasse, så gik eleven også i folkeskolen på tidligere klassetrin og i den kommune, hvor forældrene boede).

Der anvendes en statistisk analysemetode, idet der estimeres en logistisk regressionsmodel for sandsynligheden for at have gennemført en ungdomsuddannelse som 25-årig. Som forklarende (årsags-)variabel anvendes naturligvis skoleudgifter pr. elev, men derudover tages der højde for socioøkonomiske forhold i den kommune, hvor eleven gik i skole i 8. klasse (nærmere bestemt inddrages følgende variabler: andelen af børn, hvis forældre kommer fra ikke-vestlige lande, den kommunale ledighedsprocent, andelen af befolkningen med henholdsvis en erhvervsfaglig og en videregående uddannelse, og en indikatorvariabel for, om skolekommunen er en del af hovedstadsregionen), samt 58 variabler for elevernes individuelle familiebaggrund mv. (nærmere bestemt variabler for fødselsår, køn, etnicitet, søskende, familiestruktur, samt forældrenes alder, uddannelse, arbejdsmarkedstilknytning, indkomst, ledighedsgrad, formue og boligforhold), hvor oplysninger om forældrene er for det år, hvor barnet var 15 år.

Som nævnt ovenfor er det i denne type undersøgelser helt afgørende at tage højde for elevernes socioøkonomiske baggrund, når man vil estimere effekten af skoleressourcer. Sådanne variabler vides nemlig fra tidligere undersøgelser at spille en afgørende rolle for elevernes uddannelsesforløb. Til dette formål er de danske registerdata helt unikke, fordi de giver mulighed for inddragelse af en meget stor mængde centrale oplysninger om den enkelte elevs forældre.

Hovedresultater

Hovedresultatet af analyserne er, at højere skoleudgifter har en statistisk signifikant positiv effekt på chancen for at have gennemført en ungdoms-

uddannelse som 25-årig, men effekten er meget lille. En stigning i skoleudgifter pr. elev på 10% vil således ifølge estimaterne kun øge sandsynligheden for at gennemføre en ungdomsuddannelse med ca. 1% for en typisk elev. Det betyder, at det vil være særdeles dyrt for en kommune at øge andelen af elever, der gennemfører en ungdomsuddannelse, hvis midlet er at øge det generelle udgiftsniveau i kommunens folkeskoler: At få én ekstra elev pr. år til at gennemføre en ungdomsuddannelse vil koste kommunen ca. 4-5 mio.kr. pr. år i ekstra folkeskoleudgifter.

Effekterne af en stigning i ressourceforbruget er mindre, når det måles ved antallet af *lærertimer* pr. elev, end når det måles ved udgifter pr. elev. Det forhold, at udgifter pr. elev er et meget bredt ressource mål (som indeholder udgifter til skolebibliotek, bygningsvedligeholdelse, rengøring mv.), er altså ikke årsagen til, at en stigning i udgifterne pr. elev kun har små effekter på elevernes uddannelsesforløb. Mere snævert afgrænsede ressource mål baseret på antallet af lærertimer pr. elev i kommunen har mindre effekter, og effekterne er mindre (og statistisk mindre signifikante), jo mere snævert målet for lærertimer pr. elev afgrænses.

Konklusionen er altså, at selv i de tilfælde, hvor der beregnes statistisk signifikante effekter af et øget ressourceforbrug, er disse effekter meget små. Med andre ord har de eksisterende kommunale udgiftsforskelle stort set ingen praktisk betydning for elevernes senere uddannelsesforløb.

Analyserne viser i øvrigt (som et stort antal andre undersøgelser), at det først og fremmest er den sociale arv, der er afgørende for elevernes uddannelsesforløb. Forældrenes socioøkonomiske forhold, herunder deres uddannelse, har således stor betydning. Som eksempel kan nævnes, at hvis den ene forælder har en lang videregående uddannelse, vil sandsynligheden for, at eleven gennemfører en ungdomsuddannelse, være 15-20% større, end hvis den pågældende forælder ikke har nogen uddannelse ud over grundskolen.

Uddannelsesforløbet efter folkeskolen er også påvirket af sociale og økonomiske forhold i den kommune, hvor man har gået i skole, herunder arbejdsløshedsprocenten, indbyggernes uddannelsesniveau og andelen af elever, hvis forældre kommer fra ikke-vestlige lande. For eksempel vil sandsynligheden for, at en elev gennemfører en ungdomsuddannelse, være ca. 0,3% højere, hvis andelen af kommunens indbyggere med en videregå-

ende uddannelse vokser 1 procentpoint (når der er taget højde for elevens familiebaggrund mv., herunder forældrenes uddannelse).

Ved fortolkningen af resultaterne må man tage det generelle forbehold, at der kan være forhold vedr. forældrebaggrund eller socioøkonomiske forhold i kommunen, som vi ikke har inddraget i analysen, og som systematisk kan påvirke både mængden af skoleressourcer pr. elev og elevernes sandsynlighed for at gennemføre en ungdomsuddannelse. Hvis det er tilfældet, kan det betyde, at den estimerede effekt af skoleressourcer er enten over- eller undervurderet. I og med at de anvendte registerdata giver mulighed for at tage højde for væsentlig flere centrale baggrundsvariabler i denne undersøgelse end i tilsvarende udenlandske undersøgelser, er dette problem dog mindre, end det sædvanligvis er tilfældet.

En anden ting, man skal være opmærksom på er, at det generelt er vanskeligt at måle, i hvilken udstrækning forskelle i ressourceforbrug afspejler forskelle i kvalitet, da der er mange dimensioner af kvalitet. I AKF=s undersøgelse er der fokuseret på den dimension, der har at gøre med elevernes uddannelsesforløb efter afslutningen af grundskolen. Men det er klart, at ressourceforbruget i folkeskolen kan have effekter på en lang række andre forhold, som vi ikke analyserer i denne undersøgelse, fx de pædagogiske kvaliteter i skolen, elevernes trivsel, mens de går i skole, og deres tilegnelse af (fx kreative) kundskaber, som ikke har målelige effekter i relation til de resultater i uddannelsessystemet og på arbejdsmarkedet, vi har undersøgt. Det er også vigtigt at nævne, at folkeskolelovens formålsparagraf dækker et særdeles bredt spektrum af formål. Hvorvidt ressourceforbruget har effekter på opfyldelsen af disse mål, er ikke belyst i denne undersøgelse.

Omvendt må man forvente, at fx et godt miljø i skolen og stor tilfredshed med skolen samt opfyldelsen af mange af de formål, der er nævnt i formålsparagraffen, faktisk vil give de unge bedre forudsætninger for at tilegne sig færdigheder, der er relevante for at kunne gennemføre en uddannelse efter grundskolen. Derfor kan man sige, at undersøgelsens resultater er overraskende. Men de passer godt med resultaterne fra udenlandske undersøgelser, der i de fleste tilfælde har fokuseret på effekten af øgede skoleressourcer på elevernes karakterer eller testresultater. De fleste uden-

landske undersøgelser finder heller ikke betydelige positive effekter af at øge ressourceforbruget i skolerne.

Formidling og anvendelse

Resultaterne er blevet formidlet på forskellige måder. Finansministeriet anvendte resultaterne i et kapitel i Budgetredegørelse 99. Analyserne og resultaterne blev dokumenteret i AKF-rapporten Graversen et al. (1999), og formidlet til en bredere kreds i AKF-rapporten Heinesen et al. (1999). Desuden er resultaterne formidlet i kortere form i antologier (Heinesen, 2000a og 2001) og i bladene AKF Nyt, Skolebestyrelsen og Tandhjulet (Heinesen, 2000b, 2000c og 2000d). Der er sket en formidling til forskere og embedsmænd ved oplæg på en række konferencer og seminarer i Danmark og andre lande, og der er udarbejdet et arbejdspapir på engelsk, der søges optaget i et internationalt tidsskrift (Heinesen og Graversen, 2004). Endelig er resultaterne vedr. social arv beskrevet grundigt i Heinesen (1999).

7.3 Eksempel: PISA-Programme for International Student Assessment

Emne og formål

PISA-programmet – Programme for International Student Assessment – er etableret i et samarbejde blandt OECD-medlemslande, samt enkelte andre lande. I alt 32 lande har deltaget i undersøgelsen. Formålet er at belyse, hvor godt unge mennesker er forberedt til at møde udfordringerne i dagens informationssamfund, jf. Mejding (2002) og Andersen m.fl. (2001). PISA er karakteristisk ved ikke at vurdere kompetencerne ud fra specifikke læseplaners indhold, men ved i stedet at tage udgangspunkt i de livsområder, som anses for at være vigtige for unge menneskers fremtidige livssituation. Undersøgelsen kan betegnes som en makroevaluering af den danske grundskole, som også bidrager med viden om, hvorfor danske unges kompetencer er, som de er, sammenlignet med unges i andre lande.

PISA kan forsyne uddannelsespolitikere, uddannelsesadministratorer og praktikere med en omfattende vurdering af læringsresultater målt ved slutningen af den undervisningspligtige periode og kan derved bidrage til grundlaget for uddannelsespolitiske beslutninger.

Evalueringsmodel og -design

Undersøgelsen vedrører tre faglige områder – i undersøgelsen kaldet domæner: læsning, matematik og naturfag med fokus på læsning. Gode færdigheder på ét af dem betyder for de fleste, men langt fra for alle, også gode færdigheder på et andet. Ud over dette indgår, som første gang nogensinde, målinger af elevernes personlige og sociale kompetencer, områder, der betegnes som en del af Cross Curricular Competences (CCC).

PISA lægger inden for domænerne vægt på en vurdering af elevernes evne til at reflektere over deres kundskaber og erfaringer og at behandle emner i forhold til deres eget liv. Herunder vurderes evnen til at kunne »læse mellem linjerne«, at kunne gennemskue et underforstået budskab og at kunne vurdere perspektiverne i en samfundsmæssig sammenhæng. Endelig betones de kommunikative færdigheder.

Ud over domænerne indgår baggrundsoplysninger afgivet af eleverne, omfattende elevernes klassetrin, køn, familiebaggrund, social-økonomiske baggrund, sprog talt i hjemmet, immigrantstatus, fritidsaktiviteter samt holdninger til skolegang.

PISA anvender en fremgangsmåde, der med hensyn til at vurdere en bred række af kundskaber, færdigheder og sociale kompetencer er uden fortilfælde. De særlige kendetegn ved fremgangsmåden er:

- Orientering mod uddannelsespolitiske spørgsmål.
- Fokus på kundskaber og færdigheder demonstreret i opgaver med relevans for hverdagslivssituationer – både i fortsat uddannelse, arbejdsliv, familieliv og samfundsliv.
- Regelmæssighed, idet testninger gentages hvert tredje år.
- Videnskabelighed, idet et konsortium af førende institutioner med hensyn til måling af kompetencer, assisteret af ekspertgrupper bestående af førende forskere på de indgående domæner, har stået for udarbejdelse af testmaterialet. Endvidere har ekspertgrupperne været assisteret af faglige referencegrupper med deltagelse fra en række af landene i PISA.

Dataindsamling og analysemetoder

Eleverne er blevet testet for deres færdigheder i læsning, matematik og naturvidenskabelige fag. Derudover har eleverne udfyldt et spørgeskema om forældrebaggrund og om deres holdninger til og erfaringer med skolegangen.

En samlet analyse af kvaliteten af data viser, at de er valide, pålidelige og repræsentative for hele gruppen af uddannelsessøgende født i 1984. Med læsning som fokusområde er læsefærdighederne blev bedst testet. Konklusioner, der bygger på læseresultater, drages derfor med større sikkerhed end konklusioner, der drages på baggrund af resultater i matematik og naturvidenskab.

I Danmark deltog 4.242 unge, født 1984, fordelt på 226 uddannelsesinstitutioner i den internationale del af testen (OECD-PISA).

Datas pålidelighed, repræsentativitet og validitet

Datapålidelighed

Hvis datapålideligheden skal være god, må der ikke være opstået fejl, der betyder, at de indsamlede data giver en dårlig beskrivelse af virkeligheden. Høj datapålidelighed er en forudsætning for, at data kan bruges til at drage holdbare konklusioner, men er ikke en tilstrækkelig forudsætning; data skal også være valide – dvs. relevante for problemstillingen – og repræsentative, dvs. være repræsentative for den population, man ønsker at drage konklusioner om.

Det internationale konsortium anstrengte sig for at gøre data pålidelige, bl.a. udarbejdede konsortiet adskillige drejebøger/manualer, som skulle sikre en ensartet og korrekt procedure i de 32 lande.

Endelig har konsortiet lavet en meget omfattende kvalitetssikringsprocedure og dermed givet datapålideligheden meget høj prioritet.

Validitet

Konsortiet har, bl.a. via en omfattende pilotundersøgelse, testet forskellige opgaver og deres validitet i forbindelse med PISA-undersøgelsens problemstillinger. Opgaverne er udvalgt af ekspertpaneler i samarbejde med forskere i de enkelte lande. De valgte opgaver må derfor siges at væ-

re et rimeligt manifest udtryk for den latente variabel: Elevernes kompetence.

I alle lande er opgaverne oversat fra engelsk eller fransk. Den endelige oversatte version blev til på baggrund af to uafhængige, parallelle oversættelser, der blev sammenfattet af en person med et omfattende kendskab til testkonstruktion. I de danske testopgaver, hvor der har været tvivl om oversættelsen fra engelsk til dansk, er den franske version brugt til at verificere, ligesom dansk fagekspertise har været konsulteret omkring fagspecifikke formuleringer. Til sidst blev det oversatte materiale kontrolleret af OECD-ansatte eksperter.

Repræsentativitet

Et yderligere krav til data, hvis de skal være gode, er, at de udtrykker de holdninger/præstationer i den befolkningsgruppe, man ønsker at drage konklusioner om. Det betyder, at eleverne udvalgt til at deltage, og dermed repræsenterende alle elever på 15 år, skal have nogenlunde samme sammensætning/karakteristika som hele gruppen af 15-årige. Hvis man sammenligner procenterne for de enkelte skoletyper, viser det sig, at det faktiske resultat svarer rimeligt til det forventede.

Samlet må det konkluderes, at de danske OECD-PISA-data er pålidelige, repræsentative og valide.

Hovedresultater

Resultaterne fra undersøgelsen er offentliggjort i nationale rapporter (Andersen m.fl., 2001) og i en international rapport udgivet af OECD (OECD, 2001). Her omtales nogle få af hovedresultaterne.

Læsning

De danske elever klarede sig på det jævne sammenlignet med eleverne fra de andre lande i undersøgelsen, ikke mindst når vi sammenligner med de øvrige nordiske lande. Det danske gennemsnit (497 point) lå lige ved OECD-landenes gennemsnit på 500 scorepoint. Det er bekymrende, at de danske elever fordeler sig sådan, at der er relativt mange elever med *meget dårlige* læsefærdigheder, og at der ikke er særligt mange med *meget gode* læsefærdigheder. Hvis Danmark skal gøre sig håb om at klare sig i

den internationale konkurrence på baggrund af et højt uddannelsesniveau, så er dette ikke godt nok.

Matematik

Matematik var det fagområde, de danske elever klarede bedst. Gennemsnitsresultatet (514 point) placerer Danmark lidt over OECD-gennemsnittet på 500 point. Det er på niveau med lande som Schweiz, Frankrig, Østrig, Island og Sverige, men stadigvæk efter lande som Finland, Australien, Canada og Storbritannien. Grunden til det forholdsvis gode resultat er, at det især er de svagere elever, der scorer højere i Danmark. Skal vi sammenligne os med de bedst placerede lande, er der dog markant færre danske elever, som scorer topresultater. Et andet vigtigt resultat er, at der er en overrepræsentation af nydanske elever – født i Danmark af udenlandske forældre – i den lave ende af skalaen – også set i relation til den samme type elevers præstationer i fx Norge og Sverige. Netop disse områder bør, sammen med de relativt svagere resultater for de danske piger, overvejes som satsningsområder i en styrkelse af matematikundervisningen.

De naturvidenskabelige fag (science)

På dette område klarede de danske elever sig dårligst i den internationale sammenligning. Gennemsnitsresultatet på 481 point var signifikant under OECD-landenes gennemsnit, og især de danske pigers resultater var relativt svagere, end man så det i andre lande. Da afgrænsningen af de naturfaglige kompetencer, man ønsker at undersøge i PISA, er i god overensstemmelse med intentionerne for undervisningen i de naturvidenskabelige fag i folkeskolen, er dette et resultat, der må give anledning til alvorlige overvejelser – ikke mindst set i forhold til, at man med skoleloven fra 1993 ønskede at styrke det naturfaglige område. Norge og Island lå inden for gennemsnittet for OECD-landene, mens Sverige og Finland havde resultater, der lå signifikant højere. Der er derfor et godt stykke igen, før vi kan sige, at danske elever er lige så kompetente inden for naturfagsområdet som eleverne i de øvrige nordiske lande.

Personlige og sociale kompetencer

Det er vigtigt at understrege, at PISA-undersøgelsen lægger vægt på, at faglige færdigheder ikke eksisterer i et tomrum, men at de tilhører nogle personer, som skal have lyst og motivation til at beskæftige sig med dem. Det er derfor glædeligt at se, at langt den overvejende del af de danske elever er glade for at gå i skole og også langt hen ad vejen er interesseret i fagene. De er positive over for såvel samarbejde som konkurrence, og den generelt positive holdning gælder også for de fagligt svagere fungerende elever. Eleverne ser ikke ud til at motiveres af ydre forhold som den fremtidige karriere, men snarere af en mere umiddelbar interesse. Eleverne har en god fornemmelse for egne kompetencer i forhold til klassekammeraternes, selv om resultaterne i en international sammenhæng kunne antyde en vis selvovervurdering. Men det internationale perspektiv har eleverne jo af gode grunde ikke – til gengæld er de klar over, at de ofte godt kunne yde lidt mere, når talen falder på indsats og vedholdenhed i arbejdet i almindelighed. Generelt kan vi se, at den danske folkeskoles mål om at tilgodese elevernes almene udvikling synes at være lykkedes godt. Den danske skole er omsorgsfuld og møder eleverne positivt, hvor de er, men undlader måske nogle gange i tilstrækkeligt omfang at bygge videre på og styrke elevernes potentialer.

Formidling og anvendelse

PISA-undersøgelsen er formidlet i rapporterne Andersen m.fl. (2001) og OECD (2000, 2001, 2002).

PISA kan anvendes på mange niveauer. For det første kan PISA forsyne uddannelsespolitikere med et omfattende materiale til vurdering af læringsresultater målt ved slutningen af den undervisningspligtige periode. Vurderingen sker i tal, der i så høj grad som muligt er gjort sammenlignelige, så de kan danne grundlag for politiske beslutninger og ressourceallokeringer, og PISA kan give bidrag til indsigt i den blanding af faktorer, der opererer ensartet eller forskelligt hen over lande og regioner.

For det andet kan PISA hjælpe uddannelsesadministratorer og undervisere til at erkende styrkesider og svagheder i deres egne systemer, ligesom de kan vurdere, i hvilken grad variationer i uddannelseserfaringer er unikke eller afspejler forskelle, der også ses andre steder. For eksempel vi-

ser den kendsgerning, at nogle lande opnår et højt gennemsnit af elevers resultater med kun en lille spredning mellem de bedste og de dårligste elever, at et godt middelresultat ikke behøver at ske på bekostning af de svageste elever. Ligeledes illustrerer det faktum, at forholdet mellem social baggrund og læringsresultater varierer i forskellige lande, at nogle skolesystemer vurderet ud fra de rent statistiske sammenhænge synes at kunne ændre og begrænse indflydelsen af en dårlig social baggrund på elevernes resultater.

Idet der dermed kan foretages analyser af resultater fra forskellige uddannelsessystemer ud fra et fælles sæt af kompetencer, der er relevante i et livslangt perspektiv, kan PISA give en stærk og multikulturel basis for at definere mål for kundskaber og færdigheder. PISA muliggør dermed, at man på længere sigt kan indkredse de forhold, der har relation til uddannelsesmæssig succes.

7.4 **Eksempel: Kommunernes vejledning af unge om uddannelse og erhverv**

Formål, genstand og kontekst

Formålet med evalueringen er at evaluere, hvordan vejledningen af unge i alderen 15-24 år fungerer og er organiseret i kommunerne med henblik på at sikre, at flest mulig unge kommer videre i ungdomsuddannelsessystemet efter afsluttet grundskole.

I undersøgelsen indgår såvel skolernes egen vejledning (ved skolevejleder), som ungdomsvejledningen i kommunen, kommunens beskæftigelsesvejledning og uddannelsesinstitutionernes studievejledning. I undersøgelsen fokuseres på gruppen af unge som helhed, men med særlig fokus på de unge, som er uden beskæftigelse eller ikke i gang med en uddannelse. Ungdomsgruppen er i undersøgelsen defineret som unge i alderen 15-24 år (fra 9.-10. klasse).

Der har i undersøgelsen især været fokus på skole- og ungdomsvejledningen, som er begyndelsen på vejledningsforløbet, samt på ungdomsvejledningens centrale rolle som bindeled mellem skolevejledningen og

den videre vejledning i kommunens beskæftigelsesafdeling/arbejds-markedsafdeling.

Evaluerings spørgsmål

I evalueringen ses på følgende spørgsmål:

1. Hvordan er kommunens vejledning af unge om uddannelse og erhverv organiseret?
2. Hvilke faggrupper/personer varetager vejledningen af unge?
3. Hvordan spiller de forskellige vejledningssystemer for unge sammen?
4. Hvordan søger kommunerne at sikre, at alle unge i kommunen med behov for vejledning får den nødvendige vejledning i forbindelse med valg af uddannelse og erhverv?
5. Hvilket indhold, hvilke metoder og hvilken målsætning har vejledningen, og hvilke kvalitetskrav stilles til vejledningen?
6. Hvordan er de unges fremtidsforventninger og deres syn på vejledningen?

Spørgsmålet er især, hvordan kommunens vejledning fungerer over for de svagt stillede unge samt i den forbindelse vejledningens kontakt til og samarbejde med skolevejleder, studievejledningen på uddannelsessteder, som de unge »dropper ud« af, kommunale aktiveringsforanstaltninger og sagsbehandlere i socialforvaltningen i kommunen. Spørgsmålet er dels, hvilket »sikkerhedsnet« der er spændt ud under de unge, og dels i hvor høj grad og hvordan man søger at fastholde unge, som giver op over for uddannelsesmæssige og arbejdsmæssige krav, og dels hvilke alternative aktiveringstilbud der er til denne gruppe unge.

Der var i gennemførelsen af undersøgelsen opmærksomhed på, at en særlig gruppe unge, nemlig de tosprogede unge, er unge, der ofte vil stille anderledes krav til vejledningen i forbindelse med uddannelse og beskæftigelse (jf. fx Mehlbye, 1994). Det lå imidlertid uden for undersøgelsens rammer at behandle denne gruppes særlige problematik.

Design, dataindsamling og dataanalyse

Undersøgelsen indeholder:

1. En oversigtlig kortlægning af kommunernes forskellige måder at strukturere og organisere vejledningen af unge på, baseret på en landsdækkende spørgeskemaundersøgelse.
2. En dyberegående analyse og beskrivelse af forskellige vejledningsmodeller, baseret på en interviewundersøgelse i tre casekommuner, som har forskellig størrelse, forskellig geografisk beliggenhed, og som har organiseret deres ungdomsvejledning forskelligt.

Spørgeskemaundersøgelsen til alle landets kommuner

I den landsdækkende spørgeskemaundersøgelse stilet til kommunernes koordinator eller leder af ungdomsvejledningen deltog 89% af kommunerne.

Casekommuneundersøgelsen

I interviewundersøgelsen i de tre kommuner, en landkommune, en bykommune og en storbykommune, deltog 78 vejledere og ledere samt i alt 22 elever fra 9. eller 10. klasse fra en folkeskole i hver af de tre kommuner.

I dataanalysen forsøgte opstillet forskellige prototyper på vejledningsroller og vejledningsmodeller for at skærpe grundlaget for de videre drøftelser i kommunerne.

Hovedresultater

Den typiske skolevejleder og ungdomsvejleder:

Den typiske skolevejleder og ungdomsvejleder er, ifølge den landsdækkende spørgeskemaundersøgelse, en mandlig folkeskolelærer i fyrrerne eller halvtredserne. Kvinderne udgør kun en tredjedel. Kun omkring hver tiende ungdomsvejleder og skolevejleder er under 40 år, mens hver tredje er over 50 år. 10% af landets ungdomsvejledere har ikke en vejlederuddannelse. (Skolevejlederne antages at have en skolevejlederuddannelse eller at være i gang med en sådan, da det er forudsætningen for at blive skolevejleder).

Der er en stor spredning mellem kommunerne i antallet af timer afsat til henholdsvis skolevejledning og ungdomsvejledning. For skolevejlederne kan timeantallet variere fra 0,6 timer ugentligt til 16,5 timer ugentligt.

Spredningen kan skyldes forskellige kommunale overenskomster på området og forskellig praksis med hensyn til tildeling af ekstra timer til særlige opgaver samt forskelligt beregningsgrundlag (klasser, elevårge m.m.). I snit har hver skolevejleder knap syv timer (6,8 timer) ugentligt til rådighed til skolevejledningsarbejdet. Hver ungdomsvejleder har i gennemsnit lidt over fem timer ugentligt (5,4 timer) til ungdomsvejledning. Nogle ungdomsvejledere har kun et par timer ugentligt, andre har op til en halv eller en hel arbejdsuge.

Vejledningens organisering

De tre undersøgte kommuner, som udvalgte til nærmere undersøgelse, er eksempler på kommuner, som har organiseret sig forskelligt. I interviewene giver såvel medarbejdere som ledere i de tre kommuner udtryk for, at de er tilfredse med netop deres kommunes organisering.

Vejledningens organisering i de tre kommuner kan beskrives således:

A. Det tætte netværk

I landkommunen er vejledningens organisering karakteriseret ved det, der kaldes det tætte netværk, idet de forskellige vejledningssystemer omkring unge er tæt forbundne gennem den fysiske og organisatoriske placering af ungdomsvejledningen sammen med skolevejleder og den fysiske placering sammen med AF-vejleder.

B. Det strukturerede netværk

I storbykommunen er vejledningens organisering karakteriseret ved det strukturerede netværk. Vejledningsnetværket er tæt, idet ungdomsvejlederne er organisatorisk og administrativt tilknyttet arbejdsmarkedsafdelingen og fysisk placeret på folkeskolerne. De binder dermed samarbejdet mellem to »afdelinger« sammen. Desuden er især samarbejdet mellem skolevejledere, klasselærere og ungdomsvejledere velstruktureret og veldefineret i kraft af funktionsbeskrivelser, der giver nærmere præciseringer af samarbejdsrelationer og arbejdsopgaver. Dermed er arbejdsopgaver og samarbejdsrelationer mellem de tre grupper nogenlunde ensartet på tværs af de enkelte skoler og distrikter. Ungdomsvejledningen, som er rettet mod de unge i alderen 16-19 år, er organisatorisk set placeret i arbejds-

markedsafdelingens vejlednings- og uddannelsesafdeling for de unge 18-24-årige på kontanthjælp.

C. Det åbne netværk

I bykommunen er vejledningens organisering karakteriseret ved det, man kunne kalde det åbne netværk. Ungdomsvejledningen, som er rettet mod de 16-25-årige, er organisatorisk set knyttet til skoleforvaltningen, og fysisk set er ungdomsvejlederne placeret på de fem overbygningsskoler i kommunen (derudover er der to børnehaveklasse-7. klasse-skoler alene med en skolevejledning).

Vejlederroller

Interviewundersøgelsen i de tre kommuner viser, at de forskellige måder vejlederne udøver deres vejledning på, deres prioritering af arbejdsopgaver og deres samarbejdsflader, tegner forskellige billeder af de roller, vejlederne kan påtage sig.

- *Omsorgsudøveren* er primært præget af et dybt engagement i de unge, i sit arbejde og i sine samarbejdspartnere.
- *Realisten* ser sin primære opgave som at give de unge et realistisk billede af deres muligheder og mener, at han/hun som vejleder er bedst til at vurdere, hvilke muligheder den unge har.
- *Diagnostikeren* er optaget af at få så præcise mål for den unges færdigheder og evner som muligt, og den instrumentelle vejleder er optaget af test og karakterer. Begge roller indgår ofte i realistens rolle.
- *Administratoren* ser som sin hovedopgave, at alt det administrative omkring valgfag, erhvervspraktik m.m. er i orden.
- *Sælgeren* går, formodentlig uden selv at opdage det, ind i et salgskøbs-spil med den unge, hvor vejlederen bliver ved med at give den unge forskellige forslag til og tilbud om uddannelse eller beskæftigelse, mens den unge på sin side bliver ved med at afslå tilbuddene.
- *Dynamo* er hele tiden udfarende og opmærksom på nye behov og nye idéer og er heller ikke bange for at præsentere dem.

De unge folkeskoleelevers vurdering af den vejledning, de har fået

Med hensyn til valg af uddannelse og job viser interviewene med de unge, at de er meget usikre, spekulerer meget over, hvad de vil, og svinger

meget i deres beslutninger. Det sker, at udelukkelsesmetoden bliver det udslagsgivende kriterium for valget.

De unge har meget forskellige erfaringer med den vejledning, de har modtaget. Kvalitet i vejledningen definerer de selv (indirekte via interviewene med dem) ved engagement, ihærdighed og viden fra vejlederens side. De unge foretrækker også den personlige samtale frem for skriftlige informationer i form af pjecer m.m., som de aldrig får læst. Den personlige vejledning skal være tilpasset dem og deres situation, men en alt for realitetsorienteret vejledning er ikke af det gode, der skal være mulighed for at realisere drømme og personlig udvikling. Der er bred enighed om, at skole- og ungdomsvejleder meget gerne må være den samme, for så kender de den, de skal have hjælp af, hvis de får problemer i forbindelse med uddannelse og erhverv senere hen. Det er lettest at henvende sig til et kendt ansigt i kendte omgivelser, mener flere af de unge.

De unge henter støtte fra såvel forældre, søskende, lærere, skolevejledere og kammerater. Nogle bruger alle, andre vælger ud, hvem de synes, de kan bruge bedst. Forældrene er hyppigst dem, de unge henter støtte hos, og især følelsesmæssig støtte, mens de hos lærere og evt. hos skolevejledere henter informationer om uddannelse og erhverv. Kammerater og især store søskende er gode som erfaringsformidlere, dvs. til at videregive egne erfaringer fx om gymnasiet.

Overvejelser om sikring af kvaliteten i vejledningstilbuddet

Der fremsættes på baggrund af evalueringen bl.a. følgende forslag om vejledningen i folkeskolen:

- Drøftelser med de unge om deres uddannelses- og erhvervsvalg, allerede intensivt i 9. klasse, dvs. før eleverne i 9. klasse vælger 10. klasse eller gymnasiet eller evt. vælger helt at gå ud af skolen.
- Samarbejde om uddannelsesplan og handlingsplan med de unge på deres egne præmisser.
- Vejlederne må ikke undervurdere deres egen betydning i de unges uddannelsesvalg.
- Sikring af, at de unge får mest muligt ud af deres praktikforløb i 8.-10. klasse.

Formidling og anvendelse

Undersøgelsen er formidlet i en AKF-rapport, Undervisningsministeriets Nyhedsbrev, og ved flere foredrag landet over for vejledere i kommunerne.

8 Evaluering på børn og unge-området

Jill Mehlbye

8.1 Temaer inden for børn og unge-området

Det offentliges ydelser til børn og unge er bestandig i politisk fokus. Det gælder både ydelserne til børn og unge bredt set i form af pasningstilbud, skoletilbud, fritidstilbud m.m., og det gælder særlige ydelser fx personlig støtte, specialpædagogisk bistand eller døgnanbringelse til børn og unge med særlige behov, dvs. børn og unge, der har behov for en særlig indsats. Det gælder børn og unge med store sociale og følelsesmæssige problemer, der enten gør det svært for dem at fungere blandt andre børn og unge i det normale skole- og fritidsmiljø eller betyder, at de generelt mistrives og ikke udvikler sig, som de skal, og derfor har særlig behov for specialpædagogisk eller behandlingsmæssig støtte. Det gælder børn og unge med vidtgående fysiske eller psykiske handicap, som betyder, at de må have specialtilbud i form af vidtgående specialundervisning og specialpædagogisk bistand. Endelig gælder det børn og unge fra etniske minoriteter, der har brug for særlig støtte omkring deres integration i det normale skole- og fritidsmiljø grundet deres anderledes sproglige og kulturelle baggrund.

Evalueringernes problemstillinger udspringer således typisk fra konkrete problemer, og derfor forventes det, at resultaterne er handlingsanvisende, eller i hvert fald peger hen mod, hvilke typer af handlinger der bør eller kan gennemføres. Resultaterne forventes at være handlingsorienterede både i forhold til konkrete indsatser og med henblik på eventuelle myndighedstiltag (regulering af lovgivningen eller regler og procedurer i den offentlige forvaltning).

I kommuner og amter er der således løbende behov for at gennemføre evalueringer af ydelserne over de forskellige grupper af børn og unge eller over for kommunens børn og unge generelt for at kunne vurdere, om ydelserne svarer til behovene hos børnene og de unge, hvorvidt kvalitet og effekten af tilbuddene svarer til det forventede, og om der skal ske ændringer, sættes nye i værk m.m. Det gælder både de enkelte, konkrete tilbud, det gælder tilbuddene til børn og unge generelt set, og det gælder ydelserne til børn og unge med særlige behov.

Omkring evaluering af offentlige ydelser vil det ofte være kommunen selv, der er bestiller eller ønsker denne iværksat, men det kan også være staten, der ønsker amter og kommuners tilbud, fx om døgnanbringelse evalueret.

Det kan være en kommune selv, der gennemfører evalueringen af egen indsats ved hjælp af folk ansat i kommunen, hvilket benævnes som intern evaluering. Det kan også være, at kommunen hyrer folk uden for kommunen til at gennemføre evalueringen, kaldet ekstern evaluering. Disse to former for evaluering er behandlet i kapitlet: Intern eller ekstern evaluering.

Der kan således i en kommune være behov for at evaluere de fritidstilbud, man yder til kommunens børn og unge, fx om tilbuddene er søgt af børn og unge i det forventede omfang, og om de opfylder børn og unges behov for fritidsaktiviteter. En kommune ønsker fx at få en vurdering af, hvorfor deres SFO-ordninger (skolefritidsordninger) ikke søges i det omfang som forventet (Mehlbye m.fl. 2000). En anden kommune ønsker fx at finde ud af, i hvor vid udstrækning kommunens fritidstilbud svarer til det behov og ønsker, børn og unge har, og hvorvidt der er behov for nye og anderledes tilbud (Mehlbye og Højmark Jensen 2003).

Et andet tema er kommunens behov for at vide noget om, hvordan kommunens børn og unge trives generelt, og hvordan de trives og udvikles (klarer sig) i skolen set ud fra kommunens børnepolitiske mål. Det gælder fx en kommune, hvor man ønsker at vurdere, hvordan børn og unge i kommunen har det generelt, dvs. i skolen, i fritiden og derhjemme, samt hvordan de klarer sig fagligt i skolen i forhold til deres trivsel (Mehlbye 2000).

Et tredje tema kan være, hvorvidt børn og unge med etnisk minoritetsbaggrund integreres i skolen og fritidsmiljøet ved hjælp af de særlige foranstaltninger med henblik på de tosprogede børn og unges integration, som kommunen har iværksat (Mehlbye 1994).

Ud over de mere generelle evalueringer af børn og unges situation og vilkår er der i kommunerne samt i amterne, som varetager driften af særlige specialinstitutioner for børn og unge, særlig fokus på, hvilken betydning og effekt indsatsen over for børn og unge med særlige behov har, og hvorvidt der ydes tilstrækkelige tilbud til denne gruppe børn. Det gælder i særligt omfang de dyrere foranstaltninger som døgnanbringelse af børn og unge (Boelsen m.fl. 1986), eller særlig indsats over for barnet og den unge i hjemmemiljøet (Mehlbye 1990).

8.2 **Særlige problemstillinger omkring evalueringer**

Når der skal gennemføres evalueringer, gør der sig helt særlige problemstillinger gældende. En evaluering af en indsats og effekten af denne skal holdes op imod et eller andet mål, et mål for effekt, et mål for kvalitet m.m. Imidlertid er situationen for en evaluering ofte som udgangspunkt, at målet eller kvalitetskriteriet er diffust og bredt. Typiske indledende formuleringer er, at tilbuddet skal være tilstrækkeligt i forhold til barnets behov, men spørgsmålet er, hvordan defineres tilstrækkeligt, og hvad mener vi egentlig med tilstrækkeligt? Før vi går i gang med en evaluering af en ydelse, må vi derfor vide, hvad målet med og kvaliteten af denne ydelse skal være på et langt mere konkret niveau.

Når vi fx taler om tilstrækkelige fritidstilbud til børn og unge, hvordan måler vi tilstrækkelige, og hvordan måler vi, om tilbuddene svarer til børn og unges behov? Skal alle børn og unge have adgang til et, to eller flere tilbud, skal alle børn og unge udtrykke tilfredshed med de enkelte tilbud eller måles tilfredsheden på omfanget af søgningen til tilbuddet? Der er flere muligheder.

Det samme gælder, når vi taler om indsats i forhold til børn og unge med alvorlige sociale problemstillinger. Når vi fx iværksætter døgn-

anbringelser til denne gruppe børn og unge, hvordan måler vi så, at det hjælper, og at kvaliteten af tilbuddet er den ønskede. Klart nok, at vi forventer, at børnene og de unge klarer sig bedre, eller at deres problemer »forsvinder« bliver »mindre«, og at tilbuddet lever op til professionelle standarder for socialpædagogiske tilbud. Men hvad mener vi egentlig med dette?

Når vi taler om børn og unge med handicap, hvordan måler vi fx om den vidtgående specialundervisning har haft en effekt, har børnene lært noget, og hvordan måles dette?

8.3 **Evalueringemetoder med henblik på at få en vurdering af udbyttet og effekten af indsatsen**

Spørgsmålet er, hvilke evalueringemetoder og hvilket design vi skal anvende, når vi taler om en gruppe på hvis vegne, der handles, og en gruppe, der ikke har mulighed for at udtrykke sig på samme måde som voksne brugere af en ydelse?

I de nævnte eksempler på evalueringer, hvor det enkelt sagt drejer sig om, hvilket udbytte brugerne har fået af den generelle ydelse eller den særlige indsats og en vurdering af indsatsen, er det vigtigt at kunne indsamle brugernes egen opfattelse af ydelserne og den hjælp, der er ydet.

Brugerne kan her opfattes som børnene og de unge selv, deres forældre og de professionelle, som har henvist børn og unge til særlige ydelser. Dem, der yder tilbuddet, kan ligeledes også selv have behov for at vide, om deres indsats »nytter noget«, kan gøres bedre m.m.

En særlig udfordring omkring evaluering af ydelser til børn og unge er, hvordan børnene og de unge selv inddrages i evalueringen.

I en evaluering af tilbud og ydelser til børn og unge, vil der således oftest også være behov for at inddrage børnene og de unge selv for at få inddraget deres synspunkter. Der er gennemført mange undersøgelser om og med børn, men den metodiske litteratur om, hvordan man gennemfører undersøgelser med børns deltagelse, er sparsom. I Danmark kan fremhæves Andersen og Kjærullf 2003 og Andersen og Ottesen 2002 om børns deltagelse i spørgeskemaundersøgelser og interview.

Det giver særlige metodiske overvejelser at inddrage børn og unges synspunkter. Spørgsmålet er, hvad kan børn og unge svare på? Hvor veludviklet er deres begrebsverden, hvor pålidelige er deres svar, og hvilke metoder skal anvendes, når vi spørger børn og unge?

Den hyppigst anvendte metode, når man vil have fat i en større gruppe børn i skolealderen, er at gennemføre spørgeskemaundersøgelse via skolen. Ved at gå gennem skolen er man nogenlunde sikker på både at få en repræsentativ stikprøve, samtidig med at det er let rent praktisk at lade en lærer dele spørgeskemaerne ud for efterfølgende at samle dem ind. Desuden sikrer man sig besvarelser, der ikke er drøftet igennem med forældre, der kan påvirke børnenes besvarelser.

Hvornår kan man så inddrage børn i en spørgeskemaundersøgelse?

Enkelt sagt kunne man sige, det drejer sig om børnenes modenhed, men modenhed er et vanskeligt kriterium at anvende, fordi det kræver, at man først undersøger deres modenhed. Et bedre og mere praktisk kriterium er at bruge børnenes alder.

De gængse metoder som observationsstudier, interview og spørgeskemaer kan bruges, når det gælder børn. Men man må være opmærksom på, at der gør sig særlige forhold gældende, når der er tale om interview og spørgeskemaundersøgelser med børn og unge.

Når det gælder spørgeskemaundersøgelser med børn, kan man i princippet give dem et spørgeskema, så snart de kan læse. Spørgsmålet er, hvornår de er nogenlunde sikre læsere, også hvad angår indholdsmæssig forståelse? Langt de fleste børn på fjerde klassetrin må forventes for flertallets vedkommende at være nogenlunde sikre læsere.

Det næste spørgsmål er, hvad kan børn informere om, hvad kan man stille dem som spørgsmål, og få et svar, der med nogenlunde sikkerhed er svar på det spørgsmål, man stiller?

Først og fremmest gælder det, at børn kan svare på noget, der drejer sig om dem selv, mens de vil have svært ved at besvare spørgsmål om andre fx deres forældre (jf. Andersen og Kjærulff 2003). Hvis man fx vil have baggrundsplysninger om forældrenes uddannelse viser erfaringen, at selv store børn i 15-års-alderen har svært ved at svare på dette. Hvad angår forældrenes jobtype, er det ofte lige så svært for børnene at svare på

dette. De ved måske, hvor forældrene arbejder, men ikke hvilke funktioner de varetager – det gælder også de store børn.

Spørgsmål om husstandens størrelse vil også være et problem for mange især mindre børn at svare på, idet de vil blande familiens størrelse sammen med husstandens størrelse (hvor mange søskende bor du sammen med, besvares let med, hvor mange søskende man har). Hvad angår enkle spørgsmål om dem selv, fx børnenes alder, vil de mindre børn kunne svare på, hvor gamle de er, men ikke hvornår de er født, hvilket de større godt ville kunne gøre (jf. Andersen og Kjærulff 2003).

Hvis børnene/de unge spørges om mere vurderende relative spørgsmål, fx om deres placering fagligt i klassen i forhold til de andre elever, skal de være omkring 12-13 år, før de har en realistisk vurdering af dette.

Det samme gælder hypotetiske spørgsmål af typen »hvad vil du gøre hvis...«. Hvad angår hukommelse, gælder det for børn (hvilket også gælder for voksne), at jo længere tid du går tilbage i tiden, jo større usikkerhed er bevarelsen behæftet med. Undersøgelser viser, at tre måneder er den varighed, man kan forvente, at et barn vil kunne svare nogenlunde præcist (Larsen 2002).

Generelt må man sige, især når det gælder mindre børn (12 år og derunder), at det vil give de mest sikre svar at stille spørgsmål om konkrete handlingsorienterede forhold (jf. Andersen og Kjærulff 2003). Derfor vil spørgsmål fx om fritidsaktiviteter, de går til, kunne besvares af børnene med stor sikkerhed, især når det gælder her og nu-aktiviteter.

Eksemplerne i dette kapitel repræsenterer bredden af de fremgangsmåder, som evaluering af ydelser over for børn og unge omfatter i Danmark. Det første eksempel handler om en generel evaluering af, hvordan børn og unge generelt trives og udvikler sig i skolemiljøet. Det andet eksempel omhandler en evaluering af, i hvor vid udstrækning en kommunes fritidstilbud imødekommer børns og unges behov og ønsker. Det sidste eksempel omhandler en socialforvaltnings særlige indsats over for udsatte børn og unge med alvorlige sociale og personlige problemer.

8.4 **Eksempel: Evaluering af fritidstilbuddene til børn og unge i en kommune**

Dette eksempel beskriver evaluering af en kommunes fritidsaktiviteter (Mehlbye og Højmark Jensen 2003). Evalueringen gennemførtes med henblik på at belyse børn og unges brug af fritidstilbuddene i kommunen samt børn og unges ønsker til supplerende fritidstilbud.

Overordnet formål, genstand og kontekst

I evalueringen søges dannet et nuanceret billede af forskellige grupper af børn og unge, og deres behov for tilbud i fritiden, og hvorvidt kommunens tilbud svarer til børnenes/de unges behov og interesser. Det overordnede formål var tillige at give bud på, hvilke typer af aktiviteter børn og unge ønsker, samt komme med forslag til evt. nye typer af aktiviteter.

Evalueringen omhandler kommunens børn og unge i alderen 10-18 år.

Evalueringsspørgsmål

Evalueringen fokuserede på følgende spørgsmål:

1. Hvordan anvender børn og unge deres fritid?
2. Er der særligt udsatte grupper af børn og unge uden faste fritidsaktiviteter?
3. Savner børn og unge aktivitetsmuligheder i deres fritid?
4. Hvad synes børn og unge er vigtigt i et fritidstilbud?
5. Hvilke fritidsaktiviteter efterlyses af børn og unge?

Det vil sige, at ambitionen i evalueringen var at dokumentere børn og unges aktuelle fritidsaktiviteter, deres behov for fritidstilbud samt deres bud på evt. nye typer af aktiviteter.

Design

Undersøgelsen gennemførtes indledningsvist med nøglepersoninterview med ansatte i kommunens fritidsområde. Blandt kommunens børn og unge gennemførtes to spørgeskemaundersøgelser, dels en blandt skolernes 4.- til 10.-klasses-elever, dels en blandt de unge mellem 16-18 år uden for folkeskolen. Desuden gennemførtes en interviewundersøgelse blandt forskellige udsatte grupper af unge, som ikke forventedes at deltage i spør-

geskemaundersøgelsen for at sikre en bred repræsentation af forskellige grupper af børn og unge i kommunen.

Dataindsamling

Nøglepersoninterview

Der gennemførtes indledningsvis interview med 10-12 ansatte i nøglepositioner inden for børn og unge-området, herunder bl.a. fritidskonsulenter, klubmedarbejdere, foreningsmedarbejdere, SSP (skole- og socialområdet og politi) konsulenter, boligrådgivere og nærpolti.

Spørgeskemaundersøgelser

Fire folkeskoler og to privatskoler i forskellige områder i kommunen indgik i undersøgelsen med henblik på at sikre, at børn og unge fra kommunens forskellige kvarterer indgik i undersøgelsen, dvs. børn med forskellig socioøkonomisk baggrund og fra forskellige typer af boligområder, dvs. parcelhuse og etagebyggerier.

Spørgeskemaundersøgelsen blandt børnene i skolealderen gennemførtes på skolerne via internettet, hvor skoleeleverne skulle svare på et elektronisk spørgeskema, idet alle skoler i teorien skulle kunne give deres elever mulighed for adgang til net-opkoblet pc'er. Uni-c leverede den elektroniske platform til besvarelse af spørgeskemaet.

Alle de unge i alderen 16-18 år modtog spørgeskemaet i papirudgave på deres hjemadresse og et link til besvarelse på internettet med et personligt kodenummer. De kunne således selv vælge, om de ville besvare papirudgaven eller netudgaven af spørgeskemaet. Med henblik på at motivere de unge til at besvare spørgeskemaet udloddedes tre gavekort af 500 kr. til en musikforretning.

Blandt de unge viste det sig overraskende, at de unge foretrak at besvare papirudgaven, og kun få gik på nettet og besvarede spørgeskemaet der.

I alt 1142 børn og unge deltog i spørgeskemaundersøgelsen med 694 skoleelever og 448 unge. Det svarer til en svarprocent på 40%, hvilket ligger noget i underkanten i forhold til en nogenlunde sikker repræsentativ svarprocent på minimum 70.

Den lave svarprocent blandt de 16-18-årige unge overraskede ikke, da unge i den alder er knap så opsøgende i forhold til fritidsaktiviteter, da de er en fase af livet, hvor de er optaget af uddannelse, venner og kærester, og typisk er knap så fritidsaktive som skoleeleverne, hvorfor undersøgelsen set fra de unges side kunne opfattes som knap så relevant.

Blandt skoleeleverne var den lave svarprocent mere overraskende, fordi spørgeskemaundersøgelser gennemført i skolen i skoletiden plejer at give en høj svarprocent (omkring de 90%). Baggrunden for den lave svarprocent var, at skolerne, som stillet i udsigt ved undersøgelsens indledning, i praksis ikke havde mulighed for at give eleverne adgang til net-opkoblede pc'er i det fornødne omfang.

Interview blandt grupper af unge

Der gennemførtes desuden interview med 27 unge i alderen 12-18 år, og nogenlunde lige mange piger og drenge. De unge blev interviewet i skolen, i idrætshallen og i det nærliggende butikscenter samt på en computer-café.

Analyser

Der gennemførtes korrelationsanalyse og statistiske test af de indsamlede spørgeskemadata i form af en såkaldt χ^2 -test. Ved hjælp af denne test er der mulighed for at undersøge, om de sammenhænge, der findes i datamaterialet, er signifikante. Det vil sige, at man ved hjælp af denne test kan finde ud af, om de fundne sammenhænge skyldes tilfældigheder, eller om de rent faktisk er gældende. I analyseresultaterne er alene medtaget de resultater, der er statistisk signifikante, altså afslører sikre sammenhænge, medmindre andet er anført. I analyserne arbejdedes der med et signifikansniveau på 5%, hvilket betyder, at der for alle fundne sammenhænge er mindre end 5% risiko for, at de fundne sammenhænge er tilfældige.

De kvalitative interview gennemførtes som tematiserede interview og analyseredes ud fra de relevante temaer, de repræsenterer.

Hovedresultater

Evalueringen viste, at langt de fleste børn og unge trives godt såvel i fritiden som i skolen og derhjemme. De er aktive i fritiden og synes ikke, at

de har svært ved at få tiden til at gå. Især drengene er aktive i fritiden, og især de yngste børn i undersøgelsen, dvs. de 10-12-årige er aktive i deres fritid. To tredjedele har faste fritidsaktiviteter, det gælder især skolebørnene, og især 4.-klasses-eleverne (80%), mens de 17-19-årige er mindst fritidsaktive. Der er ikke som forventet en sammenhæng mellem, hvor aktive forældrene er og børnenes aktivitetsniveau.

Børnene og de unge vurderer, at der er mange og tilstrækkelige fritidstilbud i kommunen, de savner egentlig ikke tilbud. Blandt de få, der efterlyser flere tilbud, er drengene i 8.-10. klasse samt børn og unge af arbejdsløse forældre. Børnene og de unge efterlyser især flere haller og grønne områder og helst i eget lokalområde, hvor de på eget initiativ kan dyrke boldaktiviteter. De unge efterlyser fitnessmuligheder o.l. i de kommunale klubtilbud.

I øvrigt synes børn og unge, at fritids/pasningstilbud som SFO (skolefritidsordninger) skal ligne et »hjemmemiljø«, dvs. voksne de kan snakke med, mulighed for at trække sig tilbage fra aktiviteter, se tv, spise, adgang til pc'er m.m.

En mindre gruppe børn og unge på omkring 145 børn og unge (13% af de, der deltog i undersøgelsen) syntes marginaliserede i forhold til fritidsaktiviteter og kammeratligt samvær i fritiden. Det kendetegner også denne gruppe, at de heller ikke trives så godt hjemme og i skolen. Der er således behov for særlige opsøgende fritidstilbud målrettet denne gruppe.

Formidling, anvendelse

Evalueringen formidledes i en AKF-rapport, og foredrag for fritids- og foreningslivet i kommunen samt kommunens politiske udvalg, og dannede grundlag for den videre planlægning af fritidstilbuddene i kommunen.

8.5 Eksempel: Evaluering af børns udvikling og trivsel i en kommune

Eksemplet er en evaluering gennemført af AKF i en enkelt kommune i år 2000 (Mehlbye m.fl. 2000). Evalueringen skulle bruges fremadrettet med henblik på justering af de børnepolitiske mål i kommunen, samt med hen-

blik en vurdering af, om der var behov for særlige handlingsorienterede tiltag. Målet var tillige en vurdering af den generelle trivsel blandt børn og unge i kommunen.

Overordnet formål, genstand og kontekst

I kommunen havde man opstillet et børnepolitisk mål, hvori der bl.a. stod følgende:

»Vi ser på elever som børn, der har lyst til at lære og gå i skole. Skolen skal udvikle elevernes tillid til egne muligheder... Vi lægger vægt på gensidig tolerance og på at omgås hinanden på en respektabel måde.... Der skal værnes om den daglige undervisningstid, og der bør vises respekt for helhed og ubrudthed i undervisningen.«

Den ønskede evaluering skulle relateres til de opstillede målsætninger, og evalueringen skulle danne baggrund for en drøftelse blandt kommunens børn og unge, skolens lærere, forældre og forvaltninger om miljøet og trivslen blandt børn og unge i skolerne.

I fællesskab med kommunen blev der enighed om følgende overordnede problemformulering for evalueringen:

»Hvilken sammenhæng er der mellem elevernes trivsel i skolen og deres faglige færdigheder, samt med deres fritidsanvendelse og hjemmemiljø«.

Alle børn og unge, der gik i en af folkeskolerne i kommunen, indgik i undersøgelsen.

Evalueringsspørgsmål

Den overordnede problemformulering blev nedbrudt i følgende delproblemformuleringer:

- *Hvordan oplever eleverne det faglige undervisningsmiljø i skolen? Herunder har de lyst til at gå i skole, støttes de af lærerne til at udvikle selvtillid og tro på egne evner og muligheder, og har eleverne en realistisk vurdering af egne evner og muligheder?*
- *Hvordan oplever eleverne det sociale liv i klassen, herunder mellem eleverne indbyrdes og mellem elever og lærere? Herunder oplever den enkelte elev sig accepteret af de andre elever og af lærerne? I hvor vid*

udstrækning er der tale om alvorlige drillerier (mobning) blandt eleverne?

- *Hvordan er elevernes faglige niveau set i forhold til deres generelle trivsel og oplevelse af læringsmiljøet og det sociale liv i klassen?*
- *Hvilke sammenhænge er der mellem hjemmemiljø, fritidsanvendelse og elevernes faglige miljø og deres trivsel i skolen?*
- *Hvor mange og hvilke børn og unge befinder sig i en særlig »truet« position, hvad angår deres skolefærdigheder og generelle trivsel?*

Design

Undersøgelsen bestod dels i en spørgeskemaundersøgelse (i papirudgaven) stilet til alle børn i folkeskolernes femte til niende klasse, dels i en række ens faglige færdighedsprøver i matematik, stavning og læsning til de samme børn. Type af test tilpassedes elevens klassetrin.

Spørgeskemaet var på 18 sider og indeholdt bl.a. spørgsmål om personlig baggrund, elevernes oplevelse af undervisningen, holdninger til skolen, mobning, fritidsanvendelse, venner samt psykiske og psykosomatiske lidelser og medicinforbrug,

Dataindsamling

Spørgeskemaet udformedes og afprøvedes af AKF, som derefter fremsendte det til kommunen, som stod for uddeling af spørgeskemaerne til folkeskolerne, hvor klasselærerne sørgede for, at eleverne udfyldte skemaet over en-to lektioner, alt efter hvor lang tid det tog for alle elever at udfylde spørgeskemaet. Efterfølgende sørgede skoler og kommunen for at samle alle spørgeskemaer sammen og fremsende dem til AKF.

Alle færdighedsprøver blev gennemført af skolerne selv. Såvel spørgeskemaer som færdighedsprøver blev gennemført nogenlunde samtidigt på alle skoler med henblik på at undgå, at spørgeskemaer og test blev drøftet mellem eleverne, hvilket kunne påvirke især spørgeskemabesvarelsene.

Såvel spørgeskemaer som færdighedstest blev påført skolenummer, klassebetegnelse og elevnummer, således at det blev muligt i den efterfølgende analyse at koble elevens spørgeskemabesvarelse sammen med elevens færdighedstest.

Spørgeskemaet blev besvaret af i alt 1967 elever fra 112 forskellige klasser, hvilket svarede til 90% af samtlige elever i de pågældende klasser.

Eleverne reagerede meget forskelligt på undersøgelsen, hvilket blev anskueliggjort i deres kommentarer til spørgeskemaet. Her refereres tre vidt forskellige reaktioner og budskaber:

»Hvorfor er det lige, at vi skal det her? Det er lidt fladt, at I skal vide, hvad mine forældre laver«.

»Nogle flere spørgsmål om kærester, fester og narkotika/kriminalitet – alt det her kan også påvirke skolegangen«.

»Jeg vil have en ny skoleinspektør og dermed BASTA«.

Analyser

Der gennemførtes logistiske regressionsanalyser og statistiske test af de indsamlede spørgeskemadata i form af en såkaldt χ^2 -test.

Hvad angår færdighedsprøverne, blev resultaterne fra stave- og læseprøver, som præstationsmæssigt viste sig at hænge nært sammen for den enkelte elev, angivet som henholdsvis antal point og antal fejl.

På denne måde inddeltes eleverne i tre niveauer, under middel, middel, over middel, i de enkelte prøver, alt efter hvor godt eleven klarede sig i forhold til de øvrige elever på samme klassetrin på alle skolerne. På denne måde skulle der være taget højde for evt. socioøkonomiske forskelle mellem de forskellige skolers elevgruppers sociale baggrund.

Matematikprøverne udgik af undersøgelsen, da de efterfølgende kritiseredes af et flertal af lærerne for at være utidssvarende, idet de mente, at de anvendte prøver (som udvalgte af kommunen) ikke afspejlede den undervisning, der praktiseredes i matematik.

Eleverne kategoriseredes tillige efter deres etniske baggrund, idet der var interesse for at få en viden om evt. forskelle mellem elever med dansk baggrund og elever med etnisk minoritetsbaggrund.

Hovedresultater

Evalueringen viste bl.a., at stort set alle børn og unge trives i fritiden og i skolen. De er glade for at gå i skole, får lavet deres lektier, har fritidsaktiviteter og trives med venner og kammerater. Forældrene støtter dem i deres skolegang, og familien bruger ofte fritiden sammen omkring forskelli-

ge fælles aktiviteter. En lille gruppe på omkring 100-150 elever synes at trives knap så godt.

Stort set alle elever oplever, at de generelt kommer godt ud af det med deres lærere, og at de får den hjælp og støtte fra disse, som de har behov for. Kun omkring 15-10% af børnene og de unge oplever, at de ikke har et godt forhold til deres lærere på skolen.

Især de fagligt stærke elever oplever, at de har et godt forhold til deres lærere. Flertallet af børnene/de unge synes også, at kravene i skolen passer til, hvad de kan magte. Omkring en femtedel oplever imidlertid, at lærerne undervurderer, hvad de kan rent fagligt. Omkring halvdelen af eleverne oplever, at der ofte er støj og uro i klassen i timerne, problemet er større på nogle skoler frem for andre. I de meget urolige klasser klager eleverne også oftere over hovedpine, ondt forskellige steder i kroppen og svimmelhed og de har et større forbrug af smertestillende medicin samt pjækker oftere.

Langt de fleste kommer også godt ud af det med deres kammerater. Lidt under halvdelen af alle børn og unge har imidlertid oplevet at blive mobbet en eller flere gange inden for det sidste år. En lille gruppe børn og unge (9%) oplevede at blive mobbet gentagne gange om ugen inden for det sidste skoleår.

Stort set alle børn og unge er aktive i fritiden og har mindst en fritidsinteresse, især sportslige aktiviteter dyrkes.

Hjemme oplever de, at de får god opbakning fra forældrene, og de har jævnligt aktiviteter sammen med deres forældre. De går til sport sammen med dem, overværer sportsbegivenheder, går på museum, i biografen m.m.

Formidling og anvendelse

Undersøgelsen blev formidlet i en AKF-rapport og blev brugt i de videre børnepolitiske drøftelser i kommunen.

8.6 Eksempel: En kommunes indsats over for særlig socialt udsatte børn og unge

I undersøgelsen (Mehlbye 1995) gennemføres en evaluering af indsatsen over for børn og unge med særlige svære opvækstvilkår grundet en række

alvorlige negative sociale begivenheder i familien. Evalueringen gennemførtes i tre kommuner i hovedstadsområdet.

Overordnet formål, genstand og kontekst

Undersøgelsen har fokus på børn og unge i familier, hvor brud mellem forældrene ved skilsmisse og samlivsophør er et led i en række andre sociale belastninger, såsom vold, misbrugsproblemer og arbejdsløshed. Det evalueres, hvilken hjælp børn og forældre har modtaget fra kommunens socialforvaltning, og hvorvidt de har profiteret af denne.

Det overordnede formål med evalueringen er at danne grundlag for vurderinger af, hvilken støtte og hjælp der bør ydes børn og deres forældre med henblik på at forebygge eventuelle senere eller yderligere vanskeligheder hos børnene.

Evalueringsspørgsmål

I evalueringen søgtes følgende spørgsmål besvaret:

1. Hvordan oplever forældre og børn selv familiens situation? og hvad er deres reaktioner på de hændelser, familien har været udsat for?
2. Hvilken betydning har forældrenes brud fået for børnenes dagligdag?
3. Hvilken støtte og hjælp får børn og forældre til afhjælpning af deres problemer?
4. Hvilken yderligere eller alternativ støtte bør tilbydes disse børn og deres forældre?

Design

Evalueringen gennemførtes som en række casestudier, og bestod af tre undersøgelsesdele:

1. En journalundersøgelse af et udsnit af familiesager i tre kommuners socialforvaltninger. Journalerne udvalgte ud fra følgende kriterier: familier med mindst et hjemmeboende barn under 18 år, hvor forældrene efter et længere samliv (minimum to) havde ophævet samlivet, og som over en længere række år havde været i kontakt med socialforvaltningen. Ved hjælp af journalanalysen registreredes forældres og børns beskrevne problemer og indsatsen over for forældre og børn. Alle data registreredes i et

på forhånd fastlagt skema, så journalerne kunne sammenlignes på tværs af sager.

2. *En kvalitativ interviewundersøgelse med mødre, fædre og børn fra opløste familier i kontakt til det offentlige hjælpesystem.* Det blev kun i få tilfælde muligt at etablere kontakt via familieafdelingen, fordi forældrene ikke ønskede at medvirke i interview, og især ikke ønskede, at deres børn medvirkede i interview, formodentlig fordi de oplevede interviewene som en kontrolfunktion. Derfor etableredes også kontakt via andre instanser, såsom de forskellige interventionsformer, familierne opsøgte, nemlig krisecentre, rådgivninger m.m. Forældre og børn blev interviewet om deres oplevelse af problemerne i familien, og deres erfaringer med hjælpetilbud fra det offentlige. Interviewene gennemførtes som samtaleinterview ud fra en række på forhånd fastlagte temaer.

3. *En undersøgelse af forskellige typer af interventionsformer over for familier i krise.* Undersøgelsen gennemførtes som en kvalitativ interviewundersøgelse af medarbejdere i de tre kommuners socialforvaltning, medarbejdere i en række tilbud til familier i krise, såsom krisecentre og kvindehjem, psykologisk rådgivning og behandling, familieværksted og rådgivningstilbud til børn og unge. Formålet med disse var dels at beskrive forskellige typer af hjælp til familier i krise, dels at inddrage disses erfaringer med arbejdet med familier i krise. Interviewene gennemførtes som semistrukturerede interview.

Dataindsamling

Der gennemførtes en analyse af 45 familiers sagsjournaler. I analysen indgik samlet set 90 børn og unge, som der udførtes selvstændige registreringer på. Der skete en registrering af det enkelte barns situation og iværksatte foranstaltninger et år før forældrenes brud, et år efter forældrenes brud, to år efter bruddet samt på registreringstidspunktet, såfremt bruddet lå længere end to år tilbage. Det var sagsbehandlerne i familieafdelingen, der fandt sagerne frem. Alle sager blev anonymiseret ved registreringen, således at anonymitet og fortrolighed omkring sagerne og familierne sikredes.

Der gennemførtes 12 interview med sagsbehandlere i de tre kommuners familieafdelinger samt 20 medarbejdere i samlet 15 indsatsområder over for børn og forældre i krise.

Endelig gennemførtes i alt 46 kvalitative interview med personer fra i alt 39 familier, heraf interviewedes i alt 20 børn og unge, 20 mødre og seks fædre.

Analyser og teoretisk referenceramme

Der skete en kvantitativ registrering af de 90 børns sagsforløb, idet hvert barn indgik som en selvstændig enhed i analysen.

Interviewene af medarbejderne i de forskellige indsattyper blev analyseret og beskrevet på basis af bestemte på forhånd fastlagte temaer.

Analysen af interview med børn og forældre gennemførtes inden for en hermeneutisk fortolkningsramme og formidledes i form af en række temaer, som især trådte frem i analysen af interviewene.

Analyserne gennemførtes inden for en række forskellige teorier, nemlig følgende:

1. Krise- og konfliktløsningsteorier
2. Teorier om opbrudssituationer i forbindelse med skilsmisse og samlivsophør og følelsesmæssige reaktioner i den forbindelse (elementer i en opbrudssituation)
3. Livsformer og opdragelseskulturer
4. Den sociale arvs onde cirkel.

Hovedresultater

Der er tale om en ret omfattende undersøgelse, så der gengives kun nogle udvalgte resultater i det følgende.

Journalundersøgelsen viser, at der registreres flere problemer hos forældre og børn efter bruddet set i forhold til før bruddet, og at der i den forbindelse iværksættes en række foranstaltninger. Journalanalysen viser også, at forældrene (typisk mødre) synes at profitere af hjælpen målt ud fra, at der registreres færre problemer hos forældrene efter to år og senere efter bruddet, og de selv giver udtryk for, at de har det bedre. Det samme gør sig imidlertid ikke gældende hos børnene, som udviser endnu flere typer af problemadfærd to år efter bruddet.

Dette resultat bør ses i relation til sagsbehandlingernes ofte begrænsede kendskab til børnene, ofte har de aldrig mødt dem/talt med dem, og der foreligger ofte ingen eller meget få erklæringer på børnene. Flertallet af oplysninger er i stedet baseret på forældreoplysninger (for to femtedele af børnene er beskrivelsen af børnene alene baseret på forældreoplysninger) samt mundtlige oplysninger fra samarbejdspartnere.

Børnene og de unge selv synes, at interventionsformer rettet mod familien som helhed sjældent er en hjælp til dem. Det gælder familiebehandling og familierådgivning, som de især oplever som en hjælp til de voksne i familien. Hjælpen skal i stedet være direkte målrettet mod dem som børn og unge, men uden for det etablerede regi af professionelle såsom lærere og pædagoger.

Forældrene efterlyser bl.a. især følgende typer af indsats:

- mere opmærksomhed og engagement fra sagsbehandlingernes side
- lokale anonyme rådgivningsmuligheder
- alternative samarbejdsformer mellem forældre og professionelle herunder især i form af (flere) hjemmebesøg

Børn og unge profiterer især af:

- anonyme rådgivningsmuligheder
- samtalegrupper med børn og unge i samme situation som en selv

Formidling

Evalueringen er formidlet i en AKF-rapport samt ved foredrag i de undersøgte kommuner.

8.7 Afrunding

Generelt er evaluering af indsatsen over for børn og unge ikke et særligt veludviklet felt i Danmark. Selv om der gennemføres forskning på området af fx SFI og RUC, må evalueringer og forskning inden for feltet karakteriseres som sparsom, spredt ud over hele landet på selvstændige forskningsinstitutter og universitetsafdelinger, og ofte er der tale om lokalt forankrede evalueringer og forskningsaktiviteter. Der er ikke et sær-

ligt veludviklet samarbejde mellem børn og unge-forskere, og ikke en selvstændig forskningsinstitution, der alene har til opgave at gennemføre forskning og evalueringer på børn og unge-området.

Samtidig anvendes der rigtig mange penge på tilbud til og indsats over for børn og unge – fx på skoleområdet, specialundervisningsområdet samt døgnanbringelsesområdet, hvilket kunne berettige til mere samlet og planlagt forskning og evaluering af området.

9 Evaluering af social- og sundhedsmæssig indsats over for ældre

Eigil Boll Hansen

Dette kapitel indeholder eksempler på evalueringer af en offentlig social- og sundhedsmæssig indsats over for fortrinsvis ældre. Ved en social- og sundhedsmæssig indsats over for ældre forstås her hovedsagelig en indsats, som retter sig mod nedsat funktionsevne og helbredsproblemer eller konsekvenser heraf. Det kan dreje sig om at forebygge nedsat funktionsevne, at generhverve eller vedligeholde funktionsevnen (fx genoptræning og træning) eller at kompensere for nedsat funktionsevne og helbredsproblemer (fx praktisk bistand eller pleje). Lægelige ydelser og ydelser på sygehus i øvrigt falder uden for denne fremstilling, som er afgrænset til at omfatte praktisk bistand, personlig pleje, sygepleje, hjælpemidler, særlige boliger, aktiviteter og terapi, madordninger, kørsel og forebyggende hjemmebesøg. Ikke alle typer ydelser bliver berørt i det følgende.

Der findes ikke én institution, som har domineret evalueringer inden for dette felt, men AKF har gennemført en del evalueringer og andre analyser, der vedrører den sociale og sundhedsmæssige indsats over for ældre (Frandsen, udateret). Socialministeriet har typisk været initiativtager til større evalueringer på feltet, som har været udført af forskningsinstitutioner eller konsulentfirmaer, men også af Den sociale ankestyrelses analysekontor, som synes at have fået en større rolle. Evalueringer og analyser for Socialministeriet fremgår af ministeriets hjemmeside (www.sm.dk) eller af Styrelsen for Social Service's hjemmeside (www.servicestyrelsen.dk). Evalueringer i form af tilfredshedsundersøgelser blandt brugere af fx hjemmehjælp forekommer i flere kommuner. De sidste udføres ofte i form af selvevalueringer eventuelt med ekstern bistand.

9.1 Særlige udfordringer

Effektmål

En særlig udfordring ved evaluering af en social og sundhedsmæssig indsats, der skal kompensere for nedsat funktionsevne eller helbredsproblemer, er at afgøre, hvad der er kvaliteten i eller effekten af indsatsen, og hvad der i evalueringen skal opgøres som resultat af indsatsen. Er kvalitet og effekt overhovedet det samme?

Hvis kvalitet defineres som *helheden af særpræg og kendetegn ved et produkt eller en serviceydelse, der påvirker dets evne til at opfylde borgerne/kundernes udtalte eller underforståede behov*, kan man sige, at kvalitet er nogle egenskaber ved et produkt eller en ydelse set i forhold til brugers behov. Effekten spiller altså en rolle for bedømmelse af kvalitet².

Med dette udgangspunkt kan man som Davies og Knapp (1981) tage udgangspunkt i en »input-output«-model, som de har udformet til at af-dække institutionskvalitet. De sonderer mellem tre input-former:

1. »Ressource-input«, som i stor grad er knyttet til økonomiske omkostninger, fx bemanning, bygninger, udstyr, forskellige former for forbrug, osv.
2. »Ikke-ressource-input«, dvs. ikke-materielle input, som er (eller kan være) under institutionens kontrol, såsom sider af det sociale miljø, organisering, rutiner.
3. »Ikke-påvirkbare-input« (quasi-input), som er ikke-materielle input uden for institutionens direkte påvirkning, fx beboernes og personalets personlighed, tidligere erfaring som »klient«, institutionens beliggenhed³.

Disse tre typer input vil ifølge Davies & Knapp påvirke »output«. For planlæggere vil det primært være de to første som har hovedinteresse. »Ikke-påvirkbare-input« er vigtig i forskning i institutionskvalitet, idet de kan være med til at påvirke *output*.

De to forfatter sonderer mellem »mellemliggende« (intermediate) output og »endelig« (final) output:

1. »Mellemliggende output« er konsekvenser af input, som ikke nødvendigvis har værdi i sig selv, men som bidrager til konsekvenser, der har en værdi i sig selv (dvs. for »Endelig« output). Eksempelvis aktiv deltagelse i huslige og/eller andre aktiviteter og »klientorienteret« praksis.
2. »Endelig output« er – »those consequences of care that so directly reflect aspects of welfare that they are valued in their own right« (Davies & Knapp, s. 54). Eksempelvis trivsel, synlighed og dødelighed. »Endeligt output« benævnes normalt som »outcome« i evalueringslitteraturen.

Fokus i en evaluering kan således være på kvaliteten af den umiddelbare ydelse og måden, den leveres på (output). Fx hvordan foregår visitationen til hjælp, hvordan og hvor ofte bliver gulvet vasket, hvad er indholdet i hjælp til personlig hygiejne, og hvordan optræder den medarbejder, som udfører ydelsen? Det kan sige noget om kvaliteten af indsatsen ud fra nogle antagelser om brugerens behov og ønsker eller på anden måde fastsatte standarder. En sådan evaluering siger imidlertid ikke noget om effekten af indsatsen.

Fokus kunne imidlertid også være på, hvorvidt brugeren udtrykker at have fået sine umiddelbare behov dækket gennem indsatsen (kan tilnærmelsesvis opfattes som et mellemliggende output). Fx at brugeren får gjort rent på en tilfredsstillende måde, at brugerens får bad hyppigt og grundigt nok, eller at brugeren får medicin efter forskrifterne og i et omfang, der fjerner eller dæmper symptomerne fra en konstateret sygdom. I mange tilfælde måles brugernes tilfredshed med bestemte ydelser. Det er her værd at være opmærksom på, at der som regel skal meget til, før en bruger udtrykker direkte utilfredshed. Selv om en bruger udtrykker tilfredshed, kan der således godt være forhold, som vedkommende kunne ønske sig anderledes. De fleste ældre kan selv udtrykke deres forventningsniveau og deres behovsopfyldelse på forskellige livsområder. Det ene af de efterfølgende eksempler er en evaluering, som bl.a. beror på brugernes vurdering af indsatsen.

Endelig kan fokus i evalueringen være på effekten af ydelsen på fx modtagerens funktionsevne og velbefindende (endeligt output eller outcome). Sådanne behov er ofte, men naturligvis ikke altid, underforståede, fordi modtageren, når det drejer sig om fx hjemmehjælp, primært forventer

at få hjælp til nogle praktiske gøremål i husholdningen og til personlige fornødenheder og derfor først og fremmest udtrykker sådanne behov. Det er dog ofte et mål med en social og sundhedsmæssig indsats at bidrage til, at brugerne i videst muligt omfang får styrket deres funktionsevne i dagligdagen og deres velbefindende. Der er her tale om effekter, som det kan være kompliceret at måle, og som ikke umiddelbart kan afdækkes ved at spørge til brugernes tilfredshed og oplevelse af behovsdækning. Det handler derimod om at afdække brugernes funktionsevne ved at kortlægge aktivitetsniveau på en række felter og at afdække deres trivsel og velbefindende gennem spørgsmål om fx sociale kontakter, ensomhed, tryghed og psykiske velbefindende ud fra grundige overvejelser om, hvilke aspekter af funktionsevne og velbefindende der er relevante i de konkrete tilfælde. I det andet af de efterfølgende eksempler er det forsøgt at inddrage mål for brugernes funktionsevne og velbefindende.

Risiko for bortfald

En social og sundhedsmæssig indsats vil typisk rette sig mod personer, der har lidt et funktionstab eller er i en svækket tilstand, medmindre det drejer sig om en forebyggende indsats. Det kan betyde et bortfaldsproblem for evalueringen, i det omfang evalueringen omfatter indsamling af information om eller fra brugerne. Der vil være brugere, som er i en sådan tilstand af fysisk eller psykisk svækkelse, at det ikke vil være muligt at indhente informationer om deres vurderinger af behovsdækning, velbefindende mv.

Hvis brugernes evne til at udføre en række daglige aktiviteter ønskes belyst, kan dette i mange tilfælde observeres af evaluator, eller information herom kan indhentes gennem information fra en tredjeperson, der kender brugeren godt. Derimod vil det næppe være muligt på pålidelig vis at observere brugerens tilfredshed med en bestemt ydelse eller brugerens psykiske velbefindende, ligesom det ikke vil give mening at hente information fra tredjeperson herom. Selv om vedkommende kender brugeren godt, vil der naturligvis være stor usikkerhed forbundet med denne tredjepersons fortolkning af brugerens tilfredshed eller tilstand.

Når det drejer sig om at belyse funktionsevne, aktivitetsniveau o.l., kan det således være en løsning at inddrage en tredjeperson, der kender

brugeren godt, som informant. Det bedste vil være en ægtefælle, et barn eller anden nærtstående, men kan også i nogle tilfælde være en professionel, der yder bistand til brugeren, når emnet er neutralt i forhold til den professionelles indsats. Tredjeperson kan fx inddrages ved, at pågældende er til stede under et personligt interview af brugeren, eller ved at tillade, at brugeren får hjælp fra nærtstående til at udfylde et spørgeskema.

Man må imidlertid affinde sig med, at der til belysning af emner som fx tilfredshed, holdninger og psykisk velbefindende blandt brugere af en social og sundhedsmæssig indsats vil være et uundgåeligt bortfald. »Hjælp« fra tredjeperson vil i disse tilfælde ikke give pålidelige resultater.

Under alle omstændigheder bør et spørgeskema eller et interview være så enkelt som muligt, så kort som muligt og tilpasset populationen, som det altid bør være tilfældet ved gennemførelse af spørgeskema- og interviewundersøgelser.

9.2 **Eksempel: Akutstue i Rødning Kommune**

Evalueringen af akutstuen i Rødning Kommune indgik sammen med evalueringer af andre typer af indsats i et behandlingsforløb (Hansen m.fl., 1997). Ordningen havde allerede eksisteret nogle år, da evalueringen blev udført.

Undersøgelsens formål og genstand

Formålet med akutstuen i Rødning Kommune var at afprøve foranstaltninger i den primære sundhedssektor, der kan erstatte akutte sygehusindlæggelser, som overvejende er betinget af et pleje- og aflastningsbehov. Emnet for den konkrete evaluering af akutstuen i Rødning var således at belyse, i hvilken udstrækning ordningen forebygger indlæggelse på sygehus, og i hvilken udstrækning den dækker de henvistes behov for behandling og pleje. Resultatet af evalueringen skulle tjene til inspiration for aktører i den primære og sekundære sundhedssektor vedrørende alternativer til sygehusindlæggelse betinget overvejende af et plejebestandsbehov.

Et ophold i akutstue i Rødning Kommune er et tilbud om pleje og observation, som kan sættes i værk straks ved en sygdom eller svækkelse, der

vurderes alternativt at ville have ført til indlæggelse på sygehus. Et ophold i akutstue kan foregå på et plejecenter eller i eget hjem.

Akutstuens målgruppe er alle borgere,

- hvor praktiserende læge i en akut sygdomsfase finder behov for indlæggelse til observation, pleje eller behandling ud over, hvad der normalt kan tilbydes i borgerens eget hjem,
- men hvor akut sygehusdiagnostik og -behandling ikke findes fornøden, og
- hvor den akutte fase kan forventes overstået inden for normalt otte dage, og udskrivning kan ske inden for 14 dage.

Målgruppen forventedes især at omfatte ældre. Det afgørende for at tilhøre målgruppen for et akutophold er, at alternativet skønnes at have været et sygehusophold. Patienter henvises i akutstue efter aftale mellem praktiserende læge eller vagtlæge og vagthavende hjemmesygeplejerske. Et akutstueophold indebærer sygeplejemæssig overvågning og kan oprettes i et af kommunens tre plejecentre, gennem fast vagt i hjemmet eller gennem tilsyn i hjemmet fra den integrerede ordning (døgnhjemmepleje). Det årlige antal visiterede til ordningen har svinget omkring 20.

Evalueringsmodel, design og metoder

Evalueringen bestod i en kombination af en procesevaluering og en effektevaluering. Der ønskedes en indsigt i ordningens opbygning og funktionsmåde, mens der ikke sigtedes mod at komme med forslag til ændringer i funktionsmåden. Der ønskedes endvidere en viden om, hvorledes ordningen imødekommer brugernes behov, og om ordningen forebygger indlæggelse på sygehus, så der kunne konstateres en nedgang i antal indlæggelser på sygehus af ældre borgere fra Rødning Kommune.

For at analysere, hvorledes ordningen imødekom brugernes behov, var hovedtanken at følge udviklingen i de visiteredes selvvalgte helbred og funktionsevne som følge af den behandling, de blev tilbudt. Hovedmodellen var at opgøre udviklingen i helbred og funktionsevne ved at måle helbredsstatus, når behandlingen påbegyndtes og igen tre måneder senere, hvilket vil sige tre måneder efter indskrivningen. De tre måneder er en afvejning af flere hensyn. Der skulle gå en vis periode efter behandlin-

gen, så situationen kunne antages at have stabiliseret sig. På den anden side måtte der ikke gå så lang tid, at der var stor sandsynlighed for, at ny sygdom eller problemer kunne støde til. Der skulle endvidere være gået så lang tid, at langt de fleste måtte antages at være udskrevet fra behandling.⁴ Denne generelle model blev imidlertid modificeret ved evalueringen af akutstuen i Rødning, idet der alene ved indskrivningen blev lavet en struktureret opgørelse de ældres helbredsstatus. Ved opfølgingsbesøget gennemførtes en mere kvalitativ beskrivelse af de ældres forløb og situation samt deres vurderinger heraf. Det skyldtes, at forventningerne til væsentlige forbedringer i det fysiske helbred som en direkte følge af indsatsen på forhånd var begrænsede, og at det beskedne antal brugere udelukkede en prospektiv kvantitativ analyse af udviklingen i brugernes selvvaluerede helbred og velbefindende. Evalueringsperioden strakte sig fra september 1995 til juni 1996. Det var planlagt, at indskrevne i akutstuen i en periode på 3-4 måneder skulle indgå i evalueringen, og det forventedes, at otte personer ville blive indskrevet i løbet af denne periode. Det kom ikke til at holde stik, og perioden måtte forlænges. Frem til maj 1996 var ti blevet indskrevet i akutstuen, men fire var døde kort efter indskrivningen, således at kun seks personer kunne indgå i evalueringen.

På grundlag af et spørgeskema med strukturerede spørgsmål om helbred, mobilitet, udførelse af daglige aktiviteter og velbefindende interviewede en hjemmesygeplejerske ældre, som blev indskrevet i akutstuen. Ca. 12 uger efter indskrivningen gennemførte en forsker en samtale med den ældre om udbyttet af opholdet og om oplevede ændringer i mobilitet, udførelse af daglige aktiviteter, velbefindende og tryghed. Oplysningerne ved indskrivningen blev således hovedsagelig anvendt til at beskrive de ældres tilstand ved indskrivningen og ikke til en før-efter-måling. Udbyttet af akutstueopholdet er således opgjort gennem det opfølgende besøg. Det lille antal cases satte selvfølgelig begrænsninger på muligheden for at konkludere generelt om, hvordan akutstuen imødekommer ældres behov i situationer, hvor der kan være behov for observation, pleje og behandling, men hvor sygehusindlæggelse ikke skønnes nødvendig. Det ville have været hensigtsmæssigt med en længere evalueringsperiode, så der kunne være inddraget flere cases i evalueringen. Det var der imidlertid ikke mulighed for.

En beskrivelse af ordningen blev hovedsagelig baseret på foreliggende skriftligt materiale om projektet og på interview af ledende sygeplejersker. Ordningens funktionsmåde blev belyst gennem interview af medarbejdere, som var tilknyttet akutstuen, og interview af de seks brugere om deres erfaringer med ordningen. Det var her et problem, at brugerne havde vanskeligt ved at huske, hvad der skete i den akutte situation, hvor de blev indskrevet.

En analyse af udviklingen i indlæggelser på sygehus med henblik på at belyse, om akutstuen forebygger indlæggelser på sygehus, blev baseret på foreliggende sygehusstatistik. Analysen skulle afdække, om der kunne konstateres et større fald eller en mindre stigning i indlæggelse på sygehus af ældre i Rødding end i andre kommuner. Der blev beregnet en årlig aldersrelateret indlæggelseshyppighed for årene 1990-1994 for ældre i Rødding Kommune og i to nabokommuner. Tidsserien strækker sig, fra før akutstuen blev etableret til efter dens etablering i august 1992. Sygehusstatistik fra nabokommunerne blev inddraget for at skabe et sammenligningsgrundlag. Dette skulle bidrage til at opfange generelle ændringer i indlæggelseshyppighed, som kunne skyldes ændrede procedurer mv. på sygehus eller sygdoms epidemier. Med ca. 20 indlæggelser årligt måtte forventningerne til virkningerne på indlæggelseshyppigheden naturligvis være begrænsede.

Hovedresultater

Ældre borgere som havde været indskrevet i akutstuen i efteråret 1995 og foråret 1996 var tilfredse med tilbuddet og den hjælp, de fik. De var specielt glade for at blive hjulpet i en kritisk situation, og for at det kunne ske i deres hjem eller tæt derved. Opholdet har først og fremmest hjulpet dem over en kritisk fase, mens tilbud om medicinsk behandling og genoptræning har været individuelt. Ifølge brugernes egen bedømmelse fik de altså deres behov dækket, men det kan ikke på det grundlag konkluderes, at det gælder alle, som har været eller vil blive indskrevet i akutstuen.

Ældre fra Rødding Kommune blev i den undersøgte periode indlagt mindre hyppigt på medicinske sygehusafdelinger end ældre i nabokommunen Gram, som vurderedes at være det bedste sammenligningsgrundlag, og forskellen var blevet større efter akutstuens etablering. Spørgsmålet er

imidlertid, om denne tendens kan tilskrives akutstuen. Det er sandsynligt, at Rødding Kommunes tilbud til ældre har reduceret ældre borgeres sygehusforbrug, men det er ikke muligt at opgøre omfanget, fordi andre omstændigheder kan have haft indflydelse på udviklingen.

Det er usikkert, om alternativet til indskrivning i akutstuen i alle tilfælde faktisk ville have været indlæggelse på sygehus. Alle brugerne blev visiteret ud fra den praktiserende læges og en sygeplejerskes vurdering, at de alternativt skulle have været på sygehus. Det kan imidlertid ikke afgøres med sikkerhed, om det faktisk ville være sket. Det kunne tænkes, at fx familie i nogle tilfælde ville have påtaget sig at passe den ældre. På den anden side kan det ikke udelukkes, at akutstuens eksistens for andre end de indskrevne har givet en tryghed, så indlæggelse på sygehus har kunnet undgås.

Formidling

Resultaterne af evalueringen af akutstuen i Rødding er formidlet som en del af en forskningsrapport (Hansen m.fl., 1997a). Resultaterne blev endvidere formidlet i en pjece (Hansen m.fl., 1997b) og i artikler.

9.3 Eksempel: Hjemmehjælpernes arbejdsprincipper og ældres velbefindende

Hvad betyder hjemmehjælpernes måde at arbejde på for hjælpmodtageres velbefindende? Dette er et eksempel på, hvordan forskellige metodiske tilgange kan kombineres til at fortolke effekt af en bestemt indsats; i dette tilfælde hjemmehjælperes måde at yde hjælp på (Hansen, Eskelinen og Madsen, 1999).

Evalueringens formål og genstand

Man kan kort sige, at undersøgelsens genstand var hjemmehjælp og dens udførelse i forhold til brugerne, hvor brugerne i dette tilfælde blev afgrænset til at skulle være 70 år eller derover. Det overordnede formål med undersøgelsen var at bidrage til at udvikle arbejdet inden for den kommunale hjemmepleje. Målgruppen for undersøgelsen var dermed primært le-

dende medarbejdere på ældreområdet, men også hjemmehjælpere og lokalpolitikere, der udstikker rammerne for arbejdet i hjemmeplejen. Intentionen var at afdække medarbejdernes arbejdsprincipper i form af deres holdninger og adfærd i den kommunale hjemmepleje og beskrive betydningen heraf for kvaliteten af ydelsen, hvilket i denne undersøgelse defineres ved effekten på ældres psykiske velbefindende og udnyttelse af egne ressourcer. Det var også hensigten at belyse, hvordan vilkår og forudsætninger for arbejdets udførelse slår igennem i den arbejdspraksis, som de udgående medarbejdere i den kommunale hjemmepleje udøver.

Antagelsen for evalueringen var, at hjemmehjælpernes måde at udføre hjælp på beror på nogle holdninger til, hvad der anses for vigtigt i forhold til at hjælpe brugerne. Man kan dermed sige, at måden, hjælpen ydes på, beror på nogle principper hos hjemmehjælperne. Der blev taget udgangspunkt i to overordnede principper:

- *et opgaveorienteret princip*, hvor hjemmehjælperne er orienteret mod at udføre den bevilgede ydelse
- *et personorienteret princip*, hvor hjemmehjælperne primært er orienteret mod brugernes velbefindende

I forhold til det førstnævnte princip vil hjemmehjælpernes indsats være meget bundet til de ydelser, der er bevilget, mens indsatsen i forhold til et personorienteret princip vil være mere indrettet på brugernes aktuelle situation og behov. Der var således nogle antagelser om, hvordan de to principper ville give sig udslag i hjemmehjælpernes måde at give hjælp på og på deres samspil med brugerne. Det forventedes, at et personorienteret princip ville have en positiv virkning på brugernes velbefindende og udnyttelse af egne ressourcer. Strategien i evalueringen blev at søge at afdække, hvornår et personorienteret princip dominerede, mens et opgaveorienteret princip antoges at dominere, hvis et personorienteret princip var fraværende eller kun til stede i beskeden grad.

Evalueringsmodel og design

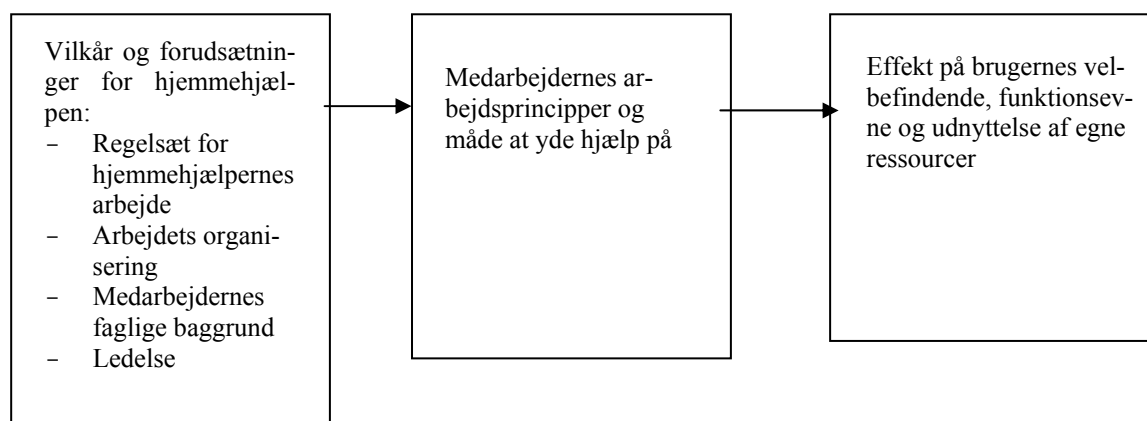
Undersøgelsen bestod af en procesevaluering af udførelsen af hjemmehjælp i to distrikter i Silkeborg Kommune og to distrikter i Svendborg Kommune og en effektevaluering af resultatet heraf på brugernes velbe-

findende og udnyttelse af egne ressourcer. Ved at vælge to kommuner og to distrikter i hver kommune var det muligt at analysere hjemmehjælpens udførelse under forskellige vilkår.

I figur 9.1 er skitseret en simpel model for den forventede sammenhæng mellem hjemmehjælpernes arbejdsprincipper og ældres velbefindende. Projektet var designet efter denne model. Den illustrerer, at medarbejdernes arbejdsprincipper og arbejdspraksis antages at have indflydelse på effekten af hjemmehjælp, som i analyserne måles ved de ældres velbefindende, funktionsevne og udnyttelse af egne ressourcer. Den illustrerer også, at arbejdsprincipper antages at være påvirket af vilkår og forudsætninger for arbejdet, hvilket kan være regler for arbejdet, arbejdets organisering, ledelsesstil og medarbejdernes faglige baggrund. Ældres velbefindende, funktionsevne og udnyttelse af egne ressourcer er også påvirket af andre faktorer fx helbredet, husstandsforhold og uformelt netværk. Disse er ikke vist i figuren, men inddraget i analyserne.

Figur 9.1 Skitse over forventede sammenhænge

9.1



Det er imidlertid vanskeligt at afgøre, i hvor høj grad ældres funktionsevne er bestemt af den ældres fysiske kapacitet, og i hvor høj grad den er et resultat af omgivelsernes krav og forventninger til den ældre. Denne tvivl betyder, at funktionsevnen ikke bruges som en variabel til at måle virkningen af et bestemt arbejdsprincip hos hjemmehjælperne, men som en baggrundsvariabel, der udtrykker brugernes fysiske kapacitet.

Undersøgelsen bestod således af tre dele svarende til indholdet i de tre kasser i figuren og en analyse af sammenhængen mellem dem. I den første

del afdækkedes vilkår og forudsætninger for hjemmehjælpernes arbejde. Den anden del har bestået i at indsamle viden om hjemmehjælpernes måde at arbejde på over for brugerne, mens den tredje del skulle kortlægge ældres funktionsevne, udnyttelse af egne ressourcer og velbefindende. Disse dele uddybes nærmere i det næste afsnit.

Datagrundlag

Den første del af projektet har haft til formål at give et billede af vilkår og forudsætninger for hjemmehjælpernes arbejde i de to udvalgte kommuner. Oplysninger blev indhentet på to måder. Et struktureret skema, som er udfyldt af ledelse og administrative medarbejdere i de to kommuner, omfattede bl.a. organisationsstrukturen i forvaltningen, organisering af hjemmehjælpen (fx graden af selvstyre i grupperne), målsætninger for hjemmehjælpernes arbejde, regelsæt og instruktioner for hjemmehjælpernes arbejde samt visitationsprocedure og -kompetence. Oplysningerne dækker hele kommunen. Desuden gennemførtes interview med ledelsesrepræsentanter fra hver af de fire områder.

Hovedvægten i projektet lå i *anden del*, som bestod i at få oplysninger om og at forstå hjemmehjælpernes måde at arbejde på over for deres ældre brugere. Disse oplysninger har været brugt til at definere forskellige arbejdsprincipper, som hjemmehjælperne anvender. Dataindsamlingen er gennemført ved en spørgeskemaundersøgelse blandt hjemmehjælperne i de fire områder og ved uddybende interview af et antal udvalgte hjemmehjælperne. Spørgeskemaundersøgelsen har skullet kortlægge udbredelsen af bestemte arbejdsprincipper, og spørgsmål om udførelsen af arbejdet, samspillet med brugerne og kontakten med brugerne har været anvendt til at afgrænse arbejdsprincipper, som danner grundlag for beskrivelsen af sammenhængen mellem arbejdsprincipper og ældres velbefindende mv.

Gennem delvis strukturerede interview af udvalgte hjemmehjælperne blev der sat fokus på at få en mere dybtgående beskrivelse af hjemmehjælpernes måde at arbejde på, herunder hvordan de forstår og fortolker deres opgave over for ældre brugere. Blandt de hjemmehjælperne, som deltog i spørgeskemaundersøgelsen, blev der udvalgt i alt otte hjemmehjælperne, som skulle repræsentere forskellige aspekter af et personorienteret arbejdsprincip. Hjemmehjælperne blev udvalgt på baggrund af deres besvarelse af

tre spørgsmål i spørgeskemaet: høj grad af fleksibilitet (udfører ofte noget andet end det aftalte), prioritering af at have kendskab til brugerens livshistorie samt prioritering af aktivering af brugeren (brugeren hjælpes til at kunne deltage i en aktivitet).

Samtidig med spørgeskemaundersøgelsen blandt hjemmehjælperne gennemførtes *en tredje del*, som bestod af en spørgeskemaundersøgelse og interview blandt ældre brugere af hjemmehjælp. Spørgeskemaundersøgelsen skulle kortlægge ældres velbefindende (fx ensomhedsfølelse og humør), funktionsevne (fx egenvurdering af helbred, selvhjulpethed ved daglige gøremål) og udnyttelse af egne ressourcer (fx aktiviteter og interesser hjemme og uden for hjemmet), der indgår som mål for kvaliteten/effekten af hjemmehjælpen. Ældres velbefindende og udnyttelse af egne ressourcer er imidlertid ikke alene påvirket af hjemmehjælpernes indsats, men også af om de ældre bor alene, om de har kontakt med familie mv. og af helbred og førlighed. Dette er også afdækket i skemaet.

I forlængelse af spørgeskemaundersøgelsen gennemførtes delvist strukturerede interview af udvalgte ældre, hvis hjemmehjælper også blev interviewet. Disse interview skulle give mere udførlige informationer om ældre brugeres livssituation og velbefindende, forventninger til hjemmehjælpen og hvad hjemmehjælpen betyder for dem samt relationen til hjemmehjælperen.

Hjemmehjælperne til undersøgelsen blev afgrænset til at omfatte alle faste hjemmehjælperne, som arbejdede i dagtimerne, og som havde været ansat i deres hjemmehjælpergruppe/sektion i mindst en måned. I undersøgelsen indgik *ikke* elever, og personer i arbejdstilbud, puljejob eller lignende. Vi ville sikre os, at populationen bestod af hjemmehjælperne, der havde en vis erfaring at bygge deres udsagn på. Populationen af *hjemmehjælpsmodtagere* omfattede alle brugere af hjemmehjælp på 70 år eller ældre, der boede i en selvstændig bolig. Undersøgelsen omfattede *ikke* ældre på plejehjem, i skærmede enheder eller bofællesskaber. Der blev udelukkende medtaget brugere, hvis hjemmehjælper også indgik i undersøgelsen. Ved udsendelsen af spørgeskemaer til hjemmehjælperne og brugere blev der anvendt et identifikationsnummer på de enkelte skemaer, som gjorde det muligt at koble besvarelsen fra den enkelte hjemmehjælper med besvarelser fra vedkommendes brugere i den efterfølgende databearbejdning.

Der blev udtaget en stikprøve på 204 hjemmehjælpere. Heraf besvarede 193 skemaet, hvilket giver en svarprocent på 95. I alt blev 1175 hjemmehjælpsmodtagere udtrukket til undersøgelsen, og 942 besvarede et spørgeskema, svarende til en besvarelsesprocent på 80. Fra 812 hjemmehjælpsmodtagere opnåedes en besvarelse fra både dem selv og deres hjemmehjælper.

Analysemetoder

Spørgeskemaundersøgelserne blev gennemført med postudsendte skemaer. Hvis økonomien havde rakt til det, ville personlige interview af brugerne have været at foretrække, fordi det ville have givet mere komplette besvarelser. Der opnåedes en tilfredsstillende samlet svarprocent, men der var enkelte spørgsmål, hvor der var et relativt stort bortfald.

Den statistiske analyse af hjemmehjælpernes arbejdsprincipper og sammenhængen mellem disse og de ældres velbefindende samt brug af egne ressourcer blev gennemført i to trin.

I første trin analyseredes spørgsmål om arbejdets tilrettelæggelse, hjemmehjælpernes samspil med brugerne og hjemmehjælpernes kontakt med brugerne for at afdække og afgrænse bestemte arbejdsprincipper, fx om hjemmehjælperne var fleksible eller havde en god social kontakt med brugerne. Til brug herfor har været gennemført en faktoranalyse for at finde frem til spørgsmål, hvor svarene varierer sammen. En faktoranalyse er en statistisk teknik til at afdække et relativt lille antal faktorer, som kan bruges til at repræsentere variabler, der samvarierer indbyrdes. Det antages således, at der bag spørgsmål, som varierer sammen, findes en uidentificeret faktor, som disse spørgsmål tilsammen beskriver. Faktoranalysen er i dette tilfælde anvendt til at udskille spørgsmål om hjemmehjælpernes arbejds måde og adfærd, som samvarierer, og som derfor tilsammen antages at beskrive et bestemt arbejdsprincip.

I det anden trin analyseredes sammenhængen mellem de fundne arbejdsprincipper og ældres velbefindende og brug af egne ressourcer. Analyserne har været gennemført i form af en *logistisk regression*, som groft sagt analyserer sandsynligheden for, at en bestemt begivenhed eller tilstand forekommer, når en række bestemte omstændigheder er til stede. Ved denne analysemetode kan betydningen af hjemmehjælpernes arbejdsprincipper

analyseres, når der samtidig tages højde for andre faktorer, som kan påvirke ældres velbefindende og brug af egne ressourcer.

Interviewene skulle uddybe spørgeskemaundersøgelsens beskrivelse af hjemmehjælpernes arbejdsprincipper og betydningen heraf for brugerne. Materialet fra interviewene blev struktureret og fortolket ud fra tre hovedemner:

- hvordan arbejdsprincipperne udmønter sig i praksis hos den enkelte bruger, og hvordan den enkelte hjemmehjælper forholder sig til sin opgave
- hvordan henholdsvis hjemmehjælpere og brugere vurderer samspillet
- hvilken betydning brugerne tillægger samspillet med hjemmehjælperen.

Hovedresultater

Hovedkonklusionerne fra undersøgelsen er følgende:

Der kan hos hjemmehjælperne observeres måder at arbejde på, der tilgodeser behov, som ligger ud over de aftalte praktiske opgaver hos brugerne. Nogle hjemmehjælpere er meget fleksible over for brugernes ønsker og behov, andre lægger vægt på at have et godt kendskab til deres brugere, at have tid til personlig kontakt med dem, eller de er gode til at snakke med deres brugere. Atter andre hjælper brugerne med at komme i gang med en aktivitet i hjemmet eller at komme ud til et arrangement. Hos nogle optræder flere af disse principper i måden at arbejde på, og det er de færreste, hvor ikke mindst ét af principperne viser sig tydeligt i arbejdsmåden.

Der er en sammenhæng mellem hjemmehjælpernes arbejdsmåde samt brugernes psykiske velbefindende og udnyttelse af egne ressourcer. Hvis hjemmehjælperne er fleksible, viser interesse for deres brugere og støtter brugerne til at være aktive, så kan det hjælpe på brugernes humør, ensomhedsfølelse og aktivitetsniveau. Fleksible og igangsættende hjemmehjælpere kan betyde, at brugerne oftere synes, at hjælpen er tilstrækkelig.

Det sociale samvær med hjemmehjælperen, og at hjemmehjælperen interesserer sig for brugerne og deres liv har stor betydning for brugerne. De sætter pris på at få hjælp til de praktiske opgaver, men de fremhæver det vigtige i en god menneskelig kontakt med hjemmehjælperen.

En bestemt arbejdsmåde er formet af regler og vilkår i det område, hvor hjemmehjælperen arbejder, men også hjemmehjælperens engage-

ment, holdninger, erfaring eller personlige egenskaber har betydning for arbejds måden.

Hovedkonklusionerne baserer sig på en fortolkning af såvel de statistiske analyser som bearbejdning af de kvalitative interview. De statistiske analyser har først og fremmest bidraget til at afdække, hvorvidt der er sammenhænge mellem hjemmehjælpernes arbejdsprincipper og ældres velbefindende, mens de kvalitative interview bl.a. har gjort det muligt at uddybe hjemmehjælpernes måde at arbejde på, beskrive, hvad brugerne finder vigtig i deres relation til hjemmehjælperne, og beskrive, hvordan de har oplevet relationen til hjemmehjælperne. De kvantitative analyser kunne afdække, hvilke arbejdsprincipper der slog igennem hos hjemmehjælperne, og som de kvalitative interview skulle fokusere på, mens de kvalitative interview har kunnet underbygge de sammenhænge, som den statistiske analyse afdækkede.

Formidling

Dokumentationen for undersøgelsen og dens resultater er offentliggjort i en forskningsrapport (Hansen, Eskelinen og Madsen 1999), men hovedresultaterne blev også formidlet i en journalistisk bearbejdet pjece (Andersen 1999). Endelig er resultaterne formidlet til relevante målgrupper gennem foredrag.

10 **Evaluering af integration af etniske minoriteter**

Jill Mehlbye og Eskil Heinesen

10.1 **Evalueringer på integrationsområdet**

Der er i samfundet meget fokus på integrationen af indvandrere og flygtninge. Det er et spørgsmål om livskvalitet, herunder om indvandrere og flygtninge lever under tilfredsstillende forhold, om de får samme muligheder på arbejdsmarkedet, uddannelsessystemet m.m. som danskere, om de føler sig generelt accepteret i samfundet, og om de rent faktisk bliver accepteret m.m. Det er også et økonomisk spørgsmål. Flygtninge og indvandrere må ofte have støtte for fuldt ud at blive integrerede i samfundet. Det gælder især den sproglige indsats, støtteordninger på arbejdsmarkedet, økonomisk støtte, fordi de ikke er i beskæftigelse, støtte i skolen og i ungdomsuddannelsessystemet.

Samtidig er der opmærksomhed på de problemstillinger, der rent faktisk gør sig gældende omkring flygtninge og indvandreres integration i det danske samfund. Det er svært for dem at komme ind på arbejdsmarkedet på grund af sproglige, uddannelsesmæssige, men også kulturelle barrierer. Det er svært for de unge at klare sig i ungdomsuddannelsessystemet, og der er et stort frafald fra ungdomsuddannelserne. Endvidere er der hele spørgsmålet om boligsituationen, hvor ghettodannelsen, dvs. koncentrationen af indvandrere og flygtninge i bestemte boligområder, drøftes med hensyn til fordele og ulemper ved sådanne koncentrationer i forhold til integrationen i det danske samfund.

De mange problemstillinger, der gør sig gældende, og de særlige indsatser, der sættes i værk, betyder også, at der er et stort behov for viden om,

hvilken betydning disse indsatser har for integrationsbestræbelserne, dvs. om de rent faktisk virker som tænkt.

Der har været gennemført en række enkeltevalueringer bl.a. på SFI, AKF og universiteterne. Men der er også behov for en mere samlet videnindsamling på området. Denne opgave har AMID (Akademiet for migrationsstudier) påtaget sig. AMID består af en række forskningsinstitutioner og forskere, der har særlig ekspertise inden for forskellige aspekter af migration, etniske spørgsmål, medborgerskab og nationalisme både i Danmark og i internationale, transnationale og globale processer. AMID gennemfører både erfaringsopsamlinger samt gennemfører ny forskning på området. AMID er etableret på basis af en femårig bevilling fra Humanistisk forskningsråd og Statens Samfundsvidenskabelige Forskningsråd, og indledtes januar 2001. Der udgives såvel arbejdsrapporter som nyhedsbreve fra AMID (se også www.amid.dk).

10.2 **Særlige udfordringer ved evalueringer på integrationsområdet**

Når der skal gennemføres evalueringer på området, hvor de berørte selv også indgår, gør der sig særlige problemstillinger gældende.

Der gælder først og fremmest sproglige barrierer. Det betyder, at der skal bruges tolkebistand. Selv om der er tolkebistand, kan der alligevel opstå problemer, fordi der kan være tale om forskellige sprog og dialekter. Fx oversatte AKF et spørgeskema til tyrkisk i en evaluering af indsatsen over for tosprogede børn og unge i en kommune (Mehlbye 1994). Det var imidlertid ikke velvalgt, fordi gruppen i den pågældende kommune, hvor evalueringen skulle gennemføres, ikke var tyrkere, men kurdere.

Endvidere gør der sig et helt andet problem gældende, som især giver problemer, hvis man vil gennemføre spørgeskemaundersøgelser, nemlig nogle indvandreres og flygtninges ringe uddannelsesniveau, hvilket betyder, at nogle er dårlige læsere eller slet ikke kan læse.

I samme evaluering var erfaringen også, at der gjorde sig en stærk social kontrol gældende, som påvirkede resultaterne. Situationen er den, at selv i storbykommuner og forstadskommuner til hovedstaden, hvor der er

en koncentration af indvandrere og flygtninge, er der ofte et stærkt og tæt socialt netværk mellem de enkelte medlemmer af samme minoritet, med de fordele, det giver i form af social kontakt, men også de ulemper, det giver i form af social kontrol, der betyder, at man langt hen ad vejen følger de fælles normer og holdninger for adfærd og meningsytringer.

10.3 **Særlige udfordringer vedr. effektforskning**

Som nævnt er der særlig interesse for at opnå en viden om, hvilken effekt de særlige indsats har for integrationen af indvandrere og flygtninge. Det betyder et behov for effektevalueringer.

Det kan gøres både ved kvalitative og kvantitative analyser. De kvalitative analyser i form af interview har deres styrke, når det gælder om at forstå mekanismer og problemstillinger omkring integrationen eller mangel på samme. De kvantitative analyser i form af spørgeskemaer og registeroplysninger kan belyse statistiske sammenhænge mellem en indsats og dens virkning.

Ved effektevalueringer baseret på statistiske analyser inden for fx arbejdsmarkeds- og uddannelsesforskningen benyttes således ofte registerdata. Det kan fx være analyser af effekten af forskellige former for aktivering på efterfølgende beskæftigelse eller analyser af effekten af skoleressourcer på unges uddannelsesvalg.

Der er med denne type af evalueringer det generelle problem, at det ikke er tilfældigt, hvilke personer der deltager i hvilke aktiveringsformer, eller hvilke elever der modtager forskellige mængder og kombinationer af skoleressourcer osv. (se også diskussionen i skolekapitlet). Der foregår nogle selektionsprocesser, som afhænger af en række faktorer vedr. det enkelte individ og vedr. det offentlige system, der visiterer forskellige personer til forskellige foranstaltninger og lignende. Nogle af disse faktorer har vi data for, andre ikke. Jo flere af de centrale faktorer, der er data for, jo bedre kan man i statistiske analyser tage højde for selektionsprocesserne.

For indvandrere er data for en række centrale faktorer mangelfulde, eller de mangler helt. Det gælder specielt vigtige baggrundsforhold som uddannelse og erhvervserfaring fra oprindelseslandet samt danskkundskaber tilegnet i Danmark. Ved effektevalueringer på integrationsområdet

spiller uobserverede faktorer derfor en større rolle, end det normalt er tilfældet.

Derfor er det i en række analyser på dette område særlig væsentligt at anvende avancerede statistiske metoder, som gør det muligt at tage højde for selektion på baggrund af uobserverede forhold.

I det følgende gives tre eksempler på evalueringer.

10.4 **Eksempel: Evaluering af kommunernes integrationsindsats – en registerundersøgelse**

Baggrund og formål

Siden 1/1 1999 har kommunerne haft ansvaret for integration af indvandrere omfattet af integrationsloven. På den baggrund bad Ministeriet for Flygtninge, Indvandrere og Integration i 2002 AKF om at gennemføre en såkaldt benchmark-analyse af kommunernes integrationsindsats, dvs. en analyse, der kan bruges som udgangspunkt for en systematisk sammenligning af de enkelte kommuners indsats, herunder identifikation af, hvilke kommuner der har succes med integrationsindsatsen, og hvilke der er mindre succesfulde.

Formålet med projektet er således at opstille en indikator for graden af succes i integrationsindsatsen i forhold til indvandrere omfattet af integrationsloven i de enkelte kommuner. Det overordnede formål er at bidrage til en forbedring af integrationsindsatsen. Indikatoren kan således benyttes som udgangspunkt for en sammenligning af integrationsindsatsen i forskellige kommuner, herunder som udgangspunkt for mere kvalitative undersøgelser af bestemte kommuners indsats (hvilke områder de prioriterer, hvordan de organiserer indsatsen osv.). En sådan dyberegående kvalitativ undersøgelse af nogle få udvalgte kommuner (foretaget af Socialforskningsinstituttet) på baggrund af den her beskrevne benchmarkanalyse er således finansieret af Ministeriet for Flygtninge, Indvandrere og Integration.

Gennem sådanne opfølgende analyser kan man få en bedre forståelse af årsagerne til, at nogle kommuner har større succes med deres integrationsindsats end andre og dermed etablere et bedre grundlag for, at de suc-

cesfulde strategier kan udbredes til flere kommuner, således at kvaliteten i integrationsindsatsen kan forbedres. Mulige brugere af resultaterne er således, foruden Integrationsministeriet, politikere, administratorer og sagsbehandlere i kommunerne, der har med integrationsområdet at gøre.

Det primære mål med integrationsindsatsen er at integrere indvandrede på arbejdsmarkedet, evt. via en periode i uddannelsessystemet. Hvis en stor andel af indvandrerne i en kommune hurtigt opnår ordinær beskæftigelse eller påbegynder en ordinær uddannelse, kan det skyldes, at kommunen yder en effektiv integrationsindsats. Det er baggrunden for at vi i projektet benytter en indikator for kommunernes indsats baseret på den gennemsnitlige varighed fra datoen for opholdstilladelse til påbegyndelse af ordinær beskæftigelse eller uddannelse. Det er imidlertid ikke hensigtsmæssigt at benytte et simpelt mål for denne gennemsnitlige varighed som indikator for kommunernes integrationsindsats, da der er stor variation over kommunerne med hensyn til indvandrernes karakteristika og kommunernes almindelige vilkår. Den opstillede indikator korrigerer derfor, så vidt det er muligt, for sådanne forskelle. Derved opnås et væsentlig bedre grundlag for sammenligning af kommunernes indsats end de simple nøgletal, der sædvanligvis benyttes på dette og andre områder.

Evalueringsmodel

Benchmark-analysen kan karakteriseres som en form for effektevaluering. Det er dog ikke effekten af en bestemt foranstaltning eller ressourceindsats, der evalueres, men effekten af de enkelte kommuners samlede integrationsindsats i forhold til indvandrere omfattet af integrationsloven, idet effekten måles i relation til, hvor lang tid der går, inden indvandrerne påbegynder ordinær beskæftigelse eller uddannelse. Eftersom kun meget få indvandrere omfattet af integrationsloven påbegynder ordinær uddannelse i den betragtede periode, svarer det stort set til, at effekten måles i forhold til, hvor hurtigt der sker en integration på arbejdsmarkedet (målt ved varigheden til første ordinære job).

Data og analysemetoder

Der er anvendt data for 1999-2001 for samtlige indvandrere omfattet af integrationsloven. Data stammer fra de administrative registre i Danmarks

Statistik, herunder AKFs indvandreregister og supplerende registerdata fra Udlændingestyrelsen vedr. bl.a. opholdsgrundlag. Foruden de individrelaterede registerdata er anvendt data på kommuneniveau fra bl.a. Danmarks Statistikbank for lokale arbejdsmarkedsforhold mv.

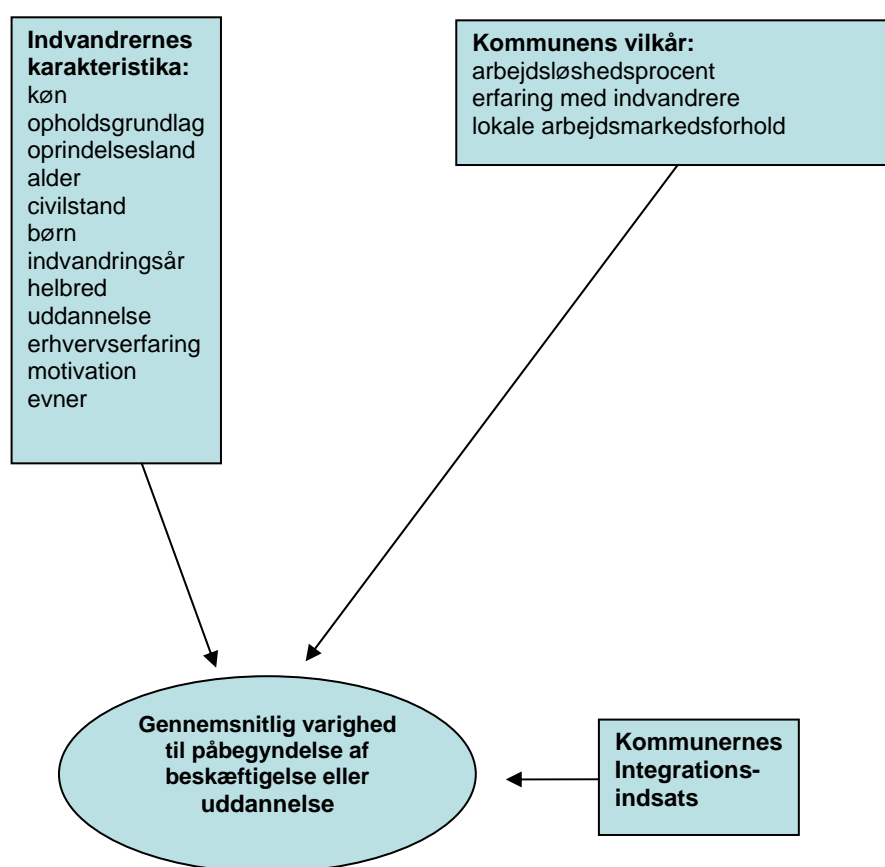
Indvandrere omfattet af integrationsloven er i undersøgelsen afgrænset som personer, der: (1) er kommet til Danmark efter 1/1 1999, (2) var mellem 18 og 61 år på tidspunktet for opholdstilladelse, (3) er kommet fra lande uden for Norden og EU, og (4) er enten flygtninge eller familiesammenførte (til flygtninge eller andre).

Metoden til opstilling af benchmark-indikatoren er baseret på estimation af en statistisk model. Ideen bag opstillingen af indikatoren er illustreret i figur 10.1. Indikatoren for en given kommune tager udgangspunkt i den gennemsnitlige varighed fra opholdstilladelse til påbegyndelse af ordinær beskæftigelse eller uddannelse for de af kommunens indvandrere, som er omfattet af integrationsloven (se den nederste boks i figuren). En kort gennemsnitlig varighed er umiddelbart tegn på en høj grad af (arbejdsmarkeds-mæssig) integration af indvandrere i kommunen. Tilsvarende er en lang gennemsnitlig varighed tegn på en mindre grad af integration. Den gennemsnitlige varighed til påbegyndelse af beskæftigelse eller uddannelse (eller andre mål for integrationsomfang) er imidlertid ikke umiddelbart velegnet som mål for effekten af kommunens integrationsindsats. Som illustreret i figur 10.1 er varigheden påvirket af både indvandrernes karakteristika, kommunens generelle vilkår og kommunens integrationsindsats. For at få en benchmark-indikator for effekten af kommunens integrationsindsats er den gennemsnitlige varighed til påbegyndelse af beskæftigelse eller uddannelse derfor korrigeret for effekten af, at gruppen af indvandrere har forskellige karakteristika i forskellige kommuner, og at kommunernes generelle vilkår er forskellige. Korrektionerne er baseret på en statistisk model for samtlige indvandrere omfattet af integrationsloven i hele landet.

Den beregnede indikator er altså baseret på varigheden fra opholdstilladelse til påbegyndelse af ordinær beskæftigelse eller ordinær uddannelse, men tager samtidig højde for, at gruppen af indvandrere har forskellige karakteristika i forskellige kommuner, og at kommunernes generelle vilkår er forskellige. De karakteristika vedr. de enkelte indvandrere, der tages højde for i den beregnede indikator, vedrører køn, opholdsgrundlag, oprindelses-

land, alder, civilstand, børn (både antal og aldersfordeling), indvandringsår og helbred. Der tages højde for forskelle i disse individuelle karakteristika på et meget detaljeret niveau. Den statistiske model, der ligger til grund for beregningerne, estimeres således særskilt for mænd og kvinder, der indgår over 40 variabler for oprindelsesland, og der beregnes forskellige effekter af alder og børn afhængig af opholdsgrundlag.

Figur 10.1 Illustration af forhold, der påvirker den gennemsnitlige varighed til påbegyndelse af beskæftigelse eller uddannelse for indvandrere omfattet af integrationsloven i en given kommune



De generelle karakteristika vedr. kommunerne, der tages højde for, er den kommunale ledighedsprocent, andelen af indvandrere fra tredjelande i kommunen i forhold til indbyggertallet, og andelen af indvandrere omfattet af integrationsloven i forhold til indbyggertallet. De to sidstnævnte variabler kan fortolkes som mål for kommunens erfaring med integration af indvandrere og omfanget af etniske netværk. En række andre variabler på

kommuneniveau, bl.a. vedr. arbejdsmarkedsforhold og erhvervsstruktur, har været inddraget i analyserne, men de viste sig ikke at være statistisk signifikante.

En negativ værdi af indikatoren for en given kommune betyder, at der i kommunen går kortere tid, end man skulle forvente (givet de nævnte karakteristika), før indvandrerne første gang får ordinær beskæftigelse eller påbegynder ordinær uddannelse, dvs. at kommunen ifølge indikatoren har succes med integrationen af indvandrere. Omvendt betyder en positiv værdi af indikatoren, at der i den pågældende kommune i gennemsnit går længere tid, før indvandrerne opnår beskæftigelse eller uddannelse i forhold til forventet, dvs. at kommunen ifølge indikatoren har mindre succes med integrationen af indvandrere, end man skulle forvente.

Resultater

De estimerede effekter i den statistiske model afspejler, at de forklarende variabler har stor betydning for chancen for hurtigt at påbegynde beskæftigelse eller uddannelse. Således er chancen for hurtigt at påbegynde beskæftigelse eller uddannelse væsentlig mindre for ældre indvandrere end for yngre, den er mindre for forældre til små børn (særligt for mødre), den er mindre for flygtninge end for familiesammenførte, og den er væsentlig mindre for indvandrere fra lande som Libanon, Somalia, Afghanistan, Syrien, Pakistan, Iran og Irak end for indvandrere fra europæiske lande. Når disse (indbyrdes meget forskellige) oprindelseslande fremhæves her, er det, fordi indvandrere fra disse lande ifølge beregningerne har specielt dårlige chancer for hurtigt at få beskæftigelse, og fordi en ikke ubetydelig del af indvandrerne omfattet af integrationsloven kommer fra disse lande. Indvandrere fra Tyrkiet udgør også en stor gruppe (ca. 9%), men de har ifølge beregningerne ikke helt så store integrationsproblemer som indvandrere fra ovennævnte lande. Der kan selvfølgelig være meget forskellige årsager til, at beregningerne indikerer, at indvandrere fra disse lande har særlige problemer med hensyn til arbejdsmarkedsmessig integration. Oprindelseslandet kan til dels afspejle graden af kultur- og sprogforskelle, men også evt. diskrimination på arbejdsmarkedet samt andre forhold, som ikke indgår i analysen på grund af mangel på data, såsom uddannelse og erhvervs erfaring fra oprindelseslandet. Selv om der i analysen i et vist

omfang tages højde for helbredsforhold (nemlig målt ved antallet af kontakter til den primære sundhedssektor), kan de beregnede effekter af forskellige oprindelseslande også afspejle systematiske forskelle i helbredsforhold, fx relateret til krigstraumer.

En større andel af indvandrere fra tredjelande i kommunen i forhold til indbyggertallet og en større andel af indvandrere omfattet af integrationsloven i forhold til indbyggertallet reducerer den forventede varighed til påbegyndelse af beskæftigelse eller uddannelse. Omvendt gælder som ventet, at en højere kommunal ledighedsprocent øger den forventede varighed.

Den opstillede indikator har en væsentlig højere kvalitet end de nøgletal, der sædvanligvis bruges i sammenligninger af kommunerne på dette og andre områder, netop i kraft af, at der tages højde for en række væsentlige forskelle mellem kommunerne. Korrektionen for disse forhold betyder, at en stor del af kommunerne bliver placeret væsentlig anderledes i »rangordningen«, når der sammenlignes med resultatet af blot at benytte den ukorrigerede varighed til påbegyndelse af beskæftigelse eller uddannelse. Korrektionen betyder endvidere, at de kommunale forskelle (dvs. variationen i varigheden) reduceres betydeligt, om end der også efter korrektionen er store forskelle.

Men det skal understreges, at indikatoren ikke er et fuldt dækkende mål for kommunernes integrationsindsats i forhold til indvandrere omfattet af integrationsloven. Der kan som nævnt være faktorer, der ikke tages højde for i beregningen af indikatoren, som har betydning, og som der derfor ideelt set burde tages højde for. Tilsvarende kan der være variabler for kommunernes generelle vilkår, som er relevante, men som vi ikke har kunnet korrigere for. En anden årsag til, at indikatoren ikke er et dækkende udtryk for kommunernes integrationsindsats, er, at den på grund af datamæssige begrænsninger er beregnet for en relativ kort periode, nemlig 1999-2001. Specielt de indvandrere, som er kommet til Danmark i 2000 og 2001, har derfor kun haft kort tid inden for dataperioden til at finde beskæftigelse eller påbegynde uddannelse. En del af kommunernes integrationsindsats, specielt i forhold til personer med svage forudsætninger, vil først give resultater med hensyn til arbejdsmarkedsmæssig integration på længere sigt.

Formidling og anvendelse

Der er udarbejdet en AKF-rapport, der dokumenterer analyser og resultater (Husted og Heinesen, 2004). Ministeriet for Flygtninge, Indvandrere og Integration har anvendt denne rapport som udgangspunkt for deres offentliggørelse af resultaterne. Desuden er der udarbejdet to artikler, som fokuserer på sammenhængen mellem benchmark-indikatoren og kommune-størrelse (med henblik på at belyse evt. stordriftsfordele), se Heinesen og Husted (2004a og 2004b). Det er tanken at supplere med yderligere formidling til en bredere kreds i form af bl.a. korte artikler, der beskriver hovedresultaterne fra projektet. Endelig er det planen at udarbejde en artikel til et videnskabeligt tidsskrift.

10.5 **Eksempel: Unges overgang fra grundskole til ungdomsuddannelse – en kvalitativ og kvantitativ evaluering**

I evalueringen (Mehlbye m.fl. 2000) er der fokus på, hvordan forløbet er fra de unges sidste år i grundskolen, og den vejledning og støtte, de får her over til de første år efter afsluttet skolegang med start på ungdomsuddannelsen eller anden form for aktivitet.

Baggrunden for evalueringen var det store frafald fra ungdomsuddannelserne især blandt unge med etnisk minoritetsbaggrund.

Evalueringens formål og genstand

Formålet med evalueringen var at belyse barriererne for unges start på og gennemførelse af en ungdomsuddannelse med særlig fokus på unge med etnisk minoritetsbaggrund, herunder hvad der kan hindre de unges frafald fra ungdomsuddannelserne. Der var i evalueringen især fokus på betydningen af vejledningen af unge om uddannelse og beskæftigelse.

Målet med evalueringen af de unges forløb og støtten og vejledningen undervejs var at komme med forslag til, hvilken type støtte og hjælp der kan styrke de unges uddannelsesvalg og forbliven på en påbegyndt ungdomsuddannelse.

Evalueringsspørgsmål

Hovedproblemstillingen var:

»Hvordan klarer forskellige grupper af unge sig i ungdomsuddannelsessystemet – med særlig fokus på unge fra etniske minoriteter – og hvad er afgørende for, hvordan de klarer sig?«

Der stilledes følgende evalueringsspørgsmål:

- Hvad er afgørende for de unges valg af ungdomsuddannelse, og hvilke forudsætninger har de for videreuddannelse efter grundskolen?
- Hvilke unge afbryder et uddannelsesforløb og hvorfor?
- Hvilken rolle spiller vejledningen af de unge om uddannelse og erhverv, både for valg af uddannelse, og når de evt. overvejer eller afbryder en ungdomsuddannelse?

Design

Evalueringen tager udgangspunkt i fire kommuner med en stor andel indvandrere og flygtninge, nemlig Ishøj, Brøndby, Karlebo og Farum Kommuner. I de fire kommuner følges de unge, der i 1996 gik i 9. og 10. klasse i kommunens skoler, og fire år frem i tiden frem til 1999.

Der anvendtes både kvalitative og kvantitative metoder dels i form af fire spørgeskemaundersøgelser, dels i form af en afsluttende interviewundersøgelse blandt de unge.

De unge fulgtes over de fire år via spørgeskemaundersøgelser. Første undersøgelse foregik i skoletiden med henblik på dels at sikre en høj svarprocent på dette tidspunkt, dels for at sikre en så høj svarprocent også fremover ved, at de unge var gjort bekendt med undersøgelsen af forskerne selv, og havde ansigt og navn på forskerne, som stod for undersøgelsen.

I første spørgeskemaundersøgelse deltog 92% af de unge (825 unge). Heraf var lidt over en femtedel unge med etnisk minoritetsbaggrund. Frafaldet skyldtes enten, at de ikke var i skole, ikke ønskede at deltage eller at klassens elever ikke var orienteret om, at de ikke skulle deltage i en undersøgelse og derfor ikke var til stede eller parate til at indgå i en undersøgelse.

Det viste sig at blive meget svært at fastholde de unge i undersøgelsesforløbet. Der var flere, der i årene efter grundskolen enten flyttede, og

hvor det ikke var muligt at finde deres nye adresse, eller de tog på længere rejser til udlandet (jf. forældre, der ringede ind til AKF og orienterede om dette).

Derudover var det generelt svært at få de unge til at svare. Resultatet var, at i anden delundersøgelse deltog 64% af de unge. I tredje delundersøgelse deltog 54% af de unge, og fjerde år deltog 58% af de unge. 47% af de unge deltog i alle fire undersøgelser (388 unge). Blandt de unge, som deltog i alle undersøgelsesdele, var kun 78 unge med etnisk minoritetsbaggrund. Det betød i praksis, at evalueringen mistede sin repræsentativitet. Undersøgelsen blev afsluttet med 15 kvalitative interview.

På grund af det forholdsvise store frafald gennemførtes en udførlig frafaldsanalyse for at afdække, om der kunne være bestemte grupper af unge, der var faldet ud af undersøgelsen. Det var vigtigt at vide, fordi det også kunne give en skævvridning i analyserne og resultaterne af evalueringen. Frafallsanalysen viste, at der ikke var større frafald blandt de unge med etnisk minoritetsbaggrund i forhold til unge med dansk baggrund. Men det viste sig, at det især var de »svage unge«, uafhængigt af etnisk baggrund, der var faldet ud af undersøgelsen. I stedet indgik især:

- unge med gode faglige kvalifikationer
- unge, hvis familie boede i ejerbolig
- unge piger
- uafklarede unge, dvs. unge, som i grundskolen ikke vidste, hvilket erhverv de ønskede at få som voksne.

I spørgeskemaer og kvalitative interview udspurgtes de unge bl.a. om tilfredshed med folkeskolen, egen vurdering af egen skole/uddannelses-situation, erfaringer med uddannelses- og erhvervsvejledning, fritids- og hjemmesituation samt ved studieophør baggrunden herfor.

Resultater

Undersøgelsen viser, at der er tre afgørende faktorer for de unges valg af uddannelse, som er medvirkende til at skabe et ulige mønster i rekrutteringen til ungdomsuddannelserne nemlig, etnisk baggrund, køn og social arv.

De unge fra de etniske minoriteter vælger især de erhvervsfaglige uddannelser frem for de gymnasiale uddannelser, som er bogligt mere krævende uddannelser, hvilket også hænger sammen med, at deres faglige niveau generelt er lavere end de danske unges, og oftere ligger under middel. Således er der en klar sammenhæng mellem karakterniveau i grundskolen og valg af ungdomsuddannelse. Samtidig ses en nær sammenhæng mellem social baggrund og karakterniveauet i grundskolen.

På trods af at de unge fra etniske minoriteter generelt klarer sig fagligt dårligere i grundskolen, udtrykker de i større omfang tilfredshed med folkeskolen end de danske unge.

Valget af de praktisk orienterede erhvervsuddannelser giver imidlertid vanskeligheder, idet de unge fra etniske minoriteter angiver, at det er svært for dem at finde en praktikplads, hvilket de angiver som en hyppig grund til, at de må afbryde deres uddannelsesforløb.

Samtidig er det især de fagligt svage unge med forældre med ringe eller ingen uddannelse, der afbryder deres uddannelsesforløb. En del unge angiver, at de afbryder en uddannelse, fordi de har mistet interessen for uddannelsen, oftest har de også fundet en anden uddannelse, de ønsker at gå i gang med.

Samtidig udtrykker de unge fra etniske minoriteter oftere bekymring for fremtiden, herunder for fremtidig arbejdsløshed end de danske unge.

Hvad angår vejledningen i folkeskolen har langt de fleste modtaget uddannelses- og beskæftigelsesvejledning, men de mener alligevel, at forældrene har størst indflydelse på deres uddannelsesvalg. Der er dog meget, der tyder på, at de får en del af deres konkrete viden om de enkelte uddannelsers og erhvervs faglige indhold fra skolevejlederne.

Med hensyn til hvilken form for vejledning de unge ønsker, giver de unge udtryk for, at de især sætter pris på den personlige individuelt orienterede vejledning, der tager udgangspunkt i deres egne interesser. Især unge fra de etniske minoriteter efterlyser mere personlig individuel vejledning og synes ikke den vejledning, de har modtaget, i tilstrækkelig omfang har taget udgangspunkt i deres personlige forudsætninger.

Imidlertid viser undersøgelsen, at omfanget af vejledning ikke synes at have nogen betydning for, hvorvidt man afbryder en uddannelse eller ej,

hvilket kunne tolkes som ikke kvantiteten, men snarere kvaliteten af vejledning har betydning for afbrydelse af en ungdomsuddannelse eller ej.

Endelig viser undersøgelsen, at mange unge ikke har kendskab til ungdomsvejledningsordningen, hvilket især har betydning for dem, der afbryder deres ungdomsuddannelsesforløb.

Den samlede konklusion er, at vejledningen af de unge bør styrkes, især over for unge fra etniske minoriteter, som også har behov for en anden type vejledning end danske unge.

10.6 **Eksempel: Integration af tosprogede børn og unge i daginstitution og skole – en kvalitativ og kvantitativ evaluering**

Evalueringen (Mehlbye 1994) omhandler integrationen af børn og unge fra etniske minoriteter i samfundet. Udgangspunktet var en kommunes integrationsmodel, som enkelt sagt bestod i, at kommunen søgte at undgå en koncentration af tosprogede børn og unge i bestemte daginstitutioner og skoler, og i stedet søgte at sprede dem på forskellige institutioner og skoler, hvilket bl.a. betød, at der var etableret kørselsordninger for kommunens børn og unge med anden etnisk baggrund end dansk. Desuden havde kommunen indskolingsgrupper med tilknytning af lærere i børnehaveklasse til 1. klasse samt intensiv danskundervisning i 3. klasse. På hver skole var der en enkelt lærer, som varetog koordinationen af indsatsen over for de tosprogede børn på skolen. Med hensyn til indsatsen i førskolealderen var der etableret en pædagogisk legestue umiddelbart før børnehaveklassestart til de børn, der ikke havde gået i børnehave, med henblik på at forberede dem til skolegangen i en dansk folkeskole. Der var endvidere tilknyttet 9 sprogformidlere til daginstitutionerne til at varetage de tosprogede børn og unges særlige behov om særlig sproglig stimulering og støtte.

Med hensyn til tilflyttende børn og unge til kommunen direkte fra udlandet, og som ikke mestrede det danske sprog, var der på sprogskolen etableret et etårigt undervisningsforløb for de 7-4-årige, og som efterfølgende ved indslusning i de almindelige klasser tildeltes 140 timers intensiv

danskundervisning i løbet af det første år. Desuden var der etableret en »senior Klasse« for de unge i alderen 14-18 år.

Ved tosprogede børn og unge forstås børn og unge, som taler et andet sprog end dansk i hjemmet. I kommunen var der en klar overvægt af pakistanske og tyrkiske børn.

Evalueringens genstand og formål

Formålet med evalueringen var en vurdering af, om de afsatte midler til integration benyttes hensigtsmæssigt, og hvorvidt de kan benyttes på en anden og bedre måde.

I evalueringen sås på den samlede indsats over for kommunens tosprogede børn og unge i alderen 0-18 år. Der var tale om en ret omfattende evaluering, hvor mange institutioner skulle inddrages i evalueringen. Der var således fokus på de professionelles indsats og betydningen af denne for tosprogede børn og unge.

Evalueringsspørgsmål

De centrale spørgsmål i evalueringen var, hvordan de tosprogede børn og unge integreredes i skoler, som et mål for, hvordan integrationsindsatsen fungerede, samt hvilke forventninger, holdninger og erfaringer, såvel personalegrupper som elever og forældre havde med integrationen.

Hovedspørgsmålene var følgende:

1. I hvor vid udstrækning integreres de tosprogede børn og unge i daginstitutioner, skoler og fritidsaktiviteter set ud fra faglige, sociale og sproglige dimensioner?
2. Hvad hæmmer og fremmer en integration af de tosprogede børn og unge?

Definitionen eller forståelsen af begrebet integration tog udgangspunkt i kommunens generelle indvandrerpolicy, som understregede vigtigheden af at respektere indvandrernes egen kulturelle identitet, og med målet om, at indvandrere i dagligdagen skulle fungere i ligestilling med danskerne. Det primære mål var sproglig integration.

Evalueringen baserede sig primært på en erfaringsindsamling fra personalet i institutioner, skoler og indsatsområder, samt fra skolebørn og forældres egne erfaringer.

De to hovedevaluerings spørgsmål opdeltes i en række delspørgsmål, såsom

- Hvilket udbytte får skoleeleverne af skolegangen?
- I hvor høj grad vurderer det professionelle personale, at de kan imødekomme de tosprogede børn og unges behov?
- Hvordan er samarbejdet mellem personalerne på tværs af institutioner og indsatsområder?
- Hvordan indgår integrationsbestæbelserne i den overordnede planlægning på skolerne?
- Hvor meget er børnene og de unge sammen i institutionen/skolen på tværs af etniske grupper?
- Hvordan former dagligdagen sig for de tosprogede børn og unge set i forhold til danske børn?
- I hvor høj grad stimulerer forældrene deres børn til at være sammen på tværs af etniske grupper – det gælder både danske forældre og forældre med en anden etnisk baggrund end dansk?

I evalueringen var der især fokus på de to største grupper af indvandrere, nemlig indvandrere fra Tyrkiet og Pakistan.

Metoder og data

I evalueringen anvendtes kvalitative metoder med henblik på at opnå en forståelse af integrationsmønstre og kvantitative metoder med henblik på at vurdere omfanget af bestemte mønstre.

I de kvantitative dataindsamlinger gennemførtes sammenlignende undersøgelser mellem danske børn og unge og børn og unge fra etniske minoriteter.

Kvalitative data

I evalueringen indgik kvalitative semi-strukturerede interview med i alt 56 personer, der arbejdede inden for aktivitets- og indsatsområder i førskole, skole og unge/voksenområdet.

Desuden interviewedes politisk og administrativ ledelse i børneforvaltningen og social- og sundhedsforvaltningen.

Endelig interviewedes 20 forældre med tyrkisk eller pakistansk baggrund samt 11 tosprogede elever med pakistanske og tyrkiske forældre, gennemført som tre tematiserede gruppeinterview.

Desuden indgik alt skriftligt materiale fra kommunen om deres arbejde med integrationen.

Kvantitative data

Der gennemførtes spørgeskemaundersøgelser med følgende grupper:

- Lærerne, hvor i alt 233 lærere deltog svarende til en svarprocent på 66
- Ledelsen på folkeskolerne, hvor seks ud af de otte skolers skoleinspektører deltog
- Eleverne på 8. og 9. klassetrin på tre forskellige folkeskoler i forskellige områder af kommunen. 244 elever deltog i undersøgelsen med en svarprocent på 87%. Der var dobbelt så stort frafald blandt de tosprogede elever som blandt de danske elever
- Halvdelen af forældrene på de tre udvalgte skoler blev bedt om at deltage i spørgeskemaundersøgelsen. Spørgeskemaet stiledes til alle forældre i skolens ene klassespør (a-klasserne) fra børnehaveklassen til 10. klasse. I alt 440 forældre deltog med en svarprocent på 62% med nogenlunde lige høj svarprocent blandt danske og forældre fra etniske minoriteter
- Daginstitutionslederne fik tilsendt et skema, der skulle udfyldes på institutionen i et samarbejde med personalet. 56 institutioner svarede med en svarprocent på 85.

Desuden indsamledes statistiske data om elevernes karakterer ved afgangsprøverne i 9. og 10. klasse samt oplysninger om, hvilken uddannelse de fortsatte på efter afsluttet grundskole.

Resultater

Generelt tegnede der sig et positivt billede. Selve integrations-idégrundlaget (forsøget på at sprede børnene og at søge at undgå en koncentration i bestemte skoler og institutioner) var der generelt stor opbakning til og

positiv vurdering af. Det gælder såvel fra de professionelle som fra børnenes forældre.

Der var dog også problemer. I daginstitutionerne mente personalet ikke, at de havde tilstrækkelige ressourcer til at yde børnene den nødvendige støtte, det gjaldt især den sproglige støtte og stimulering. Samarbejdet med forældrene blev hæmmet af deres manglende sproglige færdigheder.

I øvrigt var der bedst erfaring med integrationen af de tosprogede børn i institutioner med få tosprogede elever og overvejende socialt stærke danske børn.

Såvel de tosprogede elever som deres forældre udtrykte stor tilfredshed med skolen. Såvel børn som forældre var meget opmærksomme på, at børnene lærte godt dansk. Samtidig havde såvel børn som forældre store ambitioner om fremtidig uddannelse og erhverv, og nok også større end børnene kan indfri. Det gjaldt især elever med tyrkisk baggrund, hvor forældrene havde forholdsvis ringe eller ingen uddannelse (nogle er analfabeter), og ikke kunne hjælpe deres børn med lektier og generelt havde svært ved at støtte børnenes skolegang, fordi folkeskolen og det, der foregår der, var dem meget fremmed.

Set fra lærerside oplevedes de tosprogede elevers forældre som meget uinteresserede i deres børns skole, når det gjaldt deltagelse i forældremøder o.l., mens deres deltagelse i forældrekonsultationer var noget højere.

Hvad angår de tosprogede elevers faglige udbytte, især i de boglige fag, stod de tosprogede elever ifølge lærernes vurdering svagt, både på grund af sproglige problemer, og fordi de ikke fik den samme støtte hjemme, som de danske elever fik. Derimod klarede de sig lige så godt som de danske elever i de praktiske fag. De tosprogede elevers præstationer var også noget lavere end de danske elevers, når man vurderer det ud fra deres karakterniveau ved afgangsprøverne i 10. klasse. Dette gjaldt dog ikke pakistanske elevers præstationer ved afgangsprøverne i 9. klasse på udvidet niveau, hvor de klarede sig lige så godt som de danske elever.

Med hensyn til de unges hverdag udfoldede den sig meget forskelligt for danske og tosprogede elever, idet de kun brugte lidt fritid sammen på tværs af etnisk forskellige baggrunde, samtidig med at deres anvendelse af fritiden ofte var vidt forskellig.

Evalueringen giver anledning til følgende forslag (stikord):

- Intensivering af dansk undervisningen generelt på alle alderstrin over for de tosprogede børn og unge
- Øget støtte til de professionelle, der skal udføre integrationsarbejdet
- Opbygning af et ressource- og kraftcenter, hvor der opbygges en særlig viden og specialisering på området
- Opkvalificering af kontaktlærerne på de enkelte skoler, så deres faglige profil styrkes med nedtoning af de rent administrative funktioner
- Særlige forældremøder med de tosprogede forældre
- Udarbejdelse af en mere handleorienteret og mere synlig målsætning, som kan danne en fælles ramme for de professionelles integrationsarbejde
- Specialiseret rådgivning og vejledning til de tosprogede børn og unge
- Styrkelse af det tværsektorielle og tværinstitutionelle samarbejde.

Formidling og anvendelse

Undersøgelsen blev primært formidlet til den pågældende kommune i rapportform og ved månedlige oplæg. Kommunen anvendte undersøgelsen til planlægning af det videre integrationsarbejde.

11 Evaluering på handicapområdet

Leif Olsen

11.1 Er evaluering på handicapområdet noget særligt?

Det mest rigtige og dækkende svar må være både ja og nej.

Nej, fordi udfordringerne til evaluatorene på handicapområdet er sammenlignelige med andre områder, når det fx gælder afgrænsning af evalueringens genstand, formulering af hovedspørgsmål, udformning af design, valg af metoder til systematisk indsamling af informationer og data, valg af analytiske fremgangsmåder, inkludering af teorier og eksisterende viden mv. Ja, fordi evaluatorene på flere områder stilles over for særlige udfordringer, når det gælder om: At få statistiske oplysninger om mennesker med handicap, at indsamle oplysninger fra mennesker med kognitive eller kommunikative handicap og formidle evalueringernes resultater til denne målgruppe.

Statistik, indsamling af data og formidling

Der findes af flere grunde ikke statistik, der kategoriserer mennesker med handicap. De to væsentligste grunde hertil er at: 1) handicap er et dynamisk begreb, der defineres i forhold til det konkrete forhold mellem individ og omgivelser og derfor ikke kan bruges som generel og robust statistisk kategori og 2) det er en integreret del af dansk handicappolitik, at handicap altid skal vurderes konkret, fordi faste kategoriseringer (fx diagnoser) kan medføre stigmatisering og fejlagtige vurderinger af handicappeles muligheder og behov. For forskere og evaluatorene betyder det, at mulighederne for direkte at basere evalueringer på statistisk materiale er

meget begrænsede, når det gælder handicap. I det omfang der er brug for statistisk materiale til at styrke forskning og evalueringer, er der således primært to veje at gå. Den ene vej, man kan gå, er selv at indsamle oplysninger (fx via spørgeskemaundersøgelser), der kan danne grundlag for statistiske analyser, som det fx er gjort af Bengtsson (1997) og Clausen (2003). Den anden vej, man kan gå, er at udnytte indirekte statistiske oplysninger (fx registerbaserede oplysninger om fordelingen af offentlige ydelser og forbrug af lægemidler), som det i givne evalueringer kan være relevant at inddrage, som indikatorer til at belyse problemstillinger med relation til handicap.

De særlige udfordringer i forhold til at indsamle oplysninger fra mennesker med kognitive eller kommunikative handicap har altid været der. De er dog kommet yderligere i fokus på grund af, at handicapområdet gennem de seneste tiår er blevet påvirket af internationale konventioner⁵ om ligebehandling af handicappede og ændringer i dansk lovgivning. Fx har retssikkerhedsloven og ophævelsen af institutionsbegrebet⁶ understreget, at offentlige myndigheder skal implementere idealerne om alle menneskers ret til selvbestemmelse og medindflydelse. Der er tale om en udvikling, der skærper kravene til, at der udvikles metoder og kvalifikationer til at skabe dialog med alle borgere. Det samme forhold gør sig gældende for evaluatører og forskere, når det er relevant og vigtigt, at mennesker med kognitive eller kommunikative handicap indgår i evalueringer. Denne udfordring er til en vis grad erkendt, og Socialministeriet har fx bedt Socialt Udviklingscenter SUS⁷ om at gennemføre et metodeudviklingsprojekt, der skal styrke grundlaget for kommunikation med brugere, der ikke har et ekspressivt sprog⁸. I Bengtsson (2003) præsenteres flere eksempler på forskere og evaluatører, der arbejder med udvikling af metoder og evalueringsdesign, der søger at overvinde de særlige kommunikative udfordringer på handicapområdet. En anden vej til generel viden og indsigt på dette område er via Formidlingscentret på Handicapområdet, der har gennemført relevante projekter om brugerinddragelse (fx Sandø 2002 og Gruber 2002) og henviser til viden via deres hjemmeside www.ufch.dk.

Når det er relevant at formidle evalueringer til mennesker med kognitive eller kommunikative handicap, står evaluator over for særlige udfordringer, idet skriftlig formidling ikke altid er tilstrækkelig og derfor bør

suppleres med andre former for formidling fx elektroniske tekster⁹, visuel formidling og forskellige mødeformer, hvor man gør brug af kommunikative hjælpemidler og faglig ekspertise (fx tolkebistand).

Andre komplekse genstandsfelter

En evaluering bør altid begynde med afklaring af, hvad der er evalueringens genstand. På dette punkt ligner handicapområdet andre områder. Der synes derimod at være forskelle, når det kommer til mangfoldigheden og kompleksiteten i de mulige genstandsområder, som evalueringer kan beskæftige sig med. Handicapområdet er således karakteriseret ved en meget stor bredde og sammensathed, der gør det muligt såvel at finde klart afgrænsede genstande som genstande, der er meget vanskelige at afgrænse.

De relativt enkle genstande for evaluering findes, når de kan afgrænses til enten at omhandle medicinske eller biologiske eller psykologiske eller sociale fænomener. De genstande, der er relativt vanskelige at afgrænse findes i særlig grad, når flere af disse forhold spiller sammen og indgår i evalueringens genstand, hvilket også gælder for eksempelvis evalueringer af socialt arbejde. Der kan fx være tale om, at evalueringens genstand både indeholder medicinske, psykologiske og sociale forhold, hvilket gør det endnu vigtigere for evaluator at afgrænse genstanden operationelt for at kunne gennemføre en fokuseret og stringent evaluering. Krogstrup (2003 s. 17) taler i denne sammenhæng om, at mange problemstillinger på det sociale område kan betegnes som »vilde problemer«, som ikke kan defineres klart og entydigt og er vanskelige at adskille fra andre problemer.

Institutionel mangfoldighed

Handicapområdet er karakteriseret ved, at der er mange forskellige interesser og rekvirenter af evalueringer. Denne mangfoldighed modsvarer af en stor variation blandt evaluatorene, der har meget forskellige faglige baggrunde og organisatoriske placeringer. Evalueringer på handicapområdet gennemføres fx af universitetsforskere, sektorforskere, private og offentlige konsulenter og medarbejdere, der købes fri fra deres daglige arbejde til at gennemføre evalueringer.

Viften af organisationer og institutioner, der efterspørger og finansierer evalueringer på handicapområdet strækker sig fra statslige ministerier, over amter og kommuner til institutioner, hvoraf nogle er udbydere af forskellige typer af borgerservice i form af behandling, uddannelse, pleje mv. og andre er videninstitutioner¹⁰. Videninstitutionerne formidler viden og rådgivning på baggrund af deres arbejde med at udvikle, bearbejde og formidle viden på specifikke eller tematiserede dele af handicapområdet (Clausen 1993, Olsen 1998, Det Centrale Handicapråd 2001).

Der gennemføres mange og forskelligartede evalueringer på handicapområdet. Det er ikke muligt at yde den store variation retfærdighed i denne sammenhæng, hvor der i det følgende præsenteres to eksempler på forskellige typer af eksterne evalueringer. Evalueringerne er gennemført af forskere fra AKF i perioden 1997 til 2001, og der er tale om eksterne evalueringer. De to eksempler er forskellige i forhold til bl.a.: skala, genstand, indsamling af informationer og data, valg af evalueringsmodeller og – design med mere.

11.2 **To eksempler: En formativ evaluering og en brugerevaluering**

Det første eksempel er en formativ procesevaluering, der følger implementeringen af en ny kompleks behandlingsmodel for rehabilitering af mennesker med traumatiske hjerneskader på Kolonien Filadelfia (Høgsbro 2002). Eksemplet viser en række generelle udfordringer, der knytter sig til denne type af evaluering og særlige udfordringer i forhold til at indsamle oplysninger gennem interview med mennesker, der har pådraget sig en hjerneskade. Evalueringens budget var på ca. 1.2 mio.kr.

Det andet eksempel er to brugerevalueringer af amternes rådgivning på handicapområdet, hvor brugerne er henholdsvis kommunale medarbejdere (Rieper 2001) og mennesker med handicap (Olsen 2001)¹¹. Eksemplet viser generelle udfordringer til brugerevalueringer og særlige udfordringer i forhold til at finde og udvælge de relevante brugere og gennemførelse af interview med mennesker om kognitive eller kommunikative handicap. Evalueringernes budgetter var på henholdsvis ca. 1.5 og 1.8 mio.kr.

11.3 **Eksempel: Rehabilitering af mennesker med traumatiske hjerneskader**

AKF blev i 1997 valgt af Kolonien Filadelfia til at være ekstern evaluuator af en holistisk og tværfaglig rehabiliteringsmodel for voksne med traumatisk hjerneskade. Der var tale om en ny behandlingsmodel, der blev indført og var under udvikling ved genoptræningscenteret Kurhus på kolonien Filadelfia i Dianalund (Høgsgbro 2002).

Der er tale om en evaluering i lille skala i den forstand, at evalueringen sætter fokus på en mindre del af en større institution og de 29 patienter, der var på Kurhus i den 4-årige periode, evalueringen forløb. Der er samtidig tale om en bred og dyb evaluering, for så vidt angår den indholdsmæssige kompleksitet, idet hjerneskader og den kombinerede behandling heraf gennem, medicinske, psykologiske og sociale indsatser indeholder meget sammensatte problemstillinger.

Formål og hovedspørgsmål

Kolonien Filadelfia ønskede en evaluering af deres nye indsats i genoptræningscentret Kurhus, hvor man havde valgt at indføre en ny holistisk behandlingsmodel til rehabilitering af mennesker med traumatiske hjerneskader.

Evalueringen havde fire mål: 1) at give rehabiliteringscentret, Kurhus, tilbagemeldinger på deres arbejde med en helhedsorienteret rehabiliteringsmodel undervejs i en 2-årig evalueringsproces, 2) at pege på justeringer og udviklingsmuligheder i rehabiliteringscentrets organisation, 3) at delagtiggøre en bredere offentlighed i de problemstillinger, der knytter sig til implementering af den holistiske og tværfaglige rehabiliteringsmodel og 4) at sætte modellen ind i den internationale diskussion omkring behandling af traumatiske hjerneskader.

Evalueringens hovedspørgsmål knyttede sig til tre af de fire formål, idet det fjerde formål sigtede mod formidling af evalueringens resultater. Hovedspørgsmålene var: Hvordan vurderer brugere, personale og pårørende behandlingsmodellen gennem forløbet? Er der behov for justeringer og udviklingsmuligheder i rehabiliteringscentrets organisation? Hvordan pas-

ser behandlingsmodellen ind i den internationale diskussion omkring behandling af traumatiske hjerneskader?

Evalueringsens genstand og kontekst

Evalueringsens genstand var en beskrevet og teoretisk begrundet model for, hvordan man kan optræne (eller genoptræne) færdigheder hos mennesker med traumatiske hjerneskader umiddelbart efter, at de har forladt hospitalets første akutte indsats over for de umiddelbare neurologiske skader. Evalueringen satte fokus på den holistiske og tværfaglige tilgang, der var udgangspunktet for etableringen af rehabiliteringscentret. Det er en tilgang, som lægger op til at møde klientens specifikke individuelle problemstillinger med et tværfagligt team af professionelle, der tager højde for den individuelle klients særlige forudsætninger og fremtidsperspektiv.

Evalueringen indgik i en kontekst, hvor moderorganisationen Kolonien Filadelfia sammen med genoptræningscentret var interesseret i at få dokumenteret og belyst, i hvilken grad det lykkedes at indføre den nye behandlingsmodel. Ud over den direkte faglige interesse i at få indsigt i, om det er muligt at implementere modellen i praksis, så havde evalueringen også den funktion, at den kunne levere legitimation i forhold til omverdenen, herunder bevillingsgivere til og brugere af genoptræningscentret. Hertil hørte også et behov for at evaluere arbejdet med modellen i en international sammenhæng, idet arbejdet på et så specialiseret område som behandling af hjerneskader får større gennemslagskraft, hvis den vurderes efter en international »målestok«.

Evalueringsmodel og -design

Det blev valgt at gennemføre en procesevaluering, der blev designet som en formativ evaluering. Det var således de processuelle aspekter af genoptræningscentrets arbejde med at implementere genoptræningsmodellen, der blev evalueret. Herudover var det sigtet med evalueringen at give rehabiliteringscentret tilbagemeldinger såvel undervejs i evalueringen som efterfølgende, hvilket er kendetegnende for og styrken ved formative evalueringer.

Der havde også været muligheder for at vælge andre evalueringsmodeller og -design. Rekvirent og evaluator kunne også have valgt at gennemføre en effektevaluering, der satte fokus på præstationer eller effekter af de konkrete terapeutiske teknikker, som anvendes i rehabiliteringsmodellen fx fysioterapi, ergoterapi, talepædagogisk træning, neuropsykologisk træning eller socialpædagogiske indsatser. Når dette ikke er gjort, er der tale om et valg, der afspejler den erkendelsesinteresse, der var på det pågældende tidspunkt. Der er intet til hinder for, at den gennemførte procesevaluering efterfølges af en eller flere effektevalueringer, så den samlede evaluering af en given indsats bliver en kombination af forskellige modeller og design.

Dataindsamling

Evalueringen indsamlede¹² primært data ved hjælp af flere på hinanden følgende spørgeskemaer (fire forskellige skemaer) til professionelle, klienter og pårørende gennem en periode på fire år. Spørgeskemaerne blev uddelt ved starten af opholdet på genoptræningscentret, ved afslutningen, samt et år efter opholdet for dem, der nåede så langt i evalueringsperioden. Evalueringen indsamlede endvidere data ved hjælp af fokuserede gruppeinterview med repræsentanter for de enkelte faggrupper i genoptræningscentret i 2001. Evaluator gennemgik desuden forundersøgelser og klientplaner for de enkelte klienter. Endelig blev der gennemført et litteraturstudie af den relevante danske og internationale litteratur med fokus på muligheden for reintegration i sociale fællesskaber. Litteraturstudiet dannede efterfølgende grundlag for den internationale perspektivering af rehabiliteringsmodellen.

Spørgeskemaerne fokuserede både på klienters og pårørendes oplevelse af deres situation og rehabiliteringstilbuddet. De fokuserede endvidere på medarbejdernes oplevelse af klienter samt de pårørendes og medarbejdernes oplevelse af ledelsesforhold og tværfagligt samarbejde. Endelig blev sagsbehandlere i kommunerne og medarbejdere på de modtagende institutioner bedt om at udfylde et spørgeskema lige efter modtagelsen af borgeren med hjerneskade og igen et år efter.

Evaluator mødte de særlige metodiske udfordringer i forbindelse med interview med klienterne (mennesker med hjerneskade), der blev gennem-

ført som personlige interview. Her var det ikke muligt alene at benytte sig af verbalt sprog, hvorfor interviewerne i visse tilfælde mere foretog observationer end egentlig interviewning. Interviewerne tog notater af observationerne, som senere indgik i analyserne.

Analyse af kvantitative og kvalitative data

Anvendelsen af de forskellige strukturerede spørgeskemaer, med faste svarkategorier resulterede i et større kvantitativt materiale, som skulle analyseres. Analyserne af de kvantitative data fra spørgeskemaerne blev analyseret ved hjælp af statistikprogrammet SPSS¹³, der er målrettet til analyser inden for samfundsvidenskaberne og er relevant for evaluatører inden for det sociale område. Der findes dog også andre velegnede programmer, ligesom regneark kan bruges til mange statistiske analyser.

Anvendelsen af personlige, åbne og kvalitative interview med enkeltpersoner og grupper resulterede i et omfattende kvalitativt materiale i form af udskrifter fra interviewene. Analysen af udskrifterne fra de kvalitative interview blev gennemført ved hjælp af det kvalitative analyseprogram N'vivo¹⁴ efter principper fra Grounded Theory (Strauss og Corbin 1990) og Spradleys etnografiske metode (Spradley 1979). Analysen blev bygget op fra grunden (induktion), idet de analytiske temaer og koder blev udviklet på baggrund af den analytiske gennemgang af interviewteksterne. Processen kan billedligt karakteriseres som »at klippe og klistre« tekster ud fra afmærkninger af, hvem der har sagt hvad om de forskellige spørgsmål, man undersøger, og andre relevante spørgsmål, der eventuelt bringes på banen af de personer, der interviewes. Fordelen ved at anvende analyseprogrammer er, at det er hurtigere og enklere at genfinde de kodede tekster og udvælge dem på forskellige måder afhængig af de forskellige analyser, der viser sig relevante. En ofte fremført kritik er, at analyserne herved bliver mekaniske og løsriver de enkelte udsagn fra deres oprindelige kontekst. Hertil er at sige, at alle dårlige analyser kan falde i disse fælder, men det afgørende for det er den forsker eller evaluator, der anvender dem. Faktisk gør N'vivo det muligt at arbejde meget dynamisk. Ligesom der hele tiden er en hurtig vej tilbage til de kontekster, hvorfra de kodede tekster kommer.

Anvendelsen af den »modsatte« analysestrategi (deduktion) ville have betydet, at temaerne havde været opstillet på forhånd og fuldt ud struktureret arbejdet med interviewene, hvorved det ikke ville have været muligt for svarpersonerne at sætte nye relevante ting på dagsordenen og evaluator ikke kunne forfølge overraskende nye relevante aspekter.

Evalueringsens overordnede analytiske strategi var, at evaluator, så vidt muligt, benyttede de forskellige data på kryds og tværs i analysen af de centrale spørgsmål. I metodelitteraturen kaldes denne strategi for data-triangulering. Princippet om triangulering kan også benyttes i forhold til: Metoder, medvirkende forskere/evaluatorer og teorier (Jensen 1990).

Hovedresultater

Evalueringen kom frem til 11 hovedkonklusioner, hvoraf to eksempler på substantielle resultater og et eksempel på et metodologisk resultat fra evalueringen skal nævnes i denne sammenhæng.

Evalueringen konkluderer bl.a., at *det er lykkedes rehabiliteringscentret at skabe et tilbud, hvor udgangspunktet er klientens eget perspektiv, og hvor et tværfagligt team stiller deres professionalisme til rådighed for at indfri de forventninger, der knytter sig til dette perspektiv så godt som muligt. Pårørende og klienter oplever en betydelig grad af empati og lydhørhed. Men da et »hospitalslignende« ophold stadig er forbundet med en betydelig autoritetstro og manglende forventning om indflydelse på problemdefinition og »behandling«, skal man stadig være opmærksom på, at det er personalet, der skal give udtryk for, at åbenhed over for klientens perspektiv og forventninger er professionelt begrundet* (Høgsbro 2002 s. 12).

Det konkluderes endvidere, at evalueringen har givet grundlag for det videre arbejde med at designe en eventuelt efterfølgende effektevaluering. Høgsbro (2002 s. 15) konkluderer således at:

Der forestår et arbejde med at afgrænse og udvikle de spørgsmål, klienter og pårørende skal stilles. Det er især væsentligt, at spørgsmålene tager udgangspunkt i en kontekstuel forståelse for rehabiliteringens mål og kan tage højde for forskellige former for funktionalitet i samspillet mellem klient og omverden.

Formidling, modtagelse og anvendelse af evalueringen

Resultaterne fra evalueringen blev formidlet i AKF-rapporten *Rehabilitering af mennesker med traumatiske hjerneskader på Kolonien Filadelfia* (Høgsbro 2002), samt ved møder mellem evaluator og rekvirent både under evalueringens gennemførelse og efter evalueringsrapportens udgivelse.

Hvilke konkrete spor evalueringen har sat sig i praksis, er der ikke fulgt op på fra evaluators side, hvilket generelt set også er en sjældenhed for eksterne evaluatore, medmindre det er aftalt på forhånd og dermed er en integreret del af evalueringen. Et eksempel på en evalueringsmodel, hvor opfølgning er en integreret del af evalueringen, er KUBI-modellen, hvor de institutioner, der evalueres, for det første forpligtiger sig til at udarbejde en handlingsplan på baggrund af evalueringens resultater. Og for det andet forpligtiger sig til at tage imod evalueringsteamet, der vender tilbage til institutionerne og ser, hvordan det er gået et år efter evalueringens afslutning (Holm m.fl. 2003)¹⁵.

Centerlederen på Kurhus gav efter evalueringen udtryk for den måde evalueringen var blevet opfattet på på institutionen. Han udtalte følgende om evalueringens betydning for det videre arbejde med rehabilitering på Kurhus:

På Kurhus har vi i videst mulig omfang forsøgt at tage kritikken til efterretning: For det første er støtten til de pårørende intensiveret, således at vi nu tilbyder individuel støtte til den enkeltes familiemedlemmer. Assistanzen gives af psykologer, socialrådgivere eller andre faggrupper – afhængigt af vanskeligheder og behov. Desuden tilbyder vi en række møder – bl.a. fire årlige internatweekender – hvor pårørende kan mødes og drøfte deres situation. På disse møder orienterer og rådgiver fagfolk om hjerneskader og om de psykiske og sociale konsekvenser for familien.

For det andet har vi forsøgt at styrke samarbejdet mellem Kurhus' mange faggrupper igennem en ændring af organiseringen af den daglige træning. Der er i dag tilknyttet et såkaldt miniteam til hver enkelt klient. I miniteamet er alle de faggrupper repræsenterede, som kan støtte klienten i at nå hans eller hendes mål. Ud over at disse miniteams muliggør en smidig og effektiv træning, synes strukturen også at sikre de små faggruppers indflydelse.

For det tredje har vi opnormeret socialrådgiverne som et led i bestræbelserne på at sikre den bedst mulige koordinering mellem Kurhus og de instanser, der skal danne ramme om klienternes fremtidige tilværelse.

Bedømt efter udtalelser fra klienter, medarbejdere, pårørende og samarbejdspartnere ser det ud til, at ændringerne har medført en styrkelse af Kurhus' rehabilitering på de nævnte punkter. Og denne formodning vil forhåbentlig blive bekræftet, når vi inden for det næste år gentager de spørgeskemaundersøgelser, som har ligget til grund for AKFs kritik.

Nytter det?

Rapporten evaluerer ikke effekten af rehabiliteringen på Kurhus. Med andre ord besvarer den ikke spørgsmålene såsom "Hjælper det noget? Får klienterne noget ud af opholdet på Kurhus? Gør Kurhus en forskel?" Sådanne spørgsmål er særdeles vanskelige at besvare. Ikke desto mindre er de altafgørende for den enkelte klient og hans eller hendes pårørende. I et forsøg på at synliggøre effekten af rehabiliteringen har vi selv igangsat undersøgelser, der kan anskueliggøre det udbytte, klienterne har af opholdet. Forhåbentlig vil disse undersøgelser også medvirke til, at vi forbedrer vores tilbud yderligere.

11.4 **Eksempel: Evaluering af amternes rådgivning på handicapområdet**

AKF blev i 1999 kontaktet af Udviklingscenter for Specialrådgivning¹⁶, der gerne ville have gennemført en landsdækkende evaluering af kommunernes brug af amternes specialrådgivning på handicapområdet. Udviklingscenterets formål med at iværksætte en evaluering på området var at skabe et videngrundlag for at forbedre den amtslige rådgivning på handicapområdet.

Det blev i første omgang aftalt at gennemføre en forundersøgelse, idet evalueringens genstand, amternes rådgivning og kommunernes brug heraf, er meget kompleks, og der ikke tidligere havde været gennemført evalueringer på området. Forundersøgelsen¹⁷ bidrog derfor til at give såvel rekvisitent som evaluator bedre indsigt i områdets kompleksitet. Herigennem kunne grundlaget for at operationalisere evalueringens genstand forbedres,

hvilket bidrog til et kvalificeret valg af evalueringsmodel og evalueringsdesign (Rieper 1999).

Forundersøgelsen førte frem til et begrundet forslag til en landsdækkende evaluering af kommunernes brug af amternes rådgivning på handicapområdet. I forlængelse heraf blev det aftalt, at AKF også skulle gennemføre en brugerundersøgelse, der evaluerede de erfaringer, som mennesker med handicap havde med at bruge amternes rådgivning på handicapområdet (Olsen m.fl., 2001). Det er disse to evalueringer, der komplementerer hinanden, som beskrives i det følgende.

Formål og hovedspørgsmål

Udviklingscenter for Specialrådgivning efterspurgte og finansierede evalueringerne med det formål at bidrage til at skabe overblik, opmærksomhed og ideer til udvikling af amternes rådgivning på handicapområdet. Brugerundersøgelserne skulle skabe viden om amternes rådgivning set ud fra de to centrale brugergruppers perspektiv: 1) De kommunale medarbejdere på handicapområdet, der har brug for viden i deres arbejde, og 2) de mennesker med handicap, der er direkte brugere af amternes rådgivning.

Ud over at Udviklingscentret skulle bruge evalueringerne som baggrund for deres initiativer, så var det også målet, at amterne og amtsrådsforeningen efterfølgende skulle kunne anvende resultaterne som grundlag for deres arbejde med forbedringer af rådgivningen på handicapområdet. Herudover bestod evalueringens målgruppe af handicaporganisationerne, videncentre, forskningsinstitutioner, kommuner, ministerier og mennesker med handicap, der hver på deres måder er interessenter og brugere af amternes rådgivning.

Der blev opstillet fire hovedspørgsmål i hver af brugerevalueringerne.

I forhold til brugergruppen *kommunale medarbejdere* var hovedspørgsmålene: 1) I hvilken grad oplever medarbejderne, at der stilles krav til deres viden? 2) I hvilket omfang kan de indfri de stillede krav? 3) I hvilken grad opfylder den amtslige rådgivning deres behov for viden? og 4) Hvordan erfarer medarbejderne, at kravene til deres viden har udviklet sig de senere år, og hvilke krav forventer de fremover?

I forhold til brugergruppen *mennesker med handicap* var hovedspørgsmålene: 1) Hvordan etableres kontakten mellem brugerne og amternes rådgivning? 2) Hvordan anvender brugerne konkret amternes rådgivning? 3) Hvordan kombinerer brugerne amternes rådgivning med anden rådgivning? og 4) Hvordan vurderer brugerne det samlede rådgivningsforløb?

Evalueringsgenstand og kontekst

Genstanden for evalueringen var amtslig rådgivning, som den realiseres i relationen mellem de to brugergrupper på den ene side og amternes rådgivere på den anden side. Den forudgående pilot-undersøgelse blandt medarbejdere i tre kommuner havde vist, at amternes rådgivning var meget kompleks såvel indholdsmæssigt som organisatorisk. Rådgivningerne er organiseret på forskellige niveauer (forvaltning, centre og institutioner) og danner et til dels uoverskueligt univers. Rådgivningerne dækker et bredt spektrum af handicaptyper, og de fleste varetager både behandlings-, undervisnings- og rådgivningsopgaver. Der er således tale om et særdeles svært gennemskueligt felt, hvilket evaluator måtte tage højde for i evalueringsdesignet.

Det viste sig også, at selve (special)rådgivningsbegrebet var uklart og ikke blev anvendt som et almindeligt begreb i kommunerne. Ej heller alle de amtslige rådgivningsenheder forstod sig selv inden for begrebet rådgivning (fx taleinstitutter, revalideringscentre). Evalueringen måtte derfor tage udgangspunkt i et rummeligt og bredt begreb, hvorfor det operationaliserede begreb om rådgivning blev til »indhentning af viden«.

Evalueringen blev gennemført i en samfundsmæssig kontekst, hvor brugerundersøgelser i stigende grad havde fået plads på den politiske og administrative dagsorden. Lovgivningsmæssigt¹⁸ var der sat en række ideelle mål og standarder for graden af brugerindflydelse og inddragelse, hvilket prægede interessen for evalueringer af brugernes erfaringer.

Evalueringsmodel og -design

Det blev valgt at gennemføre en kombineret effekt-, præstations- og procesevaluering, hvor der blev sat fokus på: 1) hvad brugerne får ud af de amtslige rådgivninger (effekt), 2) i hvilket omfang de amtslige rådgivnin-

ger leverer ydelser til brugerne (præstation) og 3) hvordan relationerne og samarbejdsprocesserne er mellem brugerne og de amtslige rådgivere (proces).

Evalueringen af de kommunale medarbejders brug af amternes rådgivning blev designet som en kombineret spørgeskema- og interviewundersøgelse blandt et bredt udvalg af medarbejdere i danske kommuner. Designet sikrede, at der fremkom informationer og data, der på den ene side kunne give en kvantitativ dokumentation af medarbejdernes brug af amternes rådgivning, hvilket skete ved hjælp af et postomdelt spørgeskema. På den anden side gav gennemførelsen af efterfølgende fokusgruppeinterview med alle de medarbejdere, der havde besvaret spørgeskemaet, grundlag for kvalitative uddybninger og beskrivelser af medarbejdernes erfaringer med amternes rådgivning.

Evalueringen af handicappedes brug af amternes rådgivning blev designet som en kvalitativ interviewundersøgelse, hvor interviewene blev foretaget i brugernes eget hjem eller på et andet selvvalgt sted. De kvalitative interview gav oplysninger om brugernes erfaringer med amternes rådgivning, fortalt gennem eksempler og i sammenhæng med brugernes brug af andre rådgivere og liv i det hele taget.

Dataindsamling

Spørgeskemaet til de kommunale medarbejdere blev udsendt til 483 medarbejdere i 31 kommuner fordelt på 4 amter, og svarprocenten blev høj, idet 91% af de adspurgte besvarede spørgeskemaet. Evaluator bearbejdede de indkomne svar og udarbejdede herefter en præsentation af svarfordelingerne for grupper på 8-12 medarbejdere fra samme kommune, der blev inviteret til at deltage i fokusgruppeinterview. I fokusgruppeinterviewene kunne medarbejderne uddybe deres egne besvarelser af spørgeskemaet set i forhold til gruppens samlede besvarelser og diskuterer dem med hinanden og evaluator.

Der blev gennemført 52 fokusgruppeinterview, hvori der deltog 305 medarbejdere, hvilket svarer til 67% af de inviterede. Der var tale om et usædvanligt stort antal fokusgruppeinterview. Baggrunden for det store antal interview var, at indsatsen (amtslig rådgivning) har en kompleks karakter, der gjorde det nødvendigt at kombinere et relativt stort antal kvantitati-

ve data med stor bredde i de kvalitative data for at opnå en nuanceret, relevant og sikker viden om området.

Der blev gennemført knap 150 kvalitative interview med mennesker med handicap, idet der også her var brug for et bredt udvalg af svarpersoner, der havde erfaringer med det brede felt af amtslig rådgivning, der var undersøgelsens genstand.

Da evaluator stod over for opgaven med at udvælge personer med handicap til interview, blev den manglende statistik på området nærværende. Det er således ikke muligt at finde frem til de relevante svarpersoner gennem statistiske oplysninger, ligesom der ikke forelægger statistiske oplysninger om, hvor mange og hvem der benytter sig af amternes forskellige tilbud om rådgivning. For at finde frem til en population af relevante svarpersoner måtte evaluator derfor søge oplysninger af andre kanaler. Den valgte fremgangsmåde blev, at evaluator tog kontakt til relevante brugere gennem 10 udvalgte rådgivningsenheder i 3 udvalgte amter. Da det ikke var etisk forsvarligt, at evaluator fik konkrete oplysninger om brugerne fra rådgivningsenhederne, så blev der via dialog mellem evaluator og rådgivningsenheder fastlagt relevante kriterier for udvalg af brugere. Ud fra disse kriterier udvalgte hver rådgivningsenhed fire gange så mange brugere (20), som der skulle gennemføres interview med fra hver enhed. Rådgivningsenheden fremsendte herefter et brev til de udvalgte, som informerede om undersøgelsen og spurgte, om de kunne tænke sig at deltage i den. I brevet var der vedlagt et spørgeskema med ti korte spørgsmål fra evaluator. På det korte spørgeskema kunne brugerne give tilsagn om at deltage eller afslå og give få grundlæggende oplysninger om sig selv og graden af tilfredshed med rådgivningen. Evaluator fik således først viden om den enkelte bruger, hvis vedkommende valgte at besvare det fremsendte spørgeskema. Besvarelserne af spørgeskemaet blev sendt direkte til evaluator i en vedlagt svarkuvert, så rådgivningsenhederne ikke fik viden om, hvem der ønskede at deltage i undersøgelsen, og hvem der blev valgt til at deltage. Herved blev det sikret, at udvælgelsen til at deltage i evalueringen ikke medførte etiske problemer og greb ind i det fortrolige forhold mellem bruger og rådgivningsenhed.

Der var i de fleste tilfælde tilstrækkeligt mange brugere, der besvarede evaluators spørgeskema, hvilket gav grundlag for at udvælge og lave afta-

ler med fem brugere fra hver enhed til interview. I enkelte tilfælde måtte proceduren gentages, før der var tilstrækkeligt med brugere, der ønskede at deltage, og det var muligt at vælge mellem forskellige kategorier (fx køn, alder og tilfredshed) af brugere.

Ved gennemførelsen af de knap 150 interview stod evaluator i enkelte tilfælde over for den særlige udfordring at gennemføre interview med mennesker med kognitive og kommunikative handicap. Det var således i nogle tilfælde nødvendigt at gennemføre interviewene med bistand fra tolke og formidlere, som i nogle tilfælde var brugernes støttepædagoger, forældre og søskende. I andre tilfælde havde interviewpersonerne et kognitivt handicap, hvor interviewererne måtte arbejde meget omhyggeligt med klart sprogbrug, anerkendende gestik og fælles fortolkning af brugernes svar, som i nogle tilfælde var meget sparsomme på ord. At møde disse særlige udfordringer er en krævende opgave og kræver kvalifikationer og metoder, som man kan finde viden om i litteraturen om kvalitative metoder generelt (fx Kvale 1998) og i den specialiserede litteratur om interview med specifikke grupper (fx Folkestad 2000 og Karpatschhof 2001).

Analyse af kvantitative og kvalitative data

Analysen af medarbejdernes erfaringer med og brug af amternes rådgivning blev gennemført på baggrund af de kvantitative data fra spørgeskemaerne og de kvalitative data fra fokusgruppeinterviewene. I lighed med det første eksempel med evalueringen af behandlingsmodellen på Kurhus, blev statistikprogrammet SPSS og det kvalitative analyseprogram N'vivo anvendt i det analytiske arbejde. At det kan være særdeles nyttigt at kombinere de to typer af data, viste sig bl.a. på den måde, at de kvalitative oplysninger afslørede, at de kvantitative oplysninger om hyppigheden af medarbejdernes brug af amtlig rådgivning var upålidelige. Det kom frem ved fokusgruppeinterviewene, hvor medarbejderne så gruppens svar på disse spørgsmål og korrigerede dem ved at sige, at de efterfølgende måtte tage forbehold for deres rapporteringer. Derfor måtte evaluator lade spørgsmålet om hyppighed udgå af analysen og i stedet fokusere på en analyse, der forholdt sig til, om medarbejderne havde været i kontakt med de pågældende rådgivninger eller ej, hvilket var pålideligt nok.

De knap 150 interview med mennesker med handicap blev analyseret ved hjælp af N'vivo ud fra de fire hovedspørgsmål og en åben kodning af relevante problemstillinger, som var fremkommet i interviewene. Da de forskellige koder var oprettet, og alle interview havde været igennem kodeprocessen, var det muligt at analysere de enkelte interview og grupper af interview på kryds og tværs af amter, rådgivningstyper med videre, ligesom det var muligt at sammenholde disse oplysninger med de svar, brugere havde givet os i det indledende spørgeskema.

Hovedresultater

Da evalueringen indgik i en sammenhæng, hvor resultaternes relevans blev bundet sammen med ønsket om at lokalisere områder, hvor der var brug for forbedringer, valgte evaluator på den ene side at fremlægge sine analytiske resultater og brugernes konkrete forslag til forbedringer af amtslig rådgivning. Det vil føre for vidt at komme ind på alle del- og hovedresultater her, hvorfor der henvises til (Rieper 2001 og Olsen 2001) for de uddybende og detaljerede resultater. Her begrænser formidlingen af resultaterne sig til evaluators formuleringer af særlige udfordringer til de amtslige rådgivninger. Ved at sammenfatte resultaterne i en række synteser og formulere dem som udfordringer var det evaluators håb at bidrage til en fremadrettet og konstruktiv debat om evalueringens kritiske resultater.

Første udfordring: Gennemskuelighed

Der er brug for nogle tekniske og organisatoriske forandringer, der gør det muligt at beskrive amternes rådgivningstilbud klart og dækkende for brugerne, så der opnås gennemskuelighed.

Anden udfordring: At tage brugerne alvorligt

Den anden udfordring knytter sig til brugernes kritik af amtslige rådgivere, der ikke i tilstrækkelig grad anerkender brugerne som kompetente videnhavere.

Tredje udfordring: Sikring af faglig kompetence fra individ til system

Den tredje udfordring knytter sig til brugernes kritik af, at der er visse problemer med at leve op til forventningerne om kompetence.

Fjerde udfordring: Bedre koordination af rådgivningen

Den fjerde udfordring knytter sig til brugernes påpegning af de væsentlige problemer, der opstår, når rådgivning ikke koordineres.

Femte udfordring: Udvikling af den nye rådgiverrolle

Den femte udfordring går ud på at sikre levende faglige diskussioner om rådgiverrollen mellem amter, kommuner og mennesker med handicap. Der er behov for at etablere og fastholde diskussion af den amtslige rådgivers rådgiverrolle med inddragelse af de kommunale medarbejdere og borgere, der benytter og er afhængige af amtslig rådgivning.

Formidling, modtagelse og anvendelse

Der fandt en løbende formidling og dialog sted i evalueringsperioden mellem evaluator (AKF) og rekvirenten (Udviklingscenter for Specialrådgivning), samt en nedsat følgegruppe med interessenter i evalueringen. Denne løbende formidling og dialog bidrog til gensidig inspiration og videndeling mellem parterne og gav evaluator mulighed for at forbedre evalueringen undervejs i forhold til nye oplysninger og afklaringer.

Evalueringerne blev afsluttet med udgivelsen af AKF-rapporterne *Kommunernes brug af amtslig rådgivning på handicapområdet* (Rieper 2001) og *Hvordan skulle vi klare os uden?* (Olsen 2001). Herudover blev der skrevet en populærartikel i AKF Nyt (Nr. 1, 2001) og en analyserende artikel i Nordisk Administrativt tidsskrift (Olsen og Rieper 2001), ligesom evalueringerne dannede grundlag for et bidrag til antologien *Handicap, kvalitetsudvikling og brugerinddragelse* (Bengtsson 2003).

Evalueringens resultater og fremgangsmåder blev også formidlet ved en række møder med nøglepersoner og ved en række seminarer og konferencer i relation til såvel forskningsmiljøer som til Udviklingscentrets formidlingsaktiviteter over for amter, kommuner og mennesker med handicap.

Evalueringernes resultater er således bragt ind i nationale¹⁹ og nordiske forskningsmiljøer²⁰.

Evalueringens resultater er formidlet særligt omfattende på grund af samarbejdet med Udviklingscentret og en indgået aftale om, at evaluator skulle medvirke ved en del af den efterfølgende formidling. Udviklingscentret har således stået for at bringe evaluator i dialog med praksismiljøerne gennem to landsdækkende konferencer²¹ og seminarer i to af de amter, der indgik i evalueringen.

Med hensyn til den særlige udfordring i at formidle evalueringernes resultater til mennesker med kognitive eller kommunikative handicap er der anvendt tolkebistand i forbindelse konferencer, hvor evalueringernes resultater er blevet formidlet.

11.5 Afrunding

Uden at vi ved det med nøjagtighed, så er det en udbredt antagelse, at omkring 500.000 mennesker lever med en eller anden form for handicap i Danmark. Mange af disse mennesker er i berøring med serviceydelser fra stat, amter eller kommuner, det være sig inden for sundhedsområdet, uddannelsesområdet, socialområdet mv. På den baggrund kan det umiddelbart konkluderes, at potentialet for evalueringer på handicapområdet er stort. At give et indblik i de varierende former for evalueringer, der er gennemført og vil blive gennemført i de kommende år, er derfor ikke muligt gennem de eksempler, der er beskrevet i dette kapitel. Her er blot blevet åbnet et par af de mange vinduer, der kan kaste lys over evalueringer på området, hvorfor mange relevante spørgsmål og problemstillinger ikke er belyst. Et lille vindue er åbnet, mens andre har hasper, der kan løsnes. God arbejdslyst, hvad enten du skal videre i dit arbejde som evaluator, som rekvirent af evalueringer eller som læser og fortolker af evalueringer på dette eller andre felter.

12 **Evaluering af socialt arbejde for udsatte grupper**

Kjeld Høgsbro

Dette kapitel handler om de specielle problemstillinger, der er knyttet til indsatsen over for de særligt udsatte grupper. Disse defineres som mennesker, der på grund af misbrug, sindslidelser, hjerneskader eller lignende er delvis ude af stand til at tage stilling til egne interesser og ofte udsættes for diskrimination og manglende lydhørhed over for egne erklærede behov.

Det sociale arbejde udfoldes generelt inden for et felt af politisk genererede værdier og mål i den ene dimension, faglig baggrund/teori/metodik i den anden dimension og lokalt definerede handlemuligheder i den tredje. Inden for dette felt udvikles det sociale/pædagogiske arbejde af praktikere på baggrund af praktisk erfaring. Dermed bliver evalueringernes opgave ofte at få tydeliggjort og formidlet den praktiske viden.

Endvidere har det været et stort problem inden for det sociale arbejde at få samordnet indsatsen over for bestemte grupper. Indsatsen bliver opsplittet på afgrænsede/specialiserede tilbud uden tilstrækkelig koordinering. Det bliver derfor ofte evalueringens opgave at skabe en forståelse for det samlede tilbuds funktion i forhold til en større gruppe af borgere med forskellige behov. Først på denne baggrund kan man meningsfuldt vurdere det enkelte tilbuds funktion inden for det samlede tilbud. Dette har været den store udfordring for en række af vore evalueringer.

12.1 Hvad kendetegner de udsatte grupper?

De udsatte grupper er generelt kendetegnet ved, at de er delvis ude af stand til at tage stilling til egne interesser og ofte udsættes for diskrimination og manglende lydhørhed over for egne erklærede behov. Det er mennesker, der på et tidspunkt af deres liv ryger ind i et alvorligt misbrug, erhverver sig en sindslidelse, en hjerneskade eller lignende, som gør dem delvist ude af stand til at tage stilling til egne interesser og behov. Når de formulerer deres ønsker, bliver de derfor indimellem overhørt, fordi omgivelserne regner dem for »utilregnelige«. De udsættes for diskrimination og manglende lydhørhed, fordi deres nærmeste, samfundet i almindelighed og de professionelle, der bemander tilbuddene, er vant til, at deres erklærede ønsker er modstridende, kan svinge fra tid til anden og indimellem er uforenelige med det, de selv og omgivelserne betragter som deres mere langsigtede behov.

Disse mennesker er således i særlig grad »udsat« for andres »planer« med dem, og de er til dels udsat for, at disse planer skifter med 10 års mellemrum og i lange perioder bliver udviklet, uden at deres egne synspunkter inddrages.

At målene for det sociale arbejde ændrer sig fra tiår til tiår, er ikke kun noget, som gælder de særligt udsatte grupper. Socialt arbejde handler om at løse sociale problemer, og sociale problemer har karakter af *vilde* problemer. Det vil sige, at der ikke kan skabes den samme grad af konsensus omkring mål og midler, som det er tilfældet ved tekniske problemstillinger. Det kan være kompliceret og kostbart at sende et lille køretøj til Mars, men når først man har besluttet sig til en økonomisk ramme for projektet, er selve beslutningsprocessen omkring målene med ekspeditionen og måden, det skal gøres på, relativt entydig. Sådan er det ikke med socialt arbejde. Er problemet med misbrug fx den kriminalitet, det medfører, eller er det sygdommene? Er det stofmisbruget i sig selv, eller er det misbrugs-kulturen, som af nogle opleves som »uværdig« eller »umenneskelig«? Er det, fordi det skæmmer gågaden, hvor mennesker gerne skulle kunne gå og nyde deres materielle muligheder i fred? Er sindslidelsen en sygdom eller et handicap? Er problemet i virkeligheden samfundets manglende tolerance for det skæve, det anderledes og det uregerlige? Er det en lidelse i sig selv eller er det kun en lidelse, fordi samfundet ikke tolererer denne form for

adfærd? Bliver mennesker voldelige på grund af manglende accept, eller bliver de voldelige, fordi omgivelserne ikke har sat klare grænser for en sådan adfærd?

Svarene på alle disse spørgsmål lægger op til forskellige løsningsstrategier. Nogle retter sig mod at isolere de pågældende mennesker, andre drejer sig om at få dem integreret i sociale sammenhænge. Nogle retter sig mod helbredelse og rehabilitering, mens andre retter sig mod at reducere de skader, deres særlige livsmønster påfører dem.

I de sidste ti år er misbruget i stigende grad blevet taget alvorligt som en situation, der er et problem for misbrugeren og ikke kun for samfundet. Misbruget er i stigende grad blevet betragtet som en afhængighed, misbrugeren er offer for og i mindre grad blevet betragtet som en selvforskyldt tilstand eventuelt affødt af »dårlige påvirkninger« eller »manglende karakterstyrke«.

Sindslidelser er i de sidste ti år i stigende grad blevet betragtet som et handicap, som man til en vis grad kan lære at kompensere for, hvis man får en tilstrækkelig social støtte. De betragtes i mindre grad som sygdomme, man vil kunne fjerne med medicinsk behandling.

Hjerneskadener er i de sidste ti år i stigende grad blevet betragtet som skader, der giver den enkelte nogle helt specifikke problemer i dagligdagen. Man er ikke tabt for evigt og har fået reduceret sine evner på alle livets områder. Tværtimod har man gode muligheder for at få optrænet sine evner til at klare lige netop den sociale sammenhæng, som man har mest behov for at kunne klare.

Generelt for de udsatte grupper gælder det således, at samfundet stadig mener, at der »er noget galt med hovedet«, men at dette »hoveds« måde at fungere på i dag opfattes som langt mere kompliceret end tidligere. Man lægger vægt på ikke at generalisere for meget, men opfatte hvert enkelt menneskes problem som meget specifikt. Både i den forstand at problemet har sin egen helt specielle historie og betydning for dette menneske, og i den forstand at man ikke kan definere problemet uden at se på netop den sociale sammenhæng, netop dette menneske indgår i eller stræber mod.

Dette kræver, at professionelle og klienter samarbejder omkring løsningen af problemet. Det kræver respekt for klientens viden om vedkommendes egen situation, og det kræver et tværfagligt samarbejde omkring

behandling, social støtte og tilrettelæggelse af eventuelle rehabiliteringsforløb.

Det forhold, at problemet i stigende grad opfattes som et handicap, gør samtidig, at der indfinder sig en *ligestillingsproblematik*. Hvis der er tale om livsvarige handicap, må samfundet give rum for, at mennesker med disse problemer kan respekteres som medborgere og så vidt muligt tildeles samme rettigheder og muligheder. *Livskvalitet* og *brugerindflydelse* er derfor blevet mål i sig selv på linje med muligheder for at klare sig selv, adgang til sociale aktiviteter og ret til rehabiliteringsforløb som tilrettelægges for det enkelte menneske.

12.2 Målene for den professionelle indsats

Målene for den professionelle indsats er derfor dels ret mangfoldige dels i indbyrdes konflikt. Hvordan inddrager man fx brugerens synspunkter i tilrettelæggelsen af behandling, social støtte og rehabilitering, når brugeren regnes for delvis eller periodevis utilregnelig? Hvilke former for brugerinddragelse harmonerer med et overordnet professionelt ansvar for, at brugeren ikke skader sig selv?

Sådanne spørgsmål er igen og igen blevet stillet til de professionelle, og der har været en konstant politisk interesse i evalueringer, der fandt frem til den bedste måde at sikre en indfrielse af de mange forskelligartede målsætninger for arbejdet på.

Samtidig kan socialt arbejde ikke evalueres, uden at man tager højde for den nærmest lokalt specifikke balance mellem de forskellige målsætninger. Det sociale arbejde udfoldes inden for et felt af politiske, genererede værdier og mål i den ene dimension, faglig baggrund/teori/metodik i den anden dimension og lokalt definerede handlemuligheder i den tredje (Ebsen og Høgsbro 2001). I visse dele af landet satser man på behandling og i andre dele af landet på social støtte. Visse steder har man muligheder for et tæt samarbejde med det lokale erhvervsliv, andre steder har man en levende underskov af frivillige initiativer. Kompetencen til at gebærde sig og udnytte de lokale betingelser udvikles hos lokale ildsjæle og professionelle initiativtagere på baggrund af deres personlige erfaring («the road we make by walking»). Den udvikles som lokal social/pædagogisk kultur.

Dermed bliver det noget af en udfordring at ville lave evalueringer, der identificerer »best practice«, idet det altid vil afhænge af de lokale betingelser, hvad der er »best practice«.

Samtidig kan målene for den enkelte institution eller det enkelte tilbud godt være relativt snævre, så længe tilbuddet indgår i en samlet indsats, der dækker hele spektret af målsætninger. Det akutte behandlingstilbud har muligvis et meget snævert fokus på en stabilisering af den medicinske støtte, men så længe tilbuddet indgår i en sammenhæng med andre sociale tilbud, er helheden af behandling, social støtte og rehabilitering stadig sikret. Et værested kan have minimal brugerindflydelse, men være placeret således at brugerne reelt er i stand til at vælge mellem dette og andre tilbud.

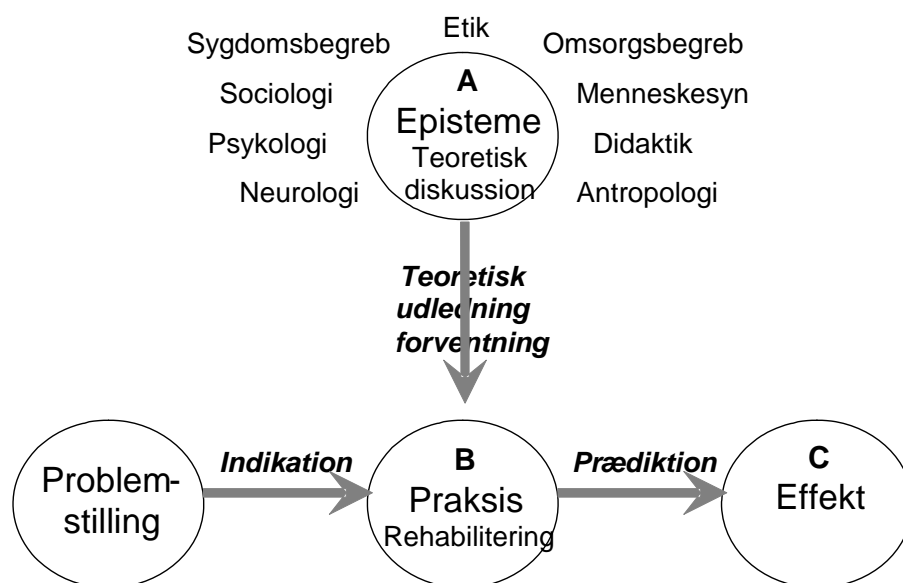
En evaluering af indsatsen må således enten rette sig mod den samlede indsats, eller den må tage højde for det aktuelle tilbuds placering inden for en samlet struktur af tilbud.

12.3 Hvad er det, vi kan evaluere?

Det professionelle perspektiv versus brugernes perspektiv

Figur 12.1 illustrerer noget af den komplekse verden, vi går ind i, når vi vil evaluere et bestemt socialt tilbud eller en bestemt form for tilbud til en bestemt gruppe.

Figur 12.1 Generel model for rehabiliteringsindsats



Figuren viser den del af verden, som de professionelle befinder sig i. Det er en verden, der er karakteriseret ved nogle bestemte forestillinger, man har om de problemer, man kan gøre noget ved, og hvad man i så tilfælde kan gøre. Man fokuserer ikke primært på det enkelte menneskes hele livssammenhæng, men på den afgrænsede del af denne livssammenhæng, som giver mennesker nogle bestemte problemer. Og blandt disse forskelligartede problemer vælger man sig et bestemt problem, som man mener, man kan gøre noget ved. Hele denne afgrænsning af problemet hænger sammen med en bestemt teoretisk forståelse af problemet (epistemet), som samtidig peger på (*indikerer*) muligheder for at afhjælpe eller fjerne problemet. Det vil samtidig sige, at man har nogle forventninger (*prædiktation*) til effekten af indsatsen. Denne helhed af årsagsforståelse, identifikation af særlige problemer, som indsatsen vil afhjælpe, og forventningerne til nogle bestemte resultater af indsatsen udgør en *model*. Og det er sådanne modeller, vi ofte er sat til at evaluere. Rekvirenten kan godt i første omgang tro, at det er den konkrete institutions arbejde, man vil have evalueret, men må i anden omgang indrømme, at det er »måden, de arbejder på« eller det »koncept«, man henholder sig til, som man ønsker evalueret. Dette kræver, at man først går ind og finder ud af, om deres praksis overhovedet er i overensstemmelse med det koncept, de henholder sig til, for eksempel i deres virksomhedsplan.

I forbindelse med evalueringen af et nyt rehabiliteringstilbud for mennesker med traumatiske hjerneskader (se kapitel 11) gik man ind og så på, hvorvidt konceptet for denne rehabilitering var i overensstemmelse med den aktuelle internationale forståelse af de problemer, der opstod i forbindelse med traumatiske hjerneskader (Høgsbro 2002). Man så på, hvordan institutionen i praksis fulgte disse anerkendte retningslinjer, hvorvidt medarbejderne havde taget konceptet til sig, og hvorvidt brugerne oplevede tilbuddet på den måde, som man antog, at de ville opleve det. På denne baggrund kunne evalueringen pege på en række forskellige skismaer:

Ifølge den internationalt anerkendte forståelse af rehabiliteringens problemstillinger (epistemet) som centeret lagde sig ind under, var det vigtigt, at rehabiliteringen opstod i et samarbejde mellem patienter og professionelle, og det var derfor en del af målgruppekriteriet, at patienterne var »motiverede« for dette samarbejde. Men dels er det svært at konstatere en

motivation, når patienterne har en stærkt svækket hukommelse, dels er det svært at tale om motivation, når skader i hjernens forreste del netop påvirker motivationen. Motivation skal tillige opfattes som en dynamisk størrelse, der selvfølgelig også er afhængig af det tilbud, patienten får. »Motivation« er således vanskeligt at bruge som adgangskriterium for et rehabiliteringstilbud til mennesker med hjerneskade, og der eksisterer den fare, at patienter bliver opgivet på grund af »manglende motivation«, hvor der i virkeligheden er tale om, at tilbuddet har været for dårligt for netop disse patienter. Konstruktionen beskytter således også institutionen over for kritik.

Epistemet har således en indbygget modsigelse knyttet til samspillet mellem patientens frontallapskade og kravet om brugerinddragelse. Et problem, som evalueringen må pege på har konsekvenser for det daglige arbejde og for de professionelles muligheder for at bringe deres praksis i overensstemmelse med det ideelle koncept.

Epistemet indeholdt også nogle krav til det tværfaglige samarbejde. I virkeligheden er det, ifølge dette episteme/paradigme ikke relevant at tale om universelle krav til kognitive og motoriske færdigheder. Målet er derfor at rehabiliterer patienten til en bestemt social sammenhæng. Her kan de lære at mestre dagligdagen, så snart man ved, hvad der karakteriserer denne dagligdag. Men dette kræver et samarbejde ikke bare mellem socialfaglige og sundhedsfaglige medarbejdere på stedet, men også med de kommunale sagsbehandlere som skal beslutte, hvilken form for hjælp patienterne kan få bagefter. Evalueringen kunne her konstatere, at de socialfaglige medarbejdere havde mindre tillid til egen faglighed og følte sig mindre inddraget i beslutningerne end det sundhedsfaglige personale. Evalueringen kunne samtidig konstatere, at samarbejdet med kommunerne var mangelfuldt.

Dette er et eksempel på, hvordan et tilsyneladende ideelt koncept kan komme i vanskeligheder, når det først skal implementeres i praksis. Evalueringen kunne ikke tillade sig at sætte spørgsmålstejn ved selve konceptet, for her måtte man ved en gennemlæsning af den internationale forskningsdiskussion konstatere, at konceptet var helt i tråd med de nyeste teoretiske overvejelser. Derimod kunne evalueringen se på, om praksis var i overensstemmelse med dette koncept og her pege på særlige strukturer, der vanskeliggjorde det tværfaglige samarbejde, vanskeliggjorde inddragelsen af

de pårørende og vanskeliggjorde samarbejdet med de instanser, der senere skulle overtage støtten til brugerne.

Den professionelle model har selvfølgelig en relation til de mennesker, indsatsen retter sig imod. Men disse mennesker kommer ikke i direkte kontakt med den professionelle model og de tanker og overvejelser, der ligger bag. Nogle terapeutiske tilbud lægger vægt på at indvie brugerne i de overvejelser, der ligger bag, men selv i disse tilfælde har brugerne deres eget liv uden for tilbuddet, som brugerne først og fremmest forholder sig til og har ansvar for. For brugerne er det i langt højere grad hele deres livssammenhæng, som er i fokus. Den gamle talemåde »operationen lykkedes, men patienten døde« udtrykker denne forskel mellem det professionelle perspektiv og brugerens perspektiv. Operationen lykkedes, fordi diagnosen viste sig at være rigtig, indgrebet korrekt foretaget, og den efterfølgende død skyldtes uundgåelige komplikationer. Men for brugeren er det ligegyldigt, om »operationen lykkedes«, hvis den ikke løste vedkommendes problemer. For brugeren er der ikke tale om et velafgrænset problem, men om mange forbundne problemer. Problemet for et menneske, der er manisk, er ikke primært, at forestillingerne er urealistiske, men at familien ikke kan holde det ud, og at vedkommende måske er ved at sætte venskaber og personlig økonomi over styr. Vedkommende vil derfor gerne have mere kontrol, uden at det går ud over den energi og det livsmod, som familien er vant til at betragte som positivt.

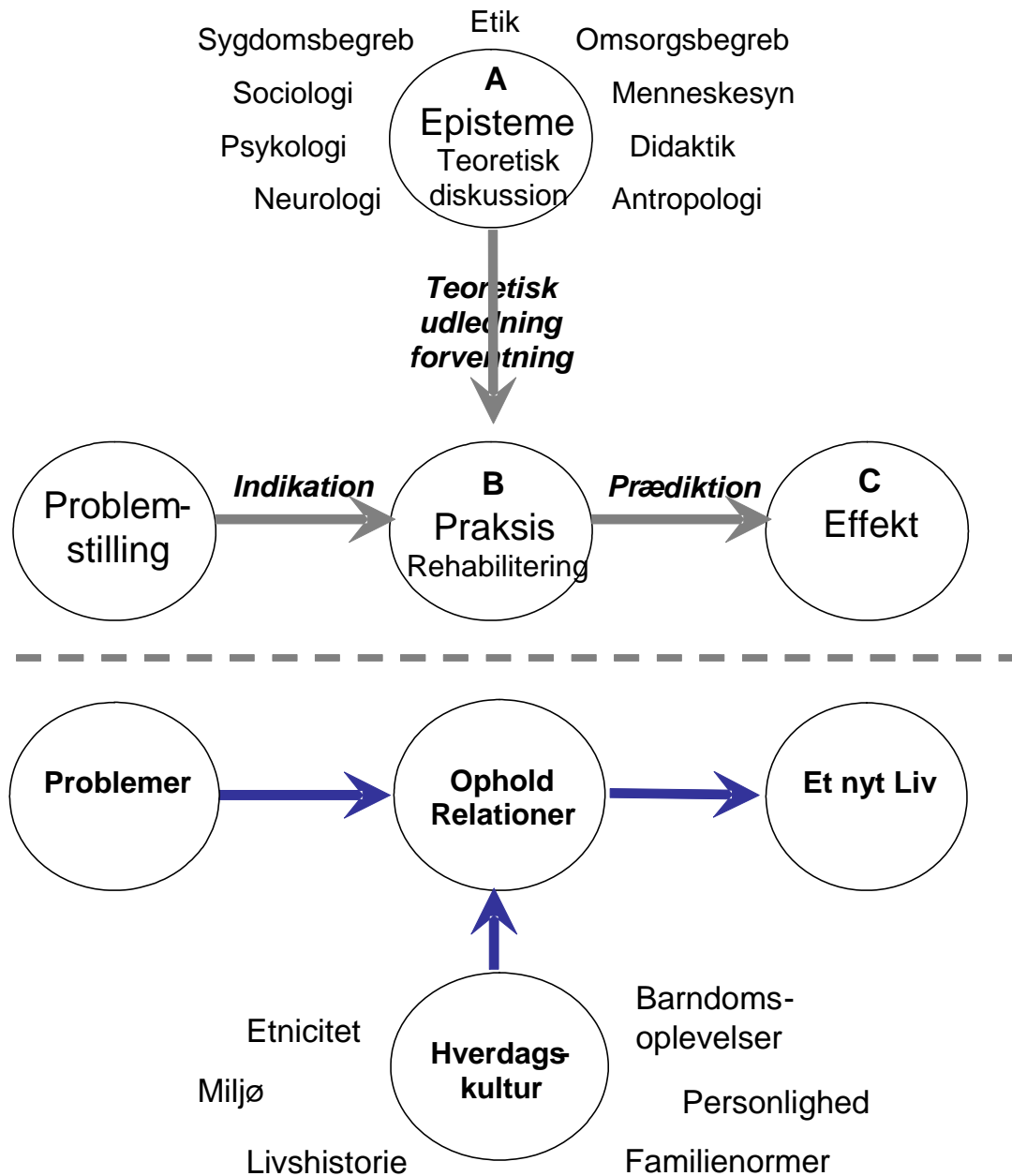
I en af mine undersøgelser indgik en kvinde, som kunne komme i kontakt med afdøde. Da hun var medlem af en lokal gruppe kvinder, som dyrkede denne egenskab sammen, var hun uinteressert i en behandling, som fjernede denne evne. Det, hun ønskede, var en behandling, som fjernede hendes periodevise manglende kontrol med disse oplevelser. Som hun selv fortalte, mente hendes psykiater selvfølgelig, at hele hendes forestilling om kontakt med afdøde var en del af hendes sygdom, men derom var de uenige. De blev i respekt for brugerens opfattelse af, hvad der var sygdom, enige om at lægge medicineringen på et niveau, der fjernede perioderne med manglende kontrol, men fastholdt evnen til at kontakte afdøde. Problemet var dermed løst, set fra patientens vinkel. Professionelt var det ikke betragtet som løst, og set fra et mere abstrakt politiske krav om frigø-

relse fra overførselsindkomster var det heller ikke løst. Kvinden fik stadig pension på grund af sin »sygdom«.

Der findes mange konkrete varianter af den problemstilling, dette eksempel illustrerer. Generelt kan man sige, at brugeren ikke opfatter den professionelle indsats som et indgreb, en praktisk realisering af et bestemt koncept, men som et tilbud om nye relationer eller fornyelse af gamle relationer til andre mennesker. Brugeren keder sig eller hygger sig med andre brugere, vælger at knytte sig til bestemte professionelle og undgår andre. Endelig har vedkommende en forestilling om et fremtidigt liv med sociale relationer, som ikke belastes af de nuværende problemer. Når vi således går ud og måler brugertilfredshed, måler vi ikke, om »operationen lykkedes«, men om brugeren fik et bedre liv, set på baggrund af vedkommendes egne forventninger. Dette mål er ikke altid sammenfaldende med det professionelle mål. Det professionelle mål for en socialpsykiatrisk indsats er for eksempel ofte reduktion af sengedagsforbruget i hospitalspsykiatrien (fordi disse pladser er meget ressourcekrævende), mens det for brugerne ofte er en kvalitet at kunne blive indlagt, når man har brug for det.

Ligesom den professionelle er optaget af, om vedkommendes forståelse af problemerne har været korrekt, er patienten optaget af, om resultatet svarer til vedkommendes forventninger, og om indsatsen svarer til vedkommendes forestillinger om, hvad der afhjælper problemet. Der er sindslidende, som mener, at de har brug for mere psykoterapi, andre, at de har brug for bedre medicin, og atter andre, som vægter sociale tilbud. Brugerne har deres egen livserfaring at tage udgangspunkt i og deres egne forestillinger om, hvad der hjælper. Ligesom de professionelle har nogle forestillinger, som knytter sig til en bestemt international diskussion, på samme måde har brugerne nogle forestillinger, som knytter sig til en bestemt hverdagskultur. Hele denne sammenhæng er illustreret i figur 12.2, hvor brugernes opfattelse af problemet og dets løsning er sat op over for den professionelle (som den nederste, nye del af figuren). Den øverste og nederste del af figuren er komplementære. De har grundlæggende samme opbygning, men de spejler ikke hinanden. Deres indhold af forestillinger og forventninger er forskelligt, og de to parter kan tale sammen og have tæt kontakt med hinanden uden nødvendigvis at overtage hinandens perspektiv.

Figur 12.2 Forholdet mellem en professionel model og brugernes livsverden



Denne forskel mellem det professionelle perspektiv og brugerperspektivet er det vigtigt at holde sig for øje, når brugertilfredshed indgår som en del af evalueringens design. Brugere er muligvis tilfredse med, at de har fået nogle smukke omgivelserne og en god kontakt til nogle venlige og engagerede mennesker, mens dette muligvis slet ikke har været målet for de

professionelle. Eller det modsatte kan være tilfældet: Tilbuddet har reduceret forbruget af dyre hospitalspladser eller reduceret antallet på overførselsindkomster, men brugerne oplever ikke, at de har fået et bedre liv.

Modelbetragtningen er særlig relevant, når vi taler om indsatsen over for de udsatte grupper, fordi disse grupper som udgangspunkt ikke formodes at kunne varetage deres egne interesser. Deres interesser varetages derfor i betydelig grad af professionelle grupper, som har deres særlige forståelse af, hvad der er bedst for brugerne. Det, brugerne kan bidrage med af input til denne forståelse, bliver fortolket eller oversat til en professionel forståelse, før der handles i forhold til det. Når misbrugeren ønsker mere metadon, skal ønsket først igennem en professionel vurdering af, hvad der er bedst for denne bruger, før der handles. Når den sindslidende ønsker mere eller mindre medicin, sker der det samme. I disse tilfælde har det stor betydning, hvordan de professionelle forstår de problemer, de har med at gøre, og hvad de mener kan hjælpe brugerne. Et sted kan man være restriktiv over for misbrug og stille store krav til deltagelse i det daglige praktiske arbejde. Et andet sted kan man have en liberal holdning til både misbrug og deltagelse. Det er vigtigt, at vi er opmærksom på, hvad denne forskellige forståelse af problemerne udspringer af, fordi det ikke kun er medicindoser og biologiske effekter, vi taler om. Det er nok så meget sociale krav og forventninger.

12.4 **Valget mellem forskellige evalueringsformer**

Før man kan vælge mellem bestemte former for evaluering, må man derfor gøre sig klart, hvad det er, man vil vide noget om.

Hvis man er tilfreds med at vide, om tilbuddet har haft den forventede effekt, kan man gennemføre en summativ evaluering. Man kan lade institutionen, de professionelle eller den bevilgende myndighed opregne en række mål, som tilbuddet skal evalueres på, og derefter følge udviklingen over tid eller gennemføre evalueringen på et nærmere fastsat tidspunkt.

Nu er problemet, at man ofte tror, at det kun er det, man vil have at vide. Og når resultatet så foreligger, begynder man pludselig at udtale sig generelt om »den slags tilbud«. Og det varer ikke længe, før man påstår, at undersøgelsen har bekræftet eller afkræftet teorien. Og det er lige netop

det, man ikke har belæg for. For hvordan kan man overhovedet vide, om der er foregået det, som de professionelle på stedet siger eller tror, der er foregået. Og hvordan kan vi vide, om praksis er i overensstemmelse med de teorier, der ligger bag. Dvs. spørgsmålet om *modeltrofasthed*.

Lige netop denne sammenhæng mellem modeltrofasthed og effekt er i fokus i *Medicinsk Teknologivurdering* (MTV). I amerikanske forsøg med opsøgende psykoseteam var der således tale om flere forskellige lokale udgaver af det fælles koncept (Nordentoft 2004, Vendsborg m.fl. 1999, Stein og Santos 1998). Det fælles bestod i, at man bemandede et team med forskellige faggrupper og gav teamet et bestemt område, hvor de skulle finde og opsøge mennesker med alvorlige sindslidelser. Derefter skulle de så i teamet finde frem til, hvordan problemerne bedst kunne afhjælpes via en samlet social og sundhedsmæssig indsats. Men teamene var derudover meget forskellige, både hvad angår de enkelte faggruppers placering i teamet, antallet af patienter, de havde ansvar for, og tidsrummet, de kunne kontaktes indenfor. Nu havde man imidlertid så mange team med i undersøgelsen, at man kunne udnytte denne forskel positivt. Man kunne nemlig se, hvilke afvigelser fra den ideale model som havde en positiv eller en negativ effekt. Set i forhold til figur 12.2 fokuserer de amerikanske undersøgelser (Nordentoft 2004) på B og C, idet de sammenligner de forskellige variationer i praksis (B) med effekten i form af et reduceret forbrug af sengedage (C). Hvis vi forventer, at effekten er den samme i Danmark, kan man derefter direkte beregne en forventet besparelse ved indførelse af modellen (Vendsborg m.fl. 1999). Som baggrund for en videnskabelig diskussion af sammenhængen mellem psykosens udvikling og den professionelle støtte har man så fremstillingen af denne sammenhæng i Stein og Santos' bog (1998), der i forhold til figur 12.2 diskuterer forholdet A→B. Resultaterne i praksis bidrager således på en og samme tid til en udvikling af tilbuddet og en udvikling af vores forståelse for, hvad der virker, og hvorfor det virker.

I visse tilfælde kan man stille sig selv det spørgsmål, om det i det hele taget giver mening at undersøge effekten. Nogle vil måske mene, at det er alfa og omega, når vi taler om evalueringer, men det er ikke helt rigtigt. Lad os fx forestille os, at vi ved alt om effekten af et givet tilbud fra internationale undersøgelser. Hvorfor så undersøge det endnu engang? Det inte-

ressante spørgsmål er måske snarere, om praksis har været i overensstemmelse med dette internationale koncept, om praksis er i overensstemmelse med det teoretiske udgangspunkt, man hævder, man har, og hvilke problemer der har været og måske typisk *vil* være, når man vil etablere et sådant tilbud i Danmark. Det kan være, der er modstand mod projektet blandt de ansatte, fordi de er vant til at gøre tingene på en anden måde, eller modstand blandt samarbejdspartnere, fordi konceptet provokerer deres selvforståelse. Hvis rekvirenten samtidig gerne vil have en evaluering, der peger på sådanne modsigelser i projektet og giver en feedback, som kan medføre en regulering af projektet undervejs, kan man pege på en *formativ evaluering*, hvor selve effektevalueringen har mindre plads. Dette var tilfældet med den nævnte undersøgelse af en hjerneskaderehabilitering. Her ville det være vanskeligt at sammenligne tilbuddet med andre tilbud inden for landets grænser, og patienterne havde i den periode, de var tilknyttet stedet en forventet »spontan« bedring af deres tilstand, hvilket gjorde en »før-efter«-evaluering omsonst. Evalueringen fokuserede derfor på de felter i figur 12.2, som er benævnt A, B, E og F. Dermed kan undersøgelsen udtale sig om eventuelle misforhold mellem teori og praksis ($A \rightarrow B$), om brugernes oplevelse af de professionelle sammenholdt med de professionelles oplevelse af brugerne ($B \leftrightarrow E$) og om brugernes oplevelse af opholdets betydning for deres eget fremtidsperspektiv ($E \rightarrow F$). Da tilbuddet vægter brugerinddragelse og motivation, er dette en række vigtige forhold at få feedback på i forbindelse med en formativ evaluering. Derimod mangler man både præcise angivelser af de professionelle mål ($B \rightarrow C$) og pålidelige mål for effekten (C).

For en del år siden lavede AKF en undersøgelse af effekten af en hjerneskaderehabilitering på Center for Hjerneskade, Københavns Universitet (Larsen, Mehlbye og Gøtz 1991). Her var der tale om rehabilitering af mennesker, hvis tilstand havde været stabil i nogle år, og man kunne derfor sammenligne deres situation efter genoptræningstilbuddet med situationen inden. Hvad angår effektmålet, var der en veldefineret interesse i at se på en mulig reduktion af behovet for offentlig forsørgelse og støtte. I dette tilfælde behøvede man således alene at fokusere på forholdet $B \rightarrow C$. I internationale artikler forholdt forskerne på KU, som ledede projektet, sig til projektets belysning af forholdet $A \rightarrow B$, hvilket bidrog til en diskussion af

genoptræningens principper og forskellige præmisser for at måle genoptræningens effekt (Teasdale et al. 1997). Dermed bidrog evalueringen indirekte til udviklingen af et nyt paradigme for forståelsen af hjernens komplekse funktioner (A). Politisk bidrog evalueringen til en mere konkret forståelse for den økonomiske gevinst knyttet til de nye muligheder for genoptræning.

12.5 **Eksempel: Undersøgelser af det samlede tilbud for hjemløse, misbrugere og sindslidende**

Baggrund og design

Ofte er der behov for at få et overblik over problemerne inden for det samlede tilbud til en bestemt gruppe. Det kan dreje sig om et behov for at vide, hvordan tilbuddene spiller sammen og samlet set opfylder brugernes behov, og om det samlede tilbud er tilstrækkeligt koordineret, planlagt og organiseret. Her er der i sagens natur tale om en mere omfattende undersøgelse, men hvis det geografiske område er velafgrænset, er evalueringer af denne art overkommelige sammenlignet med de udgifter, der er forbundet med opretholdelse og administration af tilbuddene i øvrigt.

Inden for Socialministeriets evalueringsprogram fra 2001 var der således et ønske om en mere omfattende evaluering af tilbuddene til hjemløse, misbrugere og sindslidende i Danmark. Her drejede det sig om sammenhængen i det samlede tilbud, og undersøgelsen skulle blandt andet besvare spørgsmålet om, hvilke modeller man havde for denne indsats, om tilbuddet var dækkende og om der var forskelle i den måde, forskellige grupper var dækket ind af tilbuddet på.

Evalueringen befandt sig i et typisk dilemma: Skulle man tilstræbe en repræsentativ udredning af de typer af tilbud, der gjorde sig gældende i forskellige dele af landet, eller skulle man se, hvordan tilbuddet var skruet sammen i udvalgte dele af landet. Skulle man satse på en beskrivelse af, *hvad* der foregik, eller skulle man satse på en analyse af, *hvordan* det foregik, og *hvorfor* det foregik på denne måde.

På baggrund af en prioritering valgte man det sidste og gik derfor ind og udvalgte fem forskellige regioner i landet, hvor man koncentrerede sig om samspillet mellem de forskellige tilbud og deres målgruppe inden for hver region (Høgbro et al. 2003). Evalueringen blev en undersøgelse af de *modeller*, der dannede grundlag for indsatsen i de forskellige dele af landet. Man var interesseret i både at få en viden om, hvilke forestillinger der lå bag styringen af det samlede tilbud, og hvilke forestillinger der lå til grund for den praktiske sociale indsats inden for de enkelte tilbud.

En sådan undersøgelse omfatter et væld af forskellige modeller for arbejdet. Nogle tilbud tager udgangspunkt i en religiøs forståelse for menneskers problemer og muligheder. Andre modeller lægger sig op ad den såkaldte Minnesota-model. Og ud over de modeller, der har en konkret reference til en bestemt terapeutisk tradition, vil der være en lang række tilbud, som bygger på en leders eller en medarbejdergruppes »common-sense« eller indhøstede erfaringer. Pointen med modelbetragtningen er dog, at disse former for »common-sense« og selvgjorte erfaringer stadig indeholder nogle referencer til forskellige former for systemiske, kognitive eller behavioristiske menneskesyn, som man kan afdække via interview. Selv om de frivillige og de ansatte således ikke er sig bevidst, at de refererer til en ganske bestemt forståelse af menneskelige behov og motivation, så vil dette altid være tilfældet. Epistemet (A) kommer således til udtryk i deres måde at omtale problemerne på, og det har indflydelse på deres handlinger (B) i forhold til målgruppen for indsatsen.

Metoder

Inden for hver enkelt region gennemførte man således en række undersøgelser, som skulle afsløre tilbuddets indre sammenhæng (de modeller, der indgik i arbejdet) og tilbuddets koordinering eller mangel på samme:

- En etnografisk undersøgelse (Høgsbro, Bovbjerg, Hardman Smith, Kirk og Henriksen 2003), hvor en etnograf opholdt sig i området i nogle perioder og gennemførte
 - en kortlægning af de forskellige tilbud til de forskellige målgrupper og disse tilbuds principper og målsætning

- en undersøgelse af brugernes livsverden, deres gruppemæssige tilhørsforhold, deres hverdag, brug af forskellige tilbud og personlige karriere
- en undersøgelse af samspillet mellem brugere og professionelle inden for udvalgte tilbud
- en spørgeskemaundersøgelse blandt brugere af udvalgte tilbud.
- En »auditundersøgelse«, hvor man samlede udvalgte professionelle i de forskellige regioner til en diskussion af nogle af de cases, som etnograferne havde sat fokus på (Brandt og Kirk 2003). Diskussionerne, der blev båndoptaget, berørte spørgsmål som professionel etik, begrænsningerne i tilbuddet og dilemmaerne i de professionelle krav.
- En organisationssociologisk undersøgelse af den koordinering af tilbuddet, som foregik på forvaltningsniveau i et samspil mellem forskellige sektorer og mellem amter og kommuner (Ebsen, Henriksen og Rieper, 2003).

Undersøgelsen omfattede således samtlige dele af figur 12.2. Man ville både belyse de forestillinger om problemerne og deres løsninger, som de professionelle handler ud fra, og man vil samtidig belyse brugernes forskellige livsverdenssammenhænge og se på, hvordan tilbuddene indgår som en del af deres hverdagsliv og muligvis understøtter deres personlige udvikling. Undersøgelsen så på, hvordan de professionelle tilbud har forskellige målsætninger og ambitioner, når det drejer sig om at minimere skaderne af et misbrug eller komme misbruget til livs. De har forskellige forestillinger, når det drejer sig om, hvilke krav det er nødvendigt at stille til brugerne, og hvilke krav der vil afskrække mennesker, som har brug for hjælp. De arbejder ud fra nogle forestillinger om, hvordan »en normal hverdag« ser ud, og stiller på denne baggrund forskellige krav til tidsstrukturen og arbejdsforpligtelserne i boligfællesskaberne. Nogle fokuserer på patienternes sindslidelse og mener, at kendskabet til sygdommens symptomer er den vigtigste kvalifikation for personalet, mens andre lægger vægt på de faglige udfordringer, patienterne kan få på væresteder og beskyttede værksteder. Der er således tale om et samlet tilbud, som i virkeligheden består af forskellige modeller, som er koblet sammen, hvorefter spørgsmålet er, hvad der er principperne for denne sammenkobling. Er

den koordineret af forvaltningen, som har forholdt sig bevidst til den kombination af tilbud, som er optimal? Har brugerne individuelt mulighed for at vælge de tilbud, som passer dem bedst? Eller er systemet opstået via knopskydning uden nogen egentlig prioritering eller bevidst stillingtagen til, hvad der var brug for? Har forvaltningen overhovedet en problemforståelse, som indeholder bestemte forestillinger om, hvilke tilbud der samlet set er brug for? En kombination af spørgeskemaundersøgelse blandt brugerne og kvalitative interview med professionelle, forvaltningsledere og brugere belyser disse spørgsmål.

Figur 12.2 kan således også anvendes på det overordnede forvaltningsniveau. Her er praksis knyttet til de mange forskellige tilbud, forvaltningen stiller til rådighed ud fra en idé om deres nødvendighed for en samlet brugergruppe. Man kan have en forestilling i forvaltningen om, at forskellige former for afrusning kan være lige gode, og at det vigtigste er, at brugerne får det tilbud, de hver især tror mest på. Man kan også have den forestilling, at der er ét bestemt tilbud, som har effekt, og at det gælder om at tvinge brugerne ind i netop dette tilbud.

Over for dette står en brugergruppe, der vælger tilbuddene til og fra ud fra forskellige ideer om, hvordan det kan gavne dem i hverdagen. Nogle har en idé om, at forskellige former for personlig udfordring ville hjælpe dem. Andre har afskåret sig fra forestillingen om at kvitte misbruget og skal først og fremmest bruge tilbuddene til at understøtte en krævende hverdag, der helt er underordnet kravene om at overleve i misbrugsmiljøet.

Resultater

Evalueringen konkluderede, at indsatsen kunne gennemføres på flere forskellige måder alt afhængig af de konkrete lokale betingelser for samarbejde mellem offentlige og private initiativer, men at der var en manglende forståelse for disse forskelle og et mangelfuldt lokalt overblik over de modeller, man kunne vælge imellem. Der var også en mangelfuld eller inkonsistent forståelse for betydningen af den frivillige indsats, og der manglede traditioner for tilbagevendende brugerundersøgelser, der kunne belyse det til enhver tid eksisterende behov. Desuden manglede man principper for efteruddannelse og spredning af information og viden om problemerne og nye tiltag på området, og man forholdt sig meget usystema-

tisk til spørgsmålet om brugerinddragelse, hvor man ofte manglede overordnede principper. Man manglede modeller for handleplansarbejdet i forhold til disse grupper, og man manglede modeller for mødet og samarbejdet med brugerne. Man manglede indsigt i den måde, hvorpå hver enkelt tilbud indgik i den enkelte brugers livssammenhæng, og man manglede især en integration af det sundhedsfaglige og det socialfaglige tilbud. Der var endvidere behov for en udvikling af botilbuddene, så de ikke kun udsprang af de forhåndenværende søms princip, men var målrettet denne gruppes særlige behov. Der var behov for udvikling af det opsøgende og udgående arbejde, og der var behov for udvikling af værestederne, således at de i højere grad forholdt sig til et rehabiliteringsperspektiv.

Formidling og anvendelse

Efterfølgende har evalueringen medført en del diskussion og mødevirk-somhed, fordi den, ud over ovenstående sammenfattende status over tilbuddets kvalitet, også kom med bud på, hvordan man kan forstå samspillet mellem de enkelte tilbud og brugernes livsverden og personlige udvikling. Evalueringens analyser bidrager således til at skabe overblik over indsatsen, sætte begreber på en række metoder, som har udviklet sig blandt innovative praktikere og synliggøre modeller for arbejdet, som ikke har været beskrevet eller diskuteret.

Evalueringen er på denne måde ikke en procesevaluering eller en formativ evaluering. Derimod indeholder den elementer af effektevaluering og forskningsbaseret erfaringsopsamling. Set i forhold til figuren har evalueringen fokuseret på samspillet mellem den øverste og nederste del af figur 12.2, de professionelles perspektiv og brugernes. Evalueringen har endvidere ikke kun omhandlet én bestemt model, som det var muligt i de tidligere nævnte eksempler, men flere forskellige samt den model (eller mangel på model), som gjorde sig gældende for prioriteringen af det samlede tilbud.

13 Evaluering af beskæftigelsespolitikken

Torben Pilegaard Jensen og Peter Plougmann

13.1 Introduktion og specifikke forhold

Hvad er beskæftigelsespolitik?

Beskæftigelsespolitikken er en central del af arbejdsmarkedspolitikken. Den forholder sig til alle former for initiativer, der kan forbedre beskæftigelsesmulighederne for de personer i og uden for arbejdsstyrken, der søger beskæftigelse. Beskæftigelsespolitikken er således andet og mere end aktiveringspolitik. Beskæftigelsespolitik er også effektiv arbejdsformidling, målrettet efteruddannelse og meritafklaring. Incitamentet i dagpengesystemet og målretning af vejledningsindsatsen er også vigtige virkemidler i beskæftigelsespolitikken. Endelig er beskæftigelsespolitikken en central del af integrationspolitikken over for flygtninge/indvandrere, selv om integration selvfølgelig har et meget bredere samfundsmæssigt sigte end beskæftigelsespolitikken.

De politiske rammer for beskæftigelsesindsatsen har i de seneste år fokuseret meget på at udvikle et enstrengt system, hvor der ikke længere skelnes mellem indsatsen over for kontanthjælpsmodtagere og ledige med dagpengere. Alle ledige har principielt sammen rettigheder og pligter og skal gives individuelt tilpassede tilbud.

Beskæftigelsespolitikken er ikke kun et nationalt anliggende. EU's beskæftigelsespolitiske mål (Lissabon-processen) kan klart aflæses i målsætningen for den danske beskæftigelsespolitik. Beskæftigelsespolitikken har derfor en velfærdsmålsætning, der skal realiseres ved at øge arbejdsstyrkens størrelse, øge beskæftigelsesniveauet og forbedre arbejdskraftens effektivitet, så den svarer til de krav, der stilles i det 21. århundrede.

Hvis den overordnede målsætningen for beskæftigelsespolitikken således er nogenlunde klar, så er det langt fra det samme, når man ser på delmålsætninger og krav til forskellige virkemidler. Der er her klart tale om, at evalueringer skal forholde sig til »moving targets«.

I perioder med høj beskæftigelse bliver fx effektiv formidling og mobiliseringen af den såkaldte »stille reserve«, dvs. de grupper, som befinder sig i randen af arbejdsmarkedet, helt central. I perioder med høj formel arbejdsløshed er det forebyggelse af langtidsarbejdsløshed og bekæmpelse af social udstødning, der er centrale målsætninger sammen med en indsats for at skabe flere job (à la det »rummelige arbejdsmarked«). Samtidig er der uanset konjunktur og politisk prioritering en generel forventning om, at alle de nævnte indsatsområder kan skaffe resultater.

De skiftende delmålsætninger og ofte hurtige skift i virkemidler gør det vanskeligt at se bort fra de politiske motivationer, der ligger bag beskæftigelsespolitikens valg af virkemidler. Denne politiske dimension slår igennem i evalueringsbestræbelserne. Det er således nødvendigt, at evaluator er sig de reelle magt- og interesseforhold bevidst, som ligger bag ønsket om fx skift af virkemidler, næsten uanset en evalueringens resultater og kvalitet. Anvendelsen af evalueringer i beskæftigelsespolitikken er altid underkastet de politiske interesser. Evaluator opererer, som ekstern tredjepart, derfor altid i et spændingsfelt af interesser, der stiller store krav til evaluator om såvel forståelse af kontekst, som til uafhængighed og faglig professionalisme.

Hvem anvender evalueringer?

Det er både de centrale og de regionale arbejdsmarkedsmyndigheder, der bestiller og anvender evalueringer af beskæftigelsespolitikken.

EU-Kommissionen efterspørger også evalueringer af beskæftigelsespolitikken, og der er i forhold til Socialfondsaktiviteterne næsten altid evalueringsskrav knyttet til bevillingerne.

Kommunerne optræder i stigende grad som rekvirenter af evalueringer ikke mindst i forbindelse med integrationsproblemstillinger. Det må forventes, at kommunerne, som en konsekvens af den forventede strukturreform, vil anvende evalueringer i stigende omfang.

Det er karakteristisk for evalueringer af beskæftigelsespolitikken, at arbejdsmarkedets parter (både centralt og regionalt) ofte er inddraget i formuleringen af evalueringskriterier og deltager i den efterfølgende politiske og administrative anvendelse af evalueringernes resultater.

Det er således politiske institutioner på mange niveauer, der efterspørger evalueringerne, og de inddrager ofte et bredt udsnit af de involverede »stakeholders«.

Hvad er de centrale evalueringsspørgsmål?

De centrale evalueringsspørgsmål er overordnet knyttet til de succeskriterier, der er opstillet for anvendelse af specifikke virkemidler rettet imod diverse målgrupper, se fx Jensen (1991).

Der er derfor en klar makromålsætning om en øget samlet beskæftigelse og et kvalifikationsløft af arbejdsstyrken. Men det er ikke de eneste forhold, som en evaluering forholder sig til. Der er tale om et mål-middelhierarki.

Vejen til bedre beskæftigelse for deltagerne i de forskellige beskæftigelsespolitiske initiativer er således også relevant.

Der skelnes her mellem endelige og middelbare deltagermål. Endelige deltagermål er, fx at deltagerne ud over at komme i beskæftigelse starter egen virksomhed eller kommer til en situation, hvor de reelt står til rådighed for arbejdsmarkedet.

De middelbare deltagermål er måske nok mere »bløde«, men så absolut afgørende forudsætninger for realiseringen af de endelige deltagermål og den overordnede makromålsætning.

Tabel 13.1 **Beskæftigelsespolitikens mål-middel-hierarki**

Makromål	Øge den samlede beskæftigelse Kvalifikationsløft i arbejdsstyrken
Endelige deltagermål	Beskæftigelse Etablering af egen virksomhed Rådighed
Middelbare deltagermål	Fortsat uddannelse Kvalifikationer (almene/erhvervsrettede) Personlige forudsætninger, herunder – selvtillid – uddannelsesmotivation – arbejdsmotivation Sociale forudsætninger Jobsøgningsaktiviteter

Det er fx etableringen af sociale forudsætninger for, at deltagerne kan blive aktivt jobsøgende – det er selve jobsøgningsaktiviteterne, og det kan være, at deltagerne opnår de nødvendige personlige forudsætninger for i det hele taget at arbejde (fx arbejdsmotivation og selvtillid).

Endelig er tilegnelsen af såvel almene som faglige kvalifikationer fx via efteruddannelse eller fortsat uddannelse også relevante mål og derfor retningsgivende for, hvilke evalueringsspørgsmål der opstilles.

Dette mål-middel-hierarki er bestemmende for, hvilke evalueringsspørgsmål der stilles.

Det er således væsentligt, at der ikke kun er fokus på »nøgne beskæftigelseseffekter«, men at der også kan gives svar på processuelle spørgsmål ud fra en administrativ-, en organisatorisk-, en politologisk- og interesseanalytisk tilgang.

Evalueringsspørgsmålene er ud over, hvor effektiv beskæftigelsespolitikken er på makroniveau, mindst lige så meget knyttet til processuelle mål-middel-spørgsmål. Hvorfor og hvordan et givent resultat er fremkommet, er ud fra et rekvirentsynspunkt mindst lige så relevant som markomål-sætninger. Derudover er det ud fra et legitimitetssynspunkt også væsentligt at give plads til deltagernes egne opfattelser og begrundelser.

Danske evalueringsmiljøer

Der er mange aktører i det danske evalueringsmiljø inden for beskæftigelsespolitikens område. Det er en sammenblanding af selvevaluering /interne evalueringer, forskningsrelaterede evalueringsmiljøer og kom-

mercielle private konsulentvirksomheder.

De beskæftigelsespolitiske myndigheder optræder ofte selv som evaluører via de mål- og rammestyringssystemer, der er opstillet for beskæftigelsessystemets ageren. Der er via utallige økonomistyringssystemer mange muligheder for løbende, kvantitative resultatopgørelser, der i realiteten kommer til at indgå i den løbende administrative selvrefleksion, og som bidrager til en samlet vurdering af beskæftigelsespolitikken effektivitet. Kvantitative benchmarkingredskaber, fx til at sammenligne den kommunale beskæftigelsesindsats, anvendes ofte af centrale myndigheder fx Beskæftigelses- og Finansministeriet, og disse redskaber vil givet vis blive mere udbredt i de nye storkommuner og nye regioner. Strukturkommissionens overvejelser om beskæftigelsespolitikken forvaltning og forankring understreger behovet for, at der fremover bliver foretaget løbende evalueringer af beskæftigelsespolitikken baseret på hurtige kvantitative benchmarkingsystemer.

Sektorforskningsmiljøer (AKF, SFI, Institut for grænseregionsforskning og universitetscentre Center for Arbejdsmarkedsforskning ved Aalborg Universitet (CARMA), Handelshøjskolen i Århus, Roskilde Universitetscenter, Syddansk Universitet o.a.) tilbyder sig som evaluører af beskæftigelsespolitikken. De foretrækkes ofte af myndighederne, da de fremstår som uafhængige og fagligt solide evaluører. De har dog det problem, at de kan have svært ved at reagere hurtigt.

Det private konsulentmarked har et meget stort udbud af virksomheder, der tilbyder evalueringer (fx Oxford Research, PLS-Rambøll Management, COWI og mange mindre ofte regionale konsulentvirksomheder). De foretrækkes, fordi de repræsenterer en kombination af faglighed, solide metoder og hurtighed.

13.2 **Eksempel 1: Anvendelsen af andre aktører i beskæftigelsespolitikken**

Overordnet formål med den konkrete evaluering

Dette eksempel belyser en endnu ikke afsluttet evaluering, der trods dette forhold er valgt, fordi den forholder sig til et aktuelt og politisk omdisku-

teret indsatsområder; brugen af andre aktører end AF i kontaktforsøg til ledige. Det er en central antagelse, at man ved at inddrage andre aktører med speciale i »de rigtige tilbud til særlige målgrupper« kan opnå bedre resultater end dem, der hidtil er opnået via Arbejdsformidlingen (AF) og kommunerne. Det overordnede krav til de aktiviteter, andre aktører kan sætte ind med, er, at indsatsen er tilpasset den enkeltes ressourcer og behov, og at der er et klart og målrettet jobsigte. Målet er, at ledige via beskæftigelsesindsatsen tilbydes individuelle skræddersyede tiltag, der fremmer den korteste vej til selvforsørgelse gennem ordinær uddannelse eller ordinært arbejde.

Der er siden slutningen af 2003 iværksat en række evalueringer af denne centrale del af den aktuelle beskæftigelsespolitik. Den evaluering, der beskrives i dette afsnit, er startet i februar 2004 og afsluttes primo 2005.

Baggrunden for evalueringen er, at 3 regioner (AF-Viborg Amt, AF-Århus Amt og AF-Ringkøbing Amt) har et tværregionalt samarbejde, hvorigennem der er etableret kontaktforsøg for udvalgte målgrupper med brug af eksterne aktører efter det samme udbudskoncept. Der er tale om kontaktforsøg, som er etableret i efteråret 2003 med en varighed indtil ultimo 2004. Deltagerne på de etablerede forsøg vil i takt med, at de opnår arbejde eller påbegynder uddannelse mv., forlade forsøgene.

Formålet med evalueringen er tosidigt. For det første skal evalueringen være med til at forberede beslutningsgrundlaget for arbejdsmarkedsrådene i AF-regionerne: Viborg, Århus og Ringkøbing Amter i forbindelse med kommende overvejelser om inddragelse af andre aktører i beskæftigelsesindsatsen. For det andet skal evalueringen være med til at fremme AF's muligheder for – på det udførende niveau – at kunne tilrettelægge en smidig og hensigtsmæssig administration i samspillet med andre aktører.

Temaet i evalueringen

Evalueringen har tre overordnede temaer:

- I. Processen/samspillet mellem AF og aktørerne.
- II. Deltagernes opfattelse af opholdet hos anden aktør.
- III. Effekten af andre aktørers indsats.

Der er under det første tema, processen, fokus på samspillet mellem AF og de aktører, der er valgt. Der er især lagt vægt på at belyse:

- Udbudsmaterialet/ydelsesbeskrivelse
- Samspillet i forbindelse med forløbsopstart og visitation
- Afregning af driftstilskud og bonus
- Den løbende informationsudveksling – samspil under forløbet
- Udslusning og evt. aktivering af den ledige
- AF's ressourceforbrug i forbindelse med samspillet

Temaet om deltagernes opfattelse af opholdet hos andre aktører belyses ved analyser af:

- De lediges opfattelse/vurdering af egne beskæftigelsesmuligheder – er de forbedret?
- Deltagernes opfattelse af egne beskæftigelsesmål
- Opfattelse/vurdering af jobsøgning
- Opfattelse af forløbet hos anden aktør
- Vurdering af kvalitet og indhold i forløbet

Deltagerne skal desuden forholde sig til aktiviteter, de har gennemført. Sådanne aktiviteter *kan eksempelvis* være:

- Afklaringsamtaler om den enkeltes kvalifikationer og jobmuligheder
- Støtte til udformning af ansøgninger
- Formidling til job
- Kurser i jobsøgning
- Besøg på arbejdspladser med henblik på afklaring af jobønsker
- Vejledning i brug af selvbetjeningsredskaber
- Samtaler om individuel handlingsplan
- Deltagelse i jobklubber e.l.

Effekten af andre aktørers indsats belyser udslusningseffekter og de mere langsigtede beskæftigelseseffekter:

- De direkte beskæftigelsesmæssige effekter af anden aktørs indsats.
- Overlevelsestavler
- De indirekte effekter af indsatsen
- Resultataflønningsmodellens betydning for resultat og aktøradfærd

Evalueringsdesign og metoder

Der anvendes et projektdesign, hvor der lægges vægt på at gennemføre en evaluering, der bygger på en flerhed af metoder – både kvalitative og kvantitative.

Ved at anvende både kvalitative og kvantitative metoder i evalueringen opnår evaluatoren indsigt i forhold, der ofte er umulige at indfange, hvis der kun anvendes kvantitative metoder. Ved at anvende både »hårde« og »bløde« data er det hensigten at øge forståelsen for processer, der kan optimere effekten af den givne indsats (her konkret kontaktførelsen) og ikke kun vurdere, om det er gået godt eller skidt.

Designet giver en mulighed for at besvare hvordan- og hvorforspørgsmål, der er af relevans for rekvirenten i den fremtidige, konkrete planlægning af beskæftigelsesindsatsen.

Evalueringsopgaven bliver gennemført på basis af en række forskellige datakilder:

- Desk research hos AF: fx om kriterierne for de udvalgte forløb, kriterier for inddeling i målgrupper, resultataflønningsmodellerne, retningslinjer mv. for samarbejdet med andre aktører o.a.
- Desk research hos »andre aktører«. En gennemgang af de enkelte forløb med henblik på at registrere og kort beskrive indhold og metoder i konkrete forløb. Der er fokus på, hvilke elementer aktørerne har lagt vægt på at inddrage og forløbenes strukturering mv.
- Interview med nøglepersoner i alle tre AF-regioner, tematiseret omkring de samme forhold som desk researchen.
- Interview med nøglepersoner hos »andre aktører«, tematiseret omkring de samme forhold som desk researchen, samt deres vurdering af resultataflønningsmodellerne mv.
- Interview med ledige, der har deltaget i forløbene. Det tilstræbes, at der sker en fordeling af interviewene, så der både interviewes ledige, der er kommet i job/uddannelse mv., og ledige, der ikke er placeret.
- Fokusgruppeinterview af hvert relevant hold af ledige.
- Spørgeskemaanalyse blandt alle ledige, der har deltaget i et af de relevante hold, der er med i evalueringen. Spørgeskemaet fokuserer på deltagernes opfattelse af opholdet hos anden aktør, herunder bl.a. de lediges opfattelse/vurdering af egne beskæftigelsesmuligheder, beskæftigel-

sesmål, jobsøgning mv., men også fortsat uddannelse, (erhvervede kvalifikationer), personlige og sociale forudsætninger vil blive belyst.

- DREAM-dataanalysesættet fra Beskæftigelsesministeriet anvendes i effektanalysen. Her er fokus især på den langsigtede beskæftigelseseffekt, der defineres ved beskæftigelsesstatus 9 måneder efter udslusningen. Ved at anvende beskæftigelsesstatus 9 måneder efter udslusningen fås en mere sikker indikation på effekten af indsatsen.

Tabel 13.2 Sammenhængen mellem delanalyser og datagrundlag

	Tema 1 Proces og samspil	Tema 2 Deltagernes opfattelse	Tema 3 Effekten
Desk research AF, fokus på bl.a.: – udbudsmateriale – kriterier for inddeling i målgrupper – retningslinjer mv. for samarbejdet med andre aktører – opgørelse af udslusning for ledige der har deltaget i disse eller andre forløb – resultataflønningsmodellerne	X	X	X
Desk research »andre aktører« – beskrivelser af de enkelte forløb	X	X	X
Interview med AF nøglepersoner	X		X
Interview med relevante nøglepersoner hos »andre aktører«	X		X
Interview med ledige, der har deltaget		X	X
Fokusgruppe med ledige, der har deltaget		X	X
Spørgeskema til alle ledige, der har deltaget		X	X
DREAM-dataanalysesættet			X

Hovedresultater, formidling, evt. anvendelse

Af naturnødvendige årsager foreligger der ved redaktionens afslutning (april 2004) endnu ingen hovedresultater fra evalueringsprocessen. Der kan dog knyttes en række kommentarer til formidling og anvendelse.

Det er som udgangspunkt for evalueringen defineret, at samspillet med opdragsgiver er helt centralt for gennemførelsen af evalueringen, og evaluator har i det første år af evalueringen en række møder med rekvirenterne om foreløbige observationer og overvejelser. Der er lagt vægt på, at der er tale om en egentlig sparringsproces, der gør evalueringen praktisk anvendelig for rekvirenterne. For at understøtte dette anvendelsesformål

udarbejdes der et antal notater, som rekvirenterne kan anvende i deres administrative praksis for at forbedre og effektivisere indsatsen.

13.3 **Eksempel 2: En evaluering af »Lokale koordinationsudvalg«**

Overordnet formål

I dette afsnit skitseres en evaluering af »Lokale koordinationsudvalg«, jf. Caswell og Jensen (2001).

Siden 1. januar 1999 havde alle kommuner pligt til at nedsætte et lokalt koordinationsudvalg for den forebyggende arbejdsmarkedsindsats, evt. i fællesskab med andre kommuner. Koordinationsudvalgene er en del af den samlede indsats for at skabe et mere rummeligt arbejdsmarked. Det er udvalgets opgave at »samordne og udvikle den lokale forebyggende indsats for personer, der har svært ved at klare sig på arbejdsmarkedet«, dvs. at medvirke til at fremme et mere rummeligt arbejdsmarked. Udvalget skal arbejde for at øge beskæftigelsesmulighederne for borgere, der er i fare for at miste tilknytningen til arbejdsmarkedet, eller borgere der ikke eller kun vanskeligt kan opnå beskæftigelse på almindelige vilkår.

Udvalgene er sammensat af repræsentanter for centrale aktører på området og skal rådgive kommunerne om den sociale arbejdsmarkedsrettede indsats. Koordinationsudvalgene skal desuden rådgive andre lokale aktører om den sociale arbejdsmarkedsrettede indsats, fx virksomheder og faglige organisationer. Sidst, men ikke mindst, har udvalgene til opgave at styrke samarbejdet i forhold til den sociale arbejdsmarkedsrettede indsats på lokalt plan.

Evalueringens formål var at »skaffe dokumentation for, om og på hvilken måde udvalgenes indsats, og herunder midlerne (de økonomiske), har medvirket til at fremme det rummelige arbejdsmarked på det lokale plan og forbedret samarbejdet mellem de berørte aktører«.

Derudover skulle skitseres muligheder og barrierer for udvalgenes videre arbejde.

Selv om det havde været ønskeligt at medtage mere objektivt orienterede mål for effekten af udvalgenes virke, er sådanne ikke inddraget. Det

kan også udtrykkes på den måde, at evalueringen ikke inddrager endelige deltagermål. Derimod tillagde opdragsgiver det meget stor betydning at få belyst, om udvalgene havde fremmet samarbejdet mellem de deltagende aktører.

Evalueringsdesign og -temaer

Evalueringen omfattede en række delundersøgelser. Designet bestod i spørgeskemaundersøgelser til samtlige medlemmer af lokale koordinationsudvalg og sekretariater. Disse gav et landsdækkende billede af, hvordan medlemmer og sekretariater vurderede, at udvalgene fungerede. Spørgeskemaundersøgelserne omfattende disse temaer:

- Samspil mellem koordinationsudvalg og kommune
- Koordinationsudvalgenes beslutningsgrundlag
- Koordinationsudvalgenes handleplaner
- Udvalgsaktivitet
- Effekter af indsatsen
- Rammebetingelser og effekt.

På grundlag af de to spørgeskemaundersøgelser blev udvalgt udvalg til kvalitative studier. For at øge værdien af de kvalitative studier blev udvalgt udvalg, som kunne defineres som henholdsvis velfungerende og mindre velfungerende ud fra følgende kriterier:

- udvalget har ændret de lokale arbejdspladsers indstilling til det rummelige arbejdsmarked i positiv retning
- udvalget har øget det lokale samarbejde mellem de i udvalget repræsenterede parter
- fraværet ved udvalgmøderne har været begrænset
- der er stort engagement blandt udvalgsmedlemmerne i og uden for udvalget
- der er opstillet målsætninger/handleplaner og udpeget målgrupper for udvalgets indsats
- udvalget har stor viden om det lokale arbejdsmarked og om kommunens sociale arbejdsmarkedsrettede indsats.

Udvalg, hvor det modsatte gjaldt, blev defineret som mindre velfungerende. Havde der været tale om tilfældig udvælgelse, ville muligheden for at studere karakteristika ved gode og dårlige udvalg netop ikke være til stede.

Andre kriterier kunne tænkes at være relevante, men i de kvalitative studier blev deres egnethed som udvalgs-kriterier bekræftet. Det viste sig således, at henholdsvis velfungerende og mindre velfungerende udvalg også på andre områder var godt, henholdsvis mindre godt fungerende, fx med hensyn til om de havde været med til at ændre den kommunale praksis i den sociale arbejdsmarkedsrettede indsats i positiv retning. Formålet med de kvalitative studier var at komme dybere ned i forståelsen af, hvilke barrierer der gjorde sig gældende i udvalgsarbejdet og i arbejdet for at virkeliggøre formålet med udvalgene: et mere rummeligt arbejdsmarked.

Dataindsamling og analysemetoder

Undersøgelsens delundersøgelser fremgår i oversigtlig form af tabel 13.2.

Tabel 13.3 **Undersøgelsens opbygning: dataindsamling og analysemetoder**

Undersøgelsesdel	Indhold	Metode
Spørgeskemaundersøgelse blandt udvalgsmedlemmer	Belyser udvalgsmedlemmernes deltagelse i og vurdering af udvalgsarbejdet	Spørgeskema til samtlige 1106 medlemmer af de lokale koordinationsudvalg. Besvarelsesprocent: 70
Spørgeskemaundersøgelse blandt sekretariater for de lokale koordinationsudvalg	Belyser koordinationsudvalgenes aktiviteter og sekretariaternes vurdering af udvalgenes arbejde	Spørgeskemaundersøgelse blandt samtlige 147 sekretariater for de lokale koordinationsudvalg. Besvarelsesprocent: 90
Kvalitative undersøgelser	Undersøger bagvedliggende institutionelle og politologiske faktorerens betydning for udvalgenes virke	Interview blandt udvalgsmedlemmer og sekretariater samt observationer af udvalgmøder

Hovedresultater

Evalueringen viste, at mange koordinationsudvalg på daværende tidspunkt var kommet godt i gang med deres arbejde. Udvalgene havde i mange lokalområder været med til at synliggøre visionen om det rummelige arbejdsmarked, og mange arbejdspladser viste åbenhed over for den sociale arbejdsmarkedsrettede indsats ifølge medlemmerne. Det gjaldt først og fremmest med hensyn til forebyggelse og arbejdsfastholdelse.

Disse resultater var bl.a. en konsekvens af, at koordinationsudvalgene havde medvirket til et bedre samarbejde mellem de lokale aktører, repræsenteret i udvalgene, til at udvalgene havde fået større indsigt i, hvem der var socialt udsatte, og til at dialogen mellem kommune og udvalg var blevet bedre. Koordinationsudvalgene havde ifølge medlemmerne desuden medvirket til en ændret kommunal praksis i den sociale arbejdsmarkedsrettede indsats.

De økonomiske midler, som udvalgene var tildelt, medvirkede til et forbedret samarbejde mellem udvalg og kommune, mellem kommuner og virksomheder og førte til anvendelse af nye metoder i den forebyggende arbejdsmarkedsindsats ifølge udvalgsmedlemmerne og sekretariatene.

Mål og handleplaner for udvalgsarbejdet, og en kommunal ramme for den sociale arbejdsmarkedsrettede indsats, havde positiv betydning for udvalgsarbejdet. Men omkring en tredjedel af henholdsvis udvalgene og kommunerne havde ikke sådanne. Hvis flere koordinationsudvalg skulle opnå den rolle, som de var tiltænkt, blev der peget på følgende tiltag:

- at alle kommuner opfylder kravet om at opstille en ramme for den forebyggende sociale arbejdsmarkedsrettede indsats.
- at samtlige udvalg formulerer mål og handleplaner for deres arbejde i overensstemmelse med den kommunale ramme for den sociale arbejdsmarkedsrettede indsats.
- at udvalgsmedlemmerne er engagerede i arbejdet med at fremme det rummelige arbejdsmarked.
- at der er dialog mellem koordinationsudvalg og kommunens(-ernes) politiske niveau.
- at flere udvalg får et bedre samarbejde med de praktiserende læger for at sikre deres aktive medvirken i den sociale arbejdsmarkedsrettede indsats.
- at udvalg, som fungerer på arbejdsmarkeder, der ikke er bæredygtige i forhold til den sociale arbejdsmarkedsrettede indsats, overvejer at indgå samarbejder med de nabokommuner, med hvem de har et fælles arbejdsmarked.

Formidling, evt. anvendelse

Undersøgelsens resultater er blevet formidlet gennem artikler og foredrag. Opdragsgiver modtog et stort antal rapporter, som blev formidlet til de centrale aktører på området. Det er forskernes opfattelse, at undersøgelsen har givet inspiration til mange udvalg, som derved har forbedret deres målrettede indsats. Undersøgelsen har også påkaldt sig interesse i forskningsverdenen, se fx Andersen og Torfing (2002).

14 Evaluering af aktivering ved hjælp af kvantitative metoder

Michael Rosholm

14.1 Indledning

Når man kvantitativt skal evaluere aktivering, må man naturligvis starte med at gøre sig klart, ud fra hvilket kriterium der skal evalueres. Og der er adskillige.

Det kan være, at man er interesseret i effekten på sandsynligheden for, at den enkelte aktiverede finder job. Dette kan vel betegnes som et kort-sigtet kriterium. På lidt længere sigt kunne man også være interesseret i at se, om aktiverede opnår mere stabil beskæftigelse og bedre lønninger end de ville have gjort, var de ikke blevet aktiverede.

Det kan også tænkes, at man er interesseret i en cost-benefit-analyse af hele – eller dele af – aktiveringsområdet. Denne analyse forudsætter naturligvis, at der foreligger pålidelige resultater fra den ovenstående analyse.

Endelig kan det tænkes, at man er mere interesseret i at evaluere på helt andre kriterier som den enkelte lediges trivsel, helbredsmæssige og social-psykologiske mål, etc. Sådanne kriterier lader sig naturligvis kun evaluere i det omfang, de kan kvantificeres og registreres i et datamateriale, som kan stilles til rådighed for forskerne. Dette kunne være i form af survey-baserede data, eller i form af registre, som er informative om disse faktorer (eksempelvis kan helbredsmæssige og i et vist omfang mentale konsekvenser af ledighed måles ved hjælp af lægemiddelstatistikregistret og/eller registre over sygehusindlæggelser, diagnoser, lægebesøg osv.).

De sidstnævnte muligheder er først for nylig blevet stillet til rådighed for forskerne, så derfor vil jeg i denne artikel tage udgangspunkt i et eksempel, som belyser den centrale problemstilling.

Central problemstilling 1: »Hvorledes påvirker deltagelse i aktivering sandsynligheden for, at den ledige finder et job?«

Der findes – efter min bedste viden – ikke håndbøger og eller vejledninger i denne svære kunst at bedrive kvantitative evalueringer af aktivering og lignende foranstaltninger. Der findes en del videnskabelig litteratur, som omhandler emnet, se for eksempel Heckman m.fl. (1999) og Schmid m.fl. (1996), mens der også findes megen international litteratur, som beskriver kvantitative evalueringer. Heckman m.fl. (1999) giver et overblik over denne litteratur, ligesom for eksempel Martin & Grubb (2001), og Kluwe & Schmidt (2002). *Review of Economic Studies* havde i 1997 et særnummer om kvantitative evalueringsteknikker, ligesom *Journal of the American Statistical Association* (serie A) havde det i 2002. Arbejdsministeriet (2000) og Bolvig m.fl. (2001) er eksempler på danske evalueringer, som samtidig har den fordel, at fremstillingen er mindre teknisk end i ovenstående bidrag. Jensen m.fl. 2002 giver en oversigt over effekterne af den aktive arbejdmarkedspolitik i Danmark.

Der findes flere institutioner, hvor man har ekspertisen til at foretage kvantitative evalueringer af aktivering. Det drejer sig især om AKF og SFI, men også på de økonomiske institutter på flere universiteter findes en sådan ekspertise.

Det helt centrale i en kvantitativ evaluering af effekten af aktivering er konstruktionen af effektmålet, og for at kunne beregne dette må man altid være i stand til at beskrive den kontrafaktiske situation;

Kontrafaktisk situation 1: »Hvor stor havde chancen for at få et job været, hvis den ledige ikke havde deltaget i aktivering?«

Denne kontrafaktiske situation er ekstra vanskelig at konstruere i den danske sammenhæng, hvor det er en pligt at deltage i aktivering, hvis man har været ledig eller på kontanthjælp tilstrækkeligt længe. Det er ofte ikke muligt at konstruere et sådant sammenligningsgrundlag, hvorfor den centrale problemstilling i nogle tilfælde må udvandes til en *relativ* sammenligning af job-sandsynligheden i forbindelse med deltagelse i forskellige typer af aktivering. Her omformuleres den centrale problemstilling til:

Central problemstilling 2: »Hvorledes påvirker deltagelse i en bestemt type aktivering sandsynligheden for, at den ledige finder et job sammenlignet med deltagelse i en anden type aktivering?«

og den kontrafaktiske situation er så:

Kontrafaktisk Situation 2: »Hvor stor havde chancen for at få et job været, hvis den ledige havde deltaget i en anden type aktivering?«

De associerede effektmål er naturligvis mindre interessante end dem, som er associerede med den centrale problemstilling, idet de kun tillader en relativ evaluering; hvilken type aktivering er mest effektiv, frem for en absolut; virker aktivering overhovedet? Effektmålene er ikke desto mindre stadig interessante.

En anden mulighed er at anvende et traditionelt eksperimentelt design, men det har der været modstand mod herhjemme, af forskellige årsager. Derfor har der – inden for aktiveringsområdet – kun været gennemført et enkelt ægte eksperiment, se Langager (1996). Den eksperimentelle metode er imidlertid heller ikke uden problemer, når den anvendes i samfundsvidenskabeligt øjemed, for hvordan sikres, at selvstændigt tænkende individer agerer, som var de deltagere i et laboratorieeksperiment. Se for eksempel diskussionen heraf i Rosholm & Skipper (2003).

Nedenfor giver jeg to eksempler på effektevaluering af aktiveringsforanstaltninger. Begge eksempler baserer sig på det samme datasæt, så evalueringerne adskiller sig ved, at den første forsøger at besvare den mindre ambitiøse problemstilling 2, mens den anden anvender mere komplicerede teknikker, men til gengæld – under passende antagelser – sætter os i stand til at besvare problemstilling 1. Begge eksempler er baseret på en evaluering af aktiveringsforanstaltninger i Århus Kommune, nemlig rapporten »Effekter af aktiveringsindsatsen i Århus Kommune« af Bolvig m.fl. (2001). Der er således tale om én evalueringsrapport, som indeholder flere evalueringer. Formålet med rapporten er at måle beskæftigelseseffekten af de foranstaltninger, Århus Kommune tilbyder sociale klienter. I denne gennemgang vil vi imidlertid udelukkende se på effekten af aktiveringsforanstaltninger rettet mod kontanthjælpsmodtagere. Der er gennemført to typer af evalueringer, en svarende til den mindre ambitiøse problemstilling og en svarende til den mere ambitiøse problemstilling. Ideen hermed var at

vurdere, om kommunen begik grove fejlskøn, hvis den baserede politikudviklingen på de mindre ambitiøse mål, som har den fordel, at de kan konstrueres og beregnes af kommunens egne medarbejdere.

Formålet med begge evalueringer er at sætte de kommunale myndigheder på området i stand til at fin-tune aktiveringsindsatsen, det vil sige at sætte dem i stand til at anvende de instrumenter, som giver den største effekt på det rette tidspunkt. Dette skulle gerne føre til, at kommunens kontanthjælpsmodtagere hurtigere end ellers kan bringes (tilbage) i selvforsørgelse.

14.2 **Eksempel 1: Et simpelt mål på effekt af aktivering**

Formålet med dette eksempel er at vise, hvorledes man kan konstruere relativt simple mål, som faktisk er informative om effekterne af aktiveringsforanstaltninger.

Datasættet består af individer, som i løbet af foråret 1998 påbegyndte eller afsluttede et forløb i kontanthjælpssystemet i Århus Kommune, eller påbegyndte eller afsluttede et aktiveringsforløb i forbindelse med et ophold i kontanthjælpssystemet i foråret 1998. Data er udtrukket fra forskellige kommunale registre, som så er kørt sammen for at danne en forløbshistorik for den enkelte kontanthjælpsmodtager. En forløbshistorik er i al sin enkelhed en kalender, som på ethvert tidspunkt angiver, i hvilken tilstand (beskæftiget, på kontanthjælp, i aktivering, uden for arbejdsstyrken, etc.) en person befinder sig, og man kan derfor se, hvornår en person påbegyndte et forløb på kontanthjælp, hvornår personen deltog i aktivering (og i hvilken type aktivering), hvor længe personen var på kontanthjælp, og hvorvidt personen efterfølgende fandt et job. Konstruktionen af forløbshistorikken er ofte det største arbejde i en analyse som denne, da det for det første kan være svært at koble de forskellige kommunale registre sammen, og da det for det andet meget ofte hænder, at oplysningerne i det ene register ikke stemmer med oplysningerne i det andet. I disse tilfælde er et godt samarbejde med kommunale embedsmænd (som kender registrene, deres funktion osv.) meget vigtigt, da de kan bidrage med værdifuld information

om, hvordan sådanne uoverensstemmelser kan opstå, således at man kan nå frem til en beslutning om, hvilket register man skal »stole på«. I denne fase af analysen er en god programmør (og gerne en med kendskab til registre) næsten uundværlig.²² I dette tilfælde er data også koblet sammen med informationer fra registre i Danmarks Statistik, men analysen kunne principielt godt gennemføres uden disse supplerende data.

Lad mig starte med at redegøre for den traditionelle tilgang til effekt-evaluering, og hvorfor den ikke lader sig anvende i den danske sammenhæng. Forestil jer en gruppe ledige, som alle efter ca. et år deltager i aktivering. Den traditionelle tilgang til effektevaluering består så i at sammenligne den gennemsnitlige forsørgelsesgrad et år før deltagelse i aktivering med den gennemsnitlige forsørgelsesgrad et år efter deltagelse i aktivering. Problemet heri er naturligvis, at forsørgelsesgraden året før deltagelse i aktivering naturligvis er meget høj, ellers var man jo aldrig endt op i aktivering. En høj forsørgelsesgrad er så at sige den udløsende faktor for aktivering. Derfor kan forsørgelsesgraden heller ikke vokse (den *kan* jo ikke blive større end 1). Derimod vil den helt sikkert falde som følge af, at nogle finder job, *uanset* om de aktiveres eller ej. Vort effektmål kommer således til at afspejle en kombination af den sande effekt af aktivering og naturlig afgang fra ledighed. Aktiveringseffekterne bliver »for store«. Traditionelt har man kontrolleret herfor ved at konstruere en kontrolgruppe med samme ledighedsomfang på bagen, som ikke deltager i aktivering, men dette er jo netop ikke muligt i den danske kontekst, hvor deltagelse i aktivering er obligatorisk efter et års ledighed (situationen er her sat på spidsen, idet nogle kommuner aktiverer længe før et år på kontanthjælp, og nogle får lov at gå længere end et år, men problemet med et obligatorisk aktiveringssystem består).

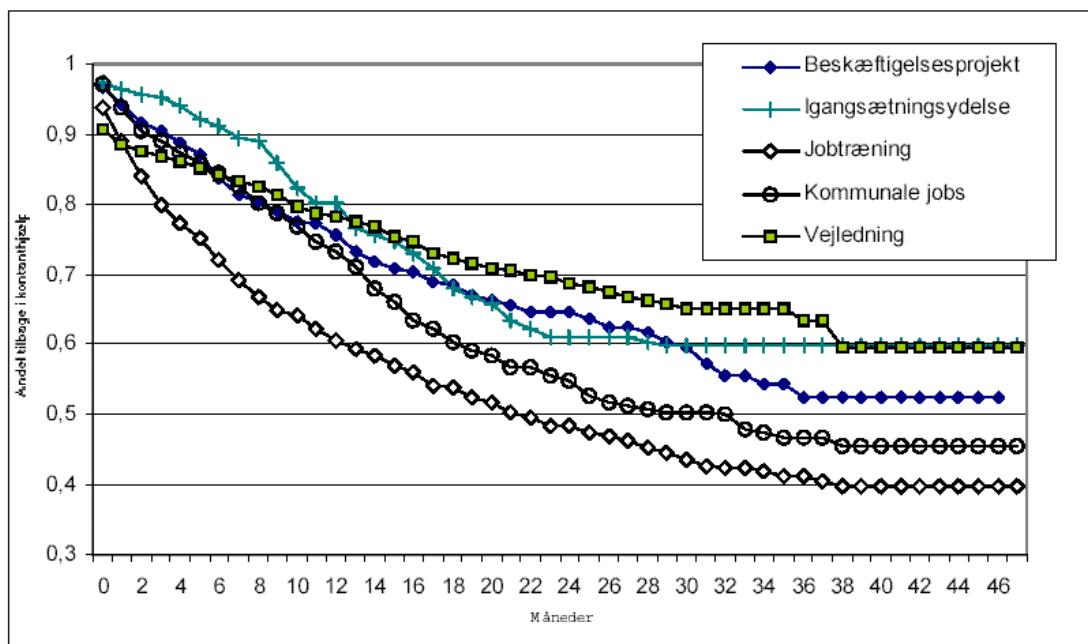
Teknikken, som beskrives ovenfor, kan dog stadig anvendes til en evaluering af de *relative* effekter af aktivering, forudsat at det er tilfældigt, hvem der deltager i hvilken type aktivering (en ikke ovenud sandsynlig forudsætning). Den samme forudsætning ligger bag konstruktionen af det effektmål, som beskrives her, og som giver et mere dynamisk billede af aktiveringseffekterne.

Lad os betragte alle individer, som netop har afsluttet et aktiveringsforløb, og så for hver type aktivering plotte *overlevelseskurven i kontant-*

hjælp. Vi definerer derfor en variabel, vi kalder T , som for hver person i vores datasæt måler tiden fra afsluttet aktivering, indtil individet forlader kontanthjælp til fordel for job eller uddannelse. Herefter beregnes for hver tidsenhed (i dette tilfælde hver måned) overlevelsesfunktionen. Lad antal individer, som afslutter en given type aktivering, være angivet ved N , og lad antal individer, som stadig er på kontanthjælp ved udgangen af den j 'te måned, være n_j . Overlevelsesfunktionen efter den j 'te måned er så ganske enkelt n_j / N .

Figur 14.1 viser fem sådanne overlevelseskurver for fem forskellige typer af aktivering, beskæftigelsesprojekter, igangsætningsydelse, jobtræning, kommunale job og vejledning (for en nærmere definition af disse grupperede aktiveringsformer, se Bolvig m.fl., 2001, som indeholder en detaljeret gennemgang af metoden).

Figur 14.1 Illustration af overlevelsesfunktioner i kontanthjælp efter forskellige typer af aktivering



Kilde: Bolvig m.fl. (2001).

Det, der mangler i figur 14.1 for (under visse forudsætninger) at kunne besvare den centrale problemstilling 1, er overlevelsesfunktionen i kontanthjælp uden aktivering. Der findes naturligvis personer, som forlader

kontanthjælpssystemet af sig selv (uden at deltage i aktivering), men de er jo (næsten) alle karakteriseret ved at have relativt korte kontanthjælpsforløb; ellers var de jo netop kommet i aktivering. At sammenligne med dem ville være uretfærdigt, idet deres overlevelseskurve per definition ville være ret stejl og således give et alt for negativt billede af aktiveringens effekt.

Det, man kan slutte fra figur 14.1, er, for eksempel, at jobtræning ser ud til at være det mest effektive instrument, mens vejledning ser ud til at være det mindst effektive. Det ses også, at hvis man »blot« lavede en »snapshot«-analyse, for eksempel ved at se på, hvor mange som stadig var i kontanthjælpssystemet et halvt år efter endt aktivering, ville man let komme til at begå fejl, idet kurverne skifter indbyrdes. Den første måned efter endt aktivering ser det således ud til, at vejledning er det mest effektive instrument. Det skyldes naturligvis, at vejledning blandt andet kan bestå af jobsøgningsassistance, som naturligt er mest effektiv, mens personen deltager. Derfor vil en relativt stor andel (dog kun ca. 10%) finde job, *mens de deltager i denne form for aktivering (og altså ikke bagefter)*.²³

Nu kan det jo også tænkes, at det heller ikke er helt tilfældigt, hvem der deltager i hvilken type aktivering. Hvis det for eksempel er således, at de, som deltager i jobtræning, i langt højere grad er arbejdsmarkedsparete end de, som modtager en anden form for aktivering, så vil dette afspejles i kurverne, som derfor ikke kun opfanger aktiveringseffekter, men også er en konsekvens af den involverede sortering af individer efter arbejdsmarkedsparethed.

I så fald er kurverne misvisende, idet forskellen mellem dem således ikke afspejler effekten af aktivering, men også en sortering. Problemet består i, at hvis man for eksempel er interesseret i effekten af vejledning i forhold til jobtræning, så har man brug for at vide, hvordan overlevelseskurven havde set ud for de, som deltager i vejledning, *hvis de havde deltaget i jobtræning i stedet*. I effektmålet ovenfor tilnærmes denne kontrafaktiske situation med overlevelsesfunktionen for de, som *faktisk* deltog i jobtræning, men hvis der er kraftige sorteringseffekter, er dette jo ikke nogen god tilnærmelse.

I næste eksempel illustreres, hvorledes nogle af disse problemstillinger kan håndteres. Dette kræver inddragelse af avanceret statistisk metode,

men jeg vil tillade mig at nøjes med at skitserer metoden og henviser til andre kilder for yderligere information herom.

14.3 **Eksempel 2: Et kompleks mål for aktivering**

Hensigten med dette eksempel er at beskrive en metode, som kan anvendes til at analysere og måle aktiveringsindsatsens effekter på varigheden af for eksempel et kontanthjælpsforløb. Varigheden af dette forløb afhænger af andet end aktiveringsindsatsen. Det gælder både for en række socioøkonomiske faktorer og på forskellige institutionelle forhold. Den anvendte model skal naturligvis kunne tage højde for sådanne faktorer, for kun herved bliver det muligt at isolere den effekt, som alene kan henføres til aktiveringsindsatsen.

Først beskrives de økonometriske modeller og metoder, som det foreslås at anvende – varighedsmodeller. Denne gennemgang indeholder en forholdsvist intuitiv forklaring af varighedsmodellen, suppleret med nogle få matematiske udtryk, som angiver de anvendte modelspecifikationer. Flere detaljer om varighedsmodeller kan findes i Lancaster (1990).

Varighedsmodeller

Varighedsmodeller udgør et vigtigt element i samfundsvidenskabelige analyser af primært ledighed, men kan også anvendes til at modellere varigheden af kontanthjælpsforløb, sygedagpengeperioder, beskæftigelsesforløb, og mange andre forløb. En varighedsmodel formuleres med udgangspunkt i afgangsraten (hazard-funktionen), som populært sagt angiver sandsynligheden for at forlade tilstanden (her: kontanthjælp) på et givet tidspunkt, givet at man ikke har forladt tilstanden før dette tidspunkt. Når der skal inddrages forklarende variabler (både indikatorer for deltagelse i aktivering og baggrundsfaktorer), kan afgangsraten antages at være givet ved en såkaldt »proportional hazard«-specifikation. Denne specifikation antager, at effekten af de forklarende variabler ikke ændres med varigheden af tilstanden.²⁴

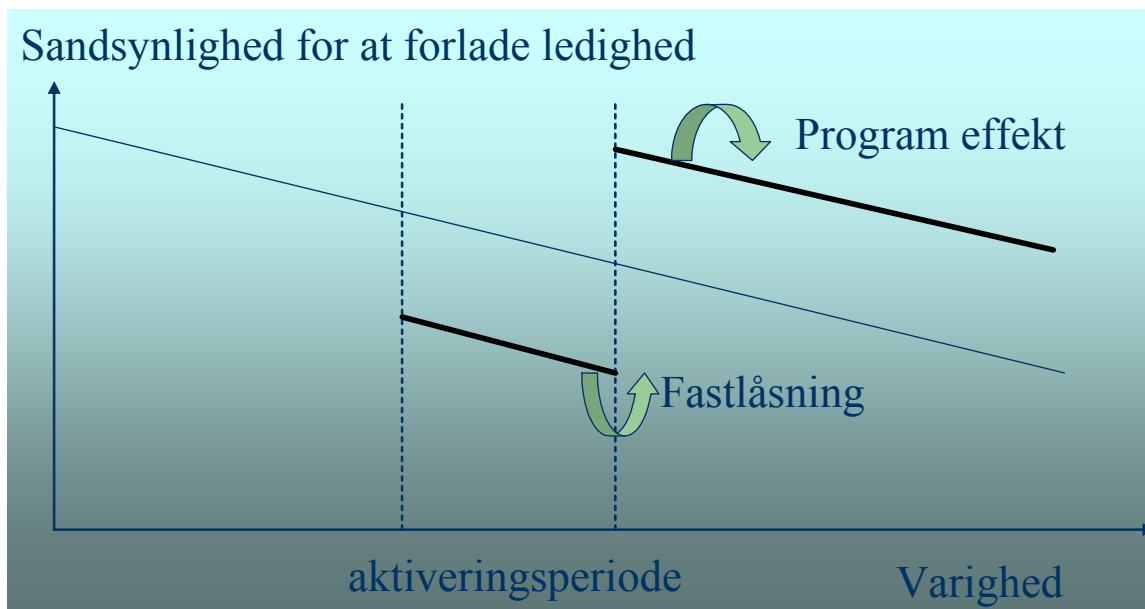
De forskellige aktiveringsindsatser kan finde sted på forskellige tidspunkter i kontanthjælpsperioden. Det er derfor naturligt at modellere aktiveringsindsatserne som tidsvarierende forklarende variabler. I alt inddrages

der dermed to sæt af aktiveringsvariabler: en vektor bestående af tidsvarierende indikatorer for deltagelse i aktivering (altså en markering af, at personen er *i* aktivering), og en vektor bestående af tidsvarierende indikatorer for afsluttet aktivering (personen befinder sig altså *efter* aktiveringen). Det første sæt indikatorer antager altså værdien 1, mens personen er i aktivering, og værdien 0 både før og efter aktiveringen, mens det andet sæt indikatorer antager værdien 1, efter at aktiveringsindsatsen er afsluttet, og værdien 0 før og under selve aktiveringen. Ved at definere aktiveringsvariablerne på denne måde bliver vi i stand til at undersøge såvel fastlåsnings effekter som efterfølgende aktiveringseffekter (sidstnævnte omtales også ofte som selve programeffekten). Fastlåsnings effekten skal forsøge at opfange, hvorvidt en person i aktivering har en lavere sandsynlighed for at finde beskæftigelse, end hvis den samme person ikke havde været i aktivering. Begrundelser herfor kunne være, at deltagelse i aktivering går ud over den tid, man alternativt kunne bruge på at søge job, eller at man ikke ønsker et job, før aktiveringen er afsluttet (hvis der er tale om et opkvalificerende forløb, giver dette naturligvis god mening). Selve programeffekten skal så opfange, at sandsynligheden for at finde job efter afsluttet aktivering gerne skulle være større, end den ville have været uden deltagelse i aktivering. For en illustration af intuitionen bag disse effekter, se figur 14.2.

Ved at beregne parametrene i den således opstillede statistiske model med udgangspunkt i den forhåndenværende forløbshistorik kan man opnå et sæt resultater, som man herefter kan begynde at analysere. Før man gør det, må man imidlertid gøre sig nogle overvejelser om validiteten af sine resultater: Selv om der i modellen er inddraget en lang række forskellige faktorer, kan der stadig være begrænsninger ved analysemetoden, som kan påvirke resultaterne. Det er for det første ikke muligt direkte at danne en kontrolgruppe, da alle – ifølge lovgivningen – før eller siden skal aktiveres. De personer, som alligevel undgår at komme i aktivering, er uden tvivl for atypiske til at kunne bruges som kontrolgruppe. Enten forlader de systemet hurtigt, eller også tilhører de gruppen af meget tunge klienter, som ikke kan placeres i et aktiveringstilbud. Fraværet af en kontrolgruppe bevirker, at der ikke umiddelbart kan måles absolutte effekter af en foranstaltning. Man kan således ikke uden yderligere restriktioner på modellen bestemme

effekten af aktivering i forhold til ikke-aktivering, men må i stedet nøjes med at måle de relative effekter af forskellige aktiveringsindsatser.

Figur 14.2 Illustration af fastlåsnings- og programeffekterne



For at komme dette problem til livs er det altså nødvendigt med yderligere restriktioner på den opstillede model. Her kommer data os til hjælp; det er heldigvis således, at der er stor variation i timingen af en foranstaltning; nogle aktiveres tidligt i denne foranstaltning, andre sent (i deres kontanthjælpsforløb). Nogle aktiveres så tidligt, at der er personer, som endnu ikke er kommet i aktivering at sammenligne med – på dette interval af kontanthjælpsvarigheder er der altså en »kontrolgruppe«. Antag nu at effekten af denne form for aktivering er den samme, uanset hvornår i kontanthjælpsforløbet den iværksættes. Vi kan så – ved at »regne baglæns« – beregne, hvordan afgangsraten for en ikke-aktiveret vil se ud senere i kontanthjælpsforløbet (hvor der egentlig ikke er nogen ikke-aktiverede at sammenligne med). Dette indses ved at se på, hvordan afgangsraten ser ud for den person, som har været aktiveret tidligt, men nu er færdig med sit aktiveringsforløb; da vi kender afgangsraten for vedkommende person (fra data) og kender effekten af aktiveringen (fra vores antagelse), kender vi således også afgangsraten for en ikke-aktiveret person. Herefter er det »en smal sag« at regne aktiveringseffekter ud for alle

øvrige aktiveringstyper. Det er altså »blot« nødvendigt at antage, at én aktiveringsform (som tilbydes kontanthjælpsmodtagere på forskellige tidspunkter) har en konstant effekt.

Selektion

Det er heller ikke uproblematisk at antage (som vi implicit har gjort det), at udvælgelsen (selektionen) til de enkelte indsatsstyper foregår tilfældigt. At jobtræning ser ud til at være et bedre aktiveringsinstrument end produktionsskoler, kan jo skyldes, at jobtræning reelt er det bedste aktiveringsinstrument. Men det kan også skyldes, at deltagerne i jobtræningsforløb generelt har flere ressourcer end deltagerne i produktionsskoleforløb, og derfor har nemmere ved at finde beskæftigelse efterfølgende. I første tilfælde er der tale om en forskel i *effekten* af de to aktiveringsinstrumenter, mens forskellen i det andet tilfælde skyldes den *selektions-skævhed*, at de mest ressourcestærke tildeles jobtræningsforløb, mens de svage sendes på produktionsskoleophold. Hvis produktionsskoledeeltagere kom i jobtræning i stedet, ville de således ikke være bedre stillet.²⁵

Man kan imidlertid ved hjælp af økonometriske metoder indarbejde selve selektionsprocessen ind i forskellige aktiveringsformer i den statistiske model, således at selektionsskævheden elimineres. Selektionsprocessen kan beskrives ved en varighedsmodel med flere afslutningsmuligheder, hvor den afhængige variabel (den variabel, som skal forklares) er tiden i kontanthjælpssystemet, indtil individet aktiveres og afslutningsmulighederne defineres af de mulige aktiveringstyper. Hvis selektionsprocessen således modelleres samtidig med den samlede varighed i kontanthjælpssystemet, som beskrevet tidligere, og hvis disse to varigheder tillades at være korrelerede, vil selektionsskævheden være fjernet. Denne model lader sig imidlertid ikke estimere, idet antallet af mulige aktiveringstyper er for stort. I Bolvig m.fl. (2003) har vi – for det samme datasæt – slået aktiveringstyperne sammen i tre grupper og foretaget denne korrektion. Konklusionerne er i store træk uændrede.

Gruppen af sociale klienter er sammensat af personer med meget forskellige personlige og faglige kompetencer. Aktiveringstilbuddene spænder fra vejlednings- og rådgivningsforløb til indsatser, som sigter mod direkte udslusning til det ordinære arbejdsmarked eller uddannelsessystem.

Man kan således ikke blot vurdere effekten af et aktiveringstilbud på grundlag af, hvor stor en andel af deltagerne som efterfølgende kommer i beskæftigelse eller uddannelse. Det forventes eksempelvis ikke, at deltagerne i et vejledningsforløb finder beskæftigelse eller påbegynder en uddannelse umiddelbart efter vejledningsforløbet, men at de afklares med hensyn til valg af efterfølgende aktiveringstilbud og formålet med dette. Succeskriteriet for et vejledningstilbud bør således baseres på, om deltageren efterfølgende vælger og fuldfører aktiviteter, der trinvist bringer vedkommende tættere på selvforsørgelse. For deltagere i forløb, som umiddelbart sigter mod udslusning til det ordinære arbejdsmarked eller uddannelsessystem, bør succeskriteriet derimod baseres på, hvor stor en andel af deltagerne der kommer i arbejde eller uddannelse bagefter. Det generelle kriterium for at betegne en aktivitet som en succes bliver således, at den ledige deltager i forløb, der i forhold til den forudgående aktivitet bringer vedkommende nærmere en ordinær beskæftigelse eller uddannelse. I dette eksempel vil vi imidlertid kun se på beskæftigelseseffekter.

I analysen indgår 14.534 kontanthjælpsmodtagere, opdelt på 6.993 personer kun med ledighed som problem og 7.541 med problemer ud over ledighed.²⁶

Nogle personer modtager slet ikke aktivering i løbet af deres kontanthjælpsforløb, primært fordi de forlader tilstanden, inden aktivering bliver aktuel. Denne gruppe udgør ca. 35%. Andre kontanthjælpsmodtagere deltager på den anden side i adskillige aktiveringsindsatser i løbet af et enkelt kontanthjælpsforløb. De forskellige typer aktivering er beskrevet i tabel 14.1.

Det ses af tabellen, at de mest anvendte typer af aktivering blandt kontanthjælpsmodtagerne er jobtræning, kurser og sprogskole. Sprogskole må i den sammenhæng antages at være et specialtilfælde, idet kun en bestemt gruppe af personer deltager i denne type aktivering, og det kan endvidere antages, at deltagelse i sprogskole for mange af disse klienters vedkommende er en forudsætning for en eventuel deltagelse i andre typer af aktivering. Jobtræning er så vidt muligt opdelt i privat og offentlig jobtræning, men der er en større gruppe, for hvem det ikke fremgår af registrene, hvorvidt de er i offentlig eller privat jobtræning. En anden type aktivering, der udgør en stor andel af de samlede aktiveringsforløb, er arbejdsafprøvning.

Dette er en erhvervmæssig aktivering, der er placeret under forrevalideringsordningen. Der findes en del forskellige kursustyper, der her i denne analyse er inddelt i tre hovedgrupper som daghøjskole, Arbejdsmarkedsuddannelserne (AMU) og anden kursusudbyder. På uddannelsesfronten findes desuden aktiveringstyper som særlig tilrettelagt uddannelse, ordinær ungdomsuddannelse og forberedende skolegang. De uddannelsesmæssige aktiveringsforløb udgør ca. 50% af al aktivering, mens de erhvervmæssige aktiveringsforløb kun udgør knap 25%, og de sidste 25% består hovedsagelig af vejledning og anden aktivering.

Tabel 14.1 Aktiveringsforløbenes fordeling på aktiveringstyper for kontanthjælpsmodtagere

Aktiveringstype	Antal aktiveringsforløb	Fordeling (procent)	Gns. varighed*
Jobtræning	692	5,4	5,3
Jobtræning, privat	241	1,9	4,3
Jobtræning, off.	436	3,4	5,7
Kommunale job	198	1,5	4,8
Arbejdsafprøvning	1.060	8,2	7,9
Igangsætningsydelse	118	0,9	12,7
Beskæftigelsesprojekt	351	2,7	4,8
Produktionsskole	28	0,2	4,6
Daghøjskole	1.221	9,4	5,7
AMU	114	0,9	3,3
Anden kursusudbyder	1.293	10,0	6,2
Særlig tilrettelagt udd.	132	1,0	4,5
Ordinær ungdomsudd.	476	3,7	15,0
Forbered. skolegang	856	6,6	7,7
Sprogskole	2.216	17,1	9,6
Gruppeforanstaltning	692	5,4	7,1
Vejledning og intro	1.297	10,0	5,0
Anden aktivering	1.518	11,7	7,2
I alt	12.939	100,0	7,2

* Målt i måneder. En lille del aktiveringsindsatserne var uafsluttede ved observationsperiodens afslutning, hvilket gør de observerede gennemsnitlige varigheder kortere end de reelle. De gennemsnitlige varigheder er udregnet ud fra samtlige aktiveringstyper, både afsluttede og uafsluttede.

Resultater

Vi vil nu vende os mod effekterne af de forskellige typer aktivering. Dis-

se bliver blot diskuteret for at illustrere, hvorledes modellens resultater kan omsættes til noget, som kan anvendes i praksis. Resultaterne findes i (Bolvig m.fl., 2001).

For alle aktiveringstyper, med undtagelse af vejledning og produktions-skole, gælder det, at der er kraftige fastholdelseeffekter. Med kraftig menes der, at afgangsraten reduceres til mellem det halve og en tiendedel, mens personen er under aktivering.

For personer med kun ledighed som problem ses det, at *vejledning og introduktionsforløb* har en fastlåsningseffekt, mens der kun er en efterfølgende positiv programeffekt for denne aktiveringstype, hvis man er over 30 år. For personer med problemer ud over ledighed ses det derimod, at vejledning har positive effekter under selve aktiveringen og også efterfølgende.

Jobtræning har generelt positive programeffekter. Opdelingen på privat og offentlig jobtræning fortæller mere, end bare hvor jobtræningen foregår, eftersom typen af jobtræning varierer for de to stedtyper. Generelt vil offentlig jobtræning bestå af individuel jobtræning, hvor den aktiverede modtager kontanthjælp samt et beskæftigelsestillæg under jobtræningsforløbet; privat jobtræning derimod består hovedsagelig af almindelig jobtræning, hvor den aktiverede modtager almindelig overenskomstløn, og arbejdsgiveren får til gengæld et løntilskud fra kommunen. For personer med problemer ud over ledighed findes der ingen signifikante effekter af jobtræning, hvilket hovedsageligt skyldes lav deltagelse i denne aktiveringsform for disse personer. For personer over 30 år med ledighed som eneste problem er privat jobtræning betydeligt mere effektivt end for personer under 25 år, hvorimod offentlig jobtræning er en anelse mere effektiv for personer under 25 år.

Mens der modtages *igangsætningsydelse*, er der en stærk fastlåsningseffekt, hvilket blot betyder, at personer, som modtager denne ydelse, ikke forlader status som selvstændig beskæftiget i løbet af den periode, hvor de modtager ydelsen. Dette kan dog ses som en positiv effekt, idet en fastholdelse i et system med igangsætningsydelse er et udtryk for, at kontanthjælpsmodtageren, i hvert tilfælde delvist, kan erhverve sig ved selvstændig virksomhed. Der er endvidere en positiv programeffekt, hvilket kan skyldes, at disse personer enten er i stand til at forblive selvstændige

eller at finde anden beskæftigelse.

Beskæftigelsesprojekter har med enkelte undtagelser ingen gunstige effekter, hverken under eller efter. Under aktivering i et beskæftigelsesprojekt er der en kraftig fastlåsningseffekt, mens programeffekten efter deltagelse er 0, bortset fra for kvinder, for hvem der findes en positiv efterfølgende effekt.

Gruppeforanstaltninger, der er et beskæftigelsesprojekt, har ligeledes kun negative effekter, med en kraftig fastlåsning og en signifikant negativ programeffekt.

Aktivering på *produktionsskoler* for personer under 25 år kun med ledighed som problem har en fastlåsningseffekt, som dog ikke findes at være signifikant, mens der er en kraftig signifikant negativ programeffekt. Der er således tale om, at ophold på en produktionsskole medfører en væsentlig forringelse af de fremtidige muligheder for at undslippe kontanthjælp. Det er dog uklart, hvorvidt dette skyldes »stigma«-effekter, demotivationseffekter eller selektionsskævhed (se diskussionen ovenfor).

For *daghøjskoler* findes en forholdsvis kraftig fastlåsningseffekt, men ingen programeffekter. Det skal i den forbindelse nævnes, at aktivering via daghøjskoler oftest iværksættes som et første trin på aktiveringstrappen, således at et kursus på en daghøjskole kan være en betingelse for videre aktivering. Beskæftigelse er derfor ikke et egentlig succeskriterium for aktivering på daghøjskole.

AMU-kurser og andre *uddannelsesaktiveringer* har en kraftig fastlåsningseffekt og ingen signifikante programeffekter.

Ordinær ungdomsuddannelse har de stærkeste effekter for kvinder og svagere effekter for mænd. Der findes en fastlåsningseffekt (hvilket ikke er overraskende) og en efterfølgende positiv programeffekt. Der er ingen effekter for personer med kun ledighed som problem, idet aktivering i ordinære ungdomsuddannelser er en del af forrevalideringsordningen og dermed kun udbydes til personer med problemer ud over ledighed. Programeffekten er til gengæld overordentligt kraftig for personer med problemer ud over ledighed, hvor afgangsraten bliver forøget med en faktor 12.

Forberedende skolegang har generelt en forholdsvis kraftig fastlåsningseffekt samt en negativ eller insignifikant programeffekt. For personer over 30 år kun med ledighed som problem er der dog en forholdsvis kraftig

positiv efterfølgende effekt.

Kommunale job har en kraftig fastholdelse og en beskedent positiv efterfølgende effekt, som dog er insignifikant for personer over 30 år.

Sprogskole har en fastlåsningseffekt og en programeffekt, som enten er insignifikant eller negativ. Dette resultat er vanskeligt at fortolke nærmere, idet der er en stor selektionsskævhed for netop denne gruppe, og det er sandsynligt, at netop denne gruppe ikke ville kunne aktiveres gennem nogen af de øvrige indsatser, førend sprogskolen er gennemført.

Arbejdsafprøvning viser moderate fastlåsningseffekter for alle grupper, men kun positive efterfølgende effekter for kvinder og for personer med problemer ud over ledighed.

Resultaterne afslører en graduering af effekterne, således at de bedste effekter findes af foranstaltninger, som har beskæftigelsesindhold, samt af ordinære uddannelser. Andre uddannelsesforanstaltninger og produktions- og daghøjskoler (muligvis med karakter af passive anbringelsesforanstaltninger) har generelt negative effekter. Der er ikke her foretaget nogen beregninger af, hvorvidt de samlede effekter er positive eller negative, idet der skal tages hensyn til, at de efterfølgende positive effekter i større eller mindre omfang opvejer fastholdelseeffekterne.

14.4 **Formidling og anvendelse**

Resultaterne fra ovenstående evaluering er blevet publiceret i et arbejdspapir fra Center for arbejdsmarked og Social forskning (CLS), ligesom det er under revision for et internationalt tidsskrift. Endvidere er resultaterne fra evalueringen blevet præsenteret i en følgegruppe bestående af kommunale embedsfolk fra Århus samt repræsentanter for ministerier samt KL. Endelig blev der i forbindelse med rapportens frigivelse afholdt et presse- og orienteringsmøde i Århus.

En af de oprindelige tanker med projektet var at undersøge, hvorvidt simple effektmål viser det samme som mere avancerede mål, og eftersom dette blev besvaret bekræftende, anvendes de simple effektmål i Århus Kommune til at justere aktiveringsindsatsen, så den er så effektiv som muligt.

15 Evaluering på arbejds miljøområdet

Leena Eskelinen

15.1 Indledning

Særlige kendetegn for arbejds miljøområdet

Arbejds miljø er et bredt begreb, og det er svært at afgrænse og definere området præcist, bl.a. fordi arbejdets karakter ændrer sig i samfundet. Klassiske arbejds miljøproblemer ændrer karakter, og der opstår nye. Den generelle tendens er, at arbejdslivets problemer bliver mere komplekse – mindre velafgrænsede – og at tyngdepunktet flytter fra det fysiske arbejds miljø til det psykosociale miljø. Arbejds miljøområdet omfatter i dag mange specialområder, og specialområderne har forskelligt genstandsfelt og forskellige metodiske tilgange. Fagligt set har arbejds miljøprofessionelle forskellig uddannelsesmæssig baggrund spændende fra naturvidenskab til psykologi.

Karakteristisk for arbejds miljøområdet synes endvidere at være, at der ikke hersker en skarp skillelinje mellem forskning og udvikling. Forsknings problemstillinger udspringer typisk fra konkrete arbejds miljøproblemer, og derfor er resultaterne af evaluerende karakter og nogle gange handlingsanvisende – handlingsanvisende både i forhold til konkrete arbejdspladser og med henblik på eventuelle myndighedstiltag (regulering af arbejds miljøet gennem lovgivningen på baggrund af nye forskningsresultater).

Arbejds miljø er imidlertid et felt, hvor systematisk evalueringsforskning ikke er udbredt hverken i Danmark eller i andre lande. Feltet er meget spredt, og udviklingen sker i forskellig takt og på forskellig måde afhængig af delområderne og disciplinerne inden for arbejds miljøforskningen (samt

disses forskningstraditioner). Den klassiske arbejdsmiljøforskning interesserer sig i høj grad for årsags-virknings-sammenhænge, dvs. hvordan forskellige arbejdsmiljøfaktorer påvirker menneskets helbred. Den bygger på naturvidenskabelig tradition med sammenligninger mellem undersøgelses- og kontrolgrupper og med dertil udviklede undersøgelsesdesign og – metodik (epidemiologi, bl.a. Hernberg 1998, Øvretveit 1999)²⁷. I en oversigt over arbejdsmiljøforskningen i Danmark er denne type forskning blevet betegnet som dosis-respons-forskning (Hasle m.fl. 1999). Skal denne type forskning karakteriseres med evalueringsterminologi, er der tale om effektevaluering (jf. Togeby, kapitel xx i denne håndbog).

Dosis-respons forskningen – eller enkeltfaktorforskning – har været dominerende igennem mange år. Men i de senere år er der sket en udvikling på dette felt, så forskningen nu omhandler bredere sammenhænge mellem eksponeringer og helbredskonsekvenser (bl.a. Rantanen 1999). En anden udviklingstendens er, i takt med samfundsudviklingen, udbredelsen af forskningen, som fokuserer på psykosocialt arbejdsmiljø og organisatoriske forhold. På disse områder er problemerne ikke velafgrænsede, og virkningerne handler ikke alene om individets reaktioner, men også organisationernes måde at fungere på. Der lægges i stigende grad vægt på udvikling af positive sider af arbejdsmiljøet, herunder samspillet mellem arbejdsvilkår og den enkelte medarbejders og arbejdsgruppers engagement, velvære og produktivitet. Repræsentanter for disse nye tilgange har påpeget vigtigheden af at udvikle evalueringsmetodik, som har andre paradigmer end den naturvidenskabelige (Griffith 1999). De efterlyser metodik både til at forske i (arbejdsmiljø)indsatsernes virkninger og processer i relation til den kontekst, hvor en intervention finder sted.

Et kendetegn for »the state of the art« af evalueringen på arbejdsmiljøområdet er, at der hverken findes oversigter over evaluering på arbejdsmiljøområdet, metodisk litteratur specifikt for området eller retningslinjer eller vejledninger til at gennemføre evalueringer.

Det vil sige, at diskussionen om evaluering i arbejdsmiljølitteraturen nærmest er ikke-eksisterende²⁸, selv om der i praksis hele tiden gennemføres evalueringer på feltet. Litteratursøgning på danske og nordiske databaser giver flere hundrede referencer til publikationer, som rapporterer resultater af forskellige indsatser på arbejdspladser.

Eksempler på tilgange og aktører i Danmark

Beskrivelsen her har til hensigt at illustrere de forskellige sammenhænge, hvor der kan tales om evaluering af arbejdsmiljøforhold eller arbejdsmiljøindsatser. Fremstillingen bygger ikke på en systematisk analyse eller opsamling på feltet. I kortlægninger af arbejdsmiljøforskningen i Danmark (Arbejdsmiljøinstituttet 1999, Hasle m.fl. 1999, Arbejdsministeriet 2001) bliver evaluering ikke nævnt som et særligt område bortset fra, at der i Arbejdsministeriets udvalgsrapport (2001) er et kort afsnit om »overvågning og effektvurdering«. Den karakteriseres som »yderst begrænset«, men påkrævet for at kunne prioritere og planlægge forebyggende indsatser.

En tilgang til at beskrive arbejdsmiljøets tilstand og arbejdsforhold i Danmark er undersøgelser, som retter sig mod befolkningens levekår, sundhedstilstand og arbejdsmiljø og bygger på større spørgeskemaundersøgelser. Her skal nævnes Socialforskningsinstituttets levekårsundersøgelse, som viser udviklingen gennem tre undersøgelsesrunder i perioden 1976-2000 (Andersen 2003), analyser af sundhed og sygelighed samt udviklingen deri i perioden 1987-2000 foretaget af Statens Institut for Folkesundhed (Kjøller & Rasmussen 2002) samt Arbejdsmiljøinstituttets lønmodtagerundersøgelse, hvor arbejdsmiljø og helbred i Danmark er blevet kortlagt hvert femte år siden 1990 (Burr m.fl. 2002).

En anden type monitorering foretages af Arbejdstilsynet dels gennem registre (anmeldte arbejdsulykker og arbejdsrelaterede lidelser, jf. Arbejdstilsynets årsstatistik), dels gennem fokuserede undersøgelser. Et eksempel på sidstnævnte tiltag er den igangværende overvågning af virksomhedernes interne forebyggende aktiviteter på syv visionsområder (Arbejdstilsynet 2003) i forbindelse med regeringens handlingsprogram for et rent arbejdsmiljø 2005. Overvågningen består bl.a. af en virksomhedsundersøgelse, der gentages tre gange i handlingsprogrammets tiårige periode.

Egentlige evalueringer på arbejdsmiljøområdet har bl.a. handlet om virkninger af handlingsprogrammer, fx evaluering af handlingsprogrammet for ensidigt, gentaget arbejde, som havde til formål at halvere omfanget af sundhedsskadeligt EGA over en syvårig periode i 1993-2000 (Hasle m.fl. 2001) og virkninger af bestemte indsatser, fx evaluering af Arbejdstilsynets pulje 'Forebyggelse af bevægeapparatskader ved tunge personløft', hvorfra

ca. 100 lokale udviklingsprojekter på social- og sundhedsområdet har fået støtte (Frederiksen og Hansen 2004.). Desuden er der gennemført evalueringer af arbejdsmiljøapparatets funktion, fx evaluering af BST-forsøg (Teknologisk Institut 2001) og evaluering af APV-ordningen. Dertil kommer de mange evalueringer, udredninger og erfaringsopsamlinger, som bl.a. sektorforskningsinstitutter, bedriftssundhedstjenester (BST'er) og konsulentvirksomheder gennemfører på arbejdspladserne i forbindelse med udviklingsprojekter og interventioner og de evalueringer, som arbejdspladserne selv foretager.

Artiklens eksempler

Eksemplerne i denne artikel repræsenterer bredden af de fremgangsmåder, som evaluering af arbejdsmiljøindsatsen omfatter i Danmark. Det første eksempel handler om monitorering (overvågning) af arbejdsmiljøindsatser på udvalgte områder på landsplan (Arbejdstilsynet 2002). Det andet eksempel beskriver evaluering af en konkret intervention på social- og sundhedsområdet, hvor formålet har været at forebygge personalets belastning på plejehjem i forbindelse med tunge personløft (BST Københavns Kommune 2003)²⁹.

15.2 Eksempel 1: Udviklingstendenser i arbejdsmiljøet (på syv visionsområder)

Det første eksempel repræsenterer »den tredje evalueringstype«, som er almindelig i dansk evalueringspraksis, nemlig monitorering (Albæk og Rieper 2001). Monitorering defineres som »regelmæssige evalueringer af offentlige indsatsers opgavevaretagelse og output«. Disse evalueringer foretages løbende eller med jævne mellemrum som led i en overvågning af, hvordan indsatser på et givet område udvikler sig, og hvilke resultater der opnås. De har til formål at tilpasse og justere offentlige organisationer og programmer inden for et givet budget og inden for rammerne af foreliggende målsætninger.

Overordnet formål, genstand og kontekst

Arbejdstilsynet følger udviklingen på arbejdsmiljøområdet som led i regeringens handlingsprogram for et rent arbejdsmiljø 2005. Handlingsprogrammet omfatter syv delområder, som i handlingsprogrammet betegnes som visionsområder, nemlig støj, tunge løft, ensidigt gentaget arbejde, arbejdsulykker, psykisk arbejdsmiljø, kemi, indeklime og børn og unges arbejdsmiljø. Der er tale om områder, som Arbejdstilsynet har prioriteret ud fra den betragtning, at der på disse områder er behov for en særlig indsats for at forbedre arbejdsmiljøforholdene. Samtidig repræsenterer de områder, som hovedparten af arbejdsmiljøforskningen i Danmark har handlet om (Arbejdstilsynet 1999).

Handlingsprogrammet retter sig således mod bestemte temaer, og hvert tema er relevant inden for flere forskellige brancher. Dermed er et stort antal danske arbejdspladser principielt dækket af handlingsprogrammet.

Indsatsen (interventionen) er forskellig på de syv visionsområder, og den består typisk af mange forskellige aktiviteter. Aktiviteterne er ikke defineret i handlingsprogrammet; programmet er en principerklæring, dvs. det påpeger ikke bestemte indsatser og tilbyder heller ikke økonomisk støtte til aktiviteterne. I henhold til strukturen af arbejdsmiljøarbejdet i Danmark foretages en stor del af indsatserne i branchearbejdstilsynenes regi. Ud over branchearbejdstilsynene har Arbejdstilsynet haft indsatser, kampagner mv. bl.a. med fokus på ensidigt, gentaget arbejde, arbejdsulykker, tunge løft og psykisk arbejdsmiljø. Fx gennemførte Arbejdstilsynet en landsdækkende indsats »Psykisk arbejdsmiljø – alles ansvar« i 2000-2002 i tre brancher og indførte i den forbindelse en særlig tilsynsindsats.

Evaluerings spørgsmål og teori

Som sekretariat for Arbejdstilsynets overvågningsudvalg har Arbejdstilsynet som opgave at følge udviklingen i henhold til handlingsprogrammet.

Hovedspørgsmålet i overvågningen er, om der sker en udvikling på visionsområder inden for handlingsprogrammets tiårige periode. Tanken med at følge situationen undervejs er at kunne sætte ind, hvis der er indikation på udvikling i utilsigtet retning. Derudover har overvågningen et me-

to disk sigte: at udvikle redskaber, som arbejdsmiljøsystemet kan anvende til at følge det til enhver tid aktuelle stade af arbejdsmiljøarbejdet.

Model og design

Arbejdstilsynet overvåger forholdene på de syv visionsområder gennem et design, som består af tre forskellige hovedkomponenter (såkaldte søjler), som hver omfatter forskellige datakilder. Ideen hermed er at kombinere flere typer indikatorer. En søjle er baseret på oplysninger om påvirkninger og helbredsforhold i arbejdsmiljøet og bygger på registeroplysninger på den ene side og kortlægninger udført af sektorforskningsinstitutter (Arbejdstiljøinstituttet, Socialforskningsinstituttet, Statens Institut for Folkesundhed) på den anden. En anden søjle drejer sig om oplysninger fra Arbejdstilsynets interview- og observationsundersøgelse om virksomhedernes interne forebyggende arbejdsmiljøaktiviteter. En tredje søjle drejer sig om de øvrige arbejdsmiljøprofessionelles aktiviteter, herunder primært branchearbejdstiljørådene. Oplysningerne i de tre søjler kan tilsammen tegne et billede af status og udvikling i handlingsprogrammet.

Dataindsamling og analysemetoder

Som beskrevet stammer data dels fra andre kilder, dels fra særlig dataindsamling til overvågningsformål.

Oplysninger om påvirkninger og helbredsforhold (»søjle 1«) bygger bl.a. på Arbejdstilsynets registre over anmeldte arbejdsulykker og arbejdsrelaterede lidelser samt på Arbejdstiljøinstituttets lønmodtagerundersøgelse.

Den første virksomhedsundersøgelse (»søjle 2«) blev gennemført i 2001 og indeholdt dataindsamling ved telefoninterview med ca. 3.200 virksomheder, og ved arbejdspladsbesøg på ca. 10% af de virksomheder, der blev telefoninterviewet. Denne første virksomhedsundersøgelse var planlagt som basis (referencemåling) af virksomhedernes forebyggende aktiviteter, som de efterfølgende virksomhedsundersøgelser kunne blive sammenlignet med. Der blev udvalgt ca. 10 branchegrupper for hvert visionsområde til telefoninterview og arbejdspladsbesøg – branchegrupper, hvor det undersøgte arbejdsmiljøproblem var mest fremtrædende.

Arbejdsmiljøprofessionelles aktiviteter/indsatser (»søjle 3«) for at opnå de i handlingsprogrammet opstillede mål er beskrevet på basis af udtræk fra Arbejdstilsynets interne informationssystem. Branchearbejdsmiljørådernes aktiviteter i form af gennemførelse af projekter rettet mod et eller flere visionsområder er beskrevet på basis af oplysninger fra Arbejdstilsynets administration af puljen.

De væsentligste datakilder i overvågningen er helbredsoplysninger om ansatte og arbejdsforhold i brancherne (»søjle 1«) samt oplysninger om forebyggende arbejdsmiljøarbejde på arbejdspladserne (»søjle 2«). Det vil sige, at oplysningerne om indsatserne (»søjle 3«) er mindre dækkende og mindre systematiske (uddrag af et journalsystem).

Hovedresultater

I Arbejdstilsynets overvågningsrapport fra 2002 beskrives udviklingstendenser inden for hvert af de syv visionsområder på baggrund af de tre typer datakilder. Data om påvirkninger og helbredsforhold samt data om virksomhedernes interne forebyggende aktiviteter ses bl.a. i forhold til branchegruppe og køn. Derudover gives der vurderinger af betydningen af evt. indsatser for udviklingstendenserne inden for hvert visionsområde. Endvidere indeholder rapporten metodiske overvejelser, herunder overvejelser om indikatorernes brugbarhed i overvågningsøjemed.

Formidling og anvendelse

Arbejdstilsynet udgiver årligt en overvågningsrapport, som er offentligt tilgængeligt (bl.a. Arbejdstilsynet 2002). Rapporten er først og fremmest skrevet til arbejdsmiljøapparatet, i første omgang til Arbejdsmiljørådets overvågningsudvalg, som følger udviklingen inden for handlingsprogrammets områder. 2002-rapporten er en oversigt over arbejdsmiljøets tilstand i Danmark og indeholder en omfattende dokumentation samt metodiske overvejelser, der kan være af interesse for arbejdsmiljøaktører – såvel praktikere som professionelle.

I henhold til ønsket om at styrke overvågning og effektvurdering (Arbejdsministeriet 2001) bidrager projektet til metoder og modeller for arbejdsmiljøovervågning i fremtiden.

15.3 **Eksempel 2: Kvalitetsudvikling i forflytning**

Dette eksempel beskriver evaluering af et lokalt udviklingsprojekt, som er gennemført med henblik på at forebygge bevægeapparatskader. Projektet adskiller sig fra arbejdspladsinterne udviklingsprojekter på to måder. Det er gennemført som led i et større program, hvor der bl.a. er stillet krav om dokumentation af det gennemførte projekt, og projektet er blevet ledet og evalueret af eksterne konsulenter (BST). De metoder og indikatorer, som er anvendt i evalueringen, er imidlertid typiske for evalueringer af en arbejdsmiljøindsats på arbejdspladserne. Det er også ofte, at evalueringen ikke er planlagt i starten – som det er tilfældet i eksemplet – og at den ofte bygger på oplysninger, som er indsamlet, efter en intervention har fundet sted.

Overordnet formål, genstand og kontekst

BST Københavns Kommune gennemførte i 1998 projektet »Det ryg'er noget – kvalitetsudvikling i personforflytning« på fire plejehjem. Projektet var blandt de næsten 100 lokale udviklingsprojekter på social- og sundhedsområdet, der fik støtte fra puljen »Forebyggelse af bevægeapparatskader ved tunge personløft« (Frederiksen og Hansen 2004).

Formålet med udviklingsprojektet var at mindske nedslidningen i plejearbejdet gennem nye arbejdsteknikker og sikre, at de nye arbejdsmetoder bliver ført ud i livet og udviklet på plejehjemmene. Sagt med andre ord sigtede udviklingsprojektet mod at sætte arbejdsmiljø – medarbejdernes arbejdsforhold – på dagsorden på arbejdspladserne samt at indføre nye faglige redskaber bl.a. gennem uddannelse.

Baggrunden for projektet var, at BST Københavns Kommune i 1996 havde gennemført en spørgeskemaundersøgelse på samtlige 74 plejehjem i kommunen og konstateret omfattende arbejdsmiljøproblemer. I den forbindelse blev det bl.a. konstateret, 1) at fornyelse af de faglige værktøjer primært havde drejet sig om udvikling af hjælpemidler og korrekt løfte-teknik, typisk fra eksperter (fysio- og ergoterapeuter) inden for området, og 2) at arbejdsmiljøkrav og -hensyn har været underlagt hensynet til plejens udførelse og beboernes behov, og at det har været svært at skabe læring, udvikling og forankring af nye faglige værktøjer.

Udviklingsprojektet gik ud på at uddanne ressourcepersoner (faglige støttepersoner) blandt personalet for hver plejehjemsafdeling og udvikle viden om forflytninger og arbejdsmiljø lokalt på afdelingerne bl.a. gennem disse ressourcepersoner og generel uddannelse af hele personalet. Der blev arbejdet med fælles vejledninger om forflytninger (forflytningspolitik), samt medarbejdernes inddragelse i udviklingen af det daglige arbejde. Endvidere byggede uddannelserne på en særlig pædagogisk tilgang som tog afstand fra en tidligere ekspertorienteret undervisning.

Udviklingsprojektet blev gennemført på fire plejehjem efter det samme koncept. Hvert af de fire plejehjem havde ca. 60-70 ansatte, og projektet omfattede hele personalet.

Evalueringsspørgsmål og teori

Evalueringen er gennemført af BST Københavns Kommune (Zwicky-Hauschild 2003). BST's interesse for at foretage en »ekstern« evaluering³⁰ af projektet var at dokumentere »værdien af de udvalgte metoder«, dvs. at undersøge, om det ved indførelse af de nye tiltag er lykkedes at forbedre arbejdsmiljøet.

Evalueringen fokuserede på følgende spørgsmål:

1. Kan det fysiske arbejdsmiljø bedres?
2. Kan der skabes kvalitetsudvikling og forankring af viden på arbejdspladsen?

Det vil sige, at ambitionen i evalueringen var at dokumentere eventuelle forbedringer i arbejdsmiljøet samt beskrive eventuel udvikling i »videnniveauet« blandt medarbejderne.

Disse evalueringsspørgsmål udfordrer til en kombination af proces- og effektevaluering. Ud fra den betragtning, at BST igangsatte udviklingsprojektet og fulgte det samt foretog den afsluttende evaluering, havde evaluatoren mulighed for at kombinere de to tilgange i den konkrete evaluering.

Model og design

Ændringer i det fysiske arbejdsmiljø (1. evalueringsspørgsmål) er søgt be-

lyst ved ændring i besvær i bevægeapparatet (lænderygbesvær), sygefravær, arbejdsskader og personaleomsætning i observationsperioden. Kvalitetsudvikling og forankring (2. evalueringsspørgsmål) er belyst ved kendskab til og overholdelse af aftaler og politikker på arbejdspladsen, samt påvirkningen af det generelle arbejdsmiljø.

Det vil sige, at udvikling i projektperioden er forsøgt belyst ved hjælp af et før-efter-design. Hvad angår det 1. evalueringsspørgsmål, bygger målingen af ændringer i arbejdsmiljøet på medarbejdernes reaktioner på ændringer (ud fra en antagelse, at et bedre arbejdsmiljø medfører færre symptomer, sygedage og arbejdsskader). Der er ikke blevet spurgt til oplevelse af arbejdsforhold i forbindelse med den spørgeskemabaserede kortlægning af symptomer.

Dataindsamling og analysemetoder

Evalueringen bygger på forskelligt datamateriale, dels data indsamlet i selve projektet (ved kvalitetsudviklingsprojektets afslutning i 1998), dels data, som evaluatoren har indsamlet i 2001 gennem en spørgeskemaundersøgelse blandt medarbejdere, samt fokusgruppeinterview med resourcepersoner og afdelingsledelse samt oplysninger, som evaluatoren har indhentet fra registre:

- sygefravær, arbejdsskader og personaleomsætning i perioden 1997-2001 (oplysninger om sygefravær og personaleomsætning stammer fra sundhedsforvaltningens økonomikontor)
- spørgeskemaundersøgelse vedr. besvær i bevægeapparatet 1998 (blev gennemført ved afslutningen af udviklingsprojektet)
- spørgeskemaundersøgelse vedr. besvær i bevægeapparatet, resourcepersonordning, forflytningsteknik, proces og forankring i 2001
- fokusgruppeinterview med resourcepersoner og den øverste ledelse i 2001.

I analyserne sammenlignes sygefravær, arbejdsskader og personaleomsætning i den fireårige periode samt omfanget af besvær i bevægeapparatet på de to tidspunkter (1998 og 2001). I forhold til begge sammenligninger rapporterer evaluatoren dataproblemer, hvilket ikke er atypisk i arbejdsmiljøevalueringer. For det første er de statistiske oplysninger ikke

præcise, fx er sygefravær registreret samlet, dvs. ikke fx i forhold til sygefravær i forbindelse med arbejdsskader. For det andet er deltagelsen i spørgeskemaundersøgelsen meget lav – svarprocenten i 2001 er 50.

Hovedresultater

Ifølge evalueringen viser lænderygbesvær en faldende tendens efter projektets ophør, mens omfanget af sygefravær (samlet set) og personaleomsætning ikke viser »nogen sammenhæng med projektet«.

I forhold til det andet evalueringsspørgsmål konkluderer evaluatoren: »der er en tendens pegende mod, at forankring og kvalitetsudvikling hænger sammen med kendskab til og overholdelse af politikker og aftaler på arbejdspladsen og en tendens pegende mod en sammenhæng mellem forankringen og det generelle løft af arbejdsmiljøbevidstheden på den enkelte institution«.

Formidling, anvendelse

Evalueringsrapport er udgivet på nettet, og den er derudover udkommet i oplag på 300 eksemplarer.

(Zwicky-Hauschild 2003, www.bst.kk.dk/nyheder/nyhedsarkiv/)

I evalueringsrapporten konstateres: »Nærværende evaluering har bidraget til indholdet i ressourcepersonuddannelsen og været med til at sikre, at de forhold, der af deltagerne blev vurderet som vigtige, er bevaret i den nuværende uddannelse. Samtidig er det håbet, at rapporten kan tjene som inspiration for plejehjem, der arbejder systematisk med forflytning og arbejdsmiljø og overvejer deltagelse i ressourcepersonuddannelsen.« Det vil sige, at rapportens brugere er »feltet«, praktikere med plejeopgaver på social- og sundhedsområdet og BST. På baggrund af projektets koncept og erfaringer har BST udviklet »ressourcepersonuddannelsen«, som tilbydes til interesserede i pleje- og omsorgssektoren. Denne uddannelse har siden 1999 (udviklingsprojektet blev gennemført i 1998) været en del af BST Københavns Kommunes tilbud til plejehjemmene inden for forflytningsteknik og kvalitetsudvikling i plejearbejdet.

15.4 Afrunding

Generelt er evaluering ikke et særligt veludviklet felt inden for arbejdsmiljøforskningen, heller ikke i lande som Finland og Sverige, som har langt stærkere traditioner på arbejdsmiljøområdet end Danmark. En stor del af de foretagne evalueringer er casestudier, som litteratursøgninger fx i nordiske databaser tydeligt viser. Derimod mangler der i høj grad forskning og forskningsbaserede oversigter, som sætter fokus på evalueringsmodeller og metoder inden for forskellige dele af arbejdsmiljøforskningen.³¹ Dette gælder især for de »unge« områder inden for arbejdsmiljøforskningen, fx det psykosociale område.

I Danmark har arbejdsmiljøforskningen været et meget spredt felt, hvad angår emner, aktører og finansiering, som oversigterne over forskningens tilstand viser (bl.a. Arbejdsministeriet 2001). Evalueringen har fulgt to hovedlinjer: på den ene side overvågning på overordnet plan og på den anden side lille-skala-evalueringer lokalt på arbejdspladserne. Den førstnævnte har overvejende kontrolsigte, mens den sidstnævnte sigter mod læring og udvikling (»erfaringsopsamling«). Det overordnede formål med begge er at styrke og forbedre arbejdsmiljøarbejdet, men der vides meget lidt om, hvilken betydning de reelt har (haft).

Feltet indebærer imidlertid store udfordringer. I den nyeste rapport fra Socialforskningsinstituttets levkårsundersøgelse konkluderes det, at »arbejdsmiljøet har udviklet sig mere ugunstigt end noget andet enkeltområde« siden midten af 1970'erne (Andersen 2003). I Center for Alternativ Samfundsanalyse (CASA)'s beskrivelse af arbejdsmiljøindsatsen i perioden 1975-2000 konstateres det, at »til trods for 25 års udvikling af arbejdsmiljøsystemet er forskning i systemets effekt og virkningsmekanismer meget begrænset og spredt; der udvikles derfor nye strategier og metoder uden et systematisk videngrundlag« (Hasle m.fl. 2000). På denne baggrund tyder det på, at det er vigtigt at styrke systematisk evaluering af arbejdsmiljøområdet.

16 Evaluering inden for energisektoren

Anders Larsen og Lene Holm Pedersen

16.1 Indledning

Energipolitikken opstod som selvstændigt policy-område i forlængelse af den første oliekrise i 1973. På dette tidspunkt var den bærende målsætning at skabe forsyningssikkerhed, men i slutningen af 70'erne ændrede dette sig, så det i stigende grad var energiimportens negative konsekvenser for betalingsbalancen, der kom i fokus. Efter Brundtland-rapporten ændrede dette sig nok engang, da miljøhensyn i forbindelse med klimaforandringer blev det bærende hensyn. Til trods for at de politiske problemstillinger således har været skiftende, har et af midlerne til at realisere målsætningerne imidlertid permanent været energibesparelser. Energi-
besparelser har derfor også været det helt centrale emneområde inden for evalueringer på energiområdet (Owen, 1999).

Slutanvendelse, lokationer og virkemidler

Evalueringer af besparelser falder inden for to hovedgrupper. En gruppe, der retter sig mod *slutanvendelsen* af energi. Det vil sige imod fx lys, ventilation, og en anden gruppe retter sig mod *lokationer*. Her er det udbredt at sondre mellem husholdninger og virksomheder, men i de senere år er der også kommet fokus på privat handel & service og offentlige institutioner. I tillæg hertil er der en række evalueringer inden for området, som fokuserer på forskellige former for virkemidler. Her kan man sondre mellem evalueringer, der retter sig mod effekten af informative, normative eller økonomiske virkemidler. Eksempler på informative tiltag er fx informationskampagner om elsparepærer, rådgivning af erhvervsvirksomheder

eller undervisning i folkeskolen. Normative virkemidler omfatter reguleringsmæssige tiltag. Det kan for eksempel være ændringer i bygningsreglementet, mens økonomiske virkemidler fx omfatter CO₂-afgifter, samt tilskud til energibesparende investeringer i virksomhederne. Evalueringseffekten af forskellige former for virkemidler bunder i en forventning om at nogle virkemidler vil vise sig at være mere effektive i nogle sammenhænge end i andre.

Det er således i høj grad energibesparelser og virkemidler, der har været genstand for evalueringer inden for energiområdet. Men det betyder absolut ikke, at der ikke er andre områder, som kan tages op. Fx findes der en del forskning, i hvilke forhold der påvirker virkningsgraden af forskellige energikilder. Det vil sige, hvordan man får mest muligt ud af brændslerne. Ligeledes er det i høj grad relevant at evaluere de store omvæltninger, der er sket inden for sektoren i forbindelse med liberaliseringen af el-markedet. Genstandsfeltet for evalueringer inden for energiområdet har således været forholdsvis velafgrænset, men i tillæg hertil er der også andre forhold, som er karakteristiske for evalueringer på dette område.

Effektevalueringer frem for procesevalueringer

For det første er det noget særligt, at en stor del af de evalueringer, der er foretaget inden for området, er effektevalueringer. Ved procesevalueringer forstås evalueringer, der fokuserer på iværksættelsesprocessen. Det vil sige, hvad der sker, fra et tiltag vedtages, og til det møder målgruppen. Herunder hvilke instanser og faggrupper der er involveret, hvordan samspillet mellem dem fungerer osv. Heroverfor findes evalueringer, hvor der lægges vægt på effekten, det vil sige på, i hvilken grad de politiske målsætninger opfyldes. I den forbindelse er det noget særligt, at der inden for energiområdet findes en klar målbar effektvariabel; nemlig forbruget af energi. På mange områder kan det være vanskeligt at vurdere, hvordan effekten af et tiltag skal måles, men når der skal laves evalueringer inden for energiområdet, er det imidlertid en væsentlig styrke, at der findes målbar effektvariabel.

For det andet er der på energiområdet – som i øvrigt også på andre områder – problemer forbundet med, at konstruere et kontrafaktisk tilfælde. Det vil sige en vurdering af, hvordan energiforbruget ville have udvik-

let sig, hvis der ikke var blevet iværksat et tiltag. Fx vil det være relevant at vurdere, hvorvidt virksomhederne ville have foretaget de samme ændringer i produktionsprocessen, og evt. om de ville have gjort det samme blot på lidt længere sigt uden virkemidler. Desuden kan der være problemer med, hvorvidt eventuelle reduktioner i energiforbruget kan henføres til det faktiske tiltag. Andre væsentlige overvejelser er, hvor lang tid det tager, før effekten af fx en informationskampagne slår igennem, og evt. før de fortager sig igen.

Et tredje forhold, som er specielt ved energiforbruget, er, at det ikke er en del af folks primære opmærksomhedsfelt. Virksomhederne foretager ikke pr. automatik energibesparelser i det omfang, som det er økonomisk rentabelt, da energiforbrug ikke ligger inden for virksomhedens primære opmærksomhedsfelt. Tilsvarende ved husholdningerne ikke nødvendigvis, hvad deres energiforbrug er, og deres holdninger til miljøspørgsmål påvirker ikke deres energiforbrug, hvorimod holdningerne godt kan påvirke købet af miljøvenlige og økologiske varer i øvrigt. En løsning er at få tal på det faktiske energiforbrug fx på elværkerne i stedet for at arbejde med det rapporterede energiforbrug. Desuden skal det nævnes, at der kan være problemer forbundet med at indsamle oplysninger om investeringer i energibesparende foranstaltninger. Det kan skyldes, at virksomhederne vurderer, at sådanne oplysninger er følsomme og ikke ønsker at videregive dem.

Der findes en række introduktioner til evaluering inden for energiområdet, som det kan være nyttigt at have indsigt i. Evaluering af energispareaktiviteter giver en omfattende introduktion til evalueringsopgaver inden for energisektoren. Der sondres mellem forskellige evalueringsdesign og gives talrige eksempler på deres anvendelse inden for energiområdet. I bogen findes også en oversigt over litteraturen inden for området, hvoraf størstedelen er international (AKF 2002).

I det følgende har vi valgt at give to eksempler på evalueringer inden for energiområdet. Det ene projekt fokuserer på et bestemt virkemiddel; nemlig frivillige aftaler. Her er der tale om et lille og afgrænset projekt, som metodisk set er baseret på casestudier og sammenligning mellem frivillige aftaler i fem lande. Det andet eksempel er et større projekt. Her anvendes en flere metoder. Der er tale om beregninger på forholdsvis hårde data. En anden central forskel er, at der i det første eksempel er tale om en

ex post-evaluering i den forstand, at data indsamles efter, at den policy-indsats, man søger at evaluere, forventes at have udmøntet sig i forandringer på virksomhederne. I det andet eksempel er der snarere tale om en ex ante-evaluering i den forstand, at datagrundlaget i beregningerne eksisterer før og uafhængigt af virkningen af et politisk tiltag på området.

16.2 **Eksempel 1: Frivillige aftaler**

I løbet af 90'erne blev »frivillige aftaler« i stigende grad taget i brug som et virkemiddel inden for energi- og miljøområdet – ikke bare i Danmark, men også i EU som helhed. Samtidigt fandtes der en forventning om, at forskellige virkemidler generisk set har forskellige egenskaber og varierende effektivitet. I forlængelse heraf har der været en stor interesse for at evaluere anvendelsen af forskellige virkemidler med henblik på at klargøre deres fordele og ulemper – og ikke mindst hvor man får mest miljø for pengene. I tråd hermed er der på AKF blevet gennemført en række evalueringer af energi og miljøpolitiske virkemidler. Det følgende eksempel er en evaluering i lille skala i den forstand, at undersøgelsen er afgrænset til anvendelsen af et enkelt virkemiddel – frivillige aftaler – i forhold til få virksomheder inden for et kort tidsrum.

Formål og fokus

AKF's evaluering udspandt sig inden for rammerne af et større EU-projekt. Det overordnede emne for projektet var at evaluere anvendelsen og effektiviteten af frivillige aftaler i udvalgte lande. Man ønskede at komme med anbefalinger til politiske beslutningstager – særligt EU-kommissionen – om, hvordan reguleringsindsatsen på CO₂-området eventuelt ville kunne forbedres ved at optimere kombinationen af virkemidler. AKF's mere afgrænsede rolle i den forbindelse var at gennemføre en evaluering af de danske aftaler. I AKF's evaluering fokuseres der på to hovedspørgsmål. For det første hvad frivillige aftaler er, og – for det andet – hvordan de fungerer i praksis. Det vil sige, at der dels lægges vægt på at beskrive elementerne i det danske aftalesystem, hvilket skal ses i lyset af, at der er betydelige forskelle mellem de virkemidler, der kaldes frivillige

aftaler i forskellige lande. Desuden lægges der vægt på at undersøge, hvordan de iværksættes konkret.

Evalueringsgenstand og kontekst

Evalueringsgenstanden er virkemidlet »frivillige aftaler«. Frivillige aftaler anvendes som alternativ til regelstyring, men graden af frivillighed kan variere. Der kan således være sanktioner forbundet med ikke at indgå en »frivillig« aftale. Frivillige aftaler forhandles mellem parterne – fx Energistyrelsen og en virksomhed – og i aftalen fastsættes bestemte mål, som parterne skal opfylde. I løbet af 90'erne blev aftaler om energieffektivisering i kombination med afgifter på CO₂-virkemidler indført for at reducere virksomhedernes udledning af CO₂. En sådan aftale indebærer, at der gennemføres teknisk gennemgang af virksomheden – et energisyn, og at virksomheden forpligter sig til at implementere rentable besparelser, som identificeres i den forbindelse. Desuden skal der indføres energibesparelser på virksomheden.

Fra 1998 til 2000 deltog AKF i et større EU-finansieret forskningsprojekt med deltagelse fra fem forskellige medlemsstater. Inden for rammerne af dette projekt udførte AKF – som nævnt – den danske del.

Evalueringsmodel og -design

Der er tale om en evaluering, hvor der systematisk indsamles viden om gennemførelsen og virkningerne af en bestemt regulering. Systematikken ligger i høj grad på valget af cases og valget af, hvilke forhold der indsamles data om. Disse valg muliggør en sammenligning af forskelle og ligheder, hvilket igen muliggør en struktureret analyse af fordele og ulemper og giver mulighed for at diskutere, hvorvidt de ting man ser, har en mere generel gyldighed. Der er ikke tale om en generalisering i statistisk forstand, men netop om en systematisk og velbegrundet analyse, hvor der også trækkes på den eksisterende forskning inden for området. Det vil sige, at det overvejes, hvorvidt de mekanismer der er på spil i det ene tilfælde, også rimeligvis kan forventes at gøre sig gældende andre steder.

Evalueringen er en kombineret proces- og effektevaluering. Evalueringen omfatter dels den proces, hvor aftalerne indgås med Energistyrel-

sen, og energisparetiltagene implementeres på virksomheden. Desuden omfatter de forsøg på at vurdere effekten af aftalerne. Det vil sige, i hvor høj grad de har medført reduktioner i energiforbruget, eller medført ændringer – fx i virksomhedens fokus – som kan forventes at lette energisparetiltag i fremtiden.

Der opstilles en række vurderingskriterier med baggrund i den mere teoretiske litteratur om virkemidler inden for miljøreguleringen. Der gøres således forsøg på at vurdere for det første, hvorvidt aftalerne medfører energibesparelser på en omkostningseffektiv måde. For det andet, hvorvidt aftalerne giver virksomheden en dynamisk og stadig tilskyndelse til at iværksætte energisparetiltag. For det tredje vurderes de institutionelle krav, aftalerne stiller til parterne, og aftalernes politiske anvendelighed. Endeligt vurderes det, hvilke risici der er forbundet med anvendelsen af et virkemiddel som frivillige aftaler.

Evalueringen er gennemført med henblik på en praktisk anvendelse. Dels søger den at give en vurdering af, hvordan reguleringen virker, og dels kommer den med vurderinger af reguleringens styrker og svagheder, samt forslag til, hvordan den evt. kan forbedres.

Konstruktion af data og analysemetoder

Evalueringen er bygget op som et casestudie med energiintensive virksomheder inden for to forskellige brancher. Det ene casestudie omfattede papirindustrien. Denne branche blev valgt med henblik på at sikre sammenligneligheden med de øvrige lande inden for EU-projektet, som gennemførte deres casestudier inden for samme branche. Inden for papirindustrien var der kun tre virksomheder, som på det tidspunkt havde en aftale, heraf valgte man at gennemføre et virksomhedsstudie på de to. Det andet casestudie omfattede mælkekondenseringsindustrien. Inden for denne branche fandtes der i alt ni virksomheder. Fem af disse – som alle var ejet af MD Foods – var omfattet af en brancheaftale. Valget af disse afspejlede et ønske om at medtage en brancheaftale for at muliggøre en sammenligning med det hollandske aftalesystem, som netop er bygget på brancheaftaler. Det primære datagrundlag for casestudierne var 22 semi-strukturerede interview, som blev suppleret med andet dokumentarisk materiale. Interviewene er blevet udskrevet, efter at de er blevet foretaget,

og forelagt interviewpersonerne til godkendelse. Der er således mulighed for at dokumentere det empiriske grundlag for analysen.

Fordele og ulemper ved evalueringens design

Overordnet set rummer et evalueringsdesign, som er bygget på detaljerede casestudier af implementeringsprocessen såvel som effekten, både fordele som ulemper. Den væsentligste fordel er, at man kommer »tæt på«. Det vil sige, at man får indsigt i, hvordan aftalerne reelt iværksættes på virksomheder, og hvilke praktiske problemer der findes. Desuden giver det kvalitative design mulighed for at foretage vurderinger omkring virksomhederne, som sjældent ville kunne klargøres i et struktureret spørgeskema. Dette gælder fx muligheder for at vurdere, om investeringerne i energibesparelser er et resultat af aftalerne, eller om de ville være foretaget alligevel. Tilvejebringelsen af nogenlunde troværdige oplysninger om sådanne forhold øges ved face to face-interview, hvor der er mulighed for at spørge til konkrete detaljer omkring investeringsbeslutningerne. Ulemper ligger til gengæld i de begrænsede muligheder for at generalisere. Eller sagt på en anden måde så er det begrænset, hvad man kan konkludere generelt om virkemidlets effektivitet på baggrund af tre casestudier.

Hovedresultater

Evalueringen viser, at de danske aftaler generelt kombinerer specifikke energispareprojekter, som medfører ændringer i energianvendelsen på kort sigt med mere langsigtede organisatoriske ændringer i form af energiledelse. Med hensyn til effekten viser det sig, at aftalen på den ene papirfabrik havde en betydning og umiddelbart resulterede i en mere energieffektiv teknologianvendelse og et forbedret videnniveau. Det vurderes samtidigt, at de mere langsigtede effekter på denne virksomhed formentlig vil være afhængige af et forsat reguleringsmæssigt pres. På den anden papirfabrik er de tekniske ændringer begrænsede. I mælkekondenseringsindustrien er de investeringer, der er foretaget i energieffektivisering, ikke direkte et resultat af aftalen, men det vurderes, at aftalen har medført en større bevidsthed om energianvendelsen i virksomheden. Generelt konkluderes det, at de investeringer, der kan reducere energiforbruget betragteligt, ikke er et resultat af aftalen om energieffektivisering, men af virk-

somhedernes overordnede investeringsstrategi. Dog har aftalerne en effekt på investeringer i hjælpeudstyr i form af ventilatorer, pumper m.m., som ikke knytter sig direkte til produktionsprocessen.

Det vurderes, at effekten af aftalerne i høj grad støttes af, at andre virkemidler som CO₂- og SO₂-afgifter også skaber incitament til at reducere energiforbruget. I de konkrete casestudier synes de subsidier, der gives til investeringer i energibesparende udstyr imidlertid ikke at påvirke virksomhedernes investeringsbeslutning. Ligeledes synes effekten af de informative virkemidler at være begrænset. Dog synes der at være en del blødere effekter i form af en øget opmærksomhed omkring energianvendelsen på virksomhederne, og et forbedret vidennetværk på området, som kan have en effekt på virksomhedernes energianvendelse på længere sigt.

I tillæg til disse konklusioner vurderes det danske aftalesystem i forhold til de tidligere nævnte kriterier: Omkostningseffektivitet, dynamik, risiko, samt institutionelle og politisk forhold. Her konkluderes det, at de danske aftaler er udformet med henblik på at realisere de mest effektive energibesparelser først, men at søgeomkostningerne er høje – og bemærkelsesværdigt højere i de individuelle aftaler end i branche aftalerne. Desuden er de administrative omkostninger forholdsvis høje. Aftalerne er ikke specielt dynamiske, men energiledelsen udgør et dynamiske element, som er under konsolidering. Desuden ses det som en styrke, at der er få free-riderproblemer, og at der stilles forholdsvis store krav til den enkelte virksomhed. Generelt synes aftale systemet i højere grad af være rettet mod de reaktive frem for de proaktive virksomheder. De støtter således de virksomheder, der er mindst bevidste om energiaspekterne af deres produktion.

Formidling og anvendelse

Evalueringen er blevet formidlet som en AKF-rapport (Johannsen og Larsen 2000). Desuden indgår en sammenlignende analyse af aftalerne i fem forskellige EU-lande, der er blevet formidlet som en arbejdsrapport (Krarup og Ramesohl 2000). Desuden blev der holdt et seminar i Bruxelles med deltagelse af embedsmænd fra kommissionen og interesseorganisationer på EU-niveau med interesse for området. Endeligt indgår studiet i større eller mindre grad i en række internationale konferencepapirer og ar-

tikler i danske og internationale tidsskrifter (Johannsen 2002 og Krarup 2002).

16.3 **Eksempel 2: Evaluering af vindkraftudbygningen**

Formål

Udgangspunktet for evalueringen var en politisk debat om vindkraftens fordelagtighed. Debatten satte sig til stadighed spor i Folketingets Enerkipolitiske Udvalg med en række spørgsmål og svar fra og til Folketingets Enerkipolitiske Udvalg. Der var en forventning om, at »vindkraftudbygningen i Danmark næppe er rentabel på markedsmæssige vilkår, men sker på grundlag af påbud og subsidier«. Derfor var evalueringens hovedspørgsmål »om der er brede samfundsmæssige forhold og hensyn, som alligevel kan begrunde, at vindkraftudbygningen er samfundsmæssig fornuftig.« (Munksgaard og Larsen 1996). Sådanne hensyn omfatter miljøet, beskæftigelsen, betalingsbalancen og erhvervspolitiske hensyn til vindmølleproducenterne. Det var hensigten med evalueringen at belyse, kvantificere og om muligt sætte værdi på disse samfundsmæssige hensyn.

Evalueringsdesign og data

Vi sammenligner en vindkraftudbygning på 1000 MW med to referencealternativer: Et kulfyret henholdsvis et naturgasfyret kraftværk på 420 MW. Trods mindre eleffekt kan referencekraftværket producere den samme mængde el som vindmøllerne. Over et år producerer et 420 MW kulfyret eller naturgasfyret kraftværk altså det samme som 1000 MW vindkraft (hvis 1000 MW vindmøller kører en time, producerer de 1000 MW-timer (MWh) eller 1000.000 kWh). Da kraftværket også kan producere el i vindstille perioder, har vi forudsat, at vindkraftudbygningen for at være ligeværdig må suppleres af 220 MW backupkapacitet fra gasfyrede turbiner. De to typer af værker betragtes i dette projekt som de realistiske alternativer til at foretage en udbygning af elsystemet med vindmøller. I evalueringsterminologi kan man kalde vindkraftudbygningen for interventionen og det kul/naturgasfyrede kraftværk for det kontrafaktiske tilfælde.

Evalueringen tager udgangspunkt i de »rene« produktionsomkostninger for vindkraft, dvs. eksklusive moms og afgifter, og derefter værdisætter de fordele og ulemper, som vindmøllerne har i forhold til traditionelle kul- eller naturgasfyrede kraftværker.

Ambitionen om at værdisætte vindkraften så langt som muligt på et fagligt forsvarligt grundlag er i dette projekt blevet opfyldt et stykke ad vejen. Der er foretaget en beregning af elproduktionsomkostningerne for vindmøller – målt pr. produceret kWh. Derudover er der foretaget en værdisætning af vindmøllernes reduktion af udledningen af CO₂ (kuldioxid), SO₂ (svovldioxid) og NO_x (kvælstofilter) fra elsystemet og af støj og visuelle effekter fra vindmøller.

Vi har ikke foretaget en værdisætning af effekterne på beskæftigelse og betalingsbalance af en udbygning med vindmøller, men har opgjort disse effekter i de naturlige nøgletal: personer pr. år henholdsvis mio.kr. Endelig blev der foretaget en industripolitisk analyse af vindmøllebranchen på et mere kvalitativt grundlag.

Metodisk og datamæssigt kan evalueringen beskrives som i tabel 16.1.

Tabel 16.1 Metode og data i evalueringen af vindmøllers samfundsmæssige op-timalitet

Delprojekt	Metode og data
Sammenligning af produktionsomkostninger (snæver samfundsøkonomi)	Omkostningsanalyse ved hjælp af data med hensyn til teknologi og brændselspriser
Eksterne miljøeffekter	Tekniske emissionskoefficienter (fx kg SO ₂ pr. ton indfyret kul) værdisat som skadeomkostninger. Med hensyn til støj og visuelle effekter gennemfører vi en af de første danske miljøvurderinger (CVM ¹ og husprisanalyse) med det formål at sætte omkostninger på de æstetiske og støjmæssige gener ved vindmøller.
Betalingsbalance og beskæftigelseeffekter	Vi anvendte en makroøkonomisk model: ADAM
Industripolitisk vurdering	Porters diamant (Karnøe og Jørgensen 1995 og Porter 1990)

1 Note: CVM = contingent valuation method omfatter primært spørgeskemaundersøgelser af betalingsvillighed for miljøgoder.

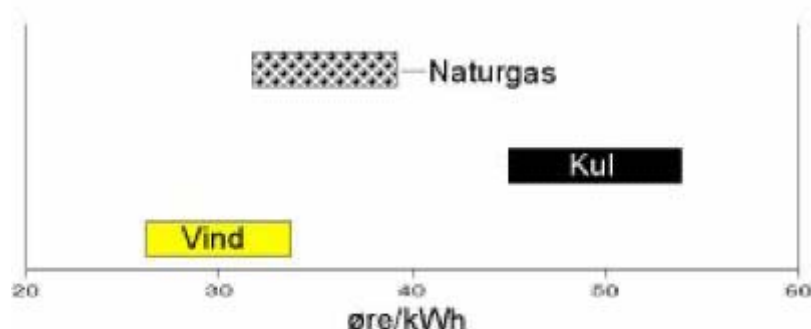
Hovedresultater

Hovedresultatet af evalueringen var, at vindkraftens miljøfordele er så betydelige, at den ud fra en bred samfundsøkonomisk vurdering er konkur-

rencedygtig med naturgasbaseret elproduktion og mere fordelagtig end el produceret på kulkraftværker.

I figur 16.1 er vist de brede, samfundsøkonomiske elproduktionsomkostninger for vind, kul og naturgas. Figuren afspejler vores forventning i 1996 med hensyn til omkostningerne omkring årtusindskiftet. Produktionsomkostningerne består af de snævre samfundsøkonomiske produktionsomkostninger og miljøomkostningerne.

Figur 16.1 Samfundsøkonomiske elproduktionsomkostninger omkring år 2000



Det fremgår af figuren, at vindkraft samfundsøkonomisk set er langt mere fordelagtig end kulkraft og i de fleste tilfælde også er mere fordelagtig end naturgas-elproduktion. Intervallerne dækker over usikkerheden på de væsentligste beregningsforudsætninger. For kul- og naturgaskraftværket drejer det sig om brændselsprisudviklingen (konstant/stigende) og værdisætningen af CO₂-, SO₂- og NO_x-udledningen. For vindkraft drejer det sig især om effektiviteten (10% henholdsvis 15%). Dette og prisen på de fremtidige vindmøller (0 til 25% billigere).

Det må stærkt understreges, at i den brede samfundsøkonomiske vurdering af vindkraften har miljøeffekterne haft afgørende betydning, herunder i særlig grad CO₂-aspektet. Vi skønner, at værdien af den reducerede CO₂-udledning i forhold til et kulfyret kraftværk udgør 18-26 øre pr. kWh – et beløb, der næsten svarer til produktionsomkostningerne pr. kWh for de mest omkostningseffektive vindmøller. Sammenlignet med et naturgasfyret kraftværk er værdien noget mindre – 9-14 øre pr. kWh – hvilket skyldes, at CO₂-udledningen kun er omkring halvt så stor på det naturgasfyrede værk.

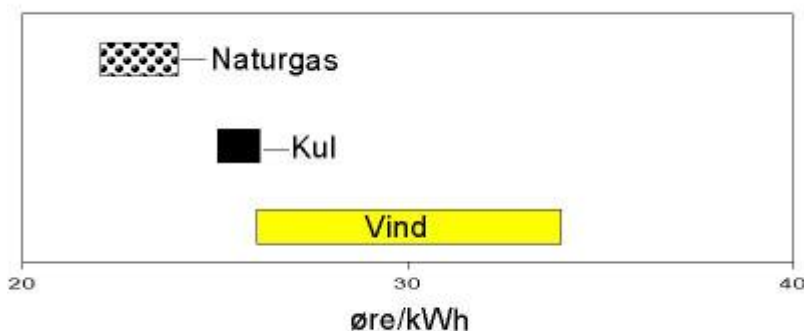
Vindmøller kan også reducere udledningen af SO₂ og NO_x fra de konventionelle kraftværker. Værdien af denne reduktion er dog væsentligt lavere end værdien af CO₂-besparelsen – kun i alt 1-2 øre pr. kWh.

Værdien af støj og visuelle gener for husstande, som bor tæt på eksisterende vindmøller, har vi opgjort til 0,04 øre pr. kWh, hvilket er forsvindende lidt i forhold til de øvrige værdisatte effekter.

De samfundsøkonomiske elproduktionsomkostninger i figur 16.1 indeholder også de snævre elproduktionsomkostninger, der er relevante ved fastlæggelsen af elpriserne over for forbrugerne. I dag er el produceret ved vindkraft i gennemsnit dyrere end el fra konventionelle kraftværker. De mest effektive vindmøller på markedet – for øjeblikket 600 kW vindmøller – kan i dag (1996) producere el til 32 øre pr. kWh (inklusive backupomkostninger på 4 øre pr. kWh), når de opstilles på steder, hvor vindforholdene svarer til den gennemsnitlige placering af de eksisterende vindmøller på land. Dette er 7-8 øre højere end produktionsprisen pr. kWh for et traditionelt kraftværk fyret med kul eller naturgas.

I figur 16.2 er vist produktionsomkostningerne pr. kWh for vind-, kul- og naturgasbaseret elproduktion, som de skønnes at være omkring år 2000.

Figur 16.2 Elproduktionsomkostninger omkring år 2000



Figur 16.2 viser, at vindmøllerne forventes at have højere produktionsomkostninger end de to referenceværker. Intervallerne i figuren afspejler usikkerheden med hensyn til brændselsprisen (kul og naturgas) og for vindkraft: effektivværdien og prisen på vindmøllen.

Vi har også undersøgt de beskæftigelses- og betalingsbalancemæssige konsekvenser af en vindkraftudbygning sammenlignet med en kulraftudbygning. Hovedresultatet af undersøgelsen er, at der ikke er nogen afgø-

rende forskel i effekterne mellem vindkraft- og kulkraftudbygningen set over en 20-årig periode. Vindkraftudbygningen skaber i forhold til en kulkraftudbygning over perioden i gennemsnit en merbeskæftigelse på godt et par hundrede mand pr. år. Omvendt giver vindkraftudbygningen anledning til en merbelastning af betalingsbalancen med i gennemsnit knap 200 mio.kr. pr. år. Dette ikke-intuitive resultat skyldes, at vindmøllerne er dyrere at opføre end kul- og gasteknologierne, samtidigt med at importkvoterne ikke er »så forskellige«, som man skulle tro. Dette opvejes ikke af brændselsindkøbet over produktionsperioden.

Refleksioner og forudsætninger

Vi har som anført benyttet ADAM-modellen til vores beregninger. Styrken ved ADAM-modellen er, at den over en længere tidsperiode kan holde rede på både de direkte og de afledte effekter i hele den danske økonomi. Ud over at tage højde for den forøgede beskæftigelse i selve vindmølleindustrien tager modellen med andre ord også højde for beskæftigelsen hos underleverandører samt i erhverv, der nyder godt af den generelle aktivitetsudvidelse i økonomien, som investeringen i elsystemet giver anledning til. Svagheden ved modellen til vores formål er, at den har et højt aggregeringsniveau, og at den ikke er indrettet specielt til at fokusere på det danske energisystem – fx findes der hverken en selvstændig elbranche eller en vindmøllebranche i modellen.

Evalueringen omfattede også en undersøgelse af den danske vindmølleindustri muligheder for at fastholde og udvikle sin internationale markedsposition i årene fremover, herunder i hvilket omfang industrien har behov for, at der føres en egentlig erhvervspolitik (Karnøe og Jørgensen 1995). Hovedkonklusionen i undersøgelsen er, at den danske vindmølleindustri har gode muligheder for at øge eksporten væsentligt i de kommende år fra 2 mia.kr. i 1994 til 4 mia.kr. i løbet af de næste 2-4 år. Dette skyldes, at vindmølleindustrien de seneste år har konsolideret deres markedsandele på verdensmarkedet og for tiden er markedsledende – både hvad angår produktteknik, etablering af brohoveder på nye markeder og referencer for pålidelig drift. Positionen på verdensmarkedet kan i løbet af nogle år blive truet af nye konkurrenter, men det skønnes, at de nye konkurrenter ikke i løbet af de første 2-3 år vil være i stand til at gøre noget væsentligt indhug i

den danske eksportvækst. Det tager nemlig tid, før udenlandske konkurrenter har etableret de tilsvarende kompetencer som de danske. Danske vindmølleproducenter kan dog sandsynligvis ikke selv dække den kraftige vækst i markedet, og derved er der en vis plads til nye producenter.

Dansk vindmølleindustriens muligheder for at fastholde den stærke position afhænger af en række faktorer, herunder af

- virksomhedernes egne konkurrencemæssige aktiver
- de vilkår, som den danske energi- og erhvervs politik giver industrien for at udvikle sig i Danmark. Disse vilkår omfatter støtte i den teknologiske udvikling og sikring af muligheder for at afprøve nye koncepter
- det finansielle systems evne og vilje til at levere den nødvendige kapital til en styrkelse af vindmølleindustrien i dens fortsatte vækst.

Det er bemærkelsesværdigt, at der ikke på noget tidspunkt er givet direkte tilskud til vindmølleproducenterne, som derfor har skullet udvikle sig i et konkurrencepræget miljø, der kun har været understøttet af de rammer, som tilskuddet til opførelse af vindmøller og de faste afregningspriser har udgjort. Der har været tale om en markedsorienteret udviklingspolitik, der har vist sig ret succesfuld ved at medvirke til at placere danske vindmøllevirksomheder stærkt blandt verdens førende.

Vi fandt det ikke rimeligt at omsætte hverken de makroøkonomiske (beskæftigelse og betalingsbalance) og industripolitiske effekter til øre pr. kWh. Dette skyldes bl.a., at det efter vores opfattelse er problematisk at lade kortsigtede konjunkturpolitiske hensyn have afgørende indflydelse på langsigtede strukturinvesteringer i elsektoren. En konjunkturberegning bygger nemlig på en række forudsætninger, der virker irrelevante i forhold til et teknologivalg for produktion af elektricitet. At præsentere disse beregninger i øre per kWh ville sløre fremstillingen af evalueringens resultat.

Til grund for konklusionerne er lagt nogle forudsætninger, som selvfølgelig kan diskuteres, ligesom der til beregningsresultaterne knytter sig nogle forbehold, som er væsentlige for fortolkningen af resultaterne. Værdisætningen af vindkraftens CO₂-fordel får afgørende indflydelse på vurderingen af, om vindkraften er samfundsøkonomisk rentabel. CO₂-fordelen er værdisat på baggrund af eksisterende danske undersøgelser, der vurderer de samfundsøkonomiske omkostninger ved at skulle opfylde den energipo-

litiske CO₂-målsætning. Der er grund til at pege på følgende væsentlige forudsætninger for resultatet:

- det blev antaget, at den danske CO₂-målsætning opfyldes på baggrund af danske bekæmpelsesforanstaltninger, dvs. der ses bort fra internationale aftaler om en optimal fordeling af CO₂-kvoterne, der tager hensyn til, i hvilke lande størst CO₂-reduktion kan opnås pr. investeret krone.
- det blev antaget, at referenceværkerne alene producerer el – ikke kraftvarme. Hvis vindkraft fortrænger kraftvarme, reduceres vindkraftens miljøgevinst væsentligt.

Derudover skal det fremhæves, at der knytter sig meget stor usikkerhed til, hvorvidt de danske, empiriske undersøgelser, der er lagt til grund for CO₂-værdisætningen, omfatter de optimale bekæmpelsesmidler og anvender dem i det optimale omfang.

Ud over værdisætningen af vindmøllernes CO₂-besparelse spiller det en afgørende rolle, hvordan energisektoren er opbygget, og hvilken sammenhæng vindmøllerne kommer til at fungere i. Det vil sige, hvilke teknologier vindproduktionen skal fungere i samspil med.

Formidling

Evalueringen er blevet formidlet i et videnskabeligt tidsskrift (Munksgaard & Larsen 1998).

Med hensyn til formidling til politikerne blev evalueringens resultater sendt til Folketingets Energipolitiske Udvalg. For den bredere offentlighed afholdt vi en workshop på Risø, hvor resultaterne blev formidlet med opponenter/kommentatorer. Evalueringens positive resultat for vindkraft gav genlyd i mere »teknologikonservative« kredse, så forfatterne til evalueringsrapporten måtte skrive en del indlæg i bl.a. Ingeniøren. Det har i øvrigt vist sig, at projektets hovedrapport har været den rapport med flest downloads af alle AKF's rapporter, der ligger i fuld tekst.

16.4 Afrunding

De to eksempler repræsenterer den del af evalueringer inden for området som retter sig mod at evaluere effekten af et energisparetiltag ex post.

Men i tillæg hertil findes der også en omfattende litteratur i form af teknologivurderinger, som retter sig mod at beregne hensigtsmæssigheden af policy-tiltag *ex ante*. Hermed menes, at de data, der indgår i beregningerne er indsamlet før og uafhængigt af, at et politisk tiltag er blevet iværksat. Desuden kan det være nødvendigt at lave antagelser om den fremtidige udvikling. Fx vil udviklingen i dollarkursen være af betydning for, hvorledes vindmølleeksporten vil udvikle sig. I *ex ante* evalueringen er der således tale om en kerne af data, der beskriver teknologier – og som er forholdsvis uforanderlige, men i tillæg hertil indgår der nogle forudsætninger og forventninger i beregningerne.

Man kan på sin vis diskutere, hvorvidt *ex ante*-evalueringer overhovedet er evalueringer for så vidt, at de data, der ligger til grund, er indsamlet før *gennemførelsen* af et tiltag. Det er jo netop et af de definerende træk ved en evaluering, at det er tale om at undersøge, hvordan en tilstand ændres, efter at et program er blevet afviklet. En *ex ante*-evaluering kan således ikke sige noget om, hvad der faktisk skete. Imidlertid kan der alligevel være gode argumenter for at gennemføre *ex ante*-evalueringer. Det er således et velkendt problem ved evalueringer, at de ikke nødvendigvis kommer til at tjene til en forbedring af grundlaget for den politiske beslutningstagning. Der kan med andre ord være langt fra opsamlingen af information om forskellige tiltag til en forbedring af videnniveauet hos de politikere, der iværksætter disse tiltag. Her har *ex ante*-evalueringer den fordel, at de kan gennemføres tidligere i den politiske proces og afspejle en aktuell politisk problemformulering, evt. allerede inden der træffes en endelig politisk beslutning. I en del tilfælde – som i vindmølleeksemplet – findes der desuden tal, som kan indgå i en ædruelig beregning af, hvilken effekt et tiltag må forventes at have. I denne situation er der således gode argumenter for at gennemføre *ex ante*-evalueringer. Imidlertid er det også tydeligt, at *ex ante*-evalueringer ikke kan vurdere, hvad der så rent faktisk blev konsekvensen. Der er således tale om to forskellige typer af evalueringer med hver deres styrker og svagheder, som begge bør findes inden for energiområdet.

17 **Cost-benefit-analyser på miljøområdet**

Emil Husted Erichsen og Camilla Damgaard

Anvendelsen af økonomiske vurderinger på miljø- og naturområdet er omdiskuteret, og der bliver ofte argumenteret med, at det ikke er rimeligt og tilmed uetisk at gøre fx rent drikkevand eller luftforurening op i kroner og øre. Det er dog værd at være opmærksom på, at også natur- og miljøområdet undergår politiske prioriteringer, og økonomiske cost-benefit-analyser kan bidrage til, at pengene bruges der, hvor de gør mest nytte. Cost-benefit-analyser vil dog ofte være forbundet med visse usikkerheder og kan derfor ikke stå alene i en beslutningsproces. Cost-benefit-analyser er et værdifuldt bidrag til at kvalificere en beslutningsproces, men de kan hverken erstatte sund fornuft eller politiske vurderinger.

17.1 **Hvad er en cost-benefit-analyse?**

I en cost-benefit-analyse opgøres forbruget af de ressourcer, som samfundet samlet set bruger og skaber ved forskellige former for miljøtiltag. Den velfærdsøkonomiske teori udgør grundlaget for cost-benefit-analyse. Når man bruger dette teoretiske grundlag resulterer det i, at befolkningens betalingsvillighed over for konsekvenserne af et tiltag kan bruges som indikatorer på den marginale nytte heraf. Konsekvenserne af et tiltag kan opfattes som ændringer i udbuddet og anvendelsen af en lang række goder, og det er dermed den værdi, som befolkningen tillægger disse ændringer, der udgør vurderingsgrundlaget. Det antages, at alle personers præferencer har lige stor vægt, og at præferencerne kan sammenlignes, når de udtrykkes som betalingsvillighed.

Intuitivt virker det rimeligt at basere en vurdering på befolkningens præferencer. Der er dog en række tilfælde hvor det ikke er rimeligt. På grund af manglende information er befolkningen måske ikke klar over, hvad der tjener den bedst. Ligeledes kan der være tilfælde, hvor de enkelte personers præferencer er i uoverensstemmelse med mere overordnede etiske principper. I sådanne tilfælde kan en cost-benefit-analyse ikke stå alene.

Når alle personers præferencer antages at have samme vægt, synes vurderingsgrundlaget grundlæggende at være demokratisk. Antagelsen har dog to u hensigtsmæssige konsekvenser. For det første har personer med mange penge umiddelbart gennem deres betalingsvillighed mulighed for at udtrykke stærkere præferencer end personer med få penge. For det andet vil et tiltag blive anset for fordelagtigt, selv om det alene kommer de velhavende til gode uden at forringe forholdene for de fattige, hvilket medfører større ulighed i samfundet. En cost-benefit-analyse bør derfor suppleres med en vurdering af tiltagets fordelingskonsekvenser.

Cost-benefit-analyser baseres på marginale, dvs. små, ændringer. Det betyder, at der ikke indgår de afledte effekter for samfundet, herunder bl.a. påvirkning af indkomstfordeling, beskæftigelse, udenrigshandel etc.

Budgetøkonomi, velfærdsøkonomi og beregningspriser

Velfærdøkonomiske cost-benefit-analyser baseres ofte på budgetøkonomiske analyser, og selv om begge metoder måler omkostninger og gevinster i monetære enheder, har de to forskellige formål. Budgetøkonomiske analyser måler pengestrømme (i form af indkomster og udgifter), mens den velfærdsøkonomiske analyse er et forsøg på at måle nytteforandringer i form af en stigning eller formindskelse af velfærden på grundlag af forbruget af ressourcer og produktionen af goder. Ud over dette basisgrundlag adskiller metoderne sig primært på to punkter; synsvinkel og priser:

Budgetøkonomiske analyser gennemføres ofte fra den enkelte virksomheds, landmands eller statens synsvinkel, mens afgrænsningen for velfærdsøkonomiske analyser er hele samfundet set under et. Den budgetøkonomiske analyse viser de direkte økonomiske konsekvenser for hver af de berørte parter. Denne analyse viser således de betalingsstrømme, som et givet tiltag giver anledning til. Hermed kan de økonomiske vindere og ta-

bere afdækkes. Til forskel medtager den velfærdsøkonomiske analyse værdien af ikke markedsomsatte goder (fx den rekreative værdi af et engområde), mens transfereringer (En transferering er en ensidig pengeoverførelse mellem to aktører på et marked – fx nationale tilskud fra stat til en virksomhed eller rentebetalingen.) i samfundet ikke medtages. Hermed belyses konsekvenserne for samfundet som en helhed.

Budgetøkonomiske priser er de priser, som bliver betalt af virksomheden eller konsumenten, fratrukket refunderbare afgifter, såsom moms for virksomheder. I en cost-benefit-analyse skal priserne som nævnt afspejle forbrugernes betalingsvilje for de producerede goder eller forbrugsmuligheder tabt gennem forbrug af ressourcer i et projekt frem for et andet. De budgetøkonomiske priser skal derfor korrigeres for eksisterende skatter og afgifter i den danske økonomi for at kunne indgå som velfærdsøkonomiske priser i analysen. Denne forskel tages der højde for ved at forhøje de budgetøkonomiske priser med såkaldte nettoafgiftsfaktorer, for henholdsvis indenlandsk og internationalt handlede goder. Nettoafgiftsfaktoren er 1,17 for indenlandsk handlede goder og 1,25 for internationalt handlede goder.

Tidsperspektiver og diskontering

Diskontering bruges til at omdanne fremtidige effekter til nutidige forbrugsmuligheder. Diskontering bygger på den forudsætning, at den nutidige værdi af omkostninger og gevinster er større end den fremtidige værdi, bl.a. fordi befolkningen generelt foretrækker forbrug nu frem for senere. Valg af kalkulationsrente kan generelt have stor betydning for udfaldet af velfærdsøkonomiske analyser. Dette gælder dog især, når omkostninger og fordele er tidsmæssigt forskudt. Ved meget lange tidshorisonter bør man overveje, om tiltaget er intergenerationelt fair.

Kvantificering og værdisætning af miljøgoder

Det kan være særdeles vanskeligt at kvantificere og værdisætte miljøeffekterne. En værdisætning afspejler, hvor meget et renere miljø værdisættes i forhold til andre ting. Men det er svært at afsløre befolkningens betalingsvillighed for miljøgoder, dels fordi der ikke eksisterer et marked for goderne, og dermed ikke en markedspris, dels fordi det i sig selv kan være svært at opgøre miljøeffekterne. Der er dog ingen tvivl om, at miljøef-

fekterne har en (evt. negativ) værdi for os, og denne værdi er det nødvendigt at sætte kroner og øre på, hvis den skal sammenlignes med markedsomsatte konsekvenser. Det er vigtigt at være opmærksom på, at prissætning generelt er forbundet med usikkerheder.

For at tage et eksempel. Skovrejsning forøger de fremtidige muligheder for en række forskellige friluftaktiviteter. En prissætning af disse muligheder må baseres på danskernes villighed til at betale for disse muligheder. For at afsløre denne betalingsvillighed kan man benytte en metode som den hypotetiske værdisætningsmetode, rejseomkostningsmetoden eller husprismetoden.

Rejseomkostningsmetoden kræver dog, at miljøgodet er af en sådan karakter, at de besøgende er villige til at afholde rejseomkostninger for at komme dertil. Derfor er denne metode mest velegnet til at værdisætte (større) kendte naturområder såsom Møns Klint eller Mols Bjerge, og må anses for at være uegnet til at værdisætte fx et skovrejsningsprojekt.

Husprismetoden afslører betalingsvilje for et miljøgode gennem merpriser for huse beliggende nær miljøgodet. Denne metode er derfor mest velegnet til at værdisætte miljøgoder, hvor de omkringboende må antages at have den største værdi af godet. Dette gør metoden velegnet til at værdisætte skovrejsningsprojekter, som må antages at have størst værdi for lokalbefolkningen.

Den hypotetiske værdisætningsmetode er baseret på interview med besøgende, hvor der spørges om deres betalingsvilje for det pågældende miljøgode, fx adgang til de danske skove. Denne metode kan i princippet bruges til værdisætning af alle miljøgoder, hvis det er muligt at få respondenterne til at afsløre deres reelle betalingsvilje.

Det er et generelt problem, at man ønsker at vurdere projekternes relative fordelagtighed ex ante, dvs. før tiltaget virkeliggøres. Ofte er det først muligt efter projekternes gennemførelse at opnå den fornødne viden om befolkningens faktiske udnyttelse af de nye muligheder og dermed deres betalingsvilje for projektet. Kun ved brug af den hypotetiske værdisætningsmetode er det muligt at værdisætte ex ante.

Alternativt kan man inddrage erfaringer fra tidligere værdisætningsundersøgelser, såkaldt benefit-transfer. Man skal imidlertid være yderst

forsigtig med at overføre observerede priser for ét specifikt miljøgode til andre goder.

En anden overordnet metode er alternativomkostningsmetoden. Hvis Danmark har vedtaget nogle reduktionsmål eller -forpligtelser som følge af fx en vandmiljøplan eller Kyotoaftalen om nedbringelse af CO₂-udslippet, så kan et tiltag erstatte et andet tiltag som del af målopfyldelsen. Prisen på det tiltag, man ikke behøver at iværksætte, kan bruges som alternativomkostningspris på et andet tiltag. Hvis man fx begynder at bruge mere CO₂-neutralt biobrændsel, behøves der ikke så mange vindmøller for at reducere CO₂-udslippet. Prisen på vindmøllealternativet kan bruges til at vurdere brændselsskiftets fordelagtighed. Denne metode siger imidlertid ikke noget om befolkningens præferencer, men er et produkt af politikernes prioriteringer.

Da der er knyttet usikkerhed til mange af de data, der bruges i en cost-benefit-analyse, er det vigtigt at vurdere, om resultatet er følsomt over for ændringer i parametre og antagelser. Dette kan gøres på flere måder, en mulighed er at regne alle størrelser i intervaller, en anden er at lave en grundberegning, og derefter ændre denne beregning med hver af de parametre og antagelser, hvortil der er knyttet usikkerhed.

Mere læsning

Både Miljøministeriet, Finansministeriet og Trafikministeriet har vejledninger i at lave cost-benefit-analyser. En lang række offentlige institutioner i Danmark laver cost-benefit-analyser på natur og miljøområdet. Ud over AKF kan nævnes Danmarks Miljøundersøgelser, Institut for Miljøvurdering, Miljøstyrelsen, Skov- og Naturstyrelsen, KVL, Københavns Universitet, Det økonomiske råd m.fl. En række private konsulentfirmaer laver også cost-benefit-analyser.

Om eksemplerne:

De to eksempler er valgt for at illustrere spændvidden i både anvendelsesområde og detaljeringsgrad. Den første analyse er en analyse af ændret arealanvendelse og er primært lavet på baggrund af eksisterende viden. Den fylder 12 sider. Den anden er en valg mellem forskellige teknikker til

håndtering af affald. Den bygger på vidtgående fuldskalaforsøg og fylder 134 sider.

Eksempel 1:

Erichsen, E. H (2003) Samfundsøkonomisk projektvurdering af skovrejsning ved Odense. AKF Forlaget

Eksempel 2:

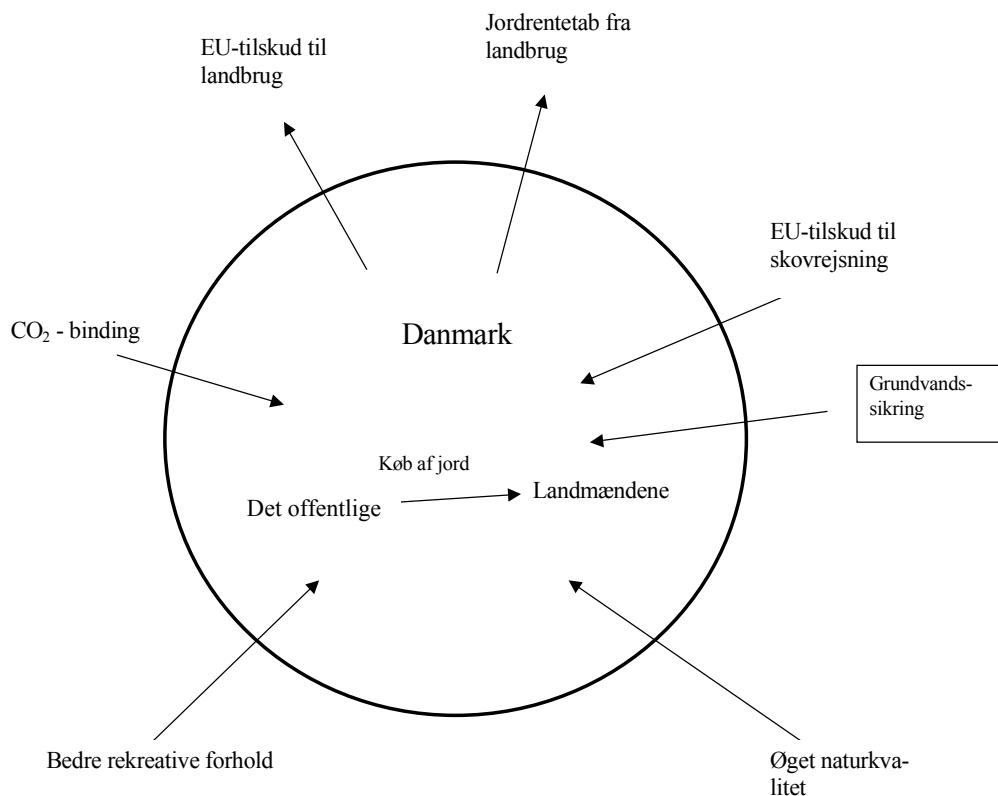
Miljøstyrelsen 2003: Skal husholdningernes madaffald brændes eller genanvendes? Samfundsøkonomisk analyse af øget genanvendelse af organisk dagrenovation Miljøprojekt nr. 814, 2003.

De refereres fra originalkilden, dvs. med de priser og den viden, der var til stede, da de blev udført.

17.2 **Eksempel 1: Cost-benefit-analyse af et skovrejsningsprojekt ved Odense**

Baggrund og formål

Odense Kommune, Odense Vandselskab og Skov- og Naturstyrelsen (SNS) indgik i 2001 en aftale, hvori der bl.a indgår et skovrejsningsprojekt på 611 ha vest for Odense by. De primære formål med projektet er at skabe friluftsmæssige tilbud til borgerne i Odense samt at beskytte grundvandet ligeledes til gavn for borgerne i Odense. Regionplanen for Fyns Amt udpeger området som Område med særlige drikkevandsinteresser, Nitratfølsomt indvindingsopland samt Område hvor skovrejsning er ønsket. I alt 26% af Fyns Areal er udpeget som nitratfølsomme indvindingsområder, og det er populært sagt disse, der er de sårbare grundvandsområder. I analysen opgøres værdierne af grundvandsbeskyttelse, rekreative anvendelsesmuligheder, CO₂-binding, mindsket næringsstofbelastning og andre miljøeffekter. På omkostningssiden er nettoomkostningerne ved overgangen fra landbrugsdrift til skovdrift beregnet. I grafisk form kan virkningerne fra projektet fremstilles således:



Odense Vandselskab a/s bestilte denne analyse hos AKF, idet de ville have et bedre grundlag for at indgå i dialog med bl.a. andre offentlige myndigheder om planens hensigtsmæssighed.

Denne cost-benefit-analyse forsøger at svare på spørgsmålet om, hvad der er mest hensigtsmæssigt ud fra en samfundsøkonomisk betragtning: at fortsætte med landbrugsdriften eller at rejse skov på området. Analysen skal ikke læses som sandheden med hensyn til samfundsøkonomien bag det konkrete skovrejsningsprojekt, men den er et ædrueligt bud på samme præsenteret med nogle overvejelser om de anvendte tal. Det er mange steder i analysen valgt at udføre beregningerne med intervaller, således at det klart fremgår, at der er usikkerhed knyttet til de fleste data.

Dataindsamling og analysemetoder

Analysen benytter de eksisterende kilder og tidligere udfærdigede analyser i stor udstrækning.

Der benyttes en kalkulationsrente på 3% for de velfærdsøkonomiske beregninger som anbefalet i Miljøministeriets vejledning i udarbejdelse af samfundsøkonomiske projektvurderinger (Møller et al. 2000).

Landbrug

Før projektstart har der været konventionelt landbrug i stort set hele projektområdet. Ved at opgøre alle input og output fra landbrugsproduktion kunne der udregnes en jordrente for netop dette område. Dette er udtryk for den forrentning af jorden, som mistes ved ændret arealanvendelse. Jordrenten er defineret som den rest, der er tilbage til aflønning af produktionsfaktoren jord, når alle andre produktionsfaktorer, der anvendes til dyrkningen – dvs. rå- og hjælpestoffer, arbejdskraft og realkapital i form af maskiner og bygninger mv. – er aflønnet svarende til disse produktionsfaktoreres alternative anvendelse.

I denne analyse er det derfor valgt at bruge den gennemsnitlige forpagtningsafgift for Fyns Amt, da en specifik jordrenteberegning er meget omfattende og ikke nødvendigvis mere præcis.

Forpagtningsafgiften i Fyns Amt er gennemsnitligt 2809 kr./ha (Fødevareøkonomisk institut 2002). Dette beløb dækker aflønningen til jorden. Jagtretten er derudover ca. 225 kr./ha/år værd. (Møller et al. 2000). Dette er en gennemsnitsværdi af jagt på øerne. Jagt på Fyn er nok billigere end Sjælland, men bynærheden til Odense trækker den anden vej. Samlet set har en landmand således en indtægt på godt 3000 kr. for hver hektar, han ejer, så det antages, at det velfærdsøkonomiske tab ved nedlæggelse af landbrugsproduktionen i området er på 60.000-100.000 kr. pr. hektar. Denne værdi inkluderer det EU-finansierede hektartilskud til landbrugsdrift, og hvis disse skulle falde bort, må det alt andet lige formodes, at prisen på jord ville falde. Det samlede tab ved ophør af landbrugsproduktion på 611 ha bliver 36,7-61,1 mio.kr.

Skovdrift

Området tilplantes med skov, der på lang sigt kan levere træprodukter. Den egentlige skovdrift i form af træproduktion giver dog ikke overskud på ret mange jorder. Det antages, at der kan opnås EU-tilskud til den kommunale del skovrejsningsprojektet, og at træproduktionen i Elme-

lund, når EU-tilskuddet til kommunal skovrejsning indregnes, balancerer. På den statslige del kan der ikke opnås EU-tilskud, og her må der påregnes en negativ jordrente fra træproduktion. På de 406 ha, som statsskovbruget står for, svarer det til et tab på 6,8 mio.kr. ved 500 kr./ha/år.

CO₂-binding

Skoven binder CO₂ i form af kulstof op gennem sin vækstperiode. Da skoven bliver fredskovspligtig, skal arealet altid være skovbevokset og skoven vil dermed fungere som permanent lager af CO₂. Dette kan indregnes i Danmarks CO₂-reduktionsforpligtigelse.

550 ha skov, dvs. den træbevoksede del af skoven, binder 8 ton CO₂/ha/år de første 90 år. Denne mængde CO₂ kan værdisættes ud fra alternativomkostningsmetoden, da skovens CO₂-binding medvirker til, at Danmark kan opfylde den fastsatte CO₂-målsætning. Energistyrelsen har i 2001 opgjort omkostningerne til at opfylde danske forpligtelser på med hensyn til CO₂-udslip. Den marginale pris på CO₂-udslip er beregnet til 243 kr. (Energistyrelsen 2001) Alternativt kan man vælge at bruge CO₂-afgiften på 100 kr./ton som pris. Denne pris afspejler den politiske betalingsvilje for at reducere CO₂-udslippet, men er heller ikke korrekt, da afgiften ikke er pålagt overalt, og at der desuden tages politiske beslutninger, der indebærer en højere pris på CO₂. Skoven binder 8 ton CO₂/hektar/år i 90 år (Møller et al. 2000), hvilket giver en nutidsværdi for CO₂-binding i skoven på 13,6-33,0 mio.kr.

Drikkevand

Odense Vandselskab vurderer, at skovrejsningen sikrer 1,2-1,4 mio. kubikmeter om året, eller hvad der svarer til en sjettedel af Odenses vandforbrug. I det følgende regnes der med gennemsnittet 1,3 mio. m³/år. Alternativerne til skovrejsningsprojektet er enten rensning af grundvandet, at flytte borerne eller at pumpe vandet fra fjernere liggende borer. Ud fra disse tiltag kan de alternative omkostninger beregnes. På baggrund af beregninger, Miljøstyrelsen foretog i 1997 af dette, antages det, at værdien af grundvandssikringen udgør mellem 1,4 og 5,5 kr./m³ (Miljøstyrelsen 1997). Herved fås en årlig værdi af grundvandsikringen på mellem 1,8 og 7,2 mio.kr. pr. år fra det år, hvor forureningen sker. Odense Vand-

selskab vurderer, at der vil gå 5-10 år, før problemerne med pesticider er så store at vandet skal renses, eller at boringen skal lukkes. Under en antagelse om, at forureningen først var indtruffet om 10 år og med en diskonteringsrate på 3% bliver nutidsværdien af grundvandssikringen dermed 45,1-177,3 mio.kr.

Rekreative værdier

Ved værdisætning af de steder, der har størst betydning for lokalområdet, er husprismetoden meget anvendelig. Husprismetoden udnytter, at prisen på et hus er afhængig af en lang række faktorer, hvoraf nogle knytter sig til huset (størrelse, alder, kvalitet) og andre til beliggenhed (Hasler et al. 2002).

Ved et skovrejsningsprojekt ved Aalborg blev der fundet store husprisstigninger allerede ved starten af skovrejsningen, og ved afslutningen af skovrejsningen blev der fundet husprisstigninger svarende til ca. 13% af husprisen. Denne undersøgelse viser klart, at de rekreative værdier kan identificeres tidligt efter skovrejsningen, og at bynære skove kan give anledning til endog meget store rekreative værdier (Hasler et al. 2002). Hvis det antages, at husene i umiddelbar nærhed af skoven stiger 10% i værdi, svarer det til 21,8 mio.kr. De 21,8 mio.kr. er derfor et underkantskøn, der jo heller ikke inkluderer værdien for de af Odenses borgere, der ikke bor tæt på skoven.

Værdien af reduceret næringsstofbelastning

Skovarealer er generelt karakteriseret ved, at der i forhold til tilsvarende landbrugsarealer udvaskes meget små mængder kvælstof (N) per arealenhed herfra. Den reducerede belastning kan have positive konsekvenser for vandkvaliteten i bl.a Ryds Å, Stavids Å, Odense Å og Odense Fjord. I den tekniske midtvejsevaluering af Vandmiljøplan II er den sparede udvaskning ved skovrejsning sat til 51 kg N/ha/år (Grant, 2000). Differencen på 51 kg N/ha/år vil derfor blive benyttet i det følgende. Den samlede udvaskningsreduktion fra skovarealet kan herefter beregnes ved at multiplicere de angivne 51 kg N/ha/år med skovarealets størrelse på 611 ha. Resultatet bliver en årlig belastningsreduktion på ca. 31,1 ton pr. år.

Det er valgt at gennemføre alternativomkostningsberegninger med det billigste tiltag i vandmiljøplan 2. Dette er anlæggelse af nye vådområder, som i midtvejsevalueringen er beregnet til omkostning pr. kg reduceret N-udvaskning på 3,6 kr. (Jacobsen 2000). Prissættes den reducerede kvælstofudvaskning på 31,1 t/år for Elmelund Skov med den nævnte pris på 3,6 kr./kg, fås en årlig værdi på ca. 112.000 kr./år svarende til ca. 184 kr./ha. Nutidsværdien af denne sparede næringstofbelastning er således 3,7 mio.kr.

Hovedresultater

Opstilles de beskrevne omkostninger og benefits i nutidsværdier for hele projektet, fås følgende udtryk for de velfærdsøkonomiske konsekvenser.

Tabel 17.1 Oversigt over velfærdsøkonomiske omkostninger og fordele

	Omkostninger i mio.kr.	Fordele i mio.kr.
Ophør af landbrugsdrift	36,7-61,1	
Træproduktion	6,8	
Grundvandssikring		45,1-177,3
CO ₂ -binding		13,6-33,0
Rekreative værdier		21,8-21,9
Indirekte effekter (mindsket udledning af stoffer med drivhuseffekt-potentiale)		4,0-9,7
Nitratudvaskning		3,7
Total	43,5-67,9	88,2-245,6
Resultat		20,3-202,1

Hvis de høje bud er lige så sandsynlige som de lave, og det er ikke en urimelig antagelse, så vil det gennemsnitlige resultat være det mest sandsynlige. Dette vil sige, at den samfundsøkonomiske værdi af skovrejsningsprojektet ved Elmelund er ca. 110 mio.kr.

Analysen viser, at det er samfundsøkonomisk dyrt at opretholde en konventionel landbrugsproduktion på bynære områder med vigtige og sårbare grundvandsressourcer. Det er meget dyrt at rense og endnu dyrere at omlægge grundvandsforsyningen ved en risiko for forurening, men relativt billigt at nedlægge landbrugsproduktionen. Dette resultat vil formentligt

også være gældende andre steder i landet, hvor de samme forhold gør sig gældende, hvilket i første omgang vil sige omkring de større byer.

Formidling og anvendelse

Odense Vandselskab sendte notatet til interessenter og aktører på vandområdet og har brugt det som en del af begrundelsen til den indsigelse, de gjorde i forhold til Fyns Amts tilladelse til etablering af en stor svineproduktion i området. Fyns Amt lavede en VVM-screening, der viste, at svineproduktionen ikke ville have særlig betydning for miljøet. Det er pesticider i grundvandet, der er problemet, og svineproduktion giver ikke anledning til nogen merbelastning i forhold til pesticider. Dette er for så vidt rigtigt, men problemet er, at med en stor svineproduktion vil landmanden ikke have mulighed for at sælge sin jord til skovrejsningsprojektet, og det vil derfor ikke blive til noget i sin nuværende form. Så svineproduktionen betyder ikke direkte en forværring af miljøtilstanden, men forhindrer en forbedring. Sagen ligger i skrivende stund (primo 2004) i Naturklagenævnet.

17.3 Eksempel: Organisk affald

Overordnet formål

I Affald 21, som er den nationale plan for affaldshåndteringen 1998-2004, er der opstillet en lang række mål for affaldshåndteringen i Danmark, blandt andet for at øge genanvendelse af affald. Inden for dagrenovation er fokus især på papir, glas, plast og organisk affald.

På EU-plan overvejer man at øge genanvendelsen af organisk affald. Det fremgår således af et arbejdsrapport fra EU-kommissionen, at der er ideer om at stille krav om separat indsamling af organisk affald i medlemslandene med henblik på genanvendelse. (EU-kommissionen, 2002)

Miljøstyrelsen har derfor analyseret de samfundsøkonomiske konsekvenser af at øge genanvendelsen af det organiske affald, herunder såvel de miljømæssige som de økonomiske konsekvenser. Formålet med analysen er at give et bud på, om det økonomisk og miljømæssigt er en god ide at anvende ressourcer på at ændre ved den gældende affaldshåndtering af organisk affald. Denne viden kan så indgå i fremtidig politikformulering på

både statsligt og kommunalt niveau. Analysen undersøger altså, hvor der fås »mere miljø for pengene« i håndteringen af organisk dagrenovation.

Evalueringsmodel og design

Fordelene og ulemperne ved forskellige metoder til håndtering af organisk affald opgøres i kroner og ører, og herved fås et mål for, hvilken håndtering der er billigst for samfundet.

I cost-benefit-analysen undersøges både de budgetøkonomiske konsekvenser og de velfærdsøkonomiske konsekvenser af at øge genanvendelsen af det organiske affald fra husholdningerne.

I den budgetøkonomiske analyse ses på de direkte økonomiske konsekvenser ved en ændret affaldshåndtering. Der tages udgangspunkt i de faktiske priser, som affaldsbehandlere står over for, dvs. tilskud og afgifter medtages. Miljøeffekterne ved de forskellige behandlingsmetoder indgår ikke. Denne type analyse viser, hvilke fordelingsmæssige konsekvenser en ændret håndtering af det organiske affald har for de forskellige aktører. I denne analyse vises kun en samlet omkostning, fordi øgede omkostninger for fx kommunen overvælttes direkte på borgerne gennem øgede affaldsgebyrer.

I den velfærdsøkonomiske analyse behandles konsekvenserne for hele samfundet under ét. Det er opgjort, hvorledes en ændret affaldshåndtering vil påvirke ressourceforbruget og miljøet for hele samfundet. Afgifter og tilskud, som i princippet blot fungerer som omfordelinger i samfundet, medtages ikke. Endvidere regnes der bl.a. i forbrugerpriser og bruges en anden diskonteringsrente end i den budgetøkonomiske analyse, jf. afsnit 1.1. Miljøkonsekvenserne er prissat i det omfang, det er muligt. Herved er det muligt at vurdere, om eventuelt øgede omkostninger ved indsamling og behandling med henblik på øget genanvendelse bliver opvejet af værdien af miljøeffekterne.

Dataindsamling og analysemetoder

Der er i analysen opstillet fem scenarier for den fremtidige håndtering af organisk dagrenovation. Disse er opstillet med forskellige kombinationer af de tre behandlingsmetoder – forbrænding, bioforgasning og central kompostering – med baggrund i målene i Affald 21 og justeret efter nye-

ste data. Der regnes i analysen dels på de tre behandlingsmetoder hver for sig og dels på kombinationer heraf i de opstillede fem scenarier, jf. tabel 17.2.

Tabel 17.2 Scenarier. Ton pr. år

	Scenario 1, Reference-Scenario	Scenario 2, Målsætning 2004, Bioforgasning	Scenario 3, Målsætning »langt sigt«, bioforgasning	Scenario 4, Målsætning 2004, Central kompostering	Scenario 5, Målsætning »langt sigt« central kompostering
Forbrænding	654.000	570.000	370.000	570.000	370.000
Bioforgasning	14.000	100.000	300.000	30.000	30.000
Central kompostering	32.000	30.000	30.000	100.000	300.000
Organisk dagrenovation i alt	700.000	700.000	700.000	700.000	700.000

Analysen fokuserer kun på den organiske del af affaldet. Både ved indsamling og behandling håndteres det organiske affald sammen med resten af husholdningsaffaldet, og omkostninger og indtægter er forsøgt opgjort specifikt for den organiske dagrenovationsdel.

Data for miljøeffekter stammer fra en livscyklusanalyse ved brug af den svenske ORWARE-model.

Datagrundlaget for den samfundsøkonomiske analyse kommer for en stor dels vedkommende fra en lang række tekniske undersøgelser og fuldskalaforsøg udført i en række kommuner suppleret med miljømæssige vurderinger af håndteringen af den organiske dagrenovation. Undersøgelserne og forsøgene blev sat i gang af Miljøstyrelsen i 1999.

Hovedresultater

Analysen viser, at det ikke kan betale sig for samfundet at genanvende organisk dagrenovation ved bioforgasning eller central kompostering. Forbrænding er samfundsøkonomisk den billigste løsning. Dernæst kommer bioforgasning, mens central kompostering er den dyreste løsning for samfundet.

Den primære årsag til, at bioforgasning og central kompostering af organisk dagrenovation er dyrere end forbrænding, er, at den nødvendige todelte indsamling af dagrenovationen er forholdsvis dyr, fordi den kræver

ekstra materiel og arbejdskraft. Selve behandlingen ved bioforgasning og central kompostering er derimod væsentlig billigere sammenlignet med forbrænding.

Den velfærdsøkonomiske analyse

Sammenlignes de velfærdsøkonomiske omkostninger for udelt indsamling og forbrænding med todelt indsamling med genanvendelse, er det 581-1.190 kr. dyrere pr. ton indsamlet organisk affald for samfundet at genanvende ved bioforgasning. Og det er 699-1.308 kr. dyrere pr. ton ved central kompostering (se tabel 17.3).

Tabel 17.3 Velfærdsøkonomiske konsekvenser ved forbrænding, bioforgasning og central kompostering af et ton organisk dagrenovation. Kr. pr. ton indsamlet affald

	Forbrænding		Bioforgasning		Central kompostering	
	Enfamilie	Etage	Enfamilie	Etage	Enfamilie	Etage
Indsamling i alt	1201	566	2063	2037	2063	2037
Behandling i alt	1060		794		585	
Miljøkonsekvenser i alt	132		202		93	
Samlede velfærdsøkonomiske omkostninger	2392	1758	3058	3033	2740	2715
Slutprodukter						
EI	154 kwh		255 kwh		23 kwh	
Varme	730kwh		579 kwh		110 kwh	
Gødning						
N (kvælstof)	0		3,63 kg		2,48 kg	
P (fosfor)	0		0,67 kg		0,97 kg	
K (Kalium)	0		1,65 kg		2,37 kg	
Værdi af produkter fra affaldsbehandlingen	464		549		113	
Samlede velfærdsøkonomiske nettoomkostninger	1928	1294	2509	2484	2627	2602
Ændring i forhold til forbrænding			581	1190	699	1308
Øvrige miljøkonsekvenser (tungmetaller i fast form)	Kg/ton affald					
Bly (Pb), (so)	i.o.		0,00178		0,00255	
Cadmium (Cd), (so)	i.o.		2,31E-05		3,32E-05	
Kviksølv (Hg), (so)	i.o.		4,97E-06		7,14E-06	
Kobber (Cu), (so)	i.o.		0,00604		0,00867	
Crom (Cr), (so)	i.o.		0,00178		0,00255	
Nikkel (Ni), (so)	i.o.		0,00124		0,00179	
Zink (Zn), (so)	i.o.		0,0142		0,0204	

i.o. ikke opgjort i livscyklusanalysen (tungmetal i slaggen)

Stort set samtlige af de miljøeffekter, som det har været muligt at kvantificere i livscyklusanalysen, er prissat. Stort set samtlige miljøeffekter,

som er opgjort kvantitativt, er prissat og indgår i den velfærdsøkonomiske analyse. Analysen viser, at værdien af de prissatte miljøeffekter kun udgør mellem fem og ti procent af de samlede velfærdsøkonomiske nettoomkostninger afhængig af, om det organiske affald bliver bioforgasset, forbrændt eller komposteret.

Værdisætningen er forbundet med store usikkerheder. Blandt andet er estimaterne for tungmetaller taget fra et norsk studie, hvilket kan være problematisk at overføre til danske forhold. Derfor skal man anskue værdisætningen som et overslag over størrelsesordenen af miljøkonsekvenserne.

Der er nogle fordele ved genanvendelse, som det ikke har været muligt at kvantificere og dermed heller ikke prissætte. Det skyldes blandt andet, at omfanget af effekterne er usikkert. Det gælder eksempelvis forbedret jordstruktur og vandbindingsevne, reduceret behov for pesticidanvendelse og eventuel forbedret slagge kvalitet fra forbrændingen af restaffaldet. Men man kan konstatere, at de forøgede omkostninger ved todelte indsamling er så store i forhold til de værdisatte miljøeffekter, at værdien af de ikke-prissatte miljøeffekter skal være overordentlig stor for at rykke resultatet.

Mål for genanvendelse vil årligt koste 70-270 mio.kr. ekstra.

Referencescenariet (Scenario 1) definerer den basisproduktion, som findes i samfundet i dag. I de scenarier, hvor der for eksempel bliver produceret mindre el og varme end i referencesituationen, bliver omkostningerne og miljøkonsekvenserne ved supplerende produktion af el og varme tilført. Ligeledes bliver omkostninger og miljøkonsekvenser ved produktion af handelsgødning trukket fra i de tilfælde, hvor affaldsbehandlingen fører til en større gødningsproduktion. Herved sikrer man, at der i alle scenarierne er den samme produktion af el, varme og gødning som i referencescenariet. Det bygger på en forudsætning om, at den nuværende produktion antages at være den nødvendige for samfundet.

Tabel 17.4 Velfærdsøkonomiske konsekvenser i de fem scenarier for den fremtidige håndtering af organisk dagrenovation

	Scenario 1 Reference	Scenario 2 100.000 t. bioforgasng.	Scenario 3 300.000 t. bioforgasng.	Scenario 4 100.000 t. c. kompst.	Scenario 5 300.000 t. c. kompst.
Samlede omkostninger, mio.kr.	1529	1606	1788	1584	1702
Samlet produktion					
• EI	378 TJ	410 TJ	483 TJ	351 TJ	257 TJ
• Varme	1761 TJ	1719 TJ	1610 TJ	1600 TJ	1153 TJ
• N (kvælstof)	130 ton	437 ton	1162 ton	357 ton	854 ton
• P (fosfor)	40 ton	97 ton	231 ton	117 ton	311 ton
• K (Kalium)	99 ton	236 ton	566 ton	287 ton	761 ton
I alt-omkostninger og miljøkonsekvenser fra den kompenserende/supplerende produktion, mio.kr.		-8,0	-25,2	22,7	92,8
Omkostninger i alt for scenariet, mio.kr.	1529	1598	1763	1606	1795
Meromkostning i forhold til scenario 1, mio.kr.	-	69	234	77	266

De samlede velfærdsøkonomiske meromkostninger ved at indføre det langsigtede mål fra Affald 21 med bioforgasning af 300.000 ton indsamlet organisk dagrenovation er 234 mio.kr. årligt (se tabel 17.4). En tilsvarende udbygning af central kompostering vil årligt indebære en meromkostning på 266 mio.kr. Affald 21's mål for bioforgasning i 2004 vil årligt koste 69 mio.kr. ekstra i forhold til den nuværende håndtering af det organiske affald. En udbygning med central kompostering vil koste 77 mio.kr. ekstra om året.

Den budgetøkonomiske analyse

Den budgetøkonomiske analyse viser de direkte økonomiske konsekvenser for borgerne af at indføre todelt indsamling af dagrenovation med efterfølgende forbehandling og bioforgasning eller central kompostering af den organiske dagrenovation.

Borgerne skal pr. ton indsamlet organisk affald betale mindst 306 kr.

mere for enfamilieboliger og mindst 722 kr. mere for etageboliger for at få genanvendt det organiske affald ved bioforgasning eller central kompostering sammenlignet med forbrænding (se tabel 17.5). Det svarer til en årlig merudgift på mindst 56 kr. pr. husstand for enfamilieboliger og 64 kroner pr. husstand for etageejendomme (med de gennemsnitlige affaldsmængder).

Tabel 17.5 Budgetøkonomiske omkostninger ved forbrænding, bioforgasning og central kompostering af et ton organisk dagrenovation. Kr. pr. ton indsamlet organisk dagrenovation

	Forbrænding		Bioforgasning		Central kompostering	
	Enfamilie	Etage	Enfamilie	Etage	Enfamilie	Etage
Indsamling i alt	1002	473	1754-3415	1640-1830	1754-3415	1640-1830
Behandling i alt	966		520		484	
Samlede budgetøkonomiske omkostninger	1968	1438	2274-3934	2160-2350	2238-3899	2134-2268
Forskel i forhold til forbrænding	-	-	306-1966	722-911	270-1931	686-876

Behandlingsomkostningerne dækker forbehandling og behandling inklusive forbrænding af rejekt for biogas og central kompostering, hvor der antages henholdsvis 35 procent og 15 procent rejekt.

Følsomhedsanalyser

Der er udført en lang række følsomhedsanalyser, hvor betydningen af centrale forudsætninger i analysen kan vurderes.

Da omkostningerne til indsamling er afgørende for, at genanvendelse samfundsøkonomisk er en dyr løsning, er der beregnet break-even omkostninger for todelt indsamling. Beregningerne belyser, hvor meget de årlige indsamlingsomkostninger skal reduceres i den velfærdsøkonomiske analyse, for at henholdsvis bioforgasning og central kompostering bliver lige så attraktivt samfundsøkonomisk som forbrænding. Det er omregnet til, hvad meromkostningen må være pr. husstand i budgetøkonomiske priser.

I basisanalysen er meromkostningen for todelt indsamling beregnet på baggrund af fuldskalaforsøgene i hovedstadsområdet, Århus og Aalborg. De viser en meromkostning på cirka 150 kr. pr. husstand om året for enfamilieboliger og cirka 110 kr. pr. husstand om året for etageejendomme. Meromkostningen for todelt indsamling skal ned under 50 kr. pr. husstand

om året for enfamilieboliger og under 20 kr. pr. husstand om året for etageejendomme for, at bioforgasning bliver mere attraktiv end forbrænding, når der ikke ændres på andre forudsætninger. For central kompostering skal meromkostningen helt ned under henholdsvis 30 kr. pr. husstand om året og 10 kr. pr. husstand om året.

Følsomhedsanalyserne, hvor der ændres på enkelte forudsætninger, ændrer gennemgående ikke på rangordenen mellem de forskellige behandlingsalternativer.

Et »optimistisk genanvendelsesscenarie« med de mest optimistiske forudsætninger om den indsamlede mængde, mængden der må bortsorteres og biogaspotentiallet bevirker, at bioforgasning bliver lidt mere attraktivt end forbrænding. Det gælder dog kun for enfamilieboliger og ikke for etageboliger.

På den baggrund må man konkludere, at rangordenen mellem behandlingsalternativerne er robust. Forbrænding er den mest attraktive løsning, efterfulgt af bioforgasning, mens central kompostering er den dyreste løsning for samfundet.

Anvendelse af analysen

Analysen og analysens resultater indgår i regeringens affaldsstrategi, som fastlægger målene i affaldshåndteringen i den kommende tid. På baggrund af analysen samt en række tekniske projekter er det valgt ikke at indføre obligatorisk todelt indsamling i Danmark, men i stedet lade det være op til den enkelte kommune at vurdere, hvilken indsamlingsmetode der vil være mest hensigtsmæssig for den pågældende kommune (Regeringen 2001). En række kommuner har, siden analysen udkom, valgt at gå væk fra den todelte indsamling af husholdningsaffald.

18 **Evalueringer på kulturområdet**

Trine Bille

Evalueringer på kulturområdet er et forholdsvis nyt fænomen, og det er samtidig et område, som inden for de sidste 10-15 år har været i kraftig vækst både med hensyn til metodeudvikling og med hensyn til konkrete gennemførte evalueringer.

Kulturen, kulturpolitikken og kulturektoren, dvs. teater, film, musik, billedkunst, kulturarv, litteratur, radio og tv etc. med tilhørende institutioner og aktører, har traditionelt ikke på samme måde som andre sektorer været underlagt evalueringer og målinger. Måske fordi kunst og kultur ikke kan »måles og vejes«, kvalitetsvurderinger er vanskelige, og det har været vanskeligt at se, hvordan og efter hvilke kriterier fx værdien og effekten af offentligt støttet kunst og kultur skal måles. Tidligere har man derfor i stor udstrækning afholdt sig fra dette, og »evalueringer« på området var stort set begrænset til humanistiske analyser af først og fremmest de enkelte værkers indhold og æstetik.

I de senere år er samfundsvidenskaberne imidlertid i større udstrækning begyndt at interessere sig for denne sektor, herunder for evalueringer af området. Dette skyldes nok primært den generelle tendens i samfundet, som går i retning af i stigende grad at skulle legitimere og argumentere for offentlige udgifter til forskellige formål, og deraf følgende krav om evalueringer af politikker, institutioner og projekter. New Public Management-bølgen i den offentlige sektor, herunder den efterhånden udbredte anvendelse af resultatkontrakter, også med kulturinstitutioner, har yderligere medvirket i denne retning.

18.1 Særlige udfordringer for evalueringer på kulturområdet

Når man begynder at tale om evalueringer på kulturområdet, må man først og fremmest gøre sig klart, at der er betragtelige grundlæggende problemer med at foretage analyser af, hvad kultur gør ved mennesker, og af, hvad denne påvirkning har af fx samfundsøkonomisk betydning. I praksis findes der vel næppe nogen umiddelbar farbar vej til at måle denne type konsekvenser, som ellers må antages at være helt væsentlige ved evaluering af kultur.

Herudover knytter vanskelighederne med at gennemføre evalueringer på kulturområdet sig først og fremmest til vanskelighederne med at vurdere *kunstnerisk kvalitet*. På det overordnede, filosofiske plan diskuteres det, om kunstnerisk kvalitet er noget »iboende«, der knytter sig til et bestemt kunstværk (jf. fx Favrholtts parameterteori, (Favrholt 2000)), eller om det i højere grad er socialt betinget og hører hjemme i kunstens sociale omgivelser (jf. fx Kjærups institutionsteori, (Kjærup 2000)). Eller er der tale om et både-og? Langsted, Hannah og Larsen (2003) indtager denne både-og-position, »at kvalitet både er noget, kunsten iboende, og noget, nogen tilskriver kunsten. Det kan lyde som en selvmodsigelse, men er egentlig udtryk for, at kvalitet udspiller og udfolder sig i mødet med kunsten. Det lader sig altså gøre på et vist abstraktionsniveau at analysere kvalitet i kunst, uden at der blot er tale om private meninger, eller at man uforbedreligt synes det ene eller det andet. Præmisserne i mødet med kunsten fastlægges både af kunstværket og af den, der oplever det« (s. 115). Hermed følger også, at kvalitet i kunsten ikke alene kan reduceres til oplevet kvalitet. Man kan altså ikke afgøre, om noget er kvalitativt godt ved blot at gå ud og spørge publikum og så sammenholde deres forventninger med det, de faktisk fik (s. 84).

Men kunstnerisk kvalitet er som en regel kun et af flere betydende elementer, når en kulturinstitution eller et kulturprojekt skal evalueres, idet der vil være andre kulturpolitiske målsætninger, herunder fx målsætninger om at nå bestemte målgrupper, som gør sig gældende. Hvis en kulturinstitutionens mål er at nå et bestemt publikum, er publikums oplevelse af kvaliteten ikke helt uvæsentlig. Hvor væsentlig den kunstneriske kvalitet er i for-

hold til andre kulturpolitiske målsætninger, vil bl.a. afhænge af, hvilken kulturinstitution der er tale om. Det er således vigtigt at gøre sig klart, hvad der er formålet med evalueringen, hvilke målsætninger, formålsbestemmelser m.m. der skal evalueres, hvilke effekter der skal måles, og om og hvordan de kan måles. På kulturområdet er det måske særligt vigtigt, fordi de kulturpolitiske mål ofte er mangfoldige og ikke særligt konkrete.

Herudover er der det økonomiske aspekt: Hvis der skal prioriteres, er det væsentligt at have et mål for, i hvilken udstrækning værdien/effekten/kvaliteten af en institution eller et projekt står mål med omkostningerne. Inddragelsen af dette aspekt stiller særlige krav til evalueringen.

18.2 **Evalueringsmodeller på kulturområdet**

Der er efterhånden udviklet en række evalueringsmodeller på kulturområdet, som alle har deres styrker og svagheder, og som alle lægger vægt på forskellige evalueringskriterier. I det følgende skal kort nævnes nogle af de væsentligste evalueringsmodeller, som der evt. kan søges inspiration i, afhængigt af hvilken type evaluering man ønsker at gennemføre. Nogle af modellerne er udviklet til at evaluere kulturpolitik, mens andre er mere rettede mod evaluering af institutioner, projekter eller enkelte forestillinger/værker. Nogle af modeller kan anvendes bredt til forskellige typer af evalueringer.

Europarådets evalueringer af forskellige landes kulturpolitik

Et af de mest omfattende evalueringsprojekter inden for kulturområdet er Europarådets evalueringer af forskellige landes kulturpolitik, der blev initieret i 1985, og som indtil nu har omfattet 18 lande, heriblandt dog ikke Danmark.³² Frankrig og Sverige var de første lande, der blev evalueret, og det svenske Kulturråd udarbejdede i denne forbindelse en omfattende evalueringsmodel med henblik på at foretage systematiske analyser af mål, midler og resultater i kulturpolitikken (Utbildningsdepartementet, 1990).³³ Det, der skulle undersøges, var: 1) Hvad ville vi? Hvad var de kulturpolitiske målsætninger? 2) Hvad gjorde vi? Hvilke konkrete indsatser blev der gjort for at opfylde målene? 3) Hvad er der sket? Hvilke effekter, resultater og udfald har indsatserne givet i forhold til målene? (jf.

Langsted m.fl., 2003, s. 49). Efter murens fald i 1989 er også en række østeuropæiske lande blevet evalueret, hvilket har medført en ændret evalueringspraksis, hvor evalueringerne i højere grad er blevet fremadrettede og udviklingsorienterede, og dermed i højere grad end tidligere har fået karakter af *procesmodellen*. Ellers kan Europarådets evalueringer vel primært karakteriseres som evalueringer efter *målopfyldelsesmodellen*, med elementer af *effektmodellen* (selv om årsags-virknings-sammenhængene – og dermed effekterne – godt kan være problematiske at måle), og de er forholdsvis deskriptive og kvantitativt orienterede, hvor bl.a. kulturvaneundersøgelser inddrages som et væsentligt element. Generelt kan man vel sige, at de fokuserer på rammer og organisering frem for indhold og kvalitet. Det sammen gælder mange andre evalueringer på kulturområdet, hvor et par danske eksempler er evalueringen af de regionale kulturforsøg (Nellemann Konsulenterne, 1999) og konsekvensanalysen af Kulturby '96 (Fridberg og Koch-Nielsen, 1997).

Det engelske Arts Council's evalueringsmodel

I England er der via Arts Council opbygget et ganske omfattende evalueringssystem, som omfatter alle kulturinstitutioner, der får støtte via Arts Council. Modellen er nøje beskrevet i *Appraisal Handbook* (1994), der mest af alt kan beskrives som en »køgebog«, og som bl.a. skal sikre ensartethed i evalueringerne. Evalueringssystemet omfatter:

- årlige kontrakter med institutionerne.
- regelmæssige show reports vedr. institutionernes produktioner, der udarbejdes af kunstfaglige eksperter.
- årlige reviews, der er Arts Councils tilsynsrapporter, og som indeholder en gennemgang af, hvorvidt institutionen har indfriet sine kunstneriske og økonomiske mål, som er fastlagt i kontrakten. De indkomne show reports indgår i vurderingen.
- apprecials, der er dybdegående evalueringer, som foretages mindst hvert 5. år, af de store institutioner og tilskudsmodtagere. Disse apprecials har til formål at vurdere institutionen på et kunstnerisk, et organisatorisk og et strategisk plan, og de udarbejdes bl.a. på baggrund af en selvevalueringsrapport, som institutionerne skal udarbejde forud for den eksterne evaluering.

I forhold til de generelle evalueringsmodeller, der er opstillet i kapitlet om evalueringssystemet, modeller og paradigmer i denne bog, kan det engelske evalueringssystem primært karakteriseres som en *målopfyldelsesmodel*, der dog indeholder elementer af andre modeller, ikke mindst *kollegiale vurderinger (peer reviews)*, men til en vis grad også *procesmodellen*. Der er både tale om evalueringer af den kunstneriske kvalitet samt en række andre målsætninger og forhold vedr. institutionernes virke. Der er således i høj grad tale om vurderinger af hele institutionernes virke, og muligheden for sammenligning med tilsvarende institutioner indgår som et væsentligt element. Et dansk eksempel på en evaluering (nok den eneste), der konkret har brugt den engelske model som inspiration, er den omfattende midtvejsevaluering af Kulturministeriets Udviklingsfond (Balling m.fl., 2001).

Det Kgl. Teaters kvalitetsprojekt

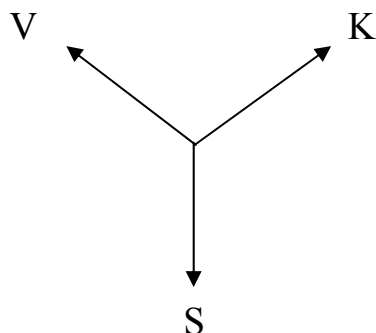
I Det Kgl. Teaters resultatkontrakt med Kulturministeriet for perioden 1996-99 var et af kravene til teatret en højnelse af den kunstneriske kvalitet, herunder et krav om, at teatret skulle gennemføre et udviklingsprojekt med henblik på en registrering og optimering af forestillingernes kvalitet. På denne baggrund satte teatret et forprojekt i gang med en række eksterne forskere, der mundede ud i en ikke offentliggjort rapport fra 1997. Der blev stillet en række forslag i rapporten, som omfattede måling og registrering af opfattelsen af forestillingernes kvalitet blandt forskellige interessenter, herunder publikum (publikumsundersøgelser), anmeldere, eksterne eksperter og teatret selv (selvevaluering), jf. Bille Hansen (1997), Jarl og Svendsen (1997) samt Jarl (2001). I forhold til de generelle evalueringsmodeller er der nærmest tale om en *interessentmodel*, hvor der er fokus på flere interessenters vurdering af kvaliteten.

Egentlige evalueringer efter *brugermodellen*, hvor der alene er fokus på brugernes vurdering af indsatsen, gennemføres også på kulturområdet, fx publikumsundersøgelser på teaterområdet, bruger- og tilfredshedsundersøgelser på biblioteksområdet o.l.

Ønskekvistmodellen

Senest har professor Jørn Langsted arbejdet med et projekt vedr. kvalitetsvurdering, som bl.a. udsprang af et behov i Århus Kommune for at evaluere sine kulturinstitutioner. Projektet har udmøntet sig i en evalueringsmodel »ønskekvistmodellen« til vurdering af kunstnerisk kvalitet i performativ kunst, dvs. teater, dans og musik (Langsted, Hannah og Larsen, 2003). Modellen er på sin vis enkel. Tre vektorer med 120 grader imellem. De tre vektorer afspejler hhv. *villen*, *kunnen* og *skullen* – eller med andre ord engagement, evner og nødvendighed – og pilenes længde indikerer styrken af de tre elementer.³⁴ Modellen er ikke normativ forstået på den måde, at det perfekte kunstværk skal rumme lige lange pile, sådan at det hele er i balance. Tværtimod er det ideen, at den kan bruges til at blive mere bevidst om, at nogle i nogle udsagn om kunstnerisk kvalitet lægger mest vægt på kunnen, at nogle i andre sammenhænge måske lægger mest vægt på villen eller på skullen, og at nogle i nogle tredje kontekster lægger vægt på sammenhængen mellem nogle af eller alle pilene (Langsted m.fl., 2003, side 154).

Figur 18.1 Ønskekvistmodellen



Villen og kunnen kan groft oversættes til kunstnerisk vilje og teknisk kvalitet (forekommer i kendt vurderingspraksis af kunstnerisk kvalitet), mens skullen afspejler værkets samfundsmæssige nødvendighed – dets kulturpolitiske relevans – og er som sådan nok den vanskeligste at operationalisere. Herudover angiver modellen nødvendigheden af at medtage det tidsmæssige perspektiv, altså at noget med tiden kan blive godt eller dårligt, at noget kan rumme potentialer. Desuden angives, at den kunstneriske

kvalitet er relateret til omgivelserne, i særlig grad til de ressourcer, man råder over, og til de kulturpolitiske målsætninger.

Ideen med modellen er at skabe et samtalerum om kvalitet, der i en dialogproces må operationaliseres. Herunder skal evalueringskriterierne aftales på forhånd i et startdokument mellem den, der skal evalueres, evaluatør og den, der iværksætter evalueringen. Der er behov for en ganske præcis diskussion og fastlæggelse af indikatorer og spørgsmål mellem parterne, inden evalueringen sættes i gang.

Modellen er således forholdsvis generel og kan anvendes til evalueringer af såvel projekter, enkeltopsætninger, institutioner og kulturpolitik. Den bygger grundlæggende på *kollegial vurdering (peer review)* af den kunstneriske kvalitet. De kommende år vil vise, om modellen består sin »praktiske prøve« og vinder udbredelse i forvaltninger o.l.

Cultural Score Card

Cultural Score Card er et forholdsvis nyudviklet system, der har til hensigt på forhånd at kunne udpege, om et kulturprojekt bliver succes eller fiasko (se fx Jacobsen, 2002 og Rosenkrands, 2002). Modellen lokaliserer forbedringspotentialer i et kulturprojekt ved at påvise uoverensstemmelser mellem målsætninger og faktiske tiltag. Den er baseret på en række spørgsmål, som projektmagerne besvarer, og den bygger på pointgivning af væsentlighed af fire faktorer, som indkredser målet med projektet: 1) kunstnerisk kvalitet/professionalisme, 2) evnen til at gennemføre, 3) markedsvurdering og 4) formål/budskab/aktualitet. Dernæst vurderes succeskriterierne på fire felter: a) kommercielt/økonomisk, b) fagligt/kulturelt, c) marked/image og d) andre mål. Svarene på spørgsmålene omdannes til tal. Disse talværdier omdannes til nogle grafiske fremstillinger (score cards), hvor man kan sammenligne det maksimalt mulige pointtal med det pågældende projekts. Hermed skulle projektets styrker og svagheder kunne identificeres, og det kan påvises, hvor projektet bør styrkes. Cultural Score Card må karakteriseres som en *målopfyldelsesmodel*, der dog ikke forholder sig til den kunstneriske kvalitet. Det er en model, der kan anvendes ex ante, og ikke som normalt ex post, og modellen henvender sig således også til fx potentielle sponsorer, der skal vurdere, om de bør vælge det ene projekt frem for det andet.

Economic impact-analyser

Economic impact-analyser anvender en kendt økonomisk metode, nemlig input-output-modeller til at måle den økonomiske gevinst i form af forbrug, indkomst og beskæftigelse i et afgrænset geografisk område af en kulturinstitution, et kulturarrangement eller kulturlivet i en by, en region eller et helt land. Den første economic impact-analyse på kulturområdet blev lavet i 1977 (Cwi og Lyall 1977), og siden er det gået stærkt. Internationalt er der efterhånden lavet en enorm mængde af economic impact analyser på kulturområdet.

I forhold til de generelle evalueringsmodeller, der er beskrevet i temadelen, er der tale om *effektmodeller*, der alene måler sideeffekterne. Problemet er, at de ikke tager udgangspunkt i formålet med kulturaktiviteter og kulturliv (herunder kvalitet og andre kulturpolitiske målsætninger). Denne type analyser, der er meget udbredte på kulturområdet, ikke mindst internationalt, uddybes i et eksempel nedenfor.

Værdisætningsanalyser

Værdisætningsanalyser har udgangspunkt i økonomisk velfærdsteori og anvender en metode, der er udviklet inden for miljøøkonomien til at værdisætte offentlige goder, der ikke handles på markedet og dermed ikke har nogen »pris«, der kan udtrykke deres værdi. Disse metoder er i stigende grad inden for de sidste 5-10 også blevet anvendt på kulturområdet. Udgangspunktet for analyserne er (som i økonomisk teori generelt), at det er befolkningens samlede præferencer og værdisætning, der lægges til grund for beregningen af den samfundsøkonomiske værdi af godet. En grundlæggende præmis, som måske i hvert fald nogle nok vil sætte spørgsmålstegn ved i forbindelse med værdisætning af kultur, idet nogle nok vil hævde, at det i langt højere grad bør være eksperter end den gennemsnitlige befolkning, der vurderer dette. Den overordnede fordel ved metoden er omvendt helt klart, at den tager udgangspunkt i selve kulturaktiviteten/institutionen og med udgangspunkt i borgernes præferencer og betalingsvillighed værdisætter alle de forskellige formål og effekter, kulturaktiviteten måtte have, hvorved resultaterne evt. kan anvendes i *økonomiske cost-benefit-analyser*. Metoden uddybes i det andet eksempel nedenfor.

Afrunding

Alt i alt viser ovenstående korte gennemgang af forskellige evalueringsmodeller på kulturområdet, at stort set alle de generelle evalueringsmodeller, der blev omtalt i temadelen, kan anvendes og bliver anvendt på kulturområdet. Som i alle andre evalueringssammenhænge afhænger valget af evalueringsmodel af formålet med evalueringen – hvad er det vi vil evaluere og ønsker at vide noget om? En af de særlige vanskeligheder på kulturområdet er evalueringen af den kunstneriske kvalitet – et område, som fx ønskevistmodellen går ind i kernen af, mens andre modeller springer let og elegant hen over ved i højere grad at fokusere på det mere umiddelbart kvantificerbare, på rammer og organisering. En evaluering af den kunstneriske kvalitet vil til gengæld ikke være tilstrækkelig i mange sammenhænge, hvor der er andre kulturpolitiske målsætninger af væsentlig betydning, og hvor prioritering af økonomiske midler indgår som et væsentligt formål med evalueringen.

I de næste afsnit beskrives to konkrete eksempler på evalueringer på kulturområdet, som AKF har gennemført. Det første eksempel er en economic impact-analyse af Brandts Klædefabrik, som AKF har gennemført for Kulturministeriet. Det er et eksempel på en effektevaluering og på en metode og en tænkning, som har været meget anvendt – og som stadig anvendes. Det næste eksempel er en undersøgelse af danskerne betalingsvillighed for Det Kgl. Teater. Et eksempel, som tager udgangspunkt i en økonomisk model og en økonomisk tænkning af en helt anden karakter, som stadig er langt mindre udbredt, men som på mange måder må anses for at være mere relevant ved evaluering af kultur.

18.3 **Economic impact-analyse af Brandts Klædefabrik i Odense**

Kulturministeriet ønskede for flere år siden at få udarbejdet en empirisk analyse af de afledte økonomiske effekter af en kulturinstitution. En undersøgelse, som AKF gennemførte (Bille Hansen 1993 samt Christoffersen og Lyk-Jensen 1994). Som case blev valgt Brandts Klædefabrik i Odense. Brandts Klædefabrik er en nedlagt klædefabrik, som efter virk-

somhedslukning i 1977 blev ledigt til andre formål. På kommunalt initiativ blev komplekset i 1987 åbnet som et center for mange kulturelle aktiviteter, herunder udstillingsstederne Kunsthallen Brandts Klædefabrik, Museet for Fotokunst samt Danmarks Grafiske Museum og Dansk Pressemuseum.

Overordnet formål med evalueringen

Det overordnede formål med evalueringen var, som nævnt, at få belyst de afledte økonomiske effekter af en kulturinstitution. Baggrunden for undersøgelsen var en udbredt opfattelse af, at kulturinvesteringer har stor betydning for den økonomiske udvikling – at kultur betaler sig. Denne opfattelse havde ikke mindst udgangspunkt i en lang række (primært udenlandske) kulturøkonomiske undersøgelser – economic impact-analyser – som påviste ikke-ubetydelig økonomiske effekter af kultur på forbrug, indkomst og beskæftigelse. Kulturministeriet ønskede derfor at få metoden og resultaterne afprøvet i dansk sammenhæng.

Evalueringsmodel og design

Evalueringsmodellen var således valgt på forhånd, nemlig en effektmodel, nærmere bestemt en economic impact-analyse, hvor Brandts Klædefabriks betydning for forbrug, indkomst og beskæftigelse skulle belyses.

Det er en grundbetragtning i velfærdsøkonomien, at en analyse af økonomiske effekter af enkelte kulturinstitutioner som Brandts Klædefabrik først for alvor får værdi, når den sættes i forhold til analyser af andre mulige anvendelser af de offentlige tilskud. Påvisning af positive efterspørgselseffekter på forbrug, indkomst og beskæftigelse er således ikke i sig selv voldsom interessante. Sådanne positive effekter vil opstå overalt i samfundet, hvor der sættes skattefinansierede aktiviteter i gang. Det afgørende må derfor være, om der med det samme beløb kan opnås en større beskæftigelsesfremgang andre steder.

Undersøgelsen af Brandts Klædefabrik var mere eller mindre den første analyse, der systematisk sammenlignede de økonomiske effekter af en kulturinstitution med alternative måder at anvende de offentlige penge på. Alternativbetragtningen blev opdyrket gennem to supplerende analyser. For det første blev effekterne af Brandts Klædefabrik sammenlignet med

effekterne af en anden kulturinstitution i Odense, nemlig H.C. Andersens Hus. For det andet blev effekterne af Brandts Klædefabrik sammenlignet med effekterne af en generel udvidelse af de offentlige aktiviteter.

Dataindsamling og analysemetode

Analysen foregik i to trin. Først blev den efterspørgsel, som Brandts Klædefabrik medfører, kortlagt ved hjælp af publikumsinterview m.m. I alt blev der gennemført ca. 2.200 personlige interview med publikum på stedet. For det andet blev de regionale og nationaløkonomiske effekter beregnet ved hjælp af samfundsøkonomiske modeller. Dette blev gjort for regionen Fyn ved anvendelse af AKF's regionaløkonomiske model (EMIL) og for hele landet ved anvendelse den nationaløkonomiske model for Danmark, ADAM.

Hovedresultater

Resultaterne af analyserne viste, at beslutningen om at etablere Brandts Klædefabrik har ført til økonomisk vækst og stigende indkomster på Fyn. Men de beregnede økonomiske effekter af Brandts Klædefabrik adskiller sig stort set ikke fra effekterne af at foretage en generel udvidelse af den offentlige sektor jævnt fordelt over samtlige offentlige udgiftsposter. Der opnås stort set den samme fremgang i forbrug, indkomst og beskæftigelse, ligesom skattetrykket stort set må forhøjes i samme grad ved de to alternative anvendelser. Derimod påvises det, at der med denne efterspørgselsorienterede betragtningsmåde kan beregnes større økonomiske effekter af H.C. Andersens Hus end af Brandts Klædefabrik. Når Brandts Klædefabrik og H.C. Andersens Hus udviser så fundamentalt forskellige samfundsøkonomiske effekter, er hovedforklaringen, at H.C. Andersens Hus er en kulturinstitution, som tiltrækker virkelig mange turister, mens Brandts Klædefabrik i højere grad henvender sig til et lokalt publikum.

Hovedkonklusionen er derfor, at med den efterspørgelseorienterede betragtningsmåde kommer en offentlig aktivitet især ud med anderledes og bedre resultater, hvis der er tale om egentlig produktion af goder til eksport, og det er netop tilfældet med H.C. Andersens Hus.

Herved er det også forklaret, hvorfor Brandts Klædefabrik i de gennemførte analyser kommer ud med stort set samme samfundsøkonomiske

effekter som et gennemsnitligt offentligt projekt. Ligesom Brandts Klædefabrik kun i ringe grad tiltrækker efterspørgsel udefra via turisme, så tiltrækker vuggestuer, skoler og plejehjem ikke turister eller andet forbrug udefra.

Formidling og anvendelse

Resultaterne blev publiceret i en AKF-rapport (Christoffersen og Lyk-Jensen 1994). Generelt har economic impact-analyserne været genstand for en omfattende debat, ikke mindst fordi de i høj grad har været anvendt legitimerende og politisk til at vise, at »kultur betaler sig«. Et udsagn, der bestemt kan diskuteres (jf. bl.a. Bille Hansen (1993) og ovenstående eksempel). Fordelen ved economic impact-analyserne er bl.a., at de anvender en velkendt metode inden for økonomisk teori. Korrekt gennemførte economic impact-analyser, vil typisk komme til det resultat, at jo flere turister der tiltrækkes udefra af en kulturinstitution/aktivitet, jo større økonomisk resultat i form af forbrug, indkomst og beskæftigelse vil analysen komme ud med (jf. ovenstående eksempel). Dermed er deres brugbarhed som udgangspunkt for en kulturpolitisk argumentation imidlertid begrænset, idet der selvsagt er mange andre kulturpolitiske målsætninger end at tiltrække turister, og det langt fra er sikkert, at det er den kultur, der på kort sigt tiltrækker besøgende udefra, der er væsentlig i forhold til en række andre kulturpolitiske målsætninger og kriterier – herunder spørgsmålet om den kunstneriske kvalitet.

Når denne 10 år gamle undersøgelse trækkes frem i denne sammenhæng, er det, fordi den på mange måder stadig er interessant. For det første indgår den tænkning og argumentation, der ligger bag economic impact-analyserne, stadig i den politiske argumentation. For det andet kan economic impact-analyser være interessante at gennemføre (og de er stadig meget udbredte, ikke mindst i udlandet), hvis de gennemføres korrekt, og hvis man holder sig for øje, hvad de kan anvendes til (og ikke anvendes til). Således kan de fx være interessante at gennemføre i forbindelse med større festivaler eller andre kulturarrangementer, hvor et af hovedformålene er at skabe opmærksomhed, og ikke mindst tiltrække besøgende og forbrug udefra. Et eksempel kunne være koncerter med Elton John eller Sir Poul McCartney i Horsens, som tiltrækker mange besøgende fra hele landet.

På idrætsområdet er denne type analyser meget anvendte (og anvendelig), og Idrætsfonden Danmark gennemfører fx jævnligt effektanalyser af større sportsarrangementer, fx EM-arrangementer o.l., og analyserne laves inden for en overskuelig økonomisk ramme.

18.4 **Værdisætning af Det Kgl. Teater**

De fleste vil formodentlig være enige i, at kulturens værdi skal findes i dens betydning for borgernes livskvalitet, identitetsdannelse, kreativitet m.m. Værdier, som ikke alene kommer den enkelte »kulturforbrugende« borger, men hele samfundet til gode. På denne måde kan kultur betragtes som et offentligt gode – eller som et mikset gode, dvs. et privat gode, hvortil der er knyttet positive eksterne effekter. Den direkte nytte, som er knyttet til kulturen, fx oplevelsen af en teaterforestilling, kan sælges på markedet som ethvert andet privat gode. De eksterne effekter – som kommer både brugere og ikke-brugere til gode – har ikke nogen pris på markedet, og kan derfor ikke gøres til genstand for en almindelig markedsvurdering.

Overordnet formål med evalueringen

AKF gennemførte for nogle år siden på eget initiativ³⁵ en evaluering af Det Kgl. Teater, som tog udgangspunkt i disse overordnede betragtninger, dvs. en samfundsmæssig værdisætning af Det Kgl. Teater som offentligt kulturgode (Bille Hansen, 1996). Denne undersøgelse var internationalt en af de første af sin slags, men siden er mange fulgt efter – dog ikke i Danmark, men i udlandet (se fx Navrud og Ready (ed.), 2002). Der var således ingen egentlig rekvirent til undersøgelsen, men det kunne der i princippet lige så godt have været.³⁶

Det Kgl. Teater blev valgt som case, da det er et meget veldefineret kulturgode. Det Kgl. Teater blev grundlagt i 1748 og er Danmarks nationalteater. Alle danskere kender det, og stort set alle har en holdning til det. På den anden side er Det Kgl. Teater også en af Danmarks mest elitære kulturinstitutioner forstået på den måde, at kun ganske få procent af befolkningen faktisk kommer i teatret. Ca. 7% havde på tidspunktet for undersøgelsen været i Det Kgl. Teater inden for det seneste år, og ca. to tred-

jedele havde aldrig været der. Teatret modtog på tidspunktet for undersøgelsen ca. 258 mio.kr. i støtte fra staten, hvilket svarede til mere end 80% af teatrets totale budget. Dette betyder alt i alt, at ikke-brugerne finansierer størstedelen af teatrets drift over deres skattebillet. Det forekommer derfor interessant at undersøge, om de danske borgere generelt synes, at Det Kgl. Teater »er pengene værd«, og hvad der i givet fald motiverer deres betalingsvillighed. En værdisætningsundersøgelse af Det Kgl. Teater kan således have en umiddelbar politisk interesse, idet den kan afsløre, om teatret er pengene værd set med skatteborgernes øjne. Hvis det er tilfældet, vil det være et langt stærkere argument for offentlig støtte end kulturaktivitetens afledte økonomiske effekter.

Evalueringsmodel og design

Teknisk set kan der anvendes forskellige metoder til at afdække borgernes betalingsvillighed for goder, som ikke handles på markedet. I denne undersøgelse anvendes contingent valuation-metoden, hvor folk interview direkte om deres villighed til at betale for et (offentligt) gode, fx over skatterne. Der findes også mere indirekte metoder, som rejseomkostningsmetoden og husprismetoden.³⁷ Disse metoder er alle meget anvendte inden for miljøøkonomien, hvor de anvendes til at bestemme den økonomiske værdi af fx mindre forurening, bevarelse af truede dyrearter etc., men contingent valuation-metoden er nok den mest udbredte. Metoden kan ideelt set tilvejebringe et økonomisk mål for den samlede værdi af et (offentligt) gode eller ændringer i udbuddet af godet, herunder værdien for såvel brugere som ikke-brugere af godet. En værdi, som ideelt set kan sammenholdes med omkostningerne ved at tilvejebringe eller opretholde godet. Der er dog en række tekniske problemer med at afsløre befolkningens »sande« betalingsvillighed, idet der nemt opstår forskellige former for skævvridning af resultaterne, fx hvis folk svarer strategisk eller på anden måde ikke angiver deres sande betalingsvillighed, men den mest udbredte opfattelse er, at disse problemer kan håndteres ved korrekt og omhyggelig anvendelse af metoden (jf. Arrow m.fl. 1993).

Der kan være flere forskellige motiver for, at folk, der ikke kommer i Det Kgl. Teater, alligevel gerne vil betale for det:

- Teatret leverer en række indirekte forbrugsydelser. Fx filmindustrien og medierne udnytter de skuespilfærdigheder og -talenter, der udvikles på Det Kgl. Teater. Derved har fx seerne af et tv-spil indirekte nytte af Det Kgl. Teater. En anden og måske mere væsentlig form for indirekte forbrug er tv-transmissioner af Det Kgl. Teaters forestillinger.
- Det Kgl. Teater har uddannelsesværdi, idet teatret indgår som en del af den almene dannelse og er med til at udvikle evner og kvalifikationer som fx æstetisk sans, social kritik, kreative ideer m.m.
- Det Kgl. Teater udgør en væsentlig del af den nationale kulturarv og identitet, og fortsat eksistens af Det Kgl. Teater har betydning for, at væsentlige dele af den nationale kulturarv, herunder ikke mindst Holberg- og Bournonvilletraditionen, bevares og er tilgængelig for fremtidige generationer.
- Det Kgl. Teater har prestige-værdi. Fx kan en international anerkendelse af balletten på Det Kgl. Teater bidrage til den nationale stolthed og identitet, gøre Danmark kendt i udlandet, tiltrække turister m.m.
- Endelig kan borgerne være villige til at betale en optionspris, dvs. de vil gerne betale for, at teatret bevares, således at de har mulighed for at komme der på et senere tidspunkt i deres liv.

Alt dette er værdier, som ikke værdisættes og sælges på et marked.

Dataindsamling og analysemetode

Undersøgelsen omfattede en tilfældigt udvalgt stikprøve af den danske befolkning over 16 år. I alt 1.843 personer er blevet interviewet pr. telefon. Disse personer er blevet spurgt om, hvad de maksimalt er villige til at betale til Det Kgl. Teater om året over deres skat for, at teatret opretholder sine aktiviteter på det nuværende niveau. Desuden er de blevet spurgt, om de mener, at Det Kgl. Teater har værdi for andre end dem, der selv kommer der, fordi det har betydning for landets kulturelle niveau, tiltrækker turister eller af andre grunde. De er blevet spurgt, om de selv kommer i Det Kgl. Teater, og i givet fald hvor ofte, samt om deres øvrige kulturelle aktiviteter. De er ligeledes blev stillet en lang række spørgsmål om deres socioøkonomiske status, herunder køn, alder, skoleuddannelse, er-

hvervsuddannelse, beskæftigelse, civilstatus, antal hjemmeboende børn, indkomst m.m.

Alle disse data blev efterfølgende analyseret ved hjælp af regressionsanalyser og forholdsvis avancerede statistiske modeller, med henblik på at undersøge bl.a. betalingsvilligheden blandt forskellige befolkningsgrupper.

Hovedresultater

Resultaterne viste, at der ikke er tvivl om, at borgernes betalingsvillighed er af en størrelsesorden, som i hvert fald står mål med det offentlige tilskud, der gives til teatret.

Ca. 18% af den danske befolkning er ikke villige til at betale noget til Det Kgl. Teater over deres skat. Umiddelbart kunne man godt have forventet, at denne andel havde været større. De direkte brugere af Det Kgl. Teater udgør en meget lille andel af befolkningen, og man kunne derfor godt forestille sig, at flere ville synes, at deres nytte af teatret er lig nul og derfor ikke er interesseret i at betale for det.

Brugerne af Det Kgl. Teater adskiller sig klart fra ikke-brugerne, der til gengæld ikke adskiller sig nævneværdigt fra gennemsnittet af befolkningen. Brugere adskiller sig markant fra gennemsnitsdanskeren først og fremmest med hensyn til, hvor i landet de bor, med hensyn til skole- og erhvervsuddannelse og med hensyn til, hvor kulturelle de i øvrigt er. Desuden er den gennemsnitlige betalingsvilje større for brugerne end for ikke-brugerne, hvilket ikke er særligt overraskende. Brugernes samlede betalingsvillighed udgør således mellem 12 og 18% af den totale betalingsvillighed, på trods af at brugerne kun udgør 7% af befolkningen. Men det betyder på den anden side også, at ikke-brugernes betalingsvillighed udgør størstedelen af den samlede betalingsvillighed. En positiv betalingsvillighed for ikke-brugerne kan som tidligere nævnt forklares med, at ikke-brugerne af forskellige årsager gerne vil betale for, at teatret eksisterer, og at de gerne vil betale en optionspris, dvs. betale for at opretholde muligheden for at kunne komme i teatret på et senere tidspunkt.

Opdeles den totale betalingsvillighed for Det Kgl. Teater i forskellige benefitkategorier, viser resultat at:

- Ca. 6-9% af den samlede betalingsvillighed kan henføres til et forbrugsoverskud for besøg på Det Kgl. Teater. Dvs., brugerne ville i vir-

keligheden gerne betale en højere billetpris, da et teaterbesøg er mere værd for dem end det, en teaterbillet koster.

- Optionsprisen udgør ca. 56-57% af den totale betalingsvillighed. Dvs. størstedelen af betalingsvilligheden kan henføres til, at borgerne gerne vil betale for at opretholde muligheden for at kunne komme i teatret en gang i fremtiden.
- Ca. 35-37% af den totale betalingsvillighed kan henføres til, at borgerne af forskellige grunde gerne vil betale for Det Kgl. Teaters eksistens, uanset om de selv kommer der, eller har til hensigt at komme der.

Hovedproblemet med denne type undersøgelser er, at borgerne stilles *hypotetiske spørgsmål*, og dermed har man ingen garanti for, at de svarer i overensstemmelse med deres sande præferencer. En måde at »tjekke« besvarelsen på er ved at undersøge, om det i datamaterialet er muligt at finde nogle sammenhænge, man på forhånd ville forvente. Fx er det i denne undersøgelse rimeligt at forvente, at en person, der kommer i Det Kgl. Teater, bor i Københavnsområdet, har en høj indkomst, en høj uddannelse og i øvrigt er meget kulturelt aktiv, vil være villig til at betale mere end en person med lav indkomst, der bor i Jylland langt væk fra Det Kgl. Teater. Dette viser sig at være tilfældet. Hvis respondenterne mener, at Det Kgl. Teater har eksistensværdi, i øvrigt er meget kulturelt aktiv, har studentereksamen og en videregående uddannelse og er kvinde, øger det sandsynligheden for at have en positiv betalingsvillighed, mens det at være arbejdsløs formindsker sandsynligheden for at have en positiv betalingsvillighed. Desuden er indkomsten og den geografiske placering af respondenterne afgørende for størrelsen af betalingsvilligheden. Med en højere indkomst øges sandsynligheden for at have en stor betalingsvilje, og jo længere væk fra hovedstadsregionen respondenterne bor, jo mindre er sandsynligheden for at have en stor betalingsvilje.

Formidling og anvendelse

Undersøgelsen indgik i en ph.d.-afhandling og er derfor publiceret der, samt i en række artikler i internationale videnskabelige tidsskrifter. Men resultaterne blev også formidlet til en bredere offentlighed bl.a. i en kronik i Berlingske Tidende, en artikel i AKF Nyt o.l.

Undersøgelsen viser, at det ved hjælp af en økonomiske værdisætningsmetode er muligt at argumentere for offentlige tilskud til en kulturinstitution. En argumentation, der tager udgangspunkt i borgernes præferencer og dermed borgernes vurdering af kvaliteten af kulturinstitutionen.

For det første viser resultaterne, at den danske befolkning faktisk gerne vil betale det, teatret koster, i offentlige tilskud. Selv om en stor del af befolkningen aldrig kommer i teatret, vil de gerne betale for teatrets eksistens og for muligheden for at kunne komme der. For det andet synes resultaterne af analysen at være forholdsvis valide, idet en række socioøkonomiske faktorer kan forklare betalingsviljen på en måde, der er i overensstemmelse med, hvad man ville forvente.

Hovedbegrundelsen for at give offentlig kulturstøtte må antages at være kulturens betydning for borgernes livskvalitet, identitetsdannelse, kreativitet m.m. Værdier, som ikke alene kommer den enkelte »kulturforbrugende« borger, men hele samfundet til gode. Contingent valuation-metoden har vist sig at være velegnet til at estimere den økonomiske størrelsesorden af disse værdier med udgangspunkt i borgernes præferencer, og dermed kan resultaterne være anvendelige i politiske prioriteringer.

Forfatterne

Trine Bille

Trine Bille er seniorforsker i AKF og ph.d. fra Økonomisk Institut ved Københavns Universitet. Hun har været ansat i AKF siden 1990 og har primært beskæftiget sig med kulturøkonomi og kulturpolitik. Har været Kultur- og Fritidschef i Helsingør Kommune fra 2001-2003, og var fra 1993-94 udvalgt af daværende kulturminister Jytte Hilden til i samarbejde med projektleder Peter Duelund at udarbejde den kulturpolitiske udredning »Kulturens Politik«. Medredaktør af Nordisk Kulturpolitisk Tidsskrift og AKF Nyt.

Camilla Damgaard

Forstkandidat med speciale i miljøøkonomi. Hun har arbejdet med miljøøkonomi og værdisætning i AKF, Skov- og Naturstyrelsen og senest Miljøstyrelsen, hvor hun sidder i den miljøøkonomiske enhed. Hun har bl.a. gennemført samfundsøkonomiske analyser inden for naturområdet og affaldsområdet.

Emil Husted Erichsen

Forstkandidat med speciale i miljøøkonomi. Han har arbejdet med værdisætning og samfundsøkonomiske analyser inden for natur- og miljøområdet i Skov- og Naturstyrelsen, AKF og Miljøstyrelsen. Han arbejder i dag som miljøøkonom i KL's kontor for teknik og miljø.

Leena Eskelinen

Leena Eskelinen er seniorforsker og ph.d. fra Helsinki universitet. Hun har forsket på det sociale og arbejdsmiljøområdet i AKF siden 1991 og har tidligere været ansat som forskningspsykolog på det finske arbejdsmiljøinstitut. Hun har gennemført undersøgelser og evalueringer på arbejdsmiljøområdet bl.a. for Arbejdsmiljørådet, Arbejdstilsynet og Beskæftigelsesministeriet, samt medvirket i forskningsevalueringer. Hun er medlem af Det videnskabelige arbejdsmiljøforskningsudvalg.

Jill Mehlbye

Jill Mehlbye er docent og uddannet cand.psych. samt har en MPA-uddannelse. Hun har især været beskæftiget inden for børn og unge-området, hvor hun har gennemført en række evalueringer af kommuners og amters indsats over for og tilbud til børn og unge. Det gælder såvel inden for socialområdet som skoleområdet. Hun har især beskæftiget sig med udsatte børn og unge med svag social baggrund og børn og unge fra etniske minoriteter, og har en omfattende publikationsvirksomhed bag sig på området.

Eigil Boll Hansen

Eigil Boll Hansen er uddannet cand.oecon. og ansat som docent i AKF. Han har gennemført flere forskningsprojekter og evalueringer især inden for kommunernes ældreområde. Forskningsfeltet har været domineret af studier af levevilkår og velbefindende blandt ældre og af social- og sundhedsydelser samt boliger til ældre. Han har bl.a. haft fokus på samspillet mellem tildeling og udførelse af ydelser til ældre og disses betydning for ældres velbefindende. Han har gennemført både kvalitativt og kvantitativt baseret forskning.

Eskil Heinesen

Eskil Heinesen er forskningsleder og ph.d. fra Økonomisk Institut ved Københavns Universitet. Han har været i AKF siden 1995 og har før da været fuldmægtig i Modelgruppen i Danmarks Statistik, kandidat- og seniorstipendiat ved Københavns Universitet og lektor ved Den Kongelige Veterinær- og Landbohøjskole. Han har gennemført forskningsprojekter

for bl.a. Finansministeriet, Integrationsministeriet og Forskningsrådene inden for bl.a. følgende områder: uddannelse, skoleressourcer, arbejdsmarked, kontanthjælp og integration.

Kjeld Høgsbro

Kjeld Høgsbro er seniorforsker, kultursociolog og ph.d. fra Københavns Universitet. Hans forskningsarbejde har især fokuseret på den sociale indsats over for særligt udsatte grupper som mennesker med alvorlige sindslidelser, handicap og misbrug. KH har dog også gennemført forskningsprojekter, der fokuserede mere generelt på samspillet mellem offentlige og frivillige initiativer bl.a. i forbindelse med evalueringer af de statsligt finansierede udviklingsprogrammer.

Torben Pilegaard Jensen

Torben Pilegaard Jensen (TPJ) er forskningsleder i AKF. TPJ har først og fremmest gennemført forskning på uddannelsesområdet og inden for det social- og arbejdsmarkedspolitiske felt. Deltager bl.a. i den internationale PISA-undersøgelse (Programme for International Student Assessment), igangsat på initiativ fra OECD. TPJ er medlem af Beskæftigelsesministeriets Offentlige Forskningsudvalg.

Anders Larsen

Anders Larsen er økonom og vicedirektør ved AKF. Han har en fortid i Undervisningsministeriet og i Gladsaxe Kommune. Han har siden midten af 1980'erne medvirket ved en lang række evalueringer inden for miljø- og energiområdet. På energiområdet typisk i form af evalueringer af energibesparelserprojekter. Hans bidrag her har ofte været at føje en cost-benefit-synsvinkel til den øvrige evaluering. For tiden er han en af de få danskere, der har overblik over forsøget på at liberalisere energisektoren.

Leif Olsen

Leif Olsen er sociolog og forsker i AKF. Hans arbejde med forskning og evaluering på handicapområdet har særligt haft fokus på det amtslige rådgivningssystem, herunder de amtslige videntcentre på handicapområdet, som han har evalueret såvel fra et medarbejder- som et brugerperspektiv.

I perioden fra 1993 til 1997 studerende han implementeringen af viden-centre i Danmark. Har senest været medredaktør af antologien: »Handicap, kvalitetsudvikling og brugerinddragelse (Bentsson 2003).

Lene Holm Pedersen

Lene Holm Pedersen er ansat som adjunkt på Institut for Statskundskab, Københavns Universitet og har en ph.d.-grad fra samme sted. Hun har tidligere været forsker i AKF og visiting scholar på Stanford, USA. Hendes primære empiriske interesse er regulering inden for det energi- og miljøpolitiske område, mens den teoretisk indfaldsvinkel primært findes inden for nyinstitutionel teori. Eksempler på forskningsområder er miljø-økonomiske ideers indflydelse på politiske beslutningsprocesser i forbindelse med indførelsen af grønne afgifter på erhvervene, europæiseringen af uafhængige reguleringsmyndigheder i energisektoren, samt analyser af livsstil og politisk forbrug.

Peter Plougmann

Peter Plougmann, Direktør New Insight A/S, Mag.scient.soc. Har i snart 25 år arbejdet med industrisociologi, arbejdsmarkedsforhold og erhvervsudvikling inden for politisk administration, forskning, erhvervsrådgivning og undervisning. PP udvikler scenarier for private virksomheder, faglige organisationer, brancheforeninger og offentlige institutioner både i Danmark og internationalt. Han er senest medforfatter til bogen »Viden og forandring« (Gyldendal 2003). Sidder i Executive Committee for European Network of Social and Economic Research Institutes (ENSR), samt i flere private forretningservicevirksomheders bestyrelse.

Olaf Rieper

Docent i AKF, uddannet i sociologi, ph.d. i organisationsteori fra Handelshøjskolen i Kbh. Har gennemført evalueringsopgaver for blandt andre Norges Forskningsråd og haft ekspertopgaver for EU-Kommissionen. Forsker i offentlig organisation og evaluering. Har arbejdet i Lyon for Center for European Evaluation Expertice med at udvikle evalueringsmetoder for de europæiske strukturfonde. Har skrevet en række bøger om

forsknings- og evalueringsmetoder. Er medstifter af Dansk Evaluerings-Selskab og dets formand fra 2001 til 2004.

Michael Rosholm

Økonom og forskningsprofessor i integration og marginalisering ved AKF og på Institut for Økonomi ved Aarhus Universitet. Forskningsprofessoratet er en del af Forskerskolen for Integration, Produktion og Velfærd. Forsker især kvantitativt i effekter af aktiveringsforanstaltninger, indvandreres integration på arbejdsmarkedet, samspillet mellem helbred og arbejdsløshed, og mellem bolig- og arbejdsmarkedet. Har skrevet en række internationalt publicerede artikler om disse emner.

Mikael Togeby

Mikael Togeby er civilingeniør og ph.d. Er ansat hos Elkraft System. Har gennemført en række evalueringer, bl.a. inden for energiområdet. Har været medforfatter til Håndbog i evaluering (1993) og Håndbog i evaluering af energispareaktiviteter (2003). Har undervist i effektevaluering og i brugen af statistiske metoder i forbindelse med evaluering.

Bilag

Vejledninger i evalueringer på nettet

Introduktion

Internettet er i løbet af få år blevet en righoldig kilde til vejledninger, guidelines, håndbøger, modeller, standarder o.l. (i det følgende er anvendt fællesbetegnelsen »vejledninger«). Man finder henvisninger til vejledninger, men i stigende omfang også vejledninger i fuld tekst. I dette kapitel skal kort omtales, hvilke værter (hjemmesider) der giver en god indgang til at finde vejledninger, forskellige typer af vejledninger, hvilke funktioner de har, og endelig omtales styrker og svagheder ved vejledninger. I øvrigt findes også mange evalueringresultater på nettet, men det er en anden historie.

Værter

Mange netsteder skifter jævnligt adresse. En generel let indgang fra Danmark til links til de forskellige værter er:

www.danskevalueringsselskab.dk, se under menupunktet »links«. Også www.akf.dk har links til evalueringssider. De netsteder, der indeholder evalueringvejledninger, kan opdeles i:

1. Evalueringfaglige foreninger
2. Nationale organisationer
3. Organisationer, der er specialiseret i evaluering
4. Internationale organisationer

De evalueringsfaglige foreninger omfatter enkeltlandes evalueringsforeninger, samt regionale foreninger, fx The European Evaluation Society (EES).

Det ældste er det amerikanske, the American Evaluation Association (AEA), med links til en rigdom af fuldtekst vejledninger:
<http://www.eval.org/EvaluationLinks/onlinehbt.txt.htm>.

Nationale organisationer, i Danmark fx KL, Amtsrådsforeningen, ministeriet, forskningsinstitutioner, videncentre, vil ofte indeholde links til vejledninger. Det gælder fx Energiministeriets link til:
www.energievaluering.dk.

Organisationer, der er specialiseret i evaluering, i Danmark fx Danmarks Evalueringsinstitut (EVA) <http://www.eva.dk/> og Danidas evalueringsenhed, DANIDA (<http://www.um.dk/danida/evalueringer/danevalpoldansk.asp>).

Internationale organisationer

OECD, www.oecd.org

Verdensbanken, <http://www.worldbank.com/oed/>, som beskriver de tilgange, Verdensbanken benytter ved evaluering af deres bistandsprogrammer.

Verdensbanken, www.worldbank.org/oed/

Den private evalueringsenhed, Center for European Evaluation Expertise, indeholder en række relevante links:

<http://www.c3e.fr/english/informations/links.htm>

Typer af vejledninger

Generelle: Vejledninger, som går på tværs af sektorer, ofte med generelle metodeanvisninger, eller overgribende, dvs. om evaluering i almindelighed, Eksempler:

Evaluation Handbook

Sponsor: WK Kellogg Foundation

Scope: Outlines a blueprint for designing and conducting evaluations, either independently or with the support of an external evaluator/consultant

Audience: Primarily for project directors on foundation grants but very broadly applicable

Format: Adobe pdf document

Link: <http://www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf>

Sektorspecifikke vejledninger: Fx om områderne uddannelse, det sociale område, sundhed, udviklingsbistand, energi og miljø mv. Fx Evaluation Cookbook (1998)

Sponsor: Learning Technology Dissemination Initiative, Heriot Watt University

Scope: Methodologies for evaluating uses of technology

Audience: Primarily educators using technology in their classrooms

Format: Webpages

Link: <http://www.icbl.hw.ac.uk/lti/cookbook/contents.html>

Håndbog i evaluering af elspareaktiviteter, Udarbejdet af SRC International, AKF og Elkraft System for Energistyrelsen m.fl.

Målgruppe: Bestillere og evaluører inden for elområdet

December 2002

http://www.energievaluering.dk/doc/Evalueringshaandbog_m_forside.pdf

Metodespecifikke vejledninger: Survey, interview mv, fx:

Evaluation Toolkit: Quasi-Experimental Design

Sponsor: Human Resource Development Canada (HRDC) - Data and Development Branch

Scope: Basics of quasi-experimental design and the need for comparison groups

Audience: General

Format: Adobe pdf document

Link: <http://www11.hrdc-drhc.gc.ca/edd/toolkit.list>

Vejledningers funktioner

Generelt er vejledningers og standarders funktion at reducere usikkerhed hos såvel evaluator som bestiller. Men vejledninger vil binde evaluator i forskellig grad. Vejledninger kan således være:

- Service for bestiller og udfører af evalueringer, dvs. det er frivilligt at bruge dem, et tilbud.
- Professionelle standarder, der i evalueringsmiljøer betragtes som »god skik«
- Obligatorisk at overholde, fx krav ved udbud af evalueringer.

Fordele og ulemper ved vejledninger

Fordele: gode vejledninger opsummerer dyrekøbte erfaringer og giver læseren mulighed for at hente inspiration og at vælge mellem en vifte af mulige fremgangsmåder og metoder.

Ulemper: Der kan ske en uhensigtsmæssig standardisering af evalueringer, som ikke passer til den konkrete situation. Det bagvedliggende paradigme er kun nævnt implicit, og det kan virke uhensigtsmæssigt styrende for valg i evalueringsprocessen. Kvaliteten i vejledningerne er svingende.

Litteratur

AKF, SRCI og Elkraft System (2002): *Håndbog i evaluering af energispareaktiviteter*. Energistyrelsen. København. www.ens.dk

Albæk Erik & Olaf Rieper (2001): Effektevaluering, monitorering og formativ evaluering. I: Dahler-Larsen P & Krogstrup H-K (2001) (red.) *Tendenser i evaluering*. Odense Universitetsforlag, s. 130-140.

Albæk, Erik (1988). *Fra sandhed til information*. Akademisk Forlag.

Albæk, Erik og Olaf Rieper (2002): Evaluation in Danish Governance. I: *the International atlas of evaluation*, eds. Jan-Eric Furubo, Ray C. Rist, Rolf Sandahl.

Andersen, Bjarne Hjorth (2003)(red.): *Udviklingen i befolkningens levekår over et kvart århundrede*. København: Socialforskningsinstituttet.

Andersen, Annemarie Møller; Niels Egelund, Torben Pilegaard Jensen, Michael Krone, Lena Lindenskov og Jan Mejding (2001): *Forventninger og færdigheder – danske unge i en international sammenligning*. København: AKF, DPU og SFI-Survey. Rapporten findes på hjemmesiderne www.akf.dk/dk2001/pdf/F&Fpisa.pdf
www.dpu.dk og www.sfi.dk/graphics/SFI/SFI-SURVEY/Danske_PISA_rapport.pdf

Andersen, Else Marie (1999): *Fleksibel hjemmehjælp er guld værd*. AKF Forlaget.

Andersen, John og Jacob Torfing (2002): *Netværksstyring i velfærdssamfundet: de lokale koordinationsudvalg*. Arbejdstekst nr. 6, CARMA, Center for arbejdsmarkedsforskning ved Aalborg Universitet.

Angrist, Joshua D. og Victor Lavy (1999): Using Maimonides= rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*: 533-575.

Arbejdsmiljøinstituttet (1999): *Dansk arbejdsmiljøforskning 1999 – en oversigt*. København, august 1999.

Arbejdsministeriet (2000): *Effekter af aktiveringsindsatsen*. Arbejdsmarkedsemner. København.

Arbejdsministeriet (2001): *Rapport fra Arbejdsministeriets udvalg om arbejdsmiljøforskning*. København, februar.

Arbejdstilsynet (2003): *Anmeldte arbejdsskader, årsopgørelse 2002* (www.at.dk)

Arbejdstilsynet (2003): *Overvågningsrapport 2001. Udviklingstendenser i arbejdsmiljøet for de syv visioner i handlingsprogrammet om et rent arbejdsmiljø 2005*. København, Arbejdstilsynet, marts 2003 (www.at.dk).

Arrow, Kenneth J.; Robert Solow, Edward Leamer, Paul Portney, Roy Radner og Howard Schumann (1993): Report of the NOAA Panel on Contingent Valuation, *US Federal Register*, vol. 58, no. 10, pp. 4602-4614, Washington D.C.

Balling, Gert; Susan Fazakerley og Dorte Skot-Hansen (2001): *KUF i det uprøvede grænsefelt. En midtvejsevaluering af Kulturministeriets Udviklingsfond*, Center for Kulturpolitiske Studier, Danmarks Biblioteksskole, København.

Bengtsson, Steen (1997): *Handicap og funktionshæmning i halvfemserne*. SFI.

Bengtsson, Steen; Inge Storgaard Bonfils og Leif Olsen (2003): *Handicap, kvalitetsudvikling og brugerinddragelse*. AKF Forlaget.

Bengtsson, Steen, Janika Wiene og Christina Bak (2003): *Lyttermødemodelen – en kvalitetsmodel for indsatsen over for børn med handicap*. Socialforskningsinstituttet.

Bille Hansen, Trine (1993): *Kulturens økonomiske betydning – state of art*, AKF Forlaget, København.

Bille Hansen, Trine (1996): *Danskernes værdisætning af Det Kgl. Teater*, AKF Forlaget, København.

Bille Hansen, Trine (1997): *Kortlægning af metoder og resultater vedrørende måling af teaterforestillinger kvalitet – den kulturøkonomiske forskningsgren, bidrag til Det Kgl. Teaters forprojekt vedrørende kvalitetsmåling* (upubliceret).

Bjørner, Thomas Bue og Henrik Holm Jensen (2002): Energy Taxes, Voluntary Agreements and Investment Subsidies – A micro panel analysis of the effect on Danish industrial companies' energy demand«. *Resource and Energy Economics* 24(3), p. 229-249.

Bolvig, Iben; Peter Jensen og Michael Rosholm (2003): *Employment Effects of Active Social Policies*. Discussion Paper No. 736, IZA Bonn.

Bolvig, Iben; Peter Jensen og Michael Rosholm (2001): *Effekter af aktiveringsindsatsen i Århus Kommune*. CLS Rapport 01-01.

Borchmann, Thomas (2003): *Om kompetenceudvikling og aktionsforskning og kompetenceudvikling gennem aktionsforskning*. Handelshøjskolen i Århus, Institut for Organisation og Ledelse.

Brandt, Preben og Mette Kirk (2003): *Billeder fra hverdagen*. AKF Forlaget.

Browning, Martin og Eskil Heinesen (2003): *Class size, teacher hours and educational attainment*. Working paper. AKF. København.

Bunnage, David; Ole Gregersen, Eigil Boll Hansen, Nanna Meilbak og Merete Platz (2001): *Kvalitet i ældreplejen*. København. Socialforskningsinstituttet. 01:3.

Burr Hermann; Elsa Bach, Vilhelm Borg og Ebbe Villadsen (2002): *Arbejds miljø i Danmark 2000. En kortlægning af lønmodtageres og selvstændiges arbejdsmiljø og helbred*. København: Arbejds miljøinstituttet.

Burtless, Gary (ed.) (1996): *Does money matter? The effect of school resources on student achievement and adult success*. Brookings Institution Press. Washington D.C.

Campbell, Donald T. og Julian C. Stanley (1963): *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Caswell, Dorte og Torben Pilegaard Jensen (2001): *Lokale koordinationsudvalg – på rette vej?* AKF Forlaget.

Christensen, Inger Marie; Klaus Jørgensen, Jørgen Jordal-Jørgensen, Anders Larsen og Lene Nielsen (1994): *Energikonsulenter i enfamiliehuse*. AKF Forlaget.

Christoffersen, Henrik og Peter Lyk-Jensen (1994): *Nogle økonomiske effekter af Brandts Klædefabrik i Odense*, AKF Forlaget.

Clausen, Hans (1993): *Center for Små Handicapgrupper. En evaluering af et videncenter*. København: Center for Små Handicapgrupper.

Clausen, Thomas (2003): *Når hørelsen svigter – Om konsekvenserne af hørenedsættelse i arbejdslivet, uddannelsessystemet og for den personlige velfærd*. SFI.

COWI A/S og Socialt Udviklingscenter SUS (2002): *Evaluering af erfaringerne med institutionsbegrebets ophævelse på handicapområdet, 1998 – 2002*. <http://www.sm.dk/netpublikationer/2002/ds5ins071002/rapport.pdf>

Cwi, David og Katherine Lyall (1977): *Economic Impact of Arts and Cultural Institutions: A Model of Assessment and a Case Study in Baltimore*, National Endowments for the Arts.

Dahler-Larsen, Peter og Hanne Katrine Krogstrup (2001) (red.): *Tendenser i evaluering*. Odense Universitetsforlag.

Dahler-Larsen, Peter (2002): *Evaluering kortlagt – en undersøgelse af amters og kommunersevalueringspraksis på det sociale område*. Systime A/S.

Dahler-Larsen, Peter og Hanne Kathrine Krogstrup (2003): *Nye veje i evaluering. Håndbog i tre evalueringsmodeller*. Systime A/S.

Dahler-Larsen, Peter og Hanne Kathrine Krogstrup (red) (2001): *Tendenser i evalueringen*. Odense Universitetsforlag.

Davies, Bleddyn og Martin Knapp (1981): *Old Peoples' Homes and the Production of Welfare*. London. Routledge & Kegan Paul.

Det Centrale Handicapråd (2000): *Værktøjsskabe til FN's standardregler. Ideer og redskaber til en kommunal handicappolitik*. Denne publikation findes på adressen: www.dch.dk/fnregl/vaerktoej

Det Centrale Handicapråd (2001): *Handicap og ligebehandling – et refleksionspapir*. København.

<http://www.dch.dk/publ/handicapogligebehandling>

Duelund, Peter (1995): Den danske kulturmodel, *Kulturens politik nr. 18*, Klim, Århus.

Duelund, Peter og Trine Bille Hansen (1994): Hvor står vi nu? *Kulturens politik nr. 1*, Klim, Århus.

Ebsen, Frank og Jesper Henriksen (2003): *Hænger det sammen*. AKF Forlaget.

Ebsen, Frank og Kjeld Høgsbro (2001): *Faglig bæredygtighed i det sociale arbejde i små kommuner*. AKF Forlaget.

Elo, Anna-Liisa (2002): Miten saada näyttöä kehittämisen vaikuttavuudesta? I: Lindström Kari & Leppänen Anneli (2002) (red.): *Työyhteisön terveys ja hyvinvointi*. Helsinki: Työterveyslaitos 2002. [...], s. 340-347.

Energistyrelsen (2000): *Energistyrelsens tilskudsordninger*. Beskrivelse og vurderinger.

- Energistyrelsen (2001): *Omkostninger ved CO₂-reduktion for udvalgte tiltag*. Midtvejsrapport.
- Erichsen, Emil Husted (2003) *Samfundsøkonomisk projektvurdering af skovrejsning ved Odense*. AKF Forlaget.
- Eriksen, Erik Oddvar og Jarle Weigård (2003): *Kommunikativt demokrati: Jürgen Habermas' teori om politik og samfund*. Hans Reitzel.
- EU-Kommissionen (2002): *Economic Analysis of options for Managing Biodegradable Municipal Waste*.
- Favrholdt, David (2000): *Æstetik og filosofi. Seks essays*, Høst Humaniora, København.
- Fetterman, David M.; Kaftarian og Wandersman (red.) (1996): *Empowerment evaluation*. Sage.
- Finansministeriet (1999): *Vejledning i udarbejdelse af samfundsøkonomiske konsekvensvurderinger*.
- FN (1993): Standardregler om Lige Muligheder for Handicappede. *Resolution 48/96*.
- Folkestad, Helge (2000): Getting the picture: Photo-assisted conversations as interview. I: *Scandinavian Journal of Disability Research*, vol. 2, no. 2, 2000.
- Foss Hansen, Hanne (2003): *Evaluering i staten: kontrol, læring eller forandring?* Samfundslitteratur.
- Frandsen, Annie Gaardsted (udateret): *Forskning inden for ældreområdet ved danske samfundsvidenskabelige forskningsinstitutioner*. Institut for Pensions- og Ældrepolitik. Upubliceret notat.
- Frederiksen, Morten og Eigil Boll Hansen (2004): *Projekter til at forebygge bevægeapparatskader ved personløft*. Delrapport 1. København: AKF Forlaget.

Fridberg, Torben og Inger Koch-Nielsen (1997): *Konsekvensanalyse af Kulturby 96*, Socialforskningsinstituttet, København.

Grant, Ruth; Gitte Blicher-Mathiesen, Villy Jørgensen, Arne Kyllingsbæk, Hanne Damgaard Poulsen, Christian Børsting, Jørgen Ole Jørgensen, Jesper S. Schou, Erik Steen Kristensen, Jesper Waage Petersen og Harald E. Mikkelsen (2000): *Vandmiljøplan II – midtvejsevaluering*. Danmarks Miljøundersøgelser.

Graversen, Brian Krogh og Eskil Heinesen (1999): *Ressourceforbrug i folkeskolen: Forskelle mellem kommuner*. AKF Forlaget.

Graversen, Brian Krogh; Eskil Heinesen og Niels Madsen (1999): *Effekter på elevernes uddannelsesforløb af ressourceforbruget i folkeskolen*. AKF Forlaget.

Griffiths, Amanda (1999): Organizational interventions: facing the limits of the natural science paradigm. *Scandinavian Journal of Work, Environment & Health* 1999; 25 (6, special issue): 589-596.

Gruber, Thomas (red.) (2002): *Brugerundersøgelser i en kommunal virkelighed*. Formidlingscenter Øst.

Guba, Egon G. og Yvonna S. Lincoln (1990): *Fourth generation evaluation*. Sage Publications.

Hallberg, Dag (2003): *Bokslut för scenekunst*, Kulturrådet, Stockholm. Findes på www.kulturradet.se

Hansen, Eigil Boll og Leena Eskelinen (2004): *Evaluering af puljen »Forebyggelse af bevægeapparatskader ved tunge personløft«*. (under udgivelse).

Hansen, Hanne Foss (2003): *Evaluering i staten. Kontrol, læring eller forandring?* København: Samfundslitteratur.

Hansen, Eigil Boll (2003): *At yde og at måle god kvalitet på ældreområdet*. København. FOKUS.

Hansen, Eigil Boll; Leena Eskelinen og Jan Kargaard Madsen (1999): *Hjemmehjælp og ældres velbefindende – en analyse af hjemmehjælpernes arbejdsprincipper i to kommuner*. AKF Forlaget.

Hansen, Eigil Boll; Leena Eskelinen, Torben Sejr og Lis Wagner (1997a): *Ældrevenlige behandlingsforløb – en analyse af fem indsattstyper*. Kbh. AKF Forlaget.

Hansen, Eigil Boll; Leena Eskelinen, Torben Sejr og Lis Wagner (1997b): *Ældres behandlingsforløb* (FOKUS pjece nr. 23). Kbh. Forlaget Kommuneinformation.

Hansen, Hanne Foss (2003): *Evaluering i staten. Kontrol, læring eller forandring?* Samfundslitteratur.

Hasle, Peter m.fl. (2001): *Evaluering af EGA-handlingsplanen*. København: CASA og DTU.

Hasle, Peter; Ole Broberg, Hans Hvenegaard og Ulla Brund Jensen (2000): *Arbejds miljøindsatsen i 25 år – succes eller fiasko?* København: CASA.

Hasle, Peter; Hans Jørgen Limborg, Pper Langaa Jensen og Annette Ledskov (1999): *Kortlægning af arbejdsmiljøforskningen i Danmark – en forundersøgelse*. København: LO. (www.lo.dk, Politik og visioner – Arbejds miljø)

Hasler, Berit; Camilla K. Damgaard, Emil Husted Erichsen, Jørgen Jordal Jørgensen og Hans Erik Kristoffersen (2002): *De rekreative værdier af skov, sø og naturgenopretning. Værdisætning af naturgoder med husprismetoden* AKF Forlaget.

Hatting, Anne og Olaf Rieper (1991): *Evaluering af lederudviklingsprogram på Vestfold Sentralsykehus*. København. AKF Forlaget.

Heckman, James; Robert Lalonde & Jeffrey Smith (1999), The Economics and Econometrics of ALMP. I *Handbook of Labor Economics, Vol. 3*, North-Holland, Amsterdam.

Heinesen, Eskil (1999): *Den sociale arvs betydning for unges valg og resultater i uddannelsessystemet*. Arbejdsrapport nr. 2 om social arv. Socialforskningsinstituttet.

Heinesen, Eskil (2000a): Er der sammenhæng mellem højt udgiftsniveau pr. elev og høj kvalitet? I: *Resultater, mål og ressourcer: Folkeskolen år 2000 B Fokuspunkt 8: Måltrettet anvendelse af ressourcerne*. Undervisningsministeriet, Kommunernes Landsforening og Danmarks Lærerforening.

Heinesen, Eskil (2000b): Ringe effekt af forskelligt ressourceforbrug. *AKF Nyt* (2): 37-40.

Heinesen, Eskil (2000c): *Ressourceforbrug og uddannelsesforløb*. Skolebestyrelsen (2).

Heinesen, Eskil (2000d): Ressourceforbruget i folkeskolen – og elevernes uddannelsesforløb. *Tandhjulet* 9 (1): 4-5.

Heinesen, Eskil (2001): Folkeskolens ressourceforbrug og effektivitet. I: Rolf Norstrand og Nils Groes (red.): *Kommunestyrets fremtid*. AKF Forlaget. København.

Heinesen, Eskil og Brian Krogh Graversen (2004): *The effect of school resources on educational attainment: Evidence from Denmark*. Arbejdsrapport, AKF og SFI.

Heinesen, Eskil og Leif Husted (2004a): Kommunernes succes med hensyn til integration af indvandrere omfattet af integrationsloven. I: Groes, Nils og Leif Olsen (red.): *Brudstykker eller sammenhæng? – En antologi om forudsætningerne for en offentlig strukturreform*. FOKUS, København.

Heinesen, Eskil og Leif Husted (2004b): Sammenligning af kommunernes indsats på integrationsområdet. *AKF Nyt*.

Heinesen, Eskil; Brian Krogh Graversen og Niels Madsen (1999): *Resourceforbrug i folkeskolen og effekter på elevernes uddannelsesforløb*. København, AKF Forlaget.

Herman, Joan L.; Lynn Lyon Morris og Carol Taylor Fitz-Gibbon (1987). *Evaluators Handbook*. Sage.

Hernberg, S. (1998): *Epidemiologia ja työterveys*. Helsinki: Työterveyslaitos. [Tidligere udgave oversat til dansk: Basisbog i arbejdsepidemiologi, København: Arbejdsmiljøinstituttet, Arbejdstilsynet 1986]

Holm, Per; Dorthe Perlt, Bernard Jensen og Birger Perlt (2003): *KUBI – en model for brugerinddragelse som led i kvalitetsudvikling i social- og sundhedssektoren*. Socialt Udviklingscenter SUS.

House, Ernest R. (1993). *Professional evaluation*. Sage.

Hoxby, Caroline M. (1999): *The effects of class size on student achievement: New evidence from population variation*. Working paper. Harvard University and National Bureau of Economic Research.

Husted, Leif og Eskil Heinesen (2004): *Benchmark-analyse af kommunernes integrationsindsats i forhold til udlændige omfattet af integrationsloven*. AKF Forlaget.

Høgsbro, Kjeld et al. (2003): *HMS-undersøgelsen*. AKF.

Høgsbro, Kjeld (1995): Sociologiens metodologiske udfordring i forbindelse med forsøgs- og udviklingsprojekter. *Dansk Sociologi nr. 4 1995*.

Høgsbro, Kjeld (1996): Fra socialteknologi til udviklingsprogrammer – udviklingen i den sociologiske sektorforskning i 80erne. I: Hansson og Aagaard *Sociologiens Historie*, Forlaget Sociologi.

Høgsbro, Kjeld (1997): *Dansk forskning vedrørende sindslidendes forhold*. ISU.

Høgsbro, Kjeld (1998): *Pandoras Æske – en undersøgelse af psykologens funktion på onkologisk afdeling*. Institut for Social Udvikling.

Høgsbro, Kjeld; Kirsten Marie Bovbjerg, Louise Hardman Smith, Mette Kirk og Jesper Henriksen (2003): *Skjulte livsverdener*. AKF Forlaget.

Høgsbro, Kjeld (2002): *Rehabilitering af mennesker med traumatiske hjerneskader på Kolonien Filadelfia*. AKF Forlaget.

Høgsbro, Kjeld og Olaf Rieper (2000): *Formative evalueringer*. I: Peter Dahler-Larsen og Hanne Krogstrup 2001.

Høgsbro, Kjeld; Birgit Kirkebæk, Sara Vafai Blom og Else Danø (1999): *Ungdom, udvikling og handicap - Erfaringer fra et udviklingsprogram for 23 unges overgang til en selvstændig voksertilværelse*. Samfundslitteratur.

Høgsbro, Kjeld; Karen Maigaard Knudsen og Birte Ravn (1990): *Kulturcenterundersøgelsen 2*, Udviklingscenter for folkeoplysning.

Jacobsen, Brian H. (2000). *Vandmiljøplan II – økonomisk midtvejsevaluering. Rapport*. Fødevarerøkonomisk Institut.

Jakobsen, Joakim (2002): Fiasko og succes på regneark, *Kulturkontakten nr, 1. s.11-12*, Kulturministeriet, København.

Jarl, Stig (2001): Elendig forestilling – jeg kunne ikke finde parkeringsplads. En studie i kvalitetsmåling på teaterfeltet, *Nordisk Kulturpolitisk Tidsskrift 2000/2*, s. 50-88.

Jarl, Stig og Randi Svendsen (1997): *Registrering og optimering af teaterforestillingerens kvalitet, bidrag til Det Kgl. Teaters forprojekt vedrørende kvalitetsmåling* (upubliceret).

Jensen, Mogens Kjær (1991): Triangulering. I: Mogens Kjær Jensen (1991): *Kvalitative metoder i anvendt samfundsforskning* kap. 5, s. 106-120. Socialforskningsinstituttet.

Jensen, Torben Pilegaard; Søren Winter, Jesper Manniche og Peter D. Ørberg (1991): *Indsatsen for langtidsledige*. Med bidrag fra Socialforskningsinstituttet ved Ivan Thaulow. AKF Forlaget.

Johannsen, Katja Sander og Anders Larsen (2000): *Voluntary Agreements – Implementation and Efficiency. The Danish Country Study*. AKF Forlaget.

Johannsen, Katja Sander (2002): Combining voluntary agreements and taxes – an evaluation of the Danish agreement scheme on energy efficiency in industry. *Journal of Cleaner Production* 10 129-141

Karnøe, Peter og Ulrik Jørgensen (1995): *Samfundsmæssig værdi af vindkraft*, Delrapport 4: Dansk vindmølleindustri internationale position og udviklingsbetingelser. AKF Forlaget.

Karpatschhof, Benny (2001): *En evaluering af brugernes oplevelse af Skp-ordningen i Københavns Kommune*. Københavns Kommune, Familie- og Arbejdsmarkedsforvaltningen.

Kjøller, Mette og Niels Kr. Rasmussen (red.) (2002): *Sundhed & sygelighed i Danmark 2000 & udviklingen siden 1987*. København: Statens Institut for Folkesundhed.

Kjørup, Søren (2000): *Kunstens filosofi – en indføring i æstetik*, Roskilde Universitetsforlag, Frederiksberg.

Klausen, Kurt Klaudi (2001): *Skulle det være noget særligt? – organisation og ledelse i det offentlige*, Børsens Forlag, København.

Kluve, Jochen og Christoph M. Schmidt (2002): Active Policy Evaluation: Problems, Methods, and Results. *Economic Policy*, October.

Koch, Anna; Leena Eskelinen, Mette Jensen og Kristian Lyk-Jensen (1997): *Sygemeldt – og i gang. Evaluering af et udviklingsprojekt om socialforvaltningens arbejde med sygedagpengemodtagere i seks kommuner*. AKF Forlaget.

Kompier, Michiel og Cary L. Cooper (1999): *Preventing stress, improving productivity: European case studies in the workplace*. London: Routledge.

Krørup, Signe (2001): Can voluntary approaches ever be efficient. *Journal of Cleaner Production* 9 135-144.

Krogstrup, Hanne Kathrine (1997): *Brugerinddragelse og organisatorisk læring i den sociale sektor*. Århus: Systime.

Krogstrup, Hanne Kathrine (2003): *Evalueringsmodeller – Evaluering på det sociale område*. Århus: Systime.

Krogstrup, Hanne Kathrine (1999): *Det handicappede samfund: om brugerinddragelse og medborgerskab*. Århus. Systime.

Krogstrup, Hanne Kathrine og Thomas Gruber (red.) (2002): *Brugerundersøgelser i en kommunal virkelighed: en antologi med bidrag fra 4 danske kommuner*. Formidlingscenter Øst.

Krogstrup, Hanne Kathrine (1996). *BIKVA-modellen – en dialog- og læringsorienteret evalueringsmodel*. Aalborg Universitet.

Krogstrup, Hanne Kathrine (1997): Brugerinddragelse i kvalitetsvurdering som kilde til generering af ny viden og læring i socialt arbejde med sindslidende. I: Høgsbro: *Dansk forskning vedrørende sindslidendes forhold*. Institut for Social Udvikling.

Krueger, Alan B. (1999): Experimental estimates of educational production functions. *Quarterly Journal of Economics*: 497-532.

Kuhn, Thomas S. (1996): *The structure of scientific revolutions*. University of Chicago Press.

Kvale, Steinar (1998): *InterView. En introduktion til det kvalitative forskningsinterview*. Hans Reitzels Forlag.

Lancaster, Tony (1990): *The Econometric Analysis of Transition Data*. Econometric Society Monographs, Cambridge University Press, 1990.

Landstingsförbundet (1994): *Från vision till verksamhet. En metod- och ideskript om utvärdering av regional kulturpolitisk*, slutrapport från Kulturmätarprojektet, Landstingsförbundet, Stockholm.

Langager, Klaus (1996): *Ledige på Kursus*. Rapport 96-9, Socialforskningsinstituttet.

Langsted, Jørn (1998): Kulturpolitisk forskning i Danmark i 1997 – legitimering af kulturpolitik? *Nordisk Kulturpolitisk Tidsskrift*, 1998/2, s.100-120.

Langsted, Jørn; Karen Hannah, Charlotte Rørdam Larsen (2003): *Ønskekivist-modellen. Kunstnerisk kvalitet i performativ kunst*, Klim, Århus

Larsen, Anders og Jesper Munksgaard (1998): Socio-economic Assessment of Wind Power – lessons from Denmark: *Energy Policy*, Vol. 26, no. 2, pp 85 - 93.

Larsen, Mehlbye og Gørtz (1991): *Kan genoptræning betale sig*. AKF.

Lindström, Kari m.fl. (2002): *Työkuormitus ja sen arviointimentelmät*. Helsinki: Työterveyslaitos 2003.

Lundkvist, Anders og Christer Hahlström (1998): *Uppföljning Utvärdering Kvalitet inom kultur- og fritidssektorn*, Svenska Kommunförbundet, Stockholm.

Martin, John P. og David Grubb (2001): *What works and for whom? A review of OECD countries' experiences with active labour market policies*, IFAU Working Paper 2001:14.

Mehlbye, Jill; Olaf Rieper og Mikael Tøgeby (1993): *Håndbog i evaluering*. København. AKF Forlaget.

Mehlbye, Jill (1994): *Tosprogede børn og unge i Albertslund Kommune*. AKF Forlaget.

Mehlbye, Jill (1996): *Kommunernes vejledning af unge om uddannelse og erhverv*. AKF Forlaget.

Mehlbye, Jill; Pauline Hagensen og Tue Halgreen (2000): *Et frit valg? Unges overgang fra grundskole til ungdomsuddannelse – særligt om unge fra etniske minoriteter*. AKF Forlaget.

Mejding, Jan (2002): *PISA 2000 – forventninger og færdigheder – danske unge i en international sammenligning*. Undervisningsministeriets tidsskrift Uddannelse (nr. 6).

Menckel, Ewa og Peter Westerholm (red.) (1999): *Evaluation in Occupational Health Practice*. Oxford: Butterworth-Heinemann.

Miljøstyrelsen (1997): *Udviklingen i den danske vandforsyningsstruktur* Arbejdsrapport fra Miljøstyrelsen, 62/1997.

Miljøstyrelsen (2003): *Skal husholdningernes madaffald brændes eller genanvendes?* Samfundsøkonomisk analyse af øget genanvendelse af organisk dagrenovation Miljøprojekt nr. 814, 2003.

Mohr, Lawrence B. (1995): *Impact Analysis for Program Evaluation*. Second edition. Thousand Oaks, Sage Publications.

Munksgaard, Jesoer og Anders Larsen (1996): *Samfundsmæssig værdi af vindkraft*. (The socio-economic value of wind power). AKF Forlaget.

Møller Pedersen, Kjeld (1979): *Effektmålingers teori og metode*. AKF Forlaget.

Møller, Flemming (2003): *Projektvurdering over tid – aspekter af diskonteringsproblemstillingen*. Danmarks Miljøundersøgelser. – Arbejdsrapporter fra DMU.

Møller, Flemming; Stig Per Andersen, Peter Grau, Henrik Huusom, Troels Madsen, Jørgen Nielsen og Lisbeth Strandmark (2000): *Samfundsøkonomisk projektvurdering af miljøprojekter*. Danmarks Miljøundersøgelser, Miljøstyrelsen og Skov- og Naturstyrelsen.

Navrud, Ståle og Richard Ready (eds.) (2002): *Valuing Cultural Heritage – Applying Environmental Valuation Techniques to Historic Buildings, Monuments and Artifacts*, Edward Elgar.

Nellemann Konsulenterne (1999): *Evaluering af de regionale kulturforsøg*, Århus.

Nielsen, Kurt Aagaard og Peter Vogelius (red.) (1996): *Aktionsforskning: medarbejderindflydelse på forskning og udvikling i arbejdsmiljøet: en antologi*. Arbejdsmiljøfondet.

Nordentoft, Merete (2004): Internationale erfaringer vedrørende udvikling af opsøgende behandlingstilbud til psykisk syge: I. Høgsbro (red.): *Socialpsykiatriens kompleksitet*. Forlaget Samfundslitteratur.

Nørgaard, Steen (1999): Evaluering, et paradoksalt redskab, *Nordisk Kulturpolitisk Tidsskrift 1999/2*, s. 81-91, Borås.

OECD (2000): *Measuring Student Knowledge and Skills: A New Framework for Assessment*, Paris.

OECD (2001): *Knowledge and Skills for Life – First Results from PISA 2000*, Paris. Publikationen kan downloades fra OECDs hjemmeside: www.pisa.oecd.org.

OECD (2002): *Sample Tasks from the PISA 2000 Assessment – Reading, Mathematical and Scientific Literacy*, Paris.

Olsen, Henning (2004): *Konstruktion og kvalitetssikring af surveydata. Et litteraturstudie om tendenser i kvantitativ metodelitteratur*. København.

Olsen, Leif (1998): *Viden der gør forskel. Erfaringsdannelse fra forsøg med videncentre på handicapområdet*. Ph.d.-afhandling. København: UCSF og Sociologisk Institut.

Olsen, Leif og Olaf Rieper (2001): Kommunernes brug af amtslig rådgivning på handicapområdet. *Nordisk Administrativt Tidsskrift Nr. 3 Oktober 2001*. (s. 285-304).

Olsen, Leif; Olaf Rieper og Christine Lindrum Iversen (2001): *Hvordan skulle vi klare os uden? – mennesker fortæller om deres brug af amtslig rådgivning på handicapområdet*. AKF Forlaget.

Owen, Gill (1999): *Public purpose or private benefit? The politics of energy conservation*, Manchester University Press. Manchester.

Patton, Michael Quinn (2002): *Qualitative Research & Evaluation Methods. 3. Edition.* Sage Publication.

Patton, Michael Quinn (1980): *Qualitativ evaluation and research methods.* Sage.

Pawson, Ray og Nich Tilley (1997): *Realistic Evaluation.* London, Sage Publications.

Porter, Michael E. (1990): *Competitive Advantage of Nations.* Free Press, New York.

Ramian, Knud og Birgitte Ahlgreen (1999): *Videnskabende netværk – Hvordan laver man praksisforskning i socialt arbejde med sindslidende.* Center for Evaluering publ. 53.

Ramian, Knud og Jónas Gústafsson (1997): *Liv i fokus.* Systime, Århus.

Ramian, Knud (2004): En fortælling om videnskabende netværk i det sociale arbejde med de sindslidende. I: Høgsbro (ed): *Socialpsykiatriens Komplexitet.* Samfundslitteratur.

Rantanen, J. (1999): Research challenges arising from changes in work-life. *Scand. j work environ & health* 1999; 25 (6, special issue): 473-483.

Rapport fra Det Kgl. Teaters forprojekt vedr: Registrering og optimering af forestillingernes kvalitet (1997), med bidrag af Søren Kjørup, Stig Jarl og Randi Svendsen, Trine Bille Hansen, Poul Erik Nielsen samt Peter Grønnegård (upubliceret).

Rieper, Olaf (1999): *Kommunernes brug af specialrådgivning på handicapområdet – resultater fra forundersøgelsen.* AKF Forlaget.

Rieper, Olaf og Ib Kristiansen (1994): *Evaluering af KS Forskning.* AKF Forlaget.

Rieper, Olaf og Jacques Toulemonde (red.) (1997): *Politics and practices of intergovernmental evaluation.* Transaction Publishers.

Rieper, Olaf; Leif Olsen, Käte Bruun Jacobsen og Vibeke Jørgensen (2001): *Kommunernes brug af amtslig rådgivning på handicapområdet*. AKF Forlaget.

Rosenkrands, Jakob (2002): Nyt analyseværktøj udpeger kulturens succeser og fiaskoer, *Mandag Morgen* nr. 40, s. 14-18.

Rosholm, Michael og Lars Skipper (2003): *Is Labour Market Training a Curse for the Unemployed? More Evidence from a Social Experiment*. Discussion Paper No. 716, IZA Bonn.

Sandø, Martin (red.) (2002): *Brugerinddragelse på flere måder*. Formidlingscenter Øst.

Schmid, Günther; Jacqueline O'Reilly og Klaus Schömann (1996): *International Handbook of Labour Market Policy and Evaluation*. Edward Elgar.

Skot-Hansen, Dorte (2002): Evalueringer som kulturpolitisk instrument, *Biblioteksarbejde* nr. 63, s. 45-55.

Socialministeriet og Finansministeriet (2002): *Web-håndbog om brugerinddragelse*. www.moderniseringsprogram.dk.

Spradley, James P. (1979): *The ethnographic interview*. Wadsworth Group/Thomson Learning.

Stein, Leonard I. og Alberto B. Santos (1998): *Assertive community treatment of persons with severe mental illness*. W.W. Norton & Co., New York.

Strauss, Anselm og Juliet Corbin (1990): *Basics of qualitative research: grounded theory, procedures and techniques*. Sage.

Taylor, Fitz-Gibbon og Lynn Lyon Morris (1987): *How to design a program evaluation*. Sage.

Teasdale, Thomas W. et al. (1997): Neuropsychological Test Scores in Relation to Return to Employment. I: *Neuropsychological Rehabilitation*, 1997, 7 (1), 23-42, Psychology Press (UK).

Teknologisk Institut, Arbejdsliv (2001): *Evaluering af forsøg med BST*. Taastrup.

The Arts Council of England (1994): *Appraisal Handbook*, The Arts Council, London.

Togebjerg, Mikael (2000): *Use of Statistical Analyses in the Evaluation of Energy Efficiency Projects – an Introduction*. København, AKF Forlaget.

Trafikministeriet (2003): *Manual for samfundsøkonomisk analyse anvendt metode og praksis på transportområdet*.

Utbildningsdepartementet (1990): *Statslig kulturpolitik i Sverige*. En svensk rapport til Europarådet, Utbildningsdepartementet, Stockholm.

Vedung, Evert (1997): *Public Policy and Program Evaluation*. New Brunswick and London. Transaction.

Vensborg, Per m.fl. (1999): *Opsøgende Psykoseteam*. DSI rapport 99.06/1999.

Zwicky-Hauschild, Birgitte (2003): *Evaluering af projektet »Det ryg'er noget! Kvalitetsudvikling i forflytning«*. BST Københavns Kommune november 2003. (www.bst.kk.dk/nyheder)

Øvretveit, John (1999): Evaluating occupational health interventions. I: Menckel, Ewa & Peter Westerholm (red.) (1999): *Evaluation in Occupational Health Practice*. Oxford: Butterworth-Heinemann, s. 38-52.

Aagaard Nielsen, Kurt (1996). *Arbejdets sociale orientering*. Forlaget Sociologi.

Emneregister

ADAM-modellen.....	264
Affaldshåndtering.....	279;287
Afværgemekanismer.....	41
Aggregerede data.....	62
AKF's forløbsregister.....	84;102
Aktionsforskning.....	26;77;330
Aktiveringsforanstaltninger.....	113;226;227
Aktiveringsprogrammer.....	37
Allokeringsbeslutninger.....	98
AMID (Akademiet for migrationsstudier).....	155
Analyseprogram.....	180;188
Arbejdsafprøvning.....	236;239
Arbejdsformidlingen (AF).....	215
Arbejdsløshed.....	59;61;133;166;211
Arbejdsmarkedsindsats.....	219;222
Arbejdsmarkedsuddannelserne (AMU).....	236
Arbejds miljøforskningen.....	240;241;242;244;251;322
Arbejds miljøområdet.....	7;240;241;242;244;251;307
Arbejdsprincipper.....	6;87;146;147;148;149;151;152;153;322
Arbejdsskader.....	92;249;316
Arbejdstilsynet.....	242;243;244;245;246;307;316;324
Auditundersøgelse.....	23;207
Behandlingsforløb.....	142;322
Behandlingsmodel.....	22;88;176;177;178

Benchmarking.....	39
Benefit-transfer	271
Beskrivende værdisætning.....	39
Beskæftigelsespolitik.....	210;215
Beskæftigelsesprojekter	229
Beslutningstager	93;255
Besparelser.....	252;256
Betalingsvillighed.....	23;94;261;268;269;270;271;295;296;301;303;304
Bevægeapparatskader	242;247;320;321
BIKVA-modellen	45
Bioforgasning	280;281;282;283;285;286;287
Black-box.....	59
Bortfald.....	141;151
Brugerevaluering	20;88
BST	243;247;248;250;333
Budgetøkonomiske analyser	269
Center for Alternativ Samfundsanalyse (CASA).....	251
Center for arbejdsmarked og Social forskning (CLS).....	239
Center for Arbejdsmarkedsforskning ved Aalborg Universitet	214
Chi ² -test	127;131
CO ₂ og SO ₂ -afgifter	259
Contingent valuation.....	305
Cost-benefit	8;20;224;268;269;270;272;274;280;295;308
Cost-benefit-analyser på miljøområdet.....	8;268
Cost-effectivness.....	20
Cost-efficiency.....	20
COWI	214;318;345
Daghøjskoler.....	238;239
Dagrenovation	92;273;279;280;281;283;285;286;329
Danmarks Evalueringsinstitut.....	95;312
Danmarks Miljøundersøgelser	272;321;329
Databanker.....	47
Deltagerobservation	47
Desk research.....	217;218
Det empirisk-analytiske paradigme	26

Det fortolkningsvidenskabelige paradigme	27
Det handlingsrettede paradigme	29
Det handlingsrettede paradigme	28
Det kommunale Momsfond	100
Det kontrollerede eksperiment	59;63
Det kritiske paradigme	28
Det psykosociale område	251
Det økonomiske råd	272
Diagnoser	173;224
Dialog	32;71;174;187;190;191;222;274;327
Dokumentation	16;18;65;73;186;219;246;247
Dosis-respons forskningen	241
Drikkevand	276
Dybdeinterview	26;47
Døgnanbringelse	119;120;121
Døgnehjemmepleje	143
Economic impact analyser	295;297;299
Effektevaluering	4;5;19;21;57;65;84;92;94;315
Effektevalueringer	5;57;63;66;253
Eksogene	99
Ekstern evaluering	51;56
Empowerment evaluering	71;72
Energiafgifter	63
Energibesparelser	20;46;58;64;252;253;254;257;258;259
Energieffektivisering	256;258
Energiledelse	256;258
Energisektoren	7;252;254;266;308;309
Erhvervspolitik	264;265
Etik	23;49;207
Etnografisk feltarbejde	47
EU-Kommissionen	93;255
EU's strukturfondsprogrammer	37
EU-tilskud	275
Evaluators rolle	5;19;26;32;78
Evalueringsgenstanden	34;35;37;46;49;173;175;183

Evalueringsbegrebet	4;15
Evalueringsenhed.....	53;55;312
Evalueringsforskning.....	69;70;78;240
Evalueringsfunktion.....	46
Evalueringskompetence	54;55
Evalueringsmodel	18;23;24;35;44;45;83;182;183;290;291;296;327
Evalueringspraksis.....	15;29;77;243;291
Evalueringsskik.....	4;49;50
Evalueringsspørgsmål.....	35;43;45;164;212;213;248;249;250
Ex ante	255;267;271;294
Faglig kompetence.....	56;189
Fastlåsning.....	238
Feedback.....	69;72;204
Finansministeriet.....	100;106;272;308;320;332;345
Fokusgruppeinterview	186;249
Fokusgruppeinterview	47;89;92;217
Foranalyse.....	50
Foreskrivende værdisætning	39
Forforståelse	27
Formativ evaluering.....	19;26;77;88
Formative evalueringer.....	68;325
Formidling	15;16;18;34;47;49;106;163;173;174;177;190;211;218;266
Formidlingsform	48
Formidlingskanal	48
Forskningsbaserede evalueringer.....	16
Forskningsbaseret erfaringsopsamling.....	74;76;209
Forundersøgelsen.....	183;184
Frankfurterskolen.....	28
Fritidsaktiviteter.....	73;85;107;120;124;125;127;128;131;168
Fritidstilbud	73;75;85;119;120;121;124;125;128
Frivillige aftale	63
Funktionsevne.....	138;139;140;141;143;148;149;150
Følsomhedsanalyser.....	286
Genanvendelse.....	92;273;279;280;282;284;286;329
Gennemførelse.....	17

Gennemskuelighed	189
Genoptræning	138;145;205;328
Grounded theory	30;47
Gruppeinterview	45;170;179
Handicaporganisationerne	184
Handleplaner.....	15;220;222
Hazard-funktionen	231
Hermeneutisk fortolkningsramme	135
Hjemløse.....	7;22;89;205
Hjemmehjælp.....	37;138;140;146;147;148;150;315
Hjemmehjælperne.....	147;148;149;150;151;152;153
Hjernesker	6;22;88;176;177;178;181;182;192;197;325
Hjælpe midler	138;175;247
Husprismetoden	271;277
Igangsætningsydelse	229;237
Impact	8;17;94;295;296;297;299
Implementere	174;178;256
Incitament	259
Indikator.....	101;157;158;159;162
Indirekte effekter	278
Individdata	62
Indsatser	16
Indsatsgruppe.....	58;59;60;61
Indvandrere	154;156;157;158;159;160;161;162;164;168;169;210;323
Informationskampagne	57;254
Institut for grænseregionsforskning	214
Institut for Miljøvurdering	272
Instrumentel	18;38
Integration på arbejdsmarkedet.....	158
Integrationsloven	157;158;159;160;161;162;323;324
Integrationsområdet	6;154;155;156;158;323
Interessent evaluering	20;87;89
Intern evaluering.....	51;56
Interventionsformer	134;136
Interventionsteori.....	41

Interviewundersøgelse	23;87;114;125;134;164;186
Investeringsstøtte	63
Jobtræning	229;230;234;235;237
Journalundersøgelse	87;133
Kausale sammenhænge	63
KL	100;239;306;312
Klimaforandringer	252
Kollegial evaluering	20
Kolonien Filadelfia	176;177;178;181;325
Kommunale job	236;239
Kompleksitet	22;36;37;46;177;183;330
Kompostering	280;281;282;283;285;286;287
Konceptuel	28;38
Konjunkturberegning	265
Kontanthjælpsmodtagere	210;226;227;234;235;236
Kontekster	27;32;180;293
Kontrol	17;32;41;46;67;72;139;155;199;320
Kontrolgruppe	27;58;59;228;232;233
Koordination	190
Koordinationsudvalgene	219;222
Kritisk teori	28
KUBI-modellen	182
Kulturinstitutioner	288;291;293;297;300
Kulturområdet	8;288;289;290;292;295;296
Kulturprojekt	289;294
Kunstnerisk kvalitet	289;293;294
Kvalitative analyser	55;156
Kvalitetsudvikling	92;95;190;247;248;250;316;324
Kvasi-eksperimenter	99
KVL	272
Kyotoaftalen	272
Københavns Universitet	204;272;306;307;308;309
Laboratorieeksperiment	226
Landbrugsproduktion	275;278
Lederudvikling	36;37;38;40;52

Ledighedsprocent.....	103;160;162
Legitimere.....	54;288
Legitimerende.....	18;299
Livscyklusanalyse.....	281
Livsformer.....	135
Logistisk regression.....	62;151
Lærertimer.....	98;104
Læring.....	17;19;46;54;247;251;320;321;322;327
Makroøkonomisk model.....	261
Marginale.....	268;269;276
Medicinsk Teknologivurdering (MTV).....	203
Meningskondensering.....	47
Miljøeffekterne.....	262;270;280
Miljøgoder.....	23;261;270;271
Miljøministeriet.....	272
Miljøpolitikken.....	28
Miljøstyrelsen.....	272;273;276;279;281;306;329
Miljøvurderinger.....	261
Ministeriet for Flygtninge, Indvandrere og Integration.....	157;163
Misbrugere.....	7;22;89;205
Misbrugsproblemer.....	133
Monitorering.....	15;91;243
Målgrupper.....	22;50;153;206;212;215;217;218;220;289
Mål-middel-sammenhænge.....	77
Målopfyldelses-evaluering.....	19
Målopfyldelsesmodel.....	46;294
Narrative analyser.....	47
New Public Management.....	288
Nøglepersoninterview.....	126
Objektiv beskrivelse.....	28
Observation.....	47
Observatør.....	26;27;32
OECD.....	21;96;106;108;109;110;111;308;312;328;330
Offentlige forvaltning.....	119
Omkostningsanalyse.....	261

Omkostningseffektiv.....	257
Oplysende	18;38
Organisationssociologisk	23;207
Organisationsteori.....	40;309
Organisk affald	8;92;279
Outcome.....	17;33;140
Output	15;17;33;97;139;140;243;275;295
Overlevelseskurven.....	228;230
Overvågning	143;242;243;246;251
Oxford Reseach	214
Paradigmer.....	4;15;24;25;26;27;30;83;241;292
Peer review	20;292;294
Personaleomsætning	92;249;250
PISA	5;21;84;96;106;107;108;109;110;111;112;308;315;329;330
Placebo	59
PLS-Rambøll Management.....	214
Policy-område.....	252
Porters diamant	261
Postspørgeskemaer	47
Praktisk anvendelse	16;17;29;257
Proces evaluering.....	19;85
Produktionsomkostningerne	23;262;263
Produktionsskoler	234;238
Programeffekten	232
Programteori	41
Psykosociale miljø	240
Publikum.....	94;289;292;298
Publikumsinterview	298
Registerdata	47;62;84;102;103;105;156;159
Registeroplysninger	47;91;92
Regressionsmodel.....	61;103
Regulering	16;39;119;204;240;256;309
Rehabilitering	6;88;177;181;325
Rejseomkostningsmetoden	271
Rekreative værdier.....	277;322

Responsiv.....	71
Risø.....	266
Rummelige arbejdsmarked	90;211;219;220;221;222
Sagsjournaler	134
Samfundsvidenskabernes filosofi	28
Samfundsøkonomisk vurdering	261
Sammenligningsgrupper	60
Sanktioner	17;256
Scenario	281;284;285
Score Card	294
Selektionsskævhed.....	234;238;239
Selvevaluering	71;72;77;213;292
Sindslidende.....	7;22;89;200;202;205;327;331
Skoleområdet.....	5;21;95;96;137
Skoleressourcer.....	5;97;98;99;102;103;105;156;308
Skolevejledningen.....	112
Skov- og Naturstyrelsen.....	272;273;306;329
Skovrejsningsprojekt	8;271;273;274;277
Sociale arv	104
Socialforskningsinstituttet	157;245;315;317;321;323;325;328
Socialkonstruktivisme.....	47
Socialministeriet	138;174;332;345
Specialpædagogisk	119
Specialundervisning.....	122
Sprogskole	235;236;239
Spørgeskemaundersøgelser.....	85;88;90;126
Statens Institut for Folkesundhed.....	242;245;326
Statistikprogrammet.....	180;188
Statistisk signifikante.....	97;104;127;161
Stigmatisering.....	173
Stordriftsfordele.....	163
Strukturinvesteringer	265
Styrelsen for Social Service.....	138
Summative evalueringer	68
Sundhedspolitikken.....	28

Survey	45;62;224
Svarprocent	126;127;151;164;170
Sygefravær	92;249;250
Sygehusstatistik	145
System modeller	19
Systematisk indsamling	17
Systemmodel	22;46
Taktisk anvendelse.....	18;38
Telefoninterview	47
Teori om indsatsen.....	40
Teori om målområdet	40;41
Theory of action.....	30;41
Tolkebistand	155;175;191
Tosprogede	6;88;113;121;155;167;168;169;170;171;172
Trafikministeriet	272;333
Trivsel.....	5;86;105;120;128;129;130;140;141;224
Udbudsmateriale	45;218
Uddannelsesvalg	88;97;117;156;163;166
Udsatte grupper.....	6;9;22;46;85;89;125;192;193;194;202;308
Udviklingscenter for Specialrådgivning	183;184;190
Ungdomsuddannelse.....	6;88;96;97;101;103;104;105;163;164;166;167; 236;238;328
Ungdomsvejledningen	85;112;114;115
Validitet	108
Varighedsmodellen	231
Velfærdøkonomiske cost-benefit-analyser	269
Videnscentre	184;308;312;330;345
Videnskabeligt paradigme	24
Vilde problemer	175;193
Vindkraftudbygningen	7;93;260;264
Virkemidler	210;211;212;252;253;254;255;256;257;259
Virksomheder	17
Virkningsmekanismer	36;46;251
Værdikriterier	20;35;39
Værdisætning.....	8;39;93;94;300;322

Værdisætningsmetode.....	271;305
Ældre	6;75;87;138;140;143;144;145;146;148;149;150;161;307
Ændringsproces	41
Æstetik	288;326
Økonomiske modeller.....	20
Årsag-virknings-forholdet	58

Noter

1. En nem og tilsyneladende tilforladelig løsning af dette dilemma er at anbefale summative evalueringer, når der er tale om veletablerede organisationer, som har haft tid til at opbygge indsatsen og anbefale formative evalueringer, når der er tale om, at organisationen af indsatsen stadig er under udvikling (Fitz-Gibbon and Morris *ibid.*).
2. En mere udførlig gennemgang af kvalitetsbegrebet findes i fx Bunnage m.fl., 2001 og Hansen, 2003.
3. Personalets personlighed kan naturligvis påvirkes over tid ved en rekruttering af personale.
4. Der var tale om en generel model for de forskellige foranstaltninger, der indgik i undersøgelsen.
5. Fx FN's konvention fra 1994: Standardregler om lige muligheder for handicappede. Se fx hjemmesiderne www.clh.dk (Center for Ligebehandling af Handicappede) eller www.handicap.dk (De Samvirkende Invalideorganisationer).
6. Denne lov er evalueret, og resultaterne er fremlagt i evalueringsrapporten: COWI A/S og Socialt Udviklingscenter SUS (2002): Evaluering af erfaringerne med institutionsbegrebets ophævelse på handicapområdet, 1998-2002
7. Projektbeskrivelsen er tilgængelig via Socialt Udviklingscenters hjemmeside: www.sus.dk. Her kan man også finde oversigter over anden relevant litteratur og link til Web-håndbog om brugerinddragelse (Socialministeriet og Finansministeriet, 2002).
8. Et videntcenter har arbejdet målrettet med denne problematik i forhold til børn og unge siden 1993. Der er tale om Videncentret om Kommunikation og Multiple Funktionsnedsættelser hos Børn og Unge uden et Talesprog, der bl.a. kan findes via hjemmesiden: www.vikom.dk.
9. Ved at gøre teksten elektronisk tilgængelig vil mennesker med synsnedsættelse få mulighed for at anvende talesyntese. Dansk Center for Tilgængelighed (www.dcf.dk) har udarbejdet anvisninger for, hvordan man opnår bedst mulig tilgængelighed for alle via hjemmesider og elektroniske tekster.
10. Der findes eksempelvis landsdækkende videntcentre, der har til opgave at indsamle, udvikle, bearbejde og formidle viden inden for følgende områder: Ordblindhed, stammen, handicapdræt, synshandicap, autisme, bevægelseshandicap, døvblevne, døve og hørehæmmede, døvblindblevne, døvblindfødte, hjerneskade, psykiatri og udviklingshæmning, socialpsykiatri, socialt udstødte, epilepsi og ældreområdet. Herudover findes en række andre centre, der også arbejder med viden på handicapområdet: Center for Lige-

behandling af handicappede, Center for små Handicapgrupper, Center for Tilgængelighed, Det Centrale Handicapråd, Formidlingscenter Øst og Hjælpemiddelinstittet.

11. AKF gennemførte efterfølgende endnu en brugerevaluering af amternes rådgivning på handicapområdet, hvor brugerne var mennesker med handicap (Olsen 2001). Denne evaluering bruges enkelte steder til at belyse problemstillingen med at få informationer og viden om og fra mennesker med handicap.
12. Det kan diskuteres om betegnelsen »at indsamle data« er den bedste, men den bruges her, idet det er en udbredt måde at omtale, hvad der mere præcist kan betegnes som »konstruktion af data«. Data er nemlig ikke noget, der bare ligger og venter på at blive samlet op. Data konstrueres i processer, hvori indgår forskere, evaluatore, teorier, begreber, metoder og leverandørerne af data fx medarbejdere, brugere, borgere, politikere mv.
13. Flere oplysninger om SPSS (Statistical Package for the Social Sciences) findes på hjemmesiden www.spss.dk.
14. Flere oplysninger om N'vivo findes på www.qsrinternational.com og i dansk sammenhæng på www.hraconsult.dk. Det skal understreges at der findes en lang række andre relevante programmer, som man kan anvende til kvalitative analyser af interviewmateriale og andre tekster.
15. En kort beskrivelse af modellen kan hentes via hjemmesiden www.sus.dk, ligesom der findes omtale og introduktion på hjemmesiden www.moderniseringsprogram.dk.
16. Udviklingscenteret blev oprettet af Amtsrådsforeningen, Københavns og Frederiksberg Kommuner i 1998 med det formål at følge udviklingen af amternes specialrådgivning på handicapområdet samt rådgive om tilrettelæggelse og dimensionering af lands- og landsdelsdækkende specialrådgivningsfunktioner. Se også www.udvcenter.dk.
17. Forundersøgelsen blev gennemført som fokusgruppeinterview i tre casekommuner, udvalgt efter varierende størrelse og beliggende i forskellige amter. Deltagere i gruppeinterviewene var medarbejdere, som arbejdede med handicapsager. Der blev afholdt fem gruppeinterview med i alt 26 medarbejdere fra social- og sundhedsforvaltningen og fra skole- og kulturforvaltningen i de respektive kommuner.
18. Flere love har betydning for myndighedernes pligt til at inddrage borgerne og give dem medindflydelse på det sociale område bl.a.: Retssikkerhedsloven, serviceloven, forvaltningsloven og offentlighedsloven, der alle indeholder regler for, hvordan kommuner og amtskommuner skal behandle sager på handicapområdet.
19. SFI (Socialforskningsinstituttet, www.sfi.dk) afholder halvårlige seminarer i »Socialpolitik og handicap«, hvor alle, der forsker i eller arbejder som evaluatore inden for området »Socialpolitik og handicap«, er velkomne til at deltage. Herudover er miljøet om-

kring FORSA (Foreningen for forskning i socialt arbejde), der samler fagfolk med interesse i området og herudover har forbindelser til nordiske miljøer, se www.forsa.dk.

20. Her er der bl.a. dannet netværket NDDR (Nordic Network on Disability Research, www.nndr.dk), ligesom Nordic social policy research meetings giver anledning til, at forskere og evaluatore kan mødes til diskussion af deres arbejde på handicapområdet.
21. Efter konferencerne udarbejdede Udviklingscentret to publikationer, der kan hentes på hjemmesiden www.udvcenter.dk: 1) Temahæfte fra Udviklingscentrets konference den 26. maj 2003 om »Større sammenhæng og kvalitet i amternes rådgivning på handicapområdet« og 2) Temahæfte om Amternes Rådgivning på Handicapområdet. Sammen drag og interview fra Udviklingscenter for Specialrådgivnings temadag den 14. november 2001.
22. I denne analyse anvendte vi en fysik-datalogi-studerende, som havde gode programmeringsegenskaber, men intet kendte til data. For at overkomme dette problem arbejdede han nært sammen med flere af de øvrige projektdeltagere, ligesom vi jævnligt holdt møder for at sikre, at konstruktionen af forløbshistorikken blev så akkurat, som det overhovedet var muligt.
23. Dette antyder også, at nogle aktiveringseffekter ikke »fanges« med denne teknik, men mere herom i næste afsnit.
24. En gennemgang af den statistiske model kan findes i Bolvig m.fl. (2001).
25. Der opstår kun selektionsskævhed i det omfang, der er systematisk selektion ind i forskellige aktiveringsformer på baggrund af *uobserverede* faktorer. Selektion ind i forskellige aktiveringsindsatser på baggrund af fx uddannelsesnivea vil derfor ikke medføre nogen selektionsskævhed. Alene af denne grund er det naturligvis hensigtsmæssigt at have et så stort sæt af baggrundsfaktorer som overhovedet muligt med i analysen.
26. En grundig beskrivelse af datamaterialet findes i Bolvig m.fl. (2001).
27. Inden for denne forskning er det – forskelligt fra psykologisk og samfundsvidenskabelig forskning – muligt at opnå alment gyldige resultater, fx vil en undersøgelse af neurotoksiske stoffers virkning på menneskets helbred være den samme i et hvilket som helst land som i Danmark.
28. Der er nogle områder, hvor evaluering er på dagsorden, fx området »occupational health service« (Menckel & Westerholm), og psykosocial/organisatorisk område (Griffith 1999, Elo 2002, Lindström 2003).
29. AKF har kendskab til begge projekter, i det førstnævnte projekt som evaluator af gennemførelsen af den første runde af undersøgelsen af virksomhedernes forebyggende ak-

tiviteter (Rieper og Eskelinen 2003) og i det sidstnævnte projekt som evaluator af virkningerne af de projekter, som fik støtte fra puljen til forebyggelse af bevægeapparatskader som følge af tunge personløft (Eskelinen og Hansen 2004; Hansen m.fl. 2004).

30. Evaluatoren er fra den samme BST som projektkonsulenterne, men hun er ekstern i forhold til udviklingsprojektet.
31. Der findes diskussion om evalueringspraksis inden for »occupational health services« (Menckel & Westerholm 1999), evidensbaseret medicin (Miettinen 1999) samt interventionsprogrammer (FIOH, NIOSH).
32. Til gengæld blev der under kulturminister Jytte Hilden igangsat en omfattende udredning og evaluering af dansk kulturpolitik, som mundende ud i et værk på 17 bind (Duelund og Bille Hansen, 1994 samt Duelund, 1995).
33. I Sverige er der i øvrigt udviklet flere andre ganske omfattende evalueringsmodeller for kulturområdet. Herunder for det første »Kulturmåtarprojektet« om evaluering af regionale kulturinstitutioner på teater-, musik-, biblioteks- og museumsområdet (Landstingsförbundet, 1994) – modellen har dog ifølge Langsted m.fl. (2003) ikke været anvendt i praksis. For det andet en håndbog om opfølgning, evaluering og kvalitet i kultur- og fritidssektoren i kommuner og regioner (Lundkvist m.fl., 1998). For det tredje har Statens Kulturråd udarbejdet en model om regnskab for scenekunst (Hallberg, 2003), som vedrører indsamling og præsentation af oplysninger om publikums opfattelse, om massemediernes opfattelse og de ansattes opfattelse af teatres kvalitet.
34. Modellen minder om en model opstillet i Klausen (2001, p. 123). Her udspænder de tre vektorer blot rummet mellem stat, marked og civilsamfund til beskrivelse af sektorer, organisationer m.m.
35. Som del af min ph.d.-afhandling.
36. Det mest berømte eksempel på en rekvireret evaluering af denne slags er Exxon Valdez-eksemplet, hvor supertankeren Exxon Valdez stødte på grund på et rev ved Alaska og spildte 11 millioner gallons råolie i havet. Denne naturkatastrofe gav anledning til en større erstatningssag, og i forbindelse med skadesvurderingen nedsatte the National Oceanic and Atmospheric Administration (NOAA) et ekspertpanel med nobelpristagerne Kenneth Arrow og Robert Solow som de ledende kræfter. Panelet skulle rådgive NOAA om contingent valuation-metoden var pålidelig nok til at kunne anvendes til økonomisk vurdering af skadernes omfang. Panelets svar var i grove træk »ja« (Arrow, K.J., m.fl. 1993).
37. Disse metoder har den fordel, at de – til forskel fra contingent valuation-metoden – tager udgangspunkt i observeret markedsadfærd. Til gengæld kan de kun anvendes til at estimere brugsværdier, ikke eksistensværdier, og dermed heller ikke den totale værdi af et offentlig gode.