



**TEMA: EFFEKTMÅLING  
METODER, ANVENDELSE OG EFFEK-  
TER**

Dette temanummers artikler er blevet til på baggrund af sessioner på DES årskonference i september 2009.

**NÅR EVALUERING NYTTER**

**EFFEKTUNDERSØGELSE INDEN FOR  
DET SPECIALPÆDAGOGISKE OMRÅDE  
- METODISKE PROBLEMSTILLINGER OG  
OVERVEJELSER**

**INNOVATIV EVALUERING  
- EVALUERING DER SKABER INDFLYDELSE**

**CONTRIBUTION ANALYSIS  
- ET ALTERNATIV TIL KLASSISKE EFFEKTDE-  
SIGNS?**

**HVORDAN SKAL EN DES KONFERENCE  
VÆRE, FOR AT DU SYNES, AT DEN ER  
GOD?**

## NÅR EVALUERING NYTTER

*Hvis evaluering skal have mening forudsætter det, at ledelsen og organisationen har gjort sig klart:*

- *Hvilke effekter på borgenes levevilkår, de har ansvaret for*
- *Hvilke ydelser de anvender for at levere effekterne*
- *Hvilke parametre, de skal måle effekterne og ydelserne på*

*Hvordan målinger og dokumentation gøres til en integreret del af organisationens virke*

*Endvidere skal lederen og organisationen selv have forholdt sig til og efterspurgt evalueringen med udgangspunkt i overstående afklaring.*

*Det er mit postulat, at disse præmisser sjældent er opfyldt og derfor bliver diverse evalueringer ofte upræcise og uanvendelige i en løbende udvikling af organisationen og dens resultater.*

Nu har der jo været fokus på resultater i mange år, kunne man sige. New Public Management bølgen, som skyllede ind over landet i 80'erne satte netop fokus på effektivitet og resultater. Det førte til et stigende antal målinger, men de blev gennemført i flæng mellem forskellige resultatparametre uden en overordnet systematik, der gjorde måleresultaterne anvendelige i et proaktivt ledelsesforløb.

### TO HOVEDPROBLEMER

Jeg tror, at vi skal bruge mere tid på at forstå, hvilke resultater organisationen skal levere til borgerne – både på kort og lang sigt – inden vi designer og gennemfører evalueringer. Desuden skal vi følge mere systematisk op på resultatleverancen – både på de samlede overordnede effekter på borgernes levevilkår og på de enkelte ydelsers bidrag til denne effekt. Der er altså to hovedproblemer:

- **At forstå hvilke resultater, der skal leveres**
- **At forstå, at det skal dokumenteres, at resultaterne leveres og at det er lederens ansvar at det gøres**

### VI DEFINERER IKKE DE ØNSKEDE RESULTATER SKARPT NOK

Kernen i resultatforståelse er at forstå, hvilke dele af borgernes liv organisationen skal varetage og hvad den konkret gør for at leve op til dette ansvar. Det er illustreret i nedenstående figur.

De fleste i den offentlige sektor ved, at det handler om borgernes levevilkår (cirkel 1). Men de enkelte organisationer må også forstå, hvilke konkrete dele af borgernes levevilkår, som den skal tage sig af (cirkel 2). De må være bevidste om, at deres fagområder er et middel til at varetage organisationens ansvar over for borgernes levevilkår (cirkel 3). Men fagdisciplinerne fører kun til effektive resultater, hvis de formidles gennem konkrete sluttydelser, som ved deres udførelse påvirker borgernes liv (cirkel 4).

Problemet er, at cirkel 2 og 4 ikke fyldes systematisk ud i de fleste offentlige organisationer. Borgernes levevilkår er italesat i den offentlige sektor, men ofte ikke konkretiseret i operationelle ansvarsområder, hvor organisationen systematisk følger med i udviklingen – i effekten af dens arbejde (cirkel 2). Fagdisciplinerne er normalt heller ikke oversat til konkrete sluttydelser, som til sammen skaber de ønskede effekter på borgernes levevilkår inden for ansvarsområdet (cirkel 4). På den måde bliver det uklart for organisationen, hvilke resultater den skal levere og hvilke ydelser den vil anvende til at producere dem. Dermed bliver produktionen af resultaterne ikke effektiv.

Min pointe er, at hvis den offentlige organisation stopper sin resultattankegang i cirkel 1 og 3 bliver arbejdet ikke målrettet. Det bliver vanskeligt at lave præcise evalueringer og dermed bliver resultaterne heller ikke så gode, som de kunne blive.

## VI DOKUMENTERER IKKE VORES RESULTATER GODT NOK

Jeg forstår godt, at resultatdokumentationen i den offentlige sektor aldrig for alvor har bevæget sig ud over den traditionelle budgetstyring garneret med stikprøver omkring effektivitet m.v. Hvordan skal man kunne være systematisk i dokumentationen af sine ydelser til borgerne, hvis man ikke klart har formuleret de ønskede effekter med tilhørende ydelser på et konkret og operationelt niveau? Det kan bare ikke lade sig gøre.

Jeg ved, at jeg bevæger mig ud i konfliktfyldt farvand, når jeg siger, at enhver offentlig leder ikke alene skal definere sine konkrete slutydelse, men også skal dokumentere eller sandsynliggøre, at ydelserne fører til resultater, som

- **Borgerne er tilfredse med på kort og lang sigt**
- **Har den nødvendige faglige kvalitet**
- **Leveres økonomisk effektivt**
- **Leveres i et godt arbejdsmiljø**
- **Er afstemt med øvrige samfundshensyn**

Måletyranni, fjernelse af ressource og fokus fra borgerne, kontrol, umuligt i en politisk styret organisation med skiftende mål, vil nogen sige. Efter min opfattelse er det blot undskyldninger for ikke at gøre det, som må forventes af enhver leder, nemlig at han dokumenterer, at han leverer resultater.

For at undgå misforståelser vil jeg gerne dvæle ved ordet "dokumentere". Jeg taler ikke om et centralt styret kontrolsystem. Jeg taler om et fleksibelt målesystem, som i hovedsagen defineres af lederen selv i forståelse med egen leder eller "bestyrelse". Jeg ved godt, at ikke alt kan måles. Men så kan det sandsynliggøres ved kvalitative vurderinger m.v.. Naturligvis skal der anvendes målemetoder, der er tilpasset faget og som er balanceret ressourcemæssigt. Det er lederens egen opgave at fastlægge sine succeskriterier og dokumentere dem. Han er den bedste til det. Gør han

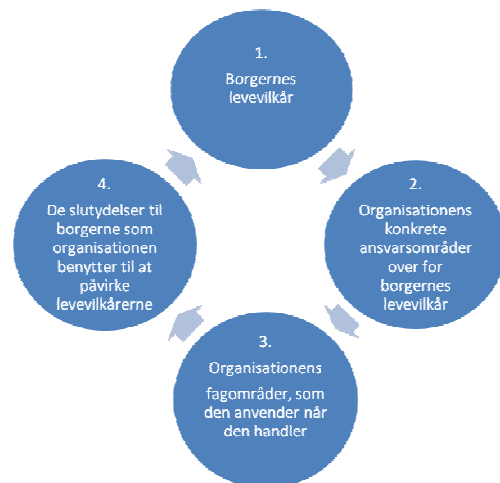
det ikke, bliver hans ledelse uforpligtigende. Og naturligvis skal han "stå på" forskernes og udviklernes resultater omkring evalueringsmetoder m.v.

## AFSLUTNING

Der er meget mere at sige om resultatforståelse. Eksempelvis håndteringen af modstridende resultato ønsker fra forskellige borger-/brugersegmenter eller underdelingen af det, jeg i denne artikel kalder slutydelse i ledelsesmæssige slutydelse. Ligeledes skal man forholde sig til de interne slutydelse, altså de leverancerne fra interne støttefunktioner, som bidrager til at organisationen effektivt kan levere varen til borgerne. Omkring disse problemstillinger og eksempler, som uddyber dem og artiklens indhold, må jeg henvise til min bog, "Når ledere skaber resultater – en bog om offentlig ledelse".

Ole Bach har egen virksomhed ved navnet Ledelsesforståelse. Ole Bach har været leder både i det private erhvervsliv og i den offentlige sektor, senest direktør for Vej og Park i Københavns Kommune i perioden fra 1999-2008. I 2008 og 2009 har Ole Bach skrevet bogen "Når ledere skaber resultater – en bog om offentlig ledelse" og rådgivet kommuner om ledelse og organisation. Bogen kan købes hos forlaget Birmar ([www.birmar.dk](http://www.birmar.dk)).

Henvendelse [olebach@ledelsesforstaaelse.dk](mailto:olebach@ledelsesforstaaelse.dk) tlf 30264268



### Den Offentlige sektors resultatcyklus

Cirkel 1 og 3 er i dag organisationernes hovedfokus. Større effektivitet i den offentlige sektor forudsætter en konkretisering af, hvilke dele af borgernes levevilkår organisationen har ansvaret for og en klar definition af, hvilke konkrete ydelser (slutydelser), som organisationen leverer til borgerne for at forbedre deres levevilkår – altså større fokus på cirkel 2 og 4.

## EFFEKTUNDERSØGELSE INDEN FOR DET SPECIALPÆDAGOGISKE OMRÅDE – METODISKE PROBLEMSTILLINGER OG OVERVEJELSER

*Strategisk Program for Velfærdsforskning blev i 2004 blev tildelt knap 100 mio. kr. på Finansloven. Programmet skulle overordnet set kaste lys over, hvilken effekt forskellige velfærdsforanstaltninger har. Et af de velfærdsområder, der skulle belyses, var specialundervisningsområdet, idet der skulle sættes særlig fokus på at identificere de komponenter, der indgår i specialpædagogikkens iværksættelse og gennemførelse og disses effekt. Der anvendes i Danmark ca. 10 mia. kr. årligt til specialundervisning. Det skal metodemæssigt præciseres, at randomiseret udvælgelse til forskellige indsatsformer, herunder brug af kontrolgrupper, ikke er hverken etisk eller praktisk mulig, da elever med særlige behov har ret til at få en specialpædagogisk indsats. Det er derfor nødvendigt at betjene sig af "blødere" quasiekperimentelle designs.*

### TEORETISK GRUNDLAG FOR UNDERSØGELSEN

Specialpædagogiske indsatser er tværfaglige og sammensatte, og de må derfor undersøges via komplekse teoretiske indfaldsvinkler og multiple metoder. Det teoretiske grundlag både i forhold til de metodiske overvejelser og de konkrete empiriske strategier er funderet i et systemisk perspektiv, et in- og eksklusions perspektiv og et aktør perspektiv med bidrag fra både læringspsykologi, udviklingspsykologi, selvpsykologien og positiv psykologi. Der må endvidere anvendes et Mixed Methods design, dvs. en undersøgelse båret af både kvantitative og kvalitative forskningsstrategier.

### PROBLEMTYPE SAMT ARTEN AF SPECIALPÆDAGOGISK BISTAND

En specialpædagogisk indsats kan finde sted i en lang række af former, ligesom de elever, der modtager specialundervisning, har vidt forskellige karakteristika. Alene dette er med til at "forplumre" en effektmåling. I den aktuelle undersøgelse er den absolut hyppigste problemtype specifikke indlæringsvanskeligheder, dvs. vanskeligheder i specifikke fagområder. Næsthypigst er socioemotionelle vanskeligheder, tæt fulgt af generelle vanskeligheder, hvorved forstås vanskeligheder over hele den boglige fagrække. ADHD udgør en relativt beskeden, mens andre kategorier er meget sparsomt forekommende. Den hyppigste form for indsats er, at eleven modtager støtte i klassen fra en

ekstra lærer fra teamet, og ved i dele af undervisningstiden være i støtte/ressourcecenteret. Derpå kommer, at man modtager støtte ved holddannelse uden ekstra lærerressource, støtte i klassen fra støtte/ressourcecenterlærer eller lignende og deltagelse i koncentreret kursus i perioder. Relativt sjældnest forekommer at man modtager støtte uden for klassen før eller efter skoletiden og at faglærer får supervision fra støtte/ressourcecenter, enten af en lærer med ekspertise i adfærd-, kontakt og trivsel eller af en læsevejleder - eller modtager hjælp fra Pædagogisk Psykologisk Rådgivning.

Det skal til ovennævnte bemærkes, at selv om indsatser typemæssigt har været relativt ens, er de fundet sted med overordentligt forskellige "doseringer", fra nogle relativt få timer til en ganske massiv indsats. Der er endvidere en meget forskellig fordeling hen over skoleåret, hvor der er eksempler på, at kursusformen bruges i bestemte måneder – eller at al støtte ophører fra maj måned, da alle ressourcer er brugt. Et andet "forstyrrende" forhold for effektundersøgelsen er, at de fleste elever har fået specialpædagogisk støtte i flere år, nogen allerede fra 1. klasse, men mere almindeligt fra starten af 2. eller 3. klasse, ligesom den specialpædagogiske støtte for de fleste elever ikke ophører ved slutningen af undersøgelsesperioden. Der er i mange tilfælde en relativt svag eller helt fraværende registreringspraksis, og det har ikke været muligt med sikkerhed at fastslå, hvilke indsatser, der er sket inden undersøgelsesåret. Konklusionen på

foranstående er, at der ikke er to elever, der har modtaget den samme indsats, hverken i undersøgelsesperioden eller over deres samlede skoleforløb indtil undersøgelsesperiodens slutning. Den specialpædagogiske indsats kan, på grund af dens varighed og individuelle forskellighed, ses som enhedsskolens løsning på et differentieringsproblem over for de ca. 10 % elever, der har særlige behov, som er nødvendige af have fokus på, hvis problemer i øvrigt er helt forskellige.

### METODERNE I UNDERSØGELSEN

Der indgår en lang række af datakilder i undersøgelsen. De kan rubriceres i følgende kategorier

- **Oplysninger fra skolernes hjemmesider, virksomhedsplaner og anden skriftlig dokumentation**
- **Semistrukturerede interview med skolernes ledelse, specialundervisningskoordinatorer, specialundervisningslærere og klasselærere gennemført af forskere og stipendiater**
- **Strukturerede interview med klasselærere gennemført af observatørene**
- **Struktureret elevspørgeskema gennemført af stipendiater**
- **Spørgeskemaer udfyldt af klasselærere**
- **Strukturerede elev-, klasse- og skolevurderinger udfyldt af observatørene**
- **Strukturerede klasserumsobservationer, "Student Membership Snapshot" gennemført af observatører og stipendiater**
- **Semistrukturerede vurderinger af arbejdet gjort i forhold til eleven med særlige behov udfyldt af observatøren**
- **Kvalitative fokusgruppeinterview med observatørene og stipendiaterne**
- **Achenbach multiinformant assessmentsystem udfyldt af elever og lærere**
- **CORS (Children Outcome Rating Scale) udfyldt af elever**

### VALG AF OUTCOME VARIABLE

Der findes ikke standardiserede outcome mål, der forholder sig til noget så komplekst som folkeskolelovens formålsparagraf samt formålene med specialpædagogisk bistand. I indeværende undersøgelse har det derfor været vigtigt at finde en pragmatisk løsning, hvor man ved en helhedsorienteret vurdering opnår en bedømmelse af, om formålet med indsatsen er lykkedes. Dette er slet og ret sket ved, at klasselæreren ved skoleårets afslutning er blevet bedt om at vurdere indsatsens effekt på en 5-trinsskala. Klasselæreren er den person, der kender eleven bedst, og klasselæreren har selv som en del af lærerteamet haft ansvaret for at skabe sammenhæng i det, der sker i almenundervisningen, og det der foregår i den specialpædagogiske indsats. Ud over dette er der anvendt to andre outcome variable, så der i alt er gjort tre forskellige forsøg på at måle effekten af indsatserne med tre variable – en bred og helhedsorienteret og to snævrere mål.

- **Klasselærerens holistiske vurdering af effekt**
- **Klasselærerens vurdering af faglig effekt via Achenbach brugt som kontrol**
- **Elevens holistiske vurdering af effekt via CORS brugt som supplement**

Klasselærerens vurdering kan karakteriseres ved at være en kvalitativ vurdering, der oversættes til en kvantitativ skala. Det samme gælder imidlertid Achenbach og CORS, hvor der også er tale om en kvalitativ vurdering, der omsættes til et kvantificeret mål. De tre outcome variable har naturligvis overlap, men der er også meget væsentlige forskelle. Først og fremmest gælder, at de faglige effekter afdækket af Achenbach er langt snævrere end lærernes overordnede vurdering, der går på såvel faglige som sociale og personlige forhold, der alle tre indgår i folkeskolelovens formålsparagraf. Det må i den forbindelse erindres, at en del af eleverne primært har socioemotionelle vanskeligheder, og nogle har ADHD, hvor adfærdstræk spiller en betydelig og måske større rolle end de rent faglige forhold. Hvad CORS angår, er der som ved klasselæreren tale om en holistisk vurdering, men det kan ikke forventes, at elever i samme grad som lærerne har

blik for folkeskolelovens formålsparagraf og for formålet med den specialpædagogiske indsats, de har været udsat for. Sidst men ikke mindst gælder, at klasselærernes vurdering dækker et helt skoleår, mens såvel Achenbach som CORS er foretaget omkring midten af efteråret 2007 og igen i løbet af foråret 2008 og dækker dermed kun godt ½ år. Med andre ord kan de to sidstnævnte outcomevariable antages at give mere beskedne resultater end klasselærernes vurdering for hele året, og de er derfor kun anvendt som en kontrolmåling eller et supplement, der forholder sig til lærerens holistiske effektvurdering.

**Tabel 1. Klasselærers holistiske vurdering af effekt**

| Grad af effekt | Procent |
|----------------|---------|
| Meget positiv  | 32 %    |
| Positiv        | 52 %    |
| Neutral        | 13 %    |
| Negativ        | 2 %     |
| Meget negativ  | 0%      |

Det fremgår af tabel 1, at for ca. en tredjedel af eleverne vurderer lærerne, at der er en meget positiv effekt. For ca. halvdelen er effekten vurderet til at være positiv. For 13 % vurderes effekten at være neutral og for kun et par procent angives den at være negativ. Der er ingen elever, der kategoriseres som have haft en meget negativ effekt af indsatserne.

Der er dermed tale om en markant positiv overvægt, og der peges dermed på, at indsatsen helt overvejende er "givet godt ud", da den vurderes til at have en positiv effekt. Det skal og kan naturligvis overvejes, om klasselærerne har været for positive, da de jo kender undersøgelsens formål – samtidig med at den specialpædagogiske indsats støtter enhedsskolens differentieringsbehov. Klasselærerne har derfor en naturlig interesse i, at den specialpædagogiske indsats for lov til at fortsætte, hvad der kan give en skævhed i positiv retning. Omvendt er det mest sandsynligt, at lærerne fokuserer på den enkelte elev, som spørgsmålene vedrører, og dermed også på de formål og intentioner, der har været med undervisningen i forhold til denne elev med særlige behov.

Undersøgelsen har med hensyn til Achenbach haft det problem, at det har været svært at få lærere til at acceptere og gennemføre udfyldelsen af skemaerne, da de som et børnepsykiatrisk screeningsmateriale indeholder spørgsmål, der virker fremmede eller ligefrem stødende, og på nogle af skolerne er det ikke lykkedes – på trods af materielle incitamenter – at opnå medvirken fra alle lærere, og der er derfor en del missing values. I tabel 2 er resultaterne fra de tre outcomevariable sammenstillet, idet de to bedste og de to dårligste kategorier er slået sammen for at opnå en tetrins skala og dermed større sammenlignelighed.

**Tabel 2. Oversigt over resultater af effektmål i projektet**

|                | Lærervurdering | Achenbach | COR    |
|----------------|----------------|-----------|--------|
| Positiv effekt | 84,4 %         | 42,7 %    | 58,9 % |
| Ingen effekt   | 13,3 %         | 33,3 %    | 12,5 % |
| Negativ effekt | 2,2%           | 24,0%     | 28,6%  |

Det fremgår, at lærernes vurdering er den udpræget mest positive, idet 84,4 % mener virkningen er meget positiv eller positiv, mens 13,3 % siger, der ikke er nogen effekt, og kun 2,2 % mener, at der er en negativ effekt. En principal component analyse har vist, at der kun er tale om én betydelig fælles komponent, som forklarer 45 % af variansen, og tilbage står 55 %, som har størst bidrag fra lærervurderingen. Den sandsynlige forklaring er, at lærerne i deres vurdering netop indtager det helhedsprægede. De snævert faglige spørgsmål, der indgår i Achenbach, har ved regressionsanalyser med baggrundsvariable vist sig at have en tæt sammenhæng med elevens indlæringssevne. Lærerne er ligeledes i deres vurdering mere holistiske end elevernes svar i CORS, der i regressionsanalyse først og fremmest viser sig at have sammenhæng med, hvordan eleven trives med lærernes praksis om eleven. Med andre ord definerer lærerne et positivt resultat under hensyntagen til væsentligt bredere forhold, end de andre to vurderingstyper. For dette taler også den væsentligt mere positive indstilling, der ligger i lærernes udsagn.

## KONKLUSION

Kan man effektmåle en specialpædagogisk indsats? Svaret er positivt eller negativt alt efter, om man går ind brugen af blødere tilgange, hvor kvalitative vurderinger anvendes kvantitativt, eller om man forkaster alt andet end hårde RCT-tilgange. Der er ingen tvivl om, at anvendelsen af RCT kan og bør anvendes i skole regi, men det kræver, at der er tale om helt specifikke og veldefinerede indsatsformer i relation til helt specifikke og veldefinerede outcome mål. Man vil kunne se, om træning i klasseledelse har effekt på graden af uro og elevers koncentrationsevne. Man vil også kunne se, om sund skolemad har indflydelse på elevernes progression i sætningslæsetests. Det specialpædagogiske felt er imidlertid alt for komplekst til at kunne anvendes i et RCT-koncept, og blødere quasiexperimentelle metoder er derfor en nødvendighed.

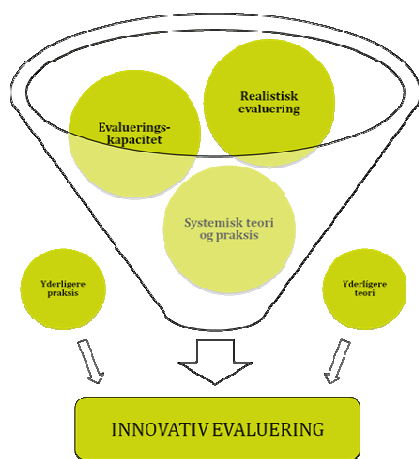
## REFERENCE

> Egelund, N. & Tetler, S. (red.)(2009): *Effekter af specialundervisningen*. København: Danmarks Pædagogiske Universitetsforlag.

## INNOVATIV EVALUERING - EVALUERING DER SKABER INDFLYDELSE OG ANVENDELSE

*Evaluering skal drive innovation af praksis og give mening for praktikerne i deres organisatoriske opgaver. Det skal skabe udvikling og ikke bare måle praksis bagudrettet eller vurdere på ikke relevante mål. Men evaluering er kompleks, og det er ikke ligetil at lave kvalificeret evaluering, der får reel indflydelse og skaber udvikling. Denne artikel omhandler en ny tænkning og tilgang, nemlig **Innovativ evaluering**, der sætter spot på, hvordan evaluering kan blive meningsfuld for organisationerne gennem anvendelse af resultaterne som driver for innovation af den professionelle praksis parallelt med, at eksterne parter får deres informationsbehov tilfredsstillet.*

Innovativ evaluering er blevet til som svar på de rammebetingelser der i disse år hersker inden for evalueringsfeltet i Danmark **1**. Sagen er, at mange af de krav der stilles om evaluering og evidensbaseret beslutninger og udvikling ofte ikke bygger på en efterspørgsel på viden fra institutionerne selv, men kommer fra deres bevillingshavende myndighed eller er en faktor for institutionens konkurrenceevne. Innovativ evaluering har fokus på at skabe evalueringer der har indflydelse på udviklingen af praksis og finder afsæt i den systemiske organisationsteori, Realistisk evaluering samt evalueringskapacitet.



**Figur 1: Grundsten i Innovativ evaluering**

Den systemiske organisationsteori bidrager med den teoretiske ramme og den grundlæggende epistemologi (se eksempelvis Hornstrup et. al, 2005)**2**. Realistisk evaluering (Pawson og Tilley, 1997)**3** kan betragtes

som en grundlæggende evalueringsmetodologi, der samtidig bidrager med interessante forklaringer på hvad det er relevant og muligt at vurdere i en evaluering, herunder kontekstens betydning. Evalueringskapacitet er afgørende for opnåelse og succes af at arbejde med Innovativ evaluering, fordi uden evalueringskapacitet i organisationen er der ikke mulighed for at opnå reel anvendelse og dermed innovation på baggrund af evalueringen (se eksempelvis Compton og Stockdill, 2002)**4**.

### SYSTEMISK BIDRAG TIL INNOVATIV EVALUERING

Det systemiske og socialkonstruktionistiske felt er oprindeligt opstået som en tilgang til familierapi (Cecchin, 1992)**5**, og har siden udviklet sig til at være en omfattende og anerkendt organisationsudvikling og -ledelsestilgang (se eksempelvis Haslebo, 2004 og Hornstrup et. al, 2005)**6**. I Innovativ evaluering betragter vi den systemiske teori og praksis i en evalueringskontekst og udvikler dermed feltet yderligere. Inspirationen fra det systemiske er omfattende, og i denne sammenhæng vil kun et mindre hjørne blive berørt (for yderligere udfoldelse se Dinesen og de Wit, kommende)**7**.

Noget af det helt centrale i Innovativ evaluering er, at evalueringer er *relationelt skabte i samarbejde mellem evaluators og praktikerne*. Denne pointe hentes fra socialkonstruktionismen med inspiration fra Kenneth Gergen. Gergen gør op med forestillingen om, at vi som mennesker har en 'selvet kerne', og præsenterer igennem socialkonstruktionismen forestillingen om, at



alting opstår i relationerne mellem mennesker (Gergen, 2004)8.

Det relationelle perspektiv har betydning for, hvordan man anskuer sin position og sin påvirkning som evaluator. Den positivistiske forestilling om, at man som evaluator kan træde ind i et felt uden at påvirke og derigennem lave evaluering "urørt af menneskehænder", betragtes i Innovativ evaluering som en umulighed, da vi altid vil påvirke. *Det, at man som evaluator påvirker feltet, betragtes i Innovativ evaluering ikke som en fejlkilde, man skal forsøge at minimere.* I stedet arbejdes der mod mest mulig hensigtsmæssig påvirkning i forhold til det, der er formålet med evalueringen.

I praksis betyder det, at evalueringsindsatsen bliver en integreret del af den evaluerede indsats, og har dermed også har til hensigt at udvikle indsatsen både under og efter forløbet. Når evalueringen er i indsatsen, har det også betydning for, hvornår evalueringen får sin indflydelse og anvendelse. Typisk ses det, at evalueringens konklusioner har til formål at bidrage til beslutningsprocesser efter endt evaluering (se eksempelvis Vedungs definition på programevaluering i Vedung, 2005)9. I Innovativ evaluering opstår *anvendelsen og indflydelsen af evaluering undervejs i evaluering*, igennem den måde evalueringen bliver designet og gennemført. Det ses fx ved brugen af de indsamlingsmetoder der anvendes, de evalueringsaktiviteter der gennemføres samt den måde evalueringens resultater løbende bliver præsenteret og diskuteret med praktikerne (Dinesen og de Wit, kommende).

Den anden centrale pointe fra det systemiske i denne sammenhæng handler om *nysgerrighed*. Nysgerrighed beskrives af Cecchin, som det systemiske alternativ til at være neutral (Cecchin, 1987)10. Jf. beskrivelserne af, at vi altid vil påvirke de relationer vi indgår i, er det i systemisk perspektiv ikke muligt at være neutral, forstået som objektiv (Hornstrup et. al., 2005 samt se beskrivelserne af autopoiese i Maturana og Varela, 1987)11 Det er være nysgerrighed handler om at være *bevidst om at åbne så mange perspektiver som muligt* (Cecchin, 1987).

Nysgerrighed har stor betydning for Innovativ evaluering, da det er en afgørende forudsætning for at kunne skabe evaluering med anvendelsespotentiale, at evaluator er nysgerrig over for den kontekst og praksis, som evalueringen gennemføres i. Humberto Maturana beskriver, at vi må betragte *multiverset*, altså de mange forskellige perspektiver, frem for universet, hvor der kun findes ét perspektiv og én sandhed (Maturana og Varela, 1987). I Innovativ evaluering er vi optagede af netop *de mange perspektiver*, og tilstræber ikke at lave endegyldige konklusioner. Som resultat af denne tilgang, er man i Innovativ evaluering ikke optaget af *kvantitet i besvarelserne*. Et perspektiv vægter altså ikke tungere udelukkende fordi mange har pointeret det, ud fra tanken om, at der kan være *helt afgørende elementer af indsatsen, som kun en enkelt respondent kan se*.

Nysgerrigheden har også afgørende betydning i dataindsamlingen og særligt den måde man gennemfører kvalitative interviews. Ved at stille spørgsmål der åbner for refleksion hos respondenterne, kan man både få nuancerede beskrivelser, samtidig med at interviewene understøtter respondenternes forståelse for deres praksis (se Tomm 1992, 1-3, Hornstrup, Tomm og Johansen, 2009 samt Dinesen og de Wit, kommende)12,13. Dermed skabes der, som beskrevet tidligere, indflydelse på baggrund af evalueringen *undervejs i evaluering* og ikke kun efterfølgende.

## REALISTISK EVALUERING

Den anden primære inspirationskilde i Innovativ evaluering er *Realistisk evaluering* (RE). Det centrale ud-sagn fra RE, der beskriver formålet med evalueringer lavet i denne tilgang, er: *hvad virker, for hvem og under hvilke omstændigheder?* Det er med andre ord lige så vigtigt at kende konteksten for en virkning, og hvilke aktører der var involveret (hvem og hvilke omstændigheder), som kendskabet til hvad der rent faktisk virkede.

I forhold til den systemiske teori tager RE udgangspunkt i, at aktørerne i en evaluering, inkl. evaluatoren og opdragsgiveren, har et billede af virkeligheden, som i deres øjne svarer til virkeligheden. Afsættet i tilgangen er, at de unikke verdensbilleder kan indfanges og samles, sættes sammen og genkendes som

mønstre i forbindelse med en evaluering, hvilket kan være afsæt for en vurdering af og forståelse af, hvad der er sket i indsatsen, og hvorfor det er sket. 'Realistisk' i denne sammenhæng henviser da til 'realisme' forstået som anerkendelse af omverdenskompleksitet repræsenteret ved aktørernes forskellige perspektiver.

Dvs. at RE anerkender at mening, resultater og effekter opstår i relationerne mellem aktører. Derfor er en anden central pointe i RE, at *resultater ikke opstår som effekter af et program, men igennem mennesker med unikke verdensopfattelser, der udfører handlinger, som, når de lægges sammen, bliver til resultater af indsatser.*

RE er en evalueringstype, der kan anvendes til at forbedre implementeringen af programmets mål og formuleringen af områdets indsatser gennem viden, forklaringer på resultater og kontinuerlig udvikling af indsatsen. Læring og videndeling er dybt forankret i denne evalueringstype. Afhængig af graden af inddragelse kan evalueringstypen også være deltagende. RE flytter sig altså fra fokus på 'virker programmet?' til 'hvad er det ved programmet, der får det til at virke?', for dermed at kunne understøtte den videre udvikling af det (Pawson og Tilley, 1997).

I praksis konceptualiseres RE ofte gennem en programteori, der har sin grundopbygning i en CMO-struktur (Context, Mechanism, Outcome). Den centrale pointe i den praktiske anvendelse af RE er at forstå konteksten for evalueringen og på baggrund af de lave evalueringer, der favner de effekter der opstår i konteksterne, frem for at lave globaliserede konklusioner af indsatser og programmets effekter (for en yderligere udfoldelse af RE i praksis se de Wit og Togeby 2005 samt Dinesen og de Wit, kommende)<sup>14</sup>.

## EVALUERINGSKAPACITET

Udviklingen af evalueringskapacitet kan abstrakt set opfattes som en måde at *professionalisere evalueringsfeltet* og den enkelte evalueringspraktiker på. I Innovativ evaluering er evalueringskapaciteten afgørende, fordi den understøtter at evaluering bliver en del af den daglige udvikling af praksis. Det handler om

at kunne se både de enkelte træer og hele skoven forstået på den måde, at den enkelte evaluering er ét træ. Dvs. at en person eller organisation med en høj evalueringskapacitet både kan lære ved at reflektere over den nuværende evaluering (træet) og samtidig reflektere over evalueringen som en enkelt hændelse i sin samlede evalueringspraksis (skoven). Det er altså med afsæt i evalueringskapaciteten det bliver muligt at inddrage det systemiske og RE (samt evalueringskapaciteten) som én fælles tilgang til evaluering, hvor man trækker på de elementer, der giver mening og kan understøtte formålet med evalueringen.

Opbygningen af evalueringskapacitet understøtter en form for ledelse, der beror på involvering af medarbejderne og internt commitment til organisationens opgaver. Internt commitment er, når medarbejdere er optaget af at udvikle deres praksis, fordi de er interesserede i at gøre deres job så godt som muligt. Modsat er eksternt commitment en holdning, hvor man udfører sit job i overensstemmelse med kriterier, som andre har bestemt (Krag & de Wit, 2003 efter Argyris)<sup>15</sup>. Den eksterne commitment følges ofte ad med eksplícite incitamentsstrukturer, som gør det vigtigere for praktikerne at opfylde målet end at reflektere over, hvorvidt målet overhovedet er meningsfuldt (Nørreklit & de Wit, 2000)<sup>16</sup>.

## Afsluttende kommentarer

Denne artikel har tilbudt et overordnet blik på de teoretiske grundsten i Innovativ evaluering, og har i det små skabt koblinger til praksis. Der er præsenteret en række tankeværktøjer, hvor de tilknyttede handleværktøjer hentes dels i de tre præsenterede perspektiver, og udvikles dels som en sammensmeltning af dem. Definitionen på Innovativ evaluering er:

*"Innovativ evaluering er en reflektiv relationel proces, hvor kontekstbestemt viden om fortiden og nutiden anvendes til fremadrettede handlinger, der understøtter aktørerne i at skabe nye muligheder og forbedret praksis parallelt med, at eksterne krav om dokumentation opfyldes."*

Innovativ evaluering handler om at være reflektiv om sin praksis, og tage kompetente og kvalificerede valg for evalueringer, der tager afsæt i den aktuelle kontekst. Udgangspunktet skal tages i evalueringskompe-

tencer frem for evalueringsfærdigheder. For at kunne skabe evalueringer, der har udviklingspotentiale skal både praktiker og evaluator forstå praksis, fordi det skaber mulighed for at vurdere, hvilke evalueringer og formål der er relevante. Samtidig giver det mulighed for at udfordre de myndighedsniveau vi indledningsvist beskrev, og som ofte stiller krav til evalueringer. I Innovativ evaluering får evaluator en rolle som *mediator* mellem praksis og myndighedsniveauet, og betragter og beskriver praksis, på en måde som myndighedsniveauet kan forstå og koble sig på. Første skridt må være at kende sin praksis indgående nok til at kunne vurdere, hvordan evalueringerne skal understøtte og udvikle, og hvordan den praksis de skal have indflydelse på ser ud.

under etablering, Bukh, Christensen og Mouritsen (red.), side 173-200  
 16. Nørreklit, H & de Wit, C K (2000): 'Økonomistyringsretorik – fra radiatorstyring til sokratisk styring i *Tillid eller teknologi – Skandinaviske perspektiver på Økonomistyring*, Larsen (red.), København

## NOTER

1. Se eksempelvis *Kvalitetsreformen* på <http://www.kvalitetsreform.dk/> og <http://www.fm.dk/Temaer/Kvalitetsreform.aspx>, *Ressourceregnskabet i uddannelsessystemet* på <http://www.uvm.dk/For%20institutioner/Tilsyn%20og%20regnskab/Ressourceregnskab.aspx> og *Den Danske Kvalitetsmodel i sundhedsvæsnet* på <http://www.ikas.dk/Den-Danske-Kvalitetsmodel.aspx>
2. Hornstrup, Carsten et al. (2005): *Systemisk Ledelse – Den refleksive praktiker*, Dansk Psykologisk forlag, Gylling
3. Pawson, R & Tilley, N (1997): *Realistic Evaluation*, SAGE Publications, London
4. Compton, D W; Baizerman, M; Stockdill, S.H (eds.) (2002): *The Art, Craft and Science of Evaluation Capacity Building*, New Directions for Evaluation, Nr. 93, forår 2002, Jossey-Bass
5. Cecchin, Gianfranco (1992): *The idea of irreverence*, udgivet i: Cecchin et. al.: *Irreverence* (1992), Karnac Books.
6. Haslebo, Gitte (2004): *Relationer i organisationer – En verden til forskel*, Dansk psykologisk forlag, Gylling
7. Dinesen, Malene Skov og de Wit, Camilla Kølsen (kommende): *Innovativ evaluering*, Dansk Psykologisk forlag, udkommer forår 2010. Tilmeld dig NovaBergs nyhedsbrev på [www.novaberg.dk](http://www.novaberg.dk), hvis du ønsker at blive informeret, når bogen udkommer.
8. Gergen, Kenneth (2004): *Virkelighed og relationer*, Dansk Psykologisk forlag, Gylling
9. Vedung, Evert (2005): *Public Policy and Program Evaluation*, Transaction Publishers, London
10. Cecchin, Gianfranco (1987): *Gensyn med hypotesedannelse, cirkularitet og neutralitet: En invitation til nysgerrighed*, Family Proces; 26:405-413.
11. Maturana, Humberto og Varela, Francisco (1987): *Kundskabens træ – den menneskelig erkendelses biologiske rødder*, ASK
12. Tomm, Karl (1988): *Interventive Interviewing*, Fam. Proc., Vol 27, March, 1988; Tomm, Karl (1992): *Planlægning som 4. Retningslinje for terapeuten*, Forum, 2/92; Tomm, Karl (1992): *Refleksiv udspørgen som middel til at fremme selvhelbredelse*, Forum, 3/92
13. Hornstrup, Carsten; Tomm, Karl og Johansen, Thomas (2009): *Spørgsmål, der gør en forskel*, udgivet på [www.macmannberg.dk](http://www.macmannberg.dk)
14. De Wit, CK og Togeby, M (2005): *Realistisk evaluering i en dansk sammenhæng – hvad virker i gymnasieskolen?* Dansk EvalueringsSelskabs nyhedsblad, forår 2005.
15. Krag L. & de Wit, C. K (2003): *Ledelse af videnmedarbejdere - et spørgsmål om commitment?*, i *Videnledelse – et praksisfelt*

## CONTRIBUTION ANALYSIS – ET ALTERNATIV TIL KLASSISKE EFFEKTDESIGNS?

*“Despite the measurement difficulty, attribution is a problem that cannot be ignored when trying to assess the performance of government programs. Without an answer to this question, little can be said about the worth of the program, nor can advice be provided about future directions.” (Mayne, 1999)*

Formålet med de fleste evalueringer er at vurdere, hvorvidt en given indsats (et projekt, et program eller en politik) har haft den tilsigtede effekt. Dette giver anledning til to centrale spørgsmål:

- **Hvorvidt de tilsigtede effekter er realiseret**
- **Og hvorvidt de dokumenterede effekter kan tilskrives indsatsen**

I en klassisk tilgang til effektmåling vil disse spørgsmål blive adresseret gennem et eksperimentelt eller quasi-eksperimentelt design.

I virkelighedens verden er det dog sådan, at det i rigtig mange tilfælde ikke er muligt at anvende eksperimentelle designs – særligt inden for velfærdsområdet. Dette kan skyldes indsatsens form, timing eller målgruppe eller at det ikke er muligt at etablere en kontrafaktisk situation, fordi indsatsen er for kompliceret (bestående af mange komponenter) eller kompleks (emergent). Og det kan skyldes etiske eller økonomiske hensyn (se bl.a. Nielsen, 2007).

Dette ændrer dog ikke ved, at det grundlæggende spørgsmål for evalueringen stadig er, hvorvidt indsatsen har haft den tilsigtede effekt. Det vanskelige er imidlertid ikke at dokumentere, hvorvidt de tilsigtede effekter er realiseret, men derimod at sandsynliggøre, at de dokumenterede effekter kan tilskrives indsatsen. I disse tilfælde er der behov for en anden tilgang end den kontrafaktiske til at sandsynliggøre sammenhængen.

Contribution Analysis kan ses som et alternativ til kontrafaktiske effektmålinger, fordi modellen kan sandsynliggøre, at en effekt hidrører fra en given indsats gennem sit fokus på *både* de formodede sammenhænge mellem indsats, umiddelbare resultater, lang-

sigtede virkninger og kontekst og andre forklaringsmodeller. Dette er modellens klare styrke.

Contribution Analysis har i den form, som Mayne beskriver i sine artikler (Mayne, 1999 og 2008) – og som han præsenterede på et oplæg på Dansk Evalueringsselskabs årskonference i 2008 – dog også nogle indbyggede metodiske og operationelle svagheder, som vil blive adresseret i denne artikel. Før vi når til styrkerne og svaghederne ved modellen, er der dog lige behov for en kort skitse af Maynes tilgang til Contribution Analysis.

### CONTRIBUTION ANALYSIS SOM EVALUERINGSSTILGANG

Mayne definerer Contribution Analysis som *“[a] specific analysis undertaken to provide information on the contribution of a program to the outcomes it is trying to influence.”* (Mayne, 1999).

Udgangspunktet for en Contribution Analysis er en forandringsteori (i), som udbygges med eksplicite antagelser om de centrale trusler mod forandringsteoriens mekanisme, (ii) andre medvirkende faktorer samt (iii) de primære rivaliserede forklaringsmodeller<sup>2</sup>. Det bærende argument i Contribution Analysis er, at hvis forandringsteorien er plausibel og implementeret med fidelitet, og alternative forklaringer er belyst med henblik på udelukkelse eller medregning, så kan det sandsynliggøres, at programmet har bidraget til realiseringen af de dokumenterede effekter. Contribution Analysis bygger således ikke på en klassisk positivistisk kausalitetsforståelse, men snarere på en hermeneutisk forståelse, hvor fokus er rettet mod sandsynliggørelse af sammenhænge frem for verifikation og falsifikation af sammenhænge.

Maynes tilgang til at gennemføre en Contribution Analysis består af seks trin, som er følgende:

### Boks 1: Maynes model for Contribution Analysis 3

#### Trin 1: Afklaring af virkningsspørgsmål

Afklaring af de centrale virkningsspørgsmål, som skal besvares og hvilken udsigelseskraft, der er efterspurgt.

#### Trin 2: Udvikling af forandringsteori

Udfold forandringsteorien for programmet, inkl. trusler mod forandringsteorien, andre medvirkende faktorer samt rivaliserende forandringsteorier (dvs. en forankret forandringsteori)

#### Trin 3: Indsamling af eksisterende dokumentation for forandringsteorien

Foretag en vurdering af styrker og svagheder for hver af forandringsteoriens antagelser og inddrag eksisterende dokumentation for forandringsteoriens antagelser og for andre medvirkende faktorer.

#### Trin 4: Kritisk vurdering af virkningsmodellen

Vurder de samlede styrker og svagheder ved forandringsteori og identificer de væsentligste trusler mod forandringsteorien.

#### Trin 5: Indsamling af supplerende dokumentation

Identificer, hvad der er brug for af ny dokumentation for at styrke forandringsteoriens troværdighed.

#### Trin 6: Revurdering af virkningsmodellen

Revurder styrker og svagheder ved forandringsteori og udled svar i forhold til de centrale virkningsspørgsmål

Der er en række lighedspunkter mellem Contribution Analysis som præsenteret af Mayne (1999, 2008) og forskellige former for teoribaseret evaluering (Chen, 1990) og realistisk evaluering (Pawson, 1997a).

Evalueringsmetodisk bygger de forskellige former for teoribåren evaluering på en eksplicit teoretisk model som fx en programteori, som testes med henblik på at afgøre programteoriens plausibilitet som forklaringsmodel. Mange teoribårne evalueringer bygger på en successiv opfattelse af kausalitet, hvor den teoretiske model har fokus på antagelser om sammenhænge mellem indsats og effekter.

Realistisk evaluering har et mere eksplicit videnskabs-teoretisk afsæt i den kritiske realisme og dermed en generativ opfattelse af kausalitet (Pawson, 1997b). Det betyder, at realistisk evaluering ikke er baseret på en programteori, men i højere grad har fokus på de underliggende generative mekanismer. Spørgsmålet er således ikke, om effekten kan tilskrives indsatsen, men hvilke faktorer, der forårsager, at effekten kan tilskrives indsatsen.

Contribution Analysis bygger ligesom teoribåren evaluering på en eksplicit programteori, men har derudover også fokus på de generative mekanismer bag antagelserne, som det er tilfældet i en realistisk evaluering.

Der er således stærke lighedspunkter mellem teoribåren evaluering, realistisk evaluering og Contribution Analysis, men der er også væsentlige forskelle mellem de forskellige tilgange. Contribution Analysis, at det er en tilgang, som adresserer nogle af de metodiske og operationelle kritikpunkter, som bliver rejst mod teoribåren evaluering og realistisk evaluering<sup>4</sup>.

### EN STYRKET DOKUMENTATION AF ANTAGELSERNE I FORANDRINGSTEORIEN...

Den første styrke ved Contribution analysis er, at modellen har en tydelig analytisk brug af forandringsteorien. Således er et af de kritikpunkter, som rejses mod teoribåren evaluering er, at der i den praktiske anvendelse sjældent foretages en systematisk dokumentation af de enkelte antagelser i den teoretiske model. I kritikken bliver det påpeget, at den teoretiske model i mange evalueringer primært anvendes som ramme for at identificere indikatorer eller som redskab til forståelse af den evaluerede indsats (Davidson, 2000). Og dermed bliver det vanskeligt at konkludere, om indsatsen har virket, blot fordi teorien ikke er blevet udfordret af data.

Contribution analysis er her en mere operationel tilgang, hvor forandringsteorien fungerer som gennemgående analysestrategi. Når man arbejder med en Contribution Analysis ud fra Maynes model indsamles systematisk eksisterende og ny dokumentation for de enkelte antagelser i forandringsteorien, dvs. antagel-

serne i (i) selve forandringsteorien, (ii) de centrale trusler mod forandringsteoriens mekanismer, (iii) andre medvirkende faktorer og (iv) de rivaliserede forklaringsmodeller. Contribution Analysis tegner hermed linjerne til en mere systematisk og transparent tilgang til, hvordan brugen af en forandringsteori kan bidrage til konklusioner om sammenhængen mellem indsats og effekter – og om sondringer mellem teorifejl og implementeringsfejl.

### OG ET ØGET FOKUS PÅ RIVALISERENDE FORKLARINGSMODELLEN

Den anden styrke ved Contribution Analysis er, at modellen har et øget fokus på rivaliserende forklaringsmodeller. Et af de kritikpunkter, som rejses mod teoribårne evalueringer er, at de primært retter deres fokus mod programteorien og ikke har opmærksomhed på andre forklaringsmodeller (Dahler-Larsen, 2001; Weiss, 1997).

I Maynes tilgang til Contribution Analysis er der et særskilt fokus på at afdække og analysere alternative forklaringer – eller de rivaliserende programteorier – i forhold til de observerede resultater og virkninger. Hvis modellens sandsynlighedslogik skal kunne anvendes er det således ikke tilstrækkeligt at kunne sandsynliggøre, at effekten hidrører fra indsatsen, men derimod nødvendigt også at kunne udelukke, at effekten hidrører fra andre indsatser.

Det er dog også min erfaring fra brugen af modellen, at den fortsat kan styrkes på det operationelle plan i forhold til både en styrket dokumentation af antagelserne og et øget fokus på rivaliserende forandringsteorier. I Maynes model for Contribution Analysis er det således uklart, hvilke kilder der ligger til grund for udfoldelsen af (i) trusler mod forandringsteorien, (ii) andre medvirkende faktorer og (iii) de rivaliserede forklaringsmodeller. Dermed opstår der en vis grad af vilkårlighed i den forandringsteori, som udgør selve grundlaget for gennemførelsen af en Contribution Analysis, og denne vilkårlighed kan underminere udsagnskraften af analysen.

### MEN OGSÅ EN HØJ GRAD AF KONTEKSTAFHÆNGIGHED

Et væsentligt spørgsmål i forbindelse med evalueringer er, hvordan evalueringens konklusioner kan anvendes efterfølgende – og dermed hvilken vidensproduktion evalueringen medvirker til. Et af de forhold, der har betydning for anvendelsen er generaliserbarheden af resultaterne – eller evalueringens eksterne validitet. Det vil sige spørgsmål om, hvorvidt evalueringens konklusioner har bredere gyldighed og dermed kan udbredes til andre sammenhænge.

Maynes model har en høj grad af kontekstafhængighed i besparelsen af virkningsspørgsmålet. Dette skyldes, at virkningsspørgsmålet i Contribution Analysis – og i teoribårne evaluering – omhandler generativ opfattelse af kausalitet (contribution) og dermed samspillet mellem indsats, effekt og kontekst.

Den høje grad af kontekstafhængighed behøver dog ikke i sig selv at være et problem, så længe kontekstafhængigheden adresseres med overvejelser om generaliserbarheden af konklusionerne. I realistisk evaluering tilvejebringes således også kontekstafhængig viden om de generative mekanismer bag en given indsats, men forskellen er her, at der ikke er fokus på generaliserbarheden af den enkelte forklaringsmodel, men derimod på en akkumulering af viden om, hvordan forskellige generative mekanismer virker i forskellige sammenhænge – og dermed i højere grad en ambition om generalisering af resultaterne som led i en realistisk syntese (Pawson, 2006).

I Maynes artikler er der dog ingen refleksioner over generaliserbarheden af de konklusioner, som kan drages på baggrund af den gennemførte Contribution Analysis, hvilket er en af modellens største metodiske svagheder. Dette betyder, at Maynes model fortsat står relativt svagt i forhold til mulighederne for generalisering og dermed også står et metodisk spændingsfelt mellem ambitionerne i det klassiske effektdesign om kontekstfri viden med bred generaliserbarhed og ambitionerne i realistisk evaluering om akkumulering af kontekstafhængig viden i form af realistisk syntese.

Så er Contribution Analysis et alternativ til klassisk effektdesigns? Hvis de operationelle svagheder i forbindelse med en styrket dokumentation af antagelserne og et øget fokus på rivaliserende forandringsteorier adresseres, kan modellen give svar på det generative virkningssspørgsmål og dermed give et metodisk solidt modspil til det klassiske effektdesign i forhold til konkrete studier. I et bredere vidensperspektiv er det dog afgørende, at der i den videre udvikling af modellen også sættes fokus på, hvordan modellens høj kontekstafhængighed adresseres, således at den tilvejebragte viden også kan finde anvendelse uden for det enkelte studie.

## REFERENCER

- > Dahler-Larsen (2001): "From Programme Theory to Constructivism: On Tragic, Magic and Competing Programmes", i: *Evaluation*, 2001
- > Davidson, E. J. (2000): "Ascertaining Causality in Theory-Based Evaluation", i: *New Directions for Evaluation*, no. 87
- > Mayne, J. (1999): "Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly" (discussion paper)
- > Mayne, J. (2008): "Addressing Cause and Effect in Simple and Complex Settings through Contribution Analysis", i: *Evaluating the Complex*, R. Schwartz, K. Forss, and M. Marra (Eds.), Transaction Publishers (forthcoming).
- > Nielsen, C. et. al. (2007): "Effektmåling", SFI
- > Pawson, R. (2006): "Evidence-Based Policy. A Realist Perspective", Sage
- > Pawson, R. og Tilley, N. (1997a): "Realistic Evaluation", Sage.
- > Pawson, R. og Tilley, N. (1997b): "An Introduction to Scientific Realist Evaluation" i: Chelimsky, E. og Shadish, W. (ed): "Evaluation for the 21st Century. A Handbook", Sage
- > Weiss, C.H., (1997): "How Can Theory-Based Evaluation Make Greater Headway" i: *Evaluation Review*, vol. 21
- > Weiss, C.H. (2000): "Which Links in Which Theories Shall We Evaluate", i: *New Directions for Evaluation*, no. 87

## NOTER

1. I artiklen anvendes termen *forandringsteori* til at omfatte en samlet teori, som indeholder både en *programteori* med antagelser om de virkende mekanismer i indsatsen og en *implementeringsteori* med antagelser om sekvensen i implementeringen af programmet (Weiss, 2000; 2008)
2. Når der i det følgende anvendelse betegnelsen forandringsteori er det med reference til denne for form forankrede eller "grounded" forandringsteori.
3. Den præcise udformning af modellen varierer i forskellige artikler og præsentationer og indeholder mellem seks og ni trin, om end sekvensen er den samme på tværs af artiklerne. I denne artikel anvendes trinmodellen som præsenteret i Mayne, 2008
4. Rambøll Management Consulting har anvendt modellen i bl.a. evaluering af De utrolige År og evaluering af forældrepålæg (SEL §52) begge på vegne af Servicestyrelsen

## HVORDAN SKAL EN DES KONFERENCE VÆRE, FOR AT DU SYNES, AT DEN ER GOD?

Diagram fra demonstrations-fokusgruppe på udviklingsseminar i forbindelse med Dansk Evaluerings-selskabs årskonference 2009. Dette diagram er udarbejdet på et udviklingsseminar i DES-regi 2009: "Tag en metode med hjem - Fokusgruppeinterview som kvalitetsudvikling".

Denne fokusgruppemetode er en videreudvikling af affinitetsmodellen og fremtidsværkstedsmetoden, hvor deltagerne gennem strukturerede brainstormrunder inspirerer hinanden til at komme med svar, som bliver skrevet på sedler, der bliver hængt op på væggen, så alle kan se dem. Runderne fortsætter, indtil ingen af deltagerne længere kan komme på nye ideer. Metoden sikrer, at alle deltagere kommer til orde, og at alle synspunkter behandles ligeværdigt. Til sidst i processen markerer deltagerne ved hjælp af farvede klistermærker, hvordan de vurderer og prioriterer de fremkomne udsagn.

Metoden er særligt velegnet til at generere ideer til fremtidig kvalitetsudvikling i organisationer. Den kan også benyttes som en del af et forskningsprojekt, forud for udarbejdelse af informationsmateriale/pjecer, research i forbindelse med andre undersøgelsesmetoder, selvstændig evaluering, som opfølgning på andre undersøgelsesresultater samt i kombination med andre evalueringsmetoder.

I forbindelse med udviklingsseminaret blev gennemført et demonstrationsforløb. Formålet var formidling af metoden, hvilket bedst lader sig gøre ved at afprøve den i praksis. Da der var begrænset tid til rådighed, måtte de forskellige faser i forløbet forceres lidt, og der blev ikke tid til deltagernes prioritering. Med de forbehold, var der alligevel enighed om, at resultatet muligvis kunne være brugbart for DES-bestyrelsen. Deltagerne blev nemlig bedt om at besvare følgende spørgsmål:

### Hvordan skal en DES konference være, for at du synes, at den er god?

Målet var, at få deltagernes syn på, hvad der gør en DES konference god. Der skal gøres opmærksom på, at deltagerne således har givet deres bud på, hvordan

de mener, at den *ideelle* konference skal være. Der er altså ikke tale om tilfredshedsvurderinger, og det kan ikke læses ud af diagrammet, hvordan deltagerne har oplevet tidligere DES konferencer.

Den ideelle gruppestørrelse i en affinitets-fokusgruppe er 7-8 personer. Udviklingsseminaret havde ca. 30 deltagere. Da det var en demonstrations-fokusgruppe, valgte vi at lade fokusgruppen være på 8 personer, som blev udskiftet med andre deltagere et par gange undervejs i brainstormforløbet. De 11 deltagere, der ikke havde deltaget i brainstormforløbet, blev bedt om at vurdere de indkomne svar med farvede klistermærker.

### LÆSEVEJLEDNING TIL DIAGRAMMET PÅ MIDTERSIDERNE

I diagrammet viser plustegn (+), spørgsmålstegn (?) og minustegn (-) i parenteserne efter udsagnene, hvorvidt de 11 deltagere var enige i de enkelte udsagn. Et tal efterfulgt af et + angiver antallet af deltagere, der var enige. Et tal efterfulgt af et - viser antallet af deltagere, der var uenige. Et tal efterfulgt af et ? viser antallet af deltagere, der mente, at de hverken var enige eller uenige eller ikke kunne vurdere udsagnet (f.eks. fordi det omhandlede andre interesser eller behov end deres egne). Såfremt der ikke summeres op til 11, er det fordi enkelte deltagere er kommet til at springe markeringen af udsagnet over pga. den knappe tid.



## Hvordan skal en DES konference være, for at du synes, at den er god?

| Økonomi  | Deltagere   | Sted & fysiske rammer  | Forplejning & pauser   | Skriftligt konferencemateriale   | Fagligt indhold (plenum, workshops, udviklingsseminarer m.v.)  | Socialt indhold & networking  | Diverse  |
|--|---|--|--|--|--|---|--|
| Prisen er et vigtigt parameter, for at også små organisationer kan deltage [11+] | Godt med deltagere med forskellige baggrunde [10+, 11+]   | Snært af luksus [7+, 3?, 1 mangler]  | Frugt i pauseme [11+]  | DES-jærneskæben skal være overskuelig og nem at bruge (f.eks. til konferenceopsamlinger, og hvad den enkelte skal deltage i) [11+] | Plan skal udfoldes på det, man tror, at man ved [11+]  | Godt muligheder for networking [10+, 1 mangler]   | Herre fysisk aktivitet, f.eks. ud af huset (f.eks. job støj) [4+, 2?, 3+, 2 mangler] |
|  | Grupper/workshops med blandede fagområder [9+, 1?, 1 mangler]   | Central placering i Danmark (Fyn) [7+, 3?, 1+]                             | Godt mad [10+, 1?]   | Der skal være litteraturlister [11+]   | Undervisning i nye metoder (få inspiration) [11+]  | Til til erfaringsudveksling (pauser hvor det er muligt) [8+, 2?, 1 mangler]                           | Arrangeret fælles løbsetur [2+, 3?, 6+]  |
|  | Flange-deltagere (stor variation) [8+, 2?, 1 mangler]   | Alle skal have mulighed for at sove det, konferencen afholdes [7+, 1?, 4+] | Pauser efter max. 1½ times oplæsning/undervisning [8+, 2?, 1+] | Der skal være uddannelseskopier til deltagere [10+, 1 mangler]   | Metodeorienterede diskussioner [11+]   | Bordplan = networking (arrangeret/blandet) [7+, 2?, 2+]   |  |
|  | Deltagere skal sammentættes i grupper (f.eks. workshops) ud fra de problemområder, de arbejder med på egen arbejdsplads (område-specifikt, f.eks. social/psykiatri/teknik m.fl.) [2+, 2?, 7+] |  | Protokollen skal være på 1½ time [6+, 2?, 2+, 1 mangler]       | Materialet fra konferencen skal lægges på DES' hjemmeside bagefter [10+, 1 mangler]  | Kritisk stillingstagen til metoderne [11+]   | Organiseret netværksdannelse [7+, 1?, 3+]   |  |
| <b>Tidspunkt</b>   |   |  | Friske og saftige småkager til kaffen [5+, 2?, 4+]             | Der skal være et overblik over et dag-for-dag program med billeder osv. på (måske i diagramform) [9+, 1?, 1 mangler]               | Opleg (manchet) skal også beskrive, hvad oplegget giver deltagere med hjem [10+, 1?]                 | Uformelle, arrangerede sociale aktiviteter om aftenen [7+, 4?]  |  |
| Konferencen skal foregå på hverdage (ikke weekend) [10+, 1?]                     |   |  |  |  | Oplegsholdere skal (søge at) efterleve manchet/overskrift og indlembrik [9+, 2?]                     | Balkort, der kan (men ikke skal) anvendes [3+, 4?, 4+]  |  |
| Konferencen skal planlægges på ca. samme tidspunkt hvert år [9+, 2?]             |   |  |  |  | Temaorienteret konference giver godt program [9+, 2+]  | "Konferencebuddy"/konferencer (f.eks. 2 gange arrangeret konferencesnak med vedkommende) [2+, 5?, 4+] |  |
|  |   |  |  |  | Emner skal være anvendelsesorienterede [8+, 3?]  |   |  |
|  |   |  |  |  | Højest muligt at gå hjem med (f.eks. værktøjer) [8+, 3+]   |   |  |
|  |   |  |  |  | Deltagere skal inddrages aktivt [7+, 3?, 1+]   |   |  |
|  |   |  |  |  | Temaer i workshops og udviklingsseminarer skal målrettespecifikke faggrupper [2+, 2?, 3+, 2 mangler] |   |  |