

**E
V
A
L
U
E
R
I
N
G
S
N
Y
T** ©

Nr. 22
Februar 2009

Dansk
Evalueringsselskab

Evalueringsnyt ændrer form

Af Jesper Pedersen, redaktør, Dansk EvalueringsSelskab

Dette nummer er sandsynligvis det sidste i rækken af i alt 22 nyhedsbreve, der er udkommet 4-5 gange om året med nyheder og evalueringsartikler til selskabets medlemmer.

Bestyrelsen har besluttet at intensivere arbejdet med nyhedsbrevet. Målet er, at der skal udsendes løbende nyhedsbreve og et magasin eller tidsskrift i elektronisk form. På denne måde skal nyhedsbrevet i nuværende form løftes ved at intensivere nyhedsstrømmen og udskille artikeldelen i et selvstændigt tidsskriftmedie.

Evalueringsnyt fandt hurtigt sin form med nyheder og artikler inden for bestemte temaer. Nyhedsbrevet har henvendt sig til en bred medlemsskare, som alle har det tilfælles, at de har en travl hverdag, hvor de med Evalueringsnyts kortfattede artikler har haft mulighed for at følge med i hvad der sker på evalueringsfronten. Det har været en fornøjelse at møde læserne på konferencer og gå-hjem-møder, hvor jeg har fået mange positive tilbagemeldinger.

Det har været et stort arbejde, men heldigvis har folk været meget villige til at skrive artikler til Evalueringsnyt, også selv om bidragsyderne ikke honoreres eller krediteres for det i forskerkredse. Nogle temaer har flere gange måtte spredes ud over to numre, fordi mange har været velvillige til at levere en artikel. Jeg vil gerne benytte lejligheden til at takke alle disse mennesker, der gennem tiden har leveret artikler til Evalueringsnyt.

Find tidligere artikler her:

<http://www.danskevalueringsselskab.dk/Nyhedsbrev.asp>

En særlig tak skal sendes til Niels Matti Søndergaard, Dorte Kabell, Tanja Miller, Poul Skov Dahl, Bianca Albers, Hanne Kathrine Krogstrup, Anders Kirkedal Nielsen, som alle har været med til at forme og tilrettelægge forskellige udgaver af Evalueringsnyt.

Dette nummer bringer 3 artikler. De to første artikler er fra oplægsholdere fra selskabets seneste konference med temaet 'Evalueringsnyt som beslutningsgrundlag - teorier, metoder og sektorer.'

Hanne Dorthe Sørensen og Mads Biering-Sørensen SCKK skriver om *forskellene mellem akkreditering og selvevaluering*. Artiklen tager udgangspunkt i en diskussion af akkredite-

ring i regi af kvalitetsreformen, og om det er meningsfuldt, at kombinere akkreditering og kvalitetsudvikling i arbejdet med kvalitetsudvikling.

Morten Greve, René Buch Nielsen og Simon Østergaard Møller, Center for Kvalitetsudvikling, skriver om *måling af "det bløde" på det sociale område*. De mener at man på det sociale område modsat medicinske forsøg, befinder sig i den "bløde" ende af kontinuet. Denne erkendelse mener de kalder på kritisk refleksion: Er datagyldigheden i orden? Åbner data for reel læring? Med udgangspunkt i tre cases forsøger de at illustrere forskellige tilgange til at dokumentere resultaterne af den sociale indsats og de udfordringer, som er forbundet med det.

Virkningsevaluering, som er oversat fra den engelske model 'Realistic Evaluation', er en af de evalueringsformer som har fået stor opmærksomhed de seneste år. Evalueringsnyt nr. 7 havde i 2005 virkningsevaluering som tema med 4 artikler fra bl.a. ophavsmanden Nick Tilley. Der er dog stadig få eksempler på virkningsevaluering i Danmark og Realistic Evaluation i udlandet. Derfor er det kærkomment, at Niels Matti Søndergaard og Eva Pallesen fra Danmarks Evalueringsinstitut har leveret en artikel til dette nummer, hvor de beskriver deres erfaringer med virkningsevaluering.

Redaktion

Redaktør:

- Jesper Pedersen, jp@discus.dk, telefon 70 20 22 29

Baggrundsgruppe:

- Poul Skov Dahl, Niras Konsulenterne
- Tanja Miller, Videncenter for Evaluering i Praksis, CEPRA, University College Nordjylland

Indholdsfortegnelse

Evalueringsnyt ændrer form.....	1
Redaktion.....	3
Nyheder	4
Aktuelle informationer fra Dansk EvalueringsSelskab	4
Gå-hjem-møde: Kan evalueringerne holde til en evaluering?	5
Evert Vedung besøgte Århus	6
Artikler.....	7
Akkreditering og selvevaluering: En vej til meningsfuld udvikling?	7
Måling af "det bløde" på det sociale område	14
Virkningsevaluering in action: Virker læsevejledning?	21
Om Dansk EvalueringsSelskab	28
Medlemskab.....	28
Bestyrelsen 2008-2009	29

Nyheder

Aktuelle informationer fra Dansk EvalueringsSelskab

I dette nummer af "Evalueringsnyt" fra Dansk EvalueringsSelskab er der 3 spændende artikler, der handler om helt aktuelle temaer på evalueringsområdet: akkreditering, måling og virkningsevaluering. Artiklerne er skrevet af nogle af dem, der arbejder med temaerne i daglig praksis. Vi håber, at alle Dansk EvalueringsSelskabs medlemmer får udbytte af at læse artiklerne, og at de kan give god inspiration.

Bestyrelsen i Dansk EvalueringsSelskab planlægger at udsende løbende nyhedsbreve til medlemskredsen og at supplere med en række artikler om evaluering i et særligt elektronisk tidsskrift. Disse artikler vil efterfølgende blive opsamlet i en trykt udgave. Vi håber initiativet vil bidrage til, at vi kan udsende helt aktuelle informationer, der har tilknytning til evalueringsområdet, og at vi med artiklerne kan fortsætte traditionen i selskabet med udgivelse af relevante metodeorienterede artikler om evalueringsmetode og evalueringsanvendelse.

Det betyder, at Evalueringsnyt nr. 22 er det sidste i den nuværende form. Fra bestyrelsen skal der lyde en stor tak til Jesper Pedersen for redaktionen af Evalueringsnyt. Jesper Pedersen stopper som redaktør af Evalueringsnyt i forbindelse med vores nye initiativ.

Bestyrelsen planlægger også at indbyde Dansk EvalueringsSelskabs medlemskreds til deltagelse i en række ERFA-grupper inden for forskellige evalueringstemaer. Dette for at bidrage til inspiration, sprede viden og skabe gode rammer for netværksdannelse.

Vi er også i fuld gang med at planlægge årets konference. Konferencen finder sted på Hotel Nyborg Strand fra torsdag den 10. september til lørdag den 12. september 2009. Overskriften på årets konference er "Effektmåling – metoder, anvendelse og effekter". Husk at sætte kryds i kalenderen allerede nu. Programmet udsendes i maj måned.

På bestyrelsens vegne

Peter Rhode
Formand

Gå-hjem-møde: Kan evalueringerne holde til en evaluering?

DES afholder gå-hjem-møde tirsdag den 27. januar 2009, kl. 16.00 til 18.00 i receptionslokalet på Århus Rådhus, Rådhuspladsen 2, 8000 Århus C. Oplægsholderen er Simon Calmar Andersen, Institut for Statskundskab, Århus Universitet

Temaet

Kvalitetsudvikling, målstyring og evalueringer vinder stadig stigende udbredelse i den private og den offentlige sektor. Det eneste, som ser ud til ikke at blive evalueret, er evalueringerne selv. Hvem ved, om evaluering og kvalitetsudvikling skaber bedre kvalitet eller om konkurrence og målstyring skaber mere effektivitet?

Oplægget vil præsentere resultatet af en række store statistiske undersøgelser af effekten af moderne styringsredskaber på skoleområdet. I de fleste tilfælde peger resultaterne på, at styringsinstrumenterne mest af alt er kontraproduktive. Spørgsmålet er, om vi skal tilbage til gamle dyder i form af stabilitet og bureaukrati.



Simon Calmar Andersen

Om oplægsholderen

Post.doc og fuldmægtig Simon Calmar Andersen er ph.d. fra Statskundskab, Aarhus Universitet, og har bl.a. forsket i politisk styring af den offentlige sektor med særligt fokus på folkeskolen. Han har blandt andet skrevet artiklerne "The unintended effects of private school competition" (Journal of Public Administration Research and Theory 2007 - med Søren Serritzlew) og "The Impact of Public Management Reforms on Student Performance" (Public Administration 2008). Generelt viser hans forskning, at mange af de nye styringsformer indfrier langt mindre, end de lover. Nogle af de NPM-inspirerede former for styring virker decideret mod hensigten.

Tilmelding finder sted til sekretariatet på e-mail sekretariat@danskevalueringsselskab.dk senest tirsdag den 20. januar 2009.

Evert Vedung besøgte Århus

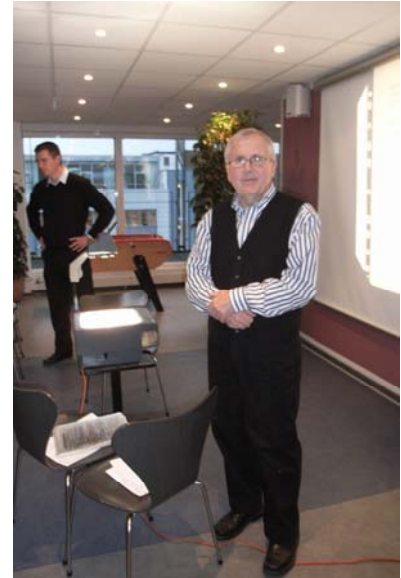
DES indbød til gå-hjem-møde med Evert Vedung den 1. december 2008 hos NIRAS Konsulenterne i Århus.

Evert Vedung holdt et oplæg om, hvad der kendetegner netværksstyring og hvordan denne form for New Public Management skal evalueres.

Evert har tidligere skrevet to artikler i Evalueringsnyt som omhandler dette tema. Du kan downloade de to artikler fra nr. 2 og nr. 9 her:

<http://www.danskevalueringsselskab.dk/Nyhedsbrev.asp>

Evert Vedung er Professor emeritus fra Uppsala og er desuden adjungeret professor ved Københavns Universitet og docent ved Åbo Akademi i Finland. I 2007 blev han æresmedlem i Dansk EvalueringsSelskab.



Evert Vedung med Poul Skov Dahl i baggrunden



Artikler

Akkreditering og selvevaluering: En vej til meningsfuld udvikling?

Af Hanne Dorthe Sørensen, Direktør, og Mads Biering-Sørensen, Specialkonsulent, SCKK

Indledning

Med regeringens kvalitetsreform er der sat fornyet fokus på kvalitetsarbejdet i den offentlige sektor. En væsentlig del af kvalitetsreformen er at give institutionerne rammerne til at tænke nyt og udvikle kvaliteten. Dette skal bl.a. ske gennem udvikling af fælles kvalitetsudviklingsværktøjer så som redskaber til brugertilfredshedsmåling, måling af faglig kvalitet (kvalitetsoplysninger) og akkrediterings-modeller. På lang sigt kan det dog vise sig, at det er målet om at udbrede akkreditering som en central metode til både kvalitetssikring og kvalitetsudvikling, der har de væsentligste følger for arbejdet med udviklingen af kvaliteten i den offentlige sektor.

SCKK (Statens Center for Kompetence- og Kvalitetsudvikling) har i de sidste 10 år understøttet statens arbejdspladser med systematisk kvalitetsudvikling bl.a. gennem selvevaluering med de helhedsorienterede kvalitetsmodeller KVIK og Excellence Modellen. Erfaringerne med selvevaluering som kvalitetsudviklingsmetode er positive, og der er særligt fra kommunalt hold en stigende interesse for anvendelsen af kvalitetsmodeller.

Artiklen vil, for det første, analysere hvad intentionerne bag akkreditering i kvalitetsreformen kan have været, for det andet, kort belyse forskellene mellem akkreditering og selvevaluering, for, til sidst, at se om det giver mening at kombinere selvevaluering og akkreditering i institutionens arbejde med kvalitetsudvikling.

Akkreditering

Lad os starte med en simpel definition af akkreditering. Akkreditering er en vurdering af, om en ydelse eller institution lever op til eksterne foruddefinerede kriterier for kvalitet. Akkrediteringen resulterer i et udsagn om, hvorvidt den evaluerede ydelse eller institution imødekommer kriterierne (EVA, 2006). Akkrediteringen vil almindeligvis blive foretaget af eksterne assessorer.

Dette giver dog stadig mulighed for at lave mange typer af akkrediteringsmodeller. Væsentlige dimensioner er fx:

- Formålet med akkreditering – herunder balancen mellem sikring og udvikling
- Akkrediteringsniveau og genstand – fx vurdering af produkt, kvalitetssystemet, organisation
- Kriteriernes udformning – fx indhold, detaljering, åbne/lukkede, generisk eller specifikke
- Resultat af akkrediteringen – godkendelse/afslag, hvilket kan have en række følger, fx anerkendelse, finansiering, tildeling af status, tildeling af ret til at foretage opgaver

I tabellen på næste side er to almindelige akkrediterings-typer, rettet mod henholdsvis produkter/ydelser og institutionsniveauer, sammenlignet med helhedsorienteret selvevaluering med kvalitetsmodeller.

	Formål	Genstand	Kriterier	Resultat
Produkt/ydelses akkreditering	<ul style="list-style-type: none"> • Sikring af tilfredsstillende niveau 	<ul style="list-style-type: none"> • Afdeling/enhed • Processer og produkt 	<ul style="list-style-type: none"> • Stor detaljering • Ofte minimumsstandarder 	<ul style="list-style-type: none"> • Godkendelse • Krav
Institutions-audit/ akkreditering	<ul style="list-style-type: none"> • Sikring og udvikling af interne styrings- og sikringsprocesser 	<ul style="list-style-type: none"> • Styrings- og kvalitetssikrings-system • Organisatoriske processer 	<ul style="list-style-type: none"> • Mindre detaljering • Kan indeholde minimumsstand. 	<ul style="list-style-type: none"> • Audit: Anbefalinger - • Akkreditering: Godkendelse og krav
Helhedsorienteret selvevaluering	<ul style="list-style-type: none"> • Udvikling (+sikring) af institutionens samlede indsatser og resultater • Skabe evalueringskultur og fælles bevidsthed om retning 	<ul style="list-style-type: none"> • Alle indsatser og resultater • Kvalitetssystematik • Resultatopnåelse (Bruger-, samfunds-, medarbejder- og nøgleresultater) 	<ul style="list-style-type: none"> • Lille detaljering • Kriterier sætter ramme for vurdering af indsatser og resultater 	<ul style="list-style-type: none"> • Styrker og forbedringer • Vurdering af kvalitetsarbejdet og resultater

Leon Lerborg har i sin artikel "Akkreditering – Bureautrusteret tillid", skelnet mellem hvad han kalder hård og blød akkreditering, der viser spændet i akkrediteringsbegrebet. Den hårde version ligger måske tættest op af den klassiske forståelse af akkreditering, og vil ofte være rettet mod ydelser (selv om få akkrediteringssystemer lever op til alle kravene). Da den hårde akkreditering er obligatorisk, og baseret på ekstern vurdering med efterfølgende myndigheds-godkendelse med væsentlige konsekvenser (fx mistet bevilling eller autorisation), skal standardernes opfyldelse være klar og konstaterbare, hvilket giver mange og detaljerede kriterier. Omvendt har den bløde variation, der ofte er institutions-rettet, almindeligvis svagere sanktioner, fx i form af positiv anerkendelse eller øget opsyn ved manglende godkendelse. Den bløde variant giver derfor ofte plads til mere dialog, åbne spørgsmål, refleksion og mere løs og kompleks vurdering etc. Den hårde variant har især til formål at sikre og kontrollere, at der er en tilfredsstillende kvalitet i det vurderede, hvorimod den bløde variant ofte har til formål, udover en vis kvalitetssikring, at understøtte kvalitetsudvikling.

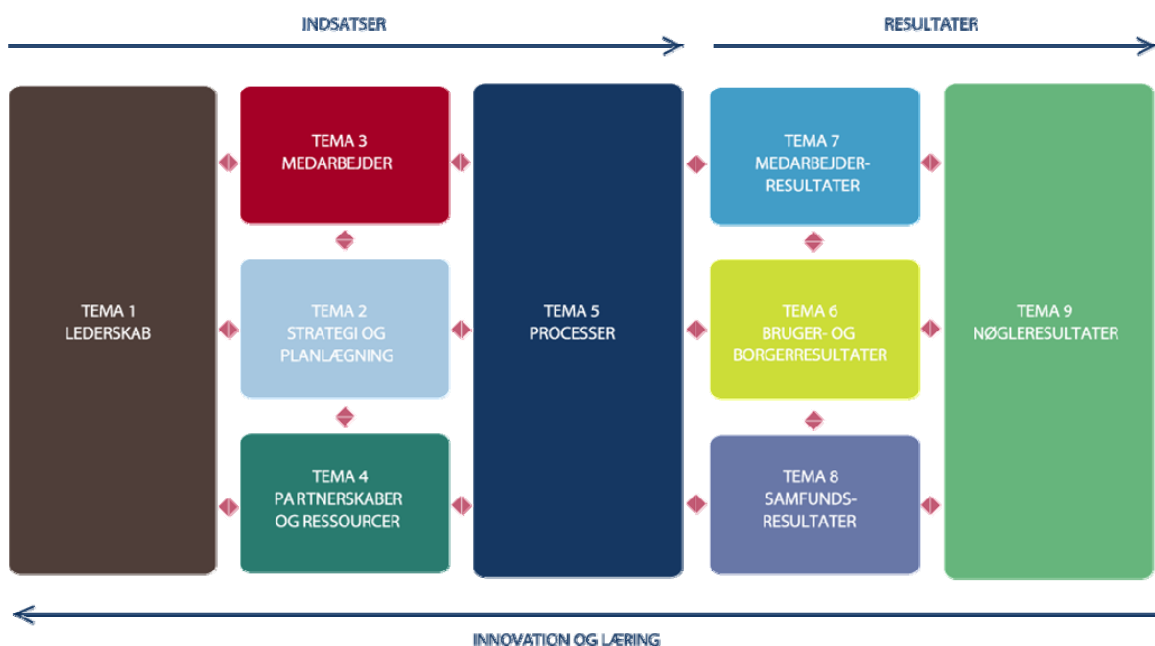
	Hårde varianter	Bløde varianter
Detaljering i standarder	Stor	Ringe
Antal standarder	Mange	Få
Fleksibilitet i standarder	Ingen	Stor
Målbarhed i standarder	Konstaterbar	Vurdering
Ekstern/intern	Ekstern akkreditering	Intern akkreditering
Forpligtethed	Obligatorisk	Frivillig
Akkrediteringsniveau	1. orden	2. Orden

Kilde: "Akkreditering - Bureautrusteret tillid". Af Leon Lerborg. Juni 2008.

Akkrediteringsmodeller i kvalitetsreformen

Akkreditering er kun et element i kvalitetsreformen, og forstås bedst i relation til reformens andre initiativer og de overordnede intentioner. Der er fx en sammenhæng mellem de forskellige redskaber til kvalitetsudvikling og målet om afbureaukratisering. Afbureaukratisering skal bl.a. opnås gennem klarere ansvarsfordeling mellem niveauer med bedre mål- og rammestyring. Dette forudsætter en klar dokumentation af målopfyldelse og sammenlignelige oplysninger om brugertilfredshed og kvalitet. Standardisering af redskaberne til brugerundersøgelser og definition af centrale kvalitetsindikatorer giver mulighed for læring om bedste praksis ved at kunne identificere de institutioner, der opnår bedst resultater og effekt på tværs af institutioner. De standardiserede dokumentationskilder udgør derudover den grundlæggende infrastruktur for indførelse af akkreditering. Cirklen slutes ved, at akkreditering på sigt skal understøtte afbureaukratisering ved at reducere behovet for tilsyn og kontrol gennem afvikling af detaljerede procesregler. Akkreditering indgår dermed i et kompleks af initiativer og mål. Eksempelvis skal akkrediteringsmodellerne "...understøtte forankringen af det lokale ansvar for kvalitetsudvikling og sikre løbende læring på de kommunale institutioner gennem ekstern sparring med uafhængige og særligt uddannede akkreditører." (Regeringen, 2007 s. 108). Her lægges vægten især på at forankre kvalitetsudviklingen bredt, og akkrediteringsaktiviteten er positivt defineret som 'ekstern sparring'. I forhold til akkreditering på daginstitutionsområdet "skal der udvikles en helt enkel akkrediteringsmodel – en "akkreditering light" – som børnehaver og vuggestuer kan bruge til at inddrage medarbejdere i at vurdere mål, ledelse, arbejdsprocesser og opnåede resultater." (Ibid, s. 108)

Akkreditering har her et væsentligt element af selvevaluering og egenkontrol, og skal derudover være 'light' og 'enkel' dvs. med få men omfangsrige kriterier. Der er selvfølgelig mange uafklarede områder i forhold til, hvordan akkrediteringsmodeller bliver udformet i praksis – men det tyder på, at modellerne hælder mod den mere bløde definition af akkreditering. Dog på en måde så akkrediteringsmodellerne stadig skal kunne afløse dele af det eksisterende kontrol- og tilsynssystem.

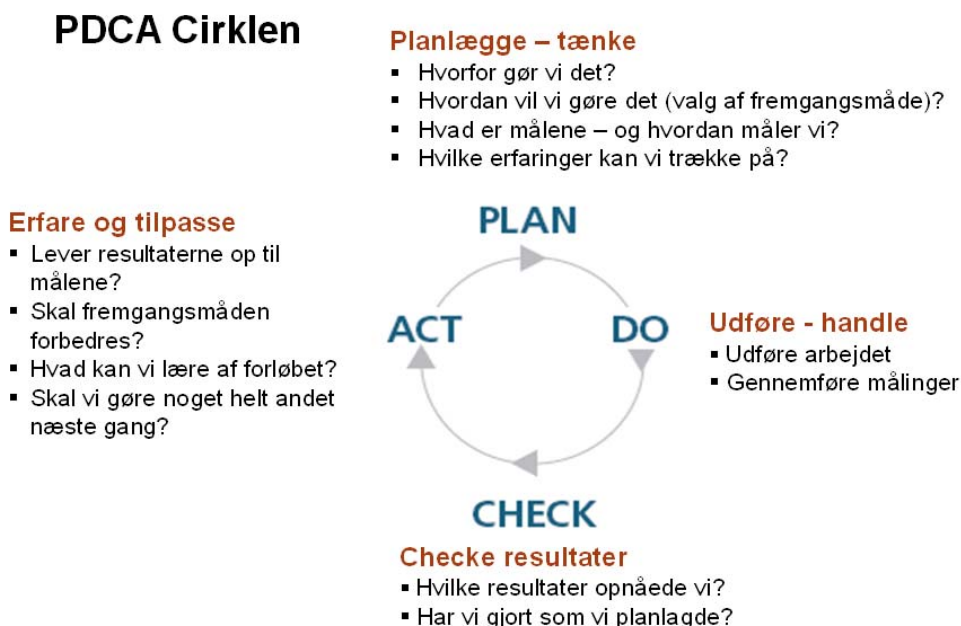


Selvevaluering med KVIK

KVIK tilbyder offentlige arbejdspladser en forholdsvis enkel og pædagogisk ramme til selvevaluering af arbejdspladsen. Selvevaluering er en kritisk intern evaluering af indsatser (det vi gør), og resultater (det vi opnår), der udføres af en repræsentativ gruppe af ledere og medarbejdere. Dermed opnås et samlet og afstemt billede af arbejdspladsens styrker (hvor vi præsterer godt), og forbedringsområder (hvor vi kan blive bedre). Kendetegnet for den gode selvevaluering er, at der skabes grobund for organisationsudvikling, der er funderet i egen praksis og forankret bredt på arbejdspladsen. KVIK Modellen er en helhedsorienteret spørgeramme, der kan bruges til at evaluere den samlede arbejdspladsen fra lederskab til nøgleresultater. Modellen rummer i alt 28 emner, som falder inden for 9 temaer. Temaerne dækker både indsatser (det vi gør), og resultater (det vi opnår). For hvert af de 28 emner er der en række spørgsmål og eksempler, der er tænkt som inspiration til at vurdere arbejdspladsens indsatser og resultater. Erfaringerne viser, at netop de praktiske eksempler gør det muligt, at alle medarbejdere, uanset uddannelse eller arbejdsområde, kan være med til at gennemføre en selvevaluering.

I selvevalueringen bliver man derudover bedt om at vurdere, hvordan man arbejder med arbejdspladsens indsatser i forhold til PDCA-cirklen.

PDCA Cirklen



PDCA-cirklen eller kvalitetshjulet er kendt som et klassisk værktøj til at arbejde systematisk med forbedringer. PDCA-cirklen handler om at indarbejde rutiner for, hvordan opgaver planlægges, gennemføres, måles og evalueres. Selvevaluering processen har altså et dobbelt sigte: At opøve kompetencer til at analysere de komplekse sammenhænge mellem indsatser, mål og resultaterne. For det andet at give kompetencer til at arbejde mere systematisk med forbedring og evaluering af

egen praksis i dagligdag. Brugt rigtig kan selvevaluering dermed, ud over at give et fælles strategisk overblik her og nu, også vær med til at skabe en evalueringskultur bredt i organisationen, der understøtter en forankret og meningsfuld udvikling af institutionens kvalitet.

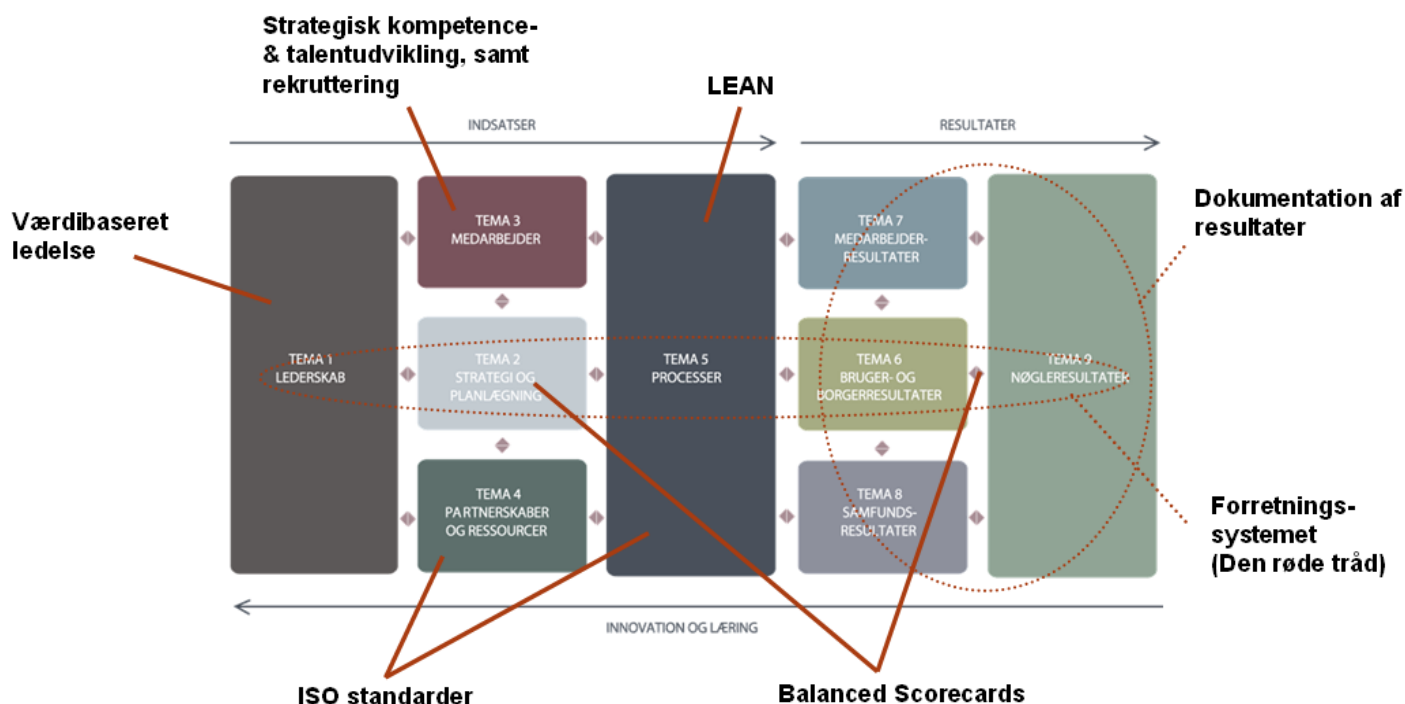
For at opsummere de oftest beskrevne udbytter af en KVIK selvevaluering:

- Skabe fælles strategisk overblik over styrker og forbedringsområder i den samlede organisation
- Fælles sprog om arbejdspladsen og ejerskab til indsatser og forbedringer
- Fokus på inddragelse af borgere og brugere, samt resultater og effekt af ydelser
- Ramme for læring – inden for og på tværs af institutioner
- Opbygning af evalueringskapacitet og læringskultur gennem kompetenceudvikling af medarbejder og ledere, der gør forandringsprocesserne meningsfulde bredt i organisationen

Akkreditering og selvevaluering

Er det meningsfuldt at kombinere selvevaluering og akkreditering i arbejdet med at sikre og udvikle kvalitet?

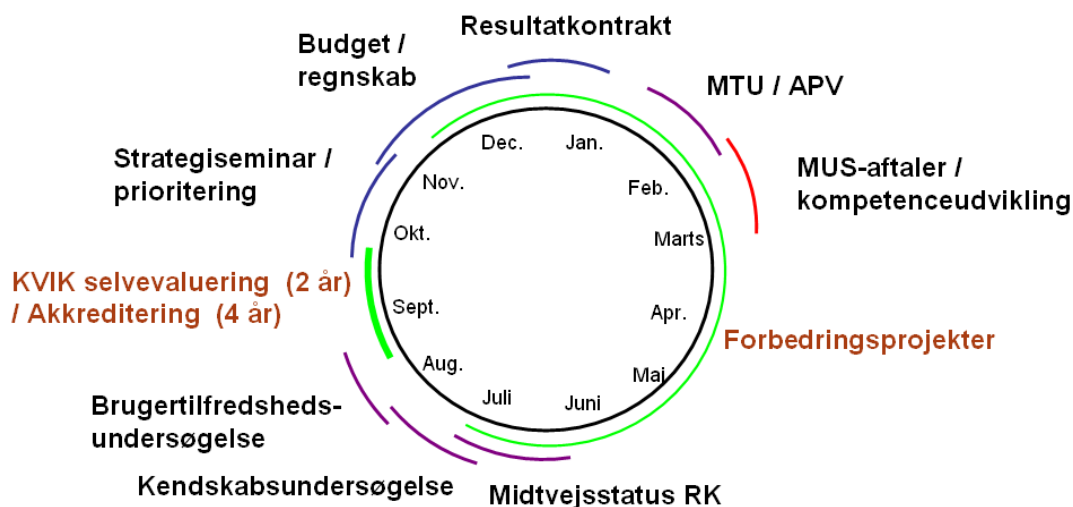
En selvevalueringsproces med KVIK giver ofte det største udbytte, når selvevalueringen ses som ét redskab blandt andre, og dermed ikke som et mirakelmiddel til at løse alle arbejdspladsens problemer.



Selvevaluering kan fx ikke i sig selv udvikle en ny strategi for arbejdspladsen. Det selvevalueringen gør, endog meget klart, er at vise hvis den røde tråd i forretnings-systemet ikke er der, dvs. at der ikke er sammenhæng mellem mission og vision, strategi og kerne-processer og de opnåede resultater. En oplagt forbedrings-aktivitet i dette tilfælde ville være at arbejde med Balanced Scorecards for at rette op på problemet. På samme måde kan selvevaluering med KVIK ikke i sig selv optimere arbejdspladsens arbejdsgange og kerneprocesser – men den kan give en fælles forståelse mellem ledelse og medarbejdere om, hvorfor det fx kunne være relevant at arbejde med LEAN for at frigive tid til andre processer.

På samme måde kan selvevaluering med KVIK hjælpe institutionen til at håndtere de udfordringer, akkreditering ofte stiller. For det første ved, gennem selvevaluering, at bruge modellens spørgeramme til at koble akkrediteringens standardiserede og generelle kriterier til institutionens konkrete og komplekse virkelighed. For det andet, ved at skabe et strategisk overblik og vise sammenhængen mellem institutionens indsatser og resultater, og hvor det er nødvendigt at ændre indsatser for at leve op til kriteriekrav til indsatserne eller resultater og effekt. For det tredje ved generelt at øge kapaciteten til at arbejde med kvalitetsudvikling og -sikring, og dermed reducere risikoen for en række af de almindelige unoder ved evaluering, der tit forstærkes ved akkreditering. Det gælder fx målforskydning (hitting the target, but missing the point), fremkomsten af statiske mål uden udviklingsperspektiv (teach to the test) og oplevelsen af meningsløse evalueringer og målinger i det daglige arbejde.

Timing af selvevaluering processen er også vigtig. Selvevalueringen skal fx tænkes ind i forhold til de styringsmæssige processer så som strategiudvikling, budgetlægning og definering af mål i resultatkontrakt, så selvevalueringen føder ind til aktiviteterne og omvendt, at der kan sættes ressourcer af til eventuelle forbedringsprojekter. Det er også en god ide at time dokumentationsindsamling og resultatopgørelse, så målingerne giver et så retvisende og aktuelt billede af institutionen som muligt.



Selvevaluering ville selvfølgelig også skulle tænkes tidsmæssigt sammen med den relevante akkrediteringscyklus. I den sammenhæng kunne man forestille sig, at der var en 2-årig cyklus på selvevalueringens aktivitet ved en 4-årig akkrediteringscyklus. To år før akkrediteringen laver institutionen en midtvejs selvevaluering for at diagnosticere de forbedringer, der er nødvendige for at kunne leve op til akkrediteringskriterierne. Umiddelbart op til akkrediteringen laves igen en selvevaluering for at klæde organisationen på ved at skabe fælles forståelse af institutionens situation og afklare, om der er forhold, man stadig skal korrigere og udbedre.

Er det meningsfuldt at kombinere selvevaluering og akkreditering i arbejdet med at sikre og udvikle kvalitet? Svaret er: Ja, i høj grad. Gjort rigtig kan selvevaluering nemlig være med til at klæde institutionen på til akkreditering, forankre akkrediteringen bredt på arbejdspladsen og gøre arbejdet med sikring og udvikling af kvalitet meningsfuldt for alle.

Mere viden om KVIK

På www.sckk.dk kan man gratis downloade KVIK Vejledningen (dvs. deltagerens spørgeramme) og Selvevalueringsguiden (der beskriver selvevalueringens proces fase for fase). På hjemmesiden vil man derudover finde mere materiale og gode råd om arbejdet med helhedsorienteret kvalitetsudvikling, information om Excellence Modellen og introduktion til Kvalitetspris for Den Offentlige Sektor og andre anerkendelser, der uddeles til institutioner der arbejder systematisk med kvalitetsudvikling på anerkendelsesværdig vis.

- EVA (2006): "Akkreditering og kriterieformulering: Erfaringer og perspektiver". Danmarks Evalueringsinstitut. 2006
- Regeringen (2007): "Bedre velfærd og større arbejdsglæde – Regeringens strategi for høj kvalitet i den offentlige service". Regeringen. August 2007.

Måling af "det bløde" på det sociale område

Af Morten Greve, René Buch Nielsen og Simon Østergaard Møller, Center for Kvalitetsudviklings

Der efterspørges i stigende grad kontant viden om resultaterne af indsatserne på det sociale område. Dette er der god ræson i. De sociale indsatser berører mange borgere og lægger beslag på en betydelig del af de offentlige udgifter.

Man skal blot gøre sig klart, at socialfaglig videnskabelse forudsætter, at måleinstrumentet yder retfærdighed til genstandens (inter)subjektive karakter. Mens det gerne er "hårde" data, der indsamles i forbindelse med medicinske forsøg, befinder vi os her i vidt omfang i den modsatte, "bløde" ende af kontinuet. Denne erkendelse kalder på kritisk refleksion: Er datagyldigheden i orden? Åbner data for reel læring?

Med udgangspunkt i tre cases hentet fra Center for Kvalitetsudviklings aktiviteter vil vi forsøge at illustrere forskellige tilgange til at dokumentere resultaterne af den sociale indsats og de udfordringer, som er forbundet med det. En første vigtig erkendelse i bestræbelserne på at få mere viden er, at der ikke findes nogen "one best way" eller "gylden standard" i forhold til socialfaglige målinger. Tværtimod er det vigtigt at åbne op for en bred vidensforståelse, som også kan indfange uventede vidensgevinster.

Case 1: Gennembrudsprojekt "Indflydelse på eget liv"

"Grafer og kurver er skudt langt over vores interesseområde – lav dem selv. Udvikling er fint, men ikke for enhver pris – arbejdet venter derhjemme!"

Ovenstående citat fra en pædagog viser en holdning dels til målinger i den pædagogiske praksis, og dels til konsulenter, som i bedste overbevisning forsøger at konkretisere og måle effekter af en specifik faglig indsats. På den anden side er anvendelsen af citatet også med til at fastholde et billede af en sektor, hvor modstand mod indførsel af måleparametre er et særligt kendetegn.

I praksis er billedet mere nuanceret, men at holdninger til måling af bløde sociale værdier aktualiseres, når pædagogisk personale skal formulere mål for den faglige indsats, er en af erfaringerne fra det landsdækkende kvalitetsudviklingsprojekt "Indflydelse på eget liv – Jubilæumsprojekt om brugerinddragelse og selvbestemmelse for mennesker med handicap" (2005-2007). Projektet blev igangsat af Amtsrådsforeningen, og baggrunden var et ønske om at markere 25 års jubilæet for statens åndssvageforsorgs og øvrige særforborgs udlægning til amterne.

Formålet med projektet var at forbedre daglig praksis – herunder samspillet og relationerne mellem brugere og medarbejdere, og derved opnå en forbedring af brugernes oplevelse af inddragelse og selvbestemmelse. Samtidigt var målet at udvikle det faglige og organisatoriske arbejde med brugerinddragelse og selvbestemmelse.

Projektet blev gennemført inden for rammerne af Gennembrudsmetoden, som er en struktureret kvalitetsudviklings- og implementeringsmetode baseret på netværkssamarbejde. Deltagerne i projektet var organiseret i team, som hver især repræsenterede et tilbud for voksne handicappede. Valg af metode var inspireret af et andet landsdækkende projekt omhandlende tvang i psykiatrien, som også var igangsat af Amtsrådsforeningen. På daværende tidspunkt viste tvangsprojektet så lovende resultater, at metodens potentiale også kunne afprøves på det sociale område.

Hvert team havde i en forberedelsesperiode afdækket deres praksis i forhold til brugernes inddragelse i eget liv, og på den baggrund formuleret en målsætning for projektdeltagelsen. Et Gennembrudsprojekt omfatter ikke benchmarking, og de lokale målsætninger vil variere fra team til team, idet afdækningen af egen praksis påviser forskellige områder, som brugere og personale ønsker at forbedre. Som inspiration til deltagerne, var der lavet et Forandringskatalog, der består af metoder og gode eksempler på brugerinddragelse.

Med afsæt i den lokalt formulerede målsætning, udarbejdede deltagerne mindst tre såkaldte resultatmål. I store træk var det ikke vanskeligt for deltagerne, at formulere hensigtserklæringer med deres deltagelse. Problemerne opstod i arbejdet med at operationalisere målsætningen. Det er et krav i den anvendte metode, at resultatmålene er målbare, hvorfor der opstod diskussioner omkring den pædagogiske virkelighed. Ifølge mange pædagoger i projektet ved konsulenter ikke noget om, hvordan pædagogisk praksis udfoldes. Hvis man ved det, vil man vide, at resultaterne af praksis ikke lader sig måle.

Det var altså ikke så meget et spørgsmål om at formulere målbare mål, men at målene implicit krævede en dataindsamling, som deltagerne anså som irrelevant ud fra en ide om, at pædagogisk praksis er kompleks og helhedsorienteret. At udtrække specifikke målbare elementer reducerer derfor den pædagogiske indsats. Flere deltagere hævdede, at det positivistiske vidensideal, som de oplevede lå bag Gennembrudsmetoden, næppe vil bundfælde sig i det socialfaglige felt.

Diskussionerne var i begyndelsen af principiel karakter, men ændrede karakter, efterhånden som deltagerne arbejdede med at formulere målbare mål. Jo mere deltagerne arbej-

dede med egne mål, jo mere blev det klart for de fleste, at der er sammenhæng mellem målsætning og dokumentation.

De følgende citater er hentet fra deltagernes selvevaluering. De viser, at deltagerne på forskellige måder har gjort sig nogle erfaringer med hensyn til måling af deres arbejde.

"Det har været givende at få indsigt i målemetoder i forhold til det pædagogiske arbejde - de bløde værdier".

"De målinger, der var indlagt som krav i projektet, var meget fremmed i den pædagogiske verden, men de viste sig at være yderst konstruktive til at bevise, at brugerne faktisk kunne mere, end vi troede i forvejen. Det har været interessant at opdage, at det ikke kun er holdningsændringer, det handler om for at få plads til mere brugerindflydelse".

"Det har været godt at have dokumentation for, at det gør en forskel, hvilket vi kunne mærke, da vi satte graferne på opslagstavlen, således at kollegaerne kunne se, at det havde gjort en forskel, når valget var givet".

Af de 16 deltagende team, som afsluttede projektet, har 10 team produceret dokumentation i form af tabeller og grafer. Tre team havde produceret andre former for dokumentation. En forsigtig konklusion er, at måling af pædagogiske indsatser i praksis giver betydelig mening, når praktikerne selv definerer, hvad der skal måles på og hvordan. Den pædagogiske verden er måske ligefrem en frugtbar muld for udvikling af nye og innovative dokumentationsformer, som supplement til den klassiske opdeling mellem kvalitative og kvantitative metoder.

Case 2: Det Sociale Indikatorprogram

Det Sociale Indikatorprogram (SIP) er et koncept for monitorering og evaluering af den sociale og socialpædagogiske indsats på døgninstitutioner for unge og botilbud for voksne handicappede.

SIP baserer sig på de metodiske principper i den såkaldte virkningsevaluering, som indebærer, at de opnåede resultater forsøges knyttet sammen med de gennemførte indsatser. For at styrke læringsperspektivet i programmet anses det således som centralt, at viden om hvilke institutioner/tilbud, der har bestemte resultater, følges af en viden om, hvordan disse resultater er opnået. SIP indeholder derfor såvel monitoreringssystemer for opnåede resultater som for gennemførte indsatser.

For yderligere at styrke læringsmulighederne for praksis af SIP og ikke mindst for at koble faglige vurderinger på monitoreringsdata, indeholder programmet også et koncept for udviklingen af audits. Audit er en proces, hvor et panel af relevante fagfolk vurderer dokumentationen i SIP med henblik på at fx at vurdere, hvorfor resultaterne ser ud, som de gør, eller hvilke forklaringer der kan ligge bag dokumenterede sammenhænge mellem resultater og processer. Tankegangen er, at auditpanelets drøftelser og konklusioner i sammenhæng med statistiske analyser på baggrund af SIPs monitoreringsdata kan danne grundlag for handlingsanvisende anbefalinger til praksis.

For at tilstræbe, at monitoreringen er relevant for praksis udvælges indikatorerne med udgangspunkt i en programteori, der formuleres i samarbejde med praktikere fra den relevante målgruppe. SIP er således opdelt i forskellige målgrupper med hver sit monitoreringsprogram. Indikatorerne for hvert program er blevet udvalgt i samarbejde med praksis på baggrund af relevans.

De udvalgte indikatorer er efterfølgende blevet operationaliseret i spørgeskemaer. Centrale principper i operationaliseringen har været enkelhed, så spørgsmålene er letforståelige og meningsfulde, og pragmatisme, hvor brugen af eksisterende validerede instrumenter er blevet brugt i så vid udstrækning som muligt (fx SDQ, Honosca).

Der er i dag programmer for sikrede døgninstitutioner, døgninstitutioner for unge med psykiske eller psykiatriske problemstillinger, botilbud for voksne med erhvervet hjerneskade, botilbud for voksne med autismespektrumforstyrrelse og botilbud for voksne med psykisk udviklingshæmning (udvikles i 2009). Disse programmer er alle rekvireret af regionerne i fællesskab. Derudover udvikles i øjeblikket et program for Region Midtjylland vedrørende socialpsykiatriske botilbud for voksne. Endelig er et program for unge anbragt på åbne døgninstitutioner fuldt implementeret i Oslo Kommune og under implementering i den norske Region Øst.

Der er endnu ikke indsamlet tilstrækkelige mængder data i programmerne til, at der er gennemført analyser og audits. Dette forventes gennemført for programmet på sikrede døgninstitutioner medio 2009. Erfaringerne fra de sikrede institutioner er imidlertid, at programmet kan bruges lokalt som afsæt for en mere systematisk udredning af de unge, som anbringes, fordi de unges problemer, ressourcer og udvikling bliver udtrykt på en overskuelig måde. Endvidere kan SIP bruges som et værktøj på institutionerne til at udvikle og synliggøre den pædagogiske faglighed, fordi det herved bliver muligt at dokumentere, hvil-

ke mål og hvilken indsats, der konkret er blevet arbejdet med i hvert enkelt anbringelsesforløb. Dette skaber samtidig et godt grundlag for at gå i dybden med et forløb.

Dermed bliver SIP et eksempel på, at dokumentationsarbejdet i sig selv kan have en vigtig refleksiv værdi for de, som dokumenterer, så længe det, der skal dokumenteres, bliver oplevet som relevant og meningsfuldt. For SIP har dette, som nævnt, været et selvstændigt succeskriterium.

Som andre monitoreringsprogrammer baseret på indikatorer, indeholder SIP dog en risiko for såkaldte konstitutive effekter for praksis, forstået som at fokus i den konkrete praksis skifter på en uhensigtsmæssig måde mod det, som der måles på, frem for det som er fagligt relevant.

SIP forsøger at imødekomme denne risiko ved at inddrage praksis tæt i udviklingsarbejdet og udvælgelsen af indikatorer, herunder at bruge ressourcer på at formulere en programteori for feltet. Dette fjerner imidlertid ikke risikoen. Da der er tale om en bred programteori, som dækker flere sociale tilbud (med indbyrdes forskellig praksis), og da der er sket en udvælgelse af hvilke indikatorer, der skal måles på, risikerer SIP fortsat, at konkrete mål og indsatser på de involverede døgninstitutioner og botilbud vil falde uden for indikatorerne. Det er særdeles vigtigt at være tydelig og opmærksom på denne risiko.

Case 3: Resultatevaluering af multisystemisk terapi

Vores tredje eksempel på et dokumentationsprojekt på det sociale område er en landsdækkende resultatevaluering af Multisystemisk Terapi (MST) gennemført i 2005-2007 af Kvalitetsafdelingen, Århus Amt og siden Center for Kvalitetsudvikling, Region Midtjylland. Undersøgelsen blev rekvireret af Jysk Socialforsknings- og Evalueringssamarbejde og Servicestyrelsen i Velfærdsministeriet.

MST er et familie- og netværksterapeutisk program, som indebærer en intensiv og kortvarig behandling af unge med adfærdsmæssige og sociale problemstillinger (normal varighed 3-5 mdr.). Metoden har været brugt i Danmark siden 2003 og har fra starten været omdiskuteret i de socialfaglige miljøer. Særligt har diskussionen kredset om, hvorvidt metodens effekt var pålideligt dokumenteret.

Metodisk var undersøgelsen delt i to. Den første del, som omfattede 213 MST-forløb, bestod af en både retrospektiv og prospektiv panelundersøgelse, hvor data indhentes fra journalmateriale og interviews med familierådgivere i kommunerne. Her blev der foretaget

opfølgingsmålinger ni og atten måneder efter udskrivningen. Undersøgelsens anden del bestod af en prospektiv panelundersøgelse og omfattede 127 familiers MST-forløb. Opfølgningen skete her seks måneder efter udskrivningen. Data blev i denne del indsamlet gennem interviews med og spørgeskemaer fra familierådgivere, de unges skoler/arbejdspladser samt de unge selv og deres familier. I forbindelse med dataindsamlingen til anden del blev bl.a. en række standardiserede måleinstrumenter anvendt, herunder CBCL og FACES IV.

Svarprocenterne fra spørgeskemaer og telefoninterviews var meget svingende i undersøgelsen. Højeste svarprocent blev opnået fra familierådgiverne (90-98 %), mens svarprocenten fra de unge var den laveste (47-71 %). Telefoninterview med forældrene gav en pæn svarandel på 85-95 %.

MST's primære resultatmål vedrører skolegang/arbejde, kriminalitet og misbrug. Undersøgelsen viste, at MST signifikant reducerede de unges fravær fra skole og arbejde, de unges kriminalitet og de unges misbrug. Det generelle billede var, at andelen af de unge, som havde problemer på disse tre områder, blev ca. halveret mellem indskrivnings- og opfølgingsmålingerne. De standardiserede instrumenter viste bl.a., at primært de unges forældre oplevede en forbedring i familiefunktionen (FACES IV), mens de unge kun i ringe grad selv oplevede en positiv forandring. Medarbejdere fra de unges skoler vurderede de unges almene psykiske befindende ved hjælp af CBCL-instrumentet, og her viste der sig en klar forbedring hos drengene, mens en mindre forbedring ved udskrivningen ikke viste sig holdbar ved opfølgningen for pigernes vedkommende. Begge gruppers problemscore ved opfølgningen lå dog klart over niveauet for aldersgruppen i befolkningen som helhed.

Bl.a. i lyset af svarprocenterne viser erfaringerne fra undersøgelsen, at der er nogle særlige udfordringer forbundet med at foretage interviews i større omfang med fx udsatte unge. For det andet viser erfaringerne fra undersøgelsen, at det er muligt at anvende standardiserede måleinstrumenter i resultatevalueringen af sociale programmer. Under de rette omstændigheder kan sådanne instrumenter udgøre en pålidelig og formentlig valid måling af, hvilke resultater en given social indsats har.

Endelig udtrykker denne evaluering også svaghederne ved mange konventionelle resultat- og effektmålingsdesign. Særligt problematisk er manglen på viden om, hvilke konkrete tiltag i MST-forløbet, der har virket over de forskellige unge, der har deltaget i undersøgelsen. Det er således ikke muligt i denne type undersøgelser at sige noget om, hvorvidt forskelle i MST's resultater på tværs af de unge hidrører fra forskelle i den konkrete behand-

lingsindsats eller i stedet fra andre ("ydre") omstændigheder. Med andre ord kan det således være vanskeligt at sige noget præcist om, hvorfor MST virker over for nogle unge, og ikke over for andre.

Som evaluering betragtet er resultatundersøgelsen af MST således behæftet med nogle svagheder og er udtryk for, at der fortsat er langt at gå i forhold til at tilvejebringe mere viden på dette område. Men det er samtidig også et vigtigt bidrag i den fortløbende proces mod at skabe mere indsigt i, hvordan det sociale arbejde virker. Mens forestillingen om, at det er muligt at tilvejebringe "sikker viden" om, hvad der virker, ikke er realistisk, er det første skridt på vejen at åbne for at et mere rummeligt vidensbegreb – også hvad angår resultatmålinger.

Nogle tværgående iagttagelser

De her nævnte eksempler er alle forsøg på at tilvejebringe konkrete målinger og undersøgelser af det sociale arbejde. Hvert af projekterne har sine styrker og svagheder og understreger, at evaluering og måling af socialt arbejde ikke kan forventes at give os en endegyldig sandhed om, hvad der virker, men i bedste fald kan tilvejebringe en foreløbig viden.

Selvom der ikke findes en "gylden standard" for, hvordan vi kan skabe sikker viden på dette område, betyder det på ingen måde, at det er umuligt at blive klogere, men det er centralt, at vi er opmærksomme på, at vi sjældent kan forudsige alle vidensgevinsterne i forbindelse med vores målinger, men i stedet må være forberedte på, at de kan tage mange – og ofte overraskende – former. Det er derfor vigtigt, at vores forståelse af, hvad viden er, er tilstrækkelig bred til, at vi kan få øje på disse gevinster.

Virkningsevaluering in action: Virker læsevejledning?

Af Niels Matti Søndergaard og Eva Pallesen, Danmarks Evalueringsinstitut

*Virkningsevaluering har været på dagsorden de seneste år i Danmark i evalueringsfaglige kredse. Men endnu er der få eksempler på evalueringer af et lidt større omfang der baserer sig på metoden. Det gælder også internationalt. Gennembladrer man de seneste numre af tidsskriftet *Evaluation*, finder man fx overraskende få eksempler på vellykket "Realistic Evaluation", "Contribution Analysis" o.l. i praksis. I en ny EVA-evaluering har vi gjort forsøget.*

I denne artikel præsenterer vi vores erfaringer med at arbejde med virkningsevaluering i forbindelse med et EVA-projekt om læsevejledere. Vi fremlægger den metode vi har fået erfaringer med, sammen med de konkrete resultater fra projektet – hvor vi har afdækket sammenhænge mellem en særlig indsats (brugen af læsevejledere på en skole) og virkningen (bliver eleverne bedre læsere af at der er en læsevejleder på deres skole?).

Baggrunden for vores projekt

Baggrunden for vores undersøgelse er at mange kommuner de senere år har uddannet en eller flere lærere pr. skole til at blive læsevejledere fordi man mener at de vil styrke arbejdet med læsning. Troen på læsevejlederfunktionens virkning kommer også til udtryk i at staten i 2006 valgte at afsætte 55 mio. kr. til at uddanne læsevejledere. Grunden til at man satser på læsevejledere, er selvfølgelig at man har en formodning om de vil have en positiv betydning for elevernes læsefærdigheder. Men medfører dét at ansætte læsevejledere at eleverne i sidste ende bliver bedre læsere – og hvis ja, hvornår og hvordan er det at læsevejlederfunktionen virker? Det ville vi gerne afdække i vores projekt.

Spørgsmålet er særligt interessant fordi læsevejledere afspejler en mere generel ide der er i vælten for tiden om faglige specialister eller spidskompetencefunktioner i folkeskolen, dvs. ideen om en specialist der udstyres med en særlig viden, og som spreder sin viden til kollegerne der dermed kan levere en mere kvalificeret undervisning. Eksempler på disse spidskompetencefunktioner eller faglige fyrtårne som de også kaldes, er mange og kommer fx fra områderne it, tosprogede og adfærd, kontakt og trivsel (AKT).

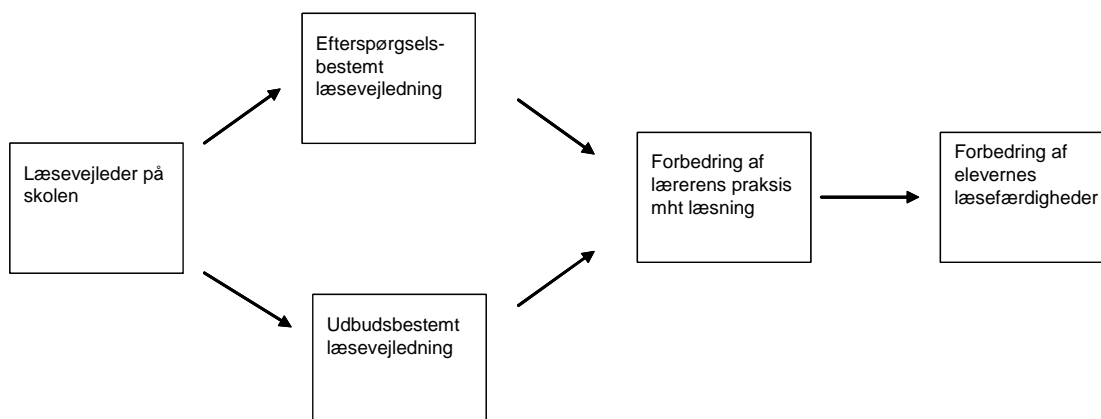
Sådan opstillede vi programteorien

Evalueringen tog udgangspunkt i Hvidovre Kommune. Hvidovre Kommune har formuleret en læsehandlingsplan hvor læsevejlederfunktionen er en krumtap.

Udgangspunktet for evalueringen var at opstille programteorien for læsevejlederindsatsen. Vi ville afdække det centrale hovedspørgsmål: Hvordan er det at læsevejlederen tænkes at virke?

Programteorien blev opstillet på baggrund af Hvidovre Kommunes læsehandlingsplan og interview på skoler og i forvaltningen. Den foreløbige programteori blev diskuteret med lærere og læsevejledere på et programteoriseminar. Programteorien havde i løbet af denne fase af projektet flere forskellige skikkelser, nogle store og komplicerede, og andre små og simple. Vi endte med at arbejde med to lag af programteorier, den forholdsvis simple illustration af programteorien som fremgår af figur 1 og to mere udfoldede versioner af programteorien som vi anvendte til næranalyser af det indsamlede datamateriale.

Figur 1. Programteori for læsevejlederindsatsen.



Figur 1 viser den mest simple fremstilling af programteorien bag læsevejlederindsatsen. Som de to pile fra boksen "Læsevejleder på skolen" viser, er der to grundlæggende forskellige måder man kan forestille sig at læsevejlederen skal indgå i opkvalificeringen af sine kolleger. I evalueringen er de døbt henholdsvis den "efterspørgselsbestemte" og den "udbudsbestemte" læsevejledning.

Hvad angår den efterspørgselsbestemte læsevejledning, er det tanken at læreren fx gennem en fast træffetid eller på mere uformelle måder skal stå til rådighed for lærerkollegerne med råd og vejledning når de har et spørgsmål om læsning. Den udbudsbestemte del af læsevejledningen knytter sig til afholdelsen af en klasselæsekonference. I Hvidovre testes eleverne på bestemte klassetrin i læsning, og klassens lærere indkaldes efterfølgende sammen med skoleinspektørerne af læsevejlederen til en konference eller et møde

hvor læseresultaterne gennemgås, og det drøftes hvilke tiltag der skal til for at udvikle klassens læsefærdigheder.

Både den efterspørgselsbestemte og udbudsbestemte læsevejledning forventes at virke gennem at opkvalificere lærerens praksis med hensyn til læsning, fx gennem øget opmærksomhed på konkrete elevers læseproblemer eller øget refleksion om egen undervisning i forhold til læsning. Og endelig er det tanken at elevernes læsefærdigheder styrkes af lærernes forbedrede praksis.

Efterforskning af virningskæden i praksis – det viste undersøgelsen

Næste skridt i vores arbejde bestod i at efterprøve om teorien fungerede i praksis. I evalueringdesignet havde vi for hvert led i programteorien formuleret nogle aktiviteter der ville være synlige hvis programteorien var sand, og ligeledes nogle resultater der skulle opnås ved hvert led af kæden. Helt banalt er fx den første betingelse for at den efterspørgselsbaserede læsevejledning kan fungere, at lærerne opsøger læsevejlederen med læserelaterede problemer, og at de oplever et udbytte af den vejledning de modtager.

For den efterspørgselsbaserede læsevejledning viste det sig at halvdelen af lærerne på mellemtrinnet på de skoler vi undersøgte, havde benyttede sig af tilbuddet om at henvende sig til læsevejlederen i skoleåret 2007/2008. Af disse havde det for hovedpartens vedkommende ført til ændringer af lærerens praksis i forhold til læsning, ligesom at hovedparten vurderede at tiltagene gavtede læseudviklingen hos eleverne.

Det viste sig altså at den efterspørgselsbaserede læsevejledning havde en effekt. Samtidig vidste resultaterne at effekten var afhængig af nogle kontekstfaktorer. Fx var faglærerne mindre tilbøjelige til at henvende sig end dansklærerne, på trods af at samtlige skoler i undersøgelsen (og flertallet af lærere) har den holdning at læsning er et anliggende for alle fag, er alle læreres ansvar, og derfor ikke kan parkeres hos dansklæreren. Ligeledes viste evalueringen at yngre lærere var mere tilbøjelige til at opsøge læsevejlederen end ældre.

I analysen af den udbudsbaserede læsevejledning sammenholdt vi resultaterne af læsetest for forskellige klasser med en analyse af klasselæsekonferencerne og de forløb der var blevet igangsat på baggrund af dem. Det viste sig at den udbudsbaserede læsevejledning havde en effekt, men også samtidig at det har stor betydning hvordan læsevejlederen agerede. Også analysen af den udbudsbaserede læsevejledning peger altså på betydningen af kontekstfaktorer for om indsatsen er virkningsfuld eller ej.

Et vigtigt resultat var at billedet af læsevejlederen som en ekspert der har en generaliseret viden om metoder og materialer mv. som hun videregiver til lærerne der så omsætter denne viden i forbedret undervisningspraksis med bedre elevresultater til følge, skulle revideres.

Det viste sig ofte for lærerne at være frustrerende at få præsenteret en generaliseret viden i form af ideer og gode råd med en underliggende forventning om at man umiddelbart kan anvende dem i praksis og dermed opnå bedre resultater. Og det viste sig faktisk også i praksis at være et svært udgangspunkt for læsevejlederen selv at det forventes at man besidder al relevant viden om læsning og kan udpege løsninger som lærerne umiddelbart kan implementere.

I stedet viste det sig at være virkningsfuldt for elevernes læsefærdigheder når læsevejlederen medtænkte sig selv som praktiker i læseindsatsen. Når læsevejlederen i stedet blot at være en der peger på relevant bøger og materiale opfatter sig selv som en "medpraktiker". Et eksempel er en læsevejleder der i en af de klasser hvor vi overværede klasselæsekonferencen, gik ud og arbejdede med at konstruere et "læsebånd" for to klasser der havde fået overraskende dårlige resultater.

Et andet fund var at det havde betydning for læsekonferencens virkning på ændringer i undervisningen om resultaterne blev opfattet som kritiske. I de klasser vi fulgte, var det afgørende for om læsevejledningen fik en virkning eller ej, at læreren skulle opleve at stå på en "brændende platform. Det vil sige at hvis resultaterne af læsetesten blev taget som udtryk for en nødvendighed af at skabe forandringer i undervisningen som helhed så var det mere sandsynligt at læsekonferencen fik en betydning for undervisningen efterfølgende. Det er et overraskende resultat fordi det i programteorien blev betonet at en vigtig kontekstfaktor for læsevejledningens virkning, var at læreren skulle føle sig tryk til klasselæsekonferencen for at kunne være tilbøjelig til at tage og råd og vejledning til sig..

Til gengæld viste undersøgelsen af virningskæden også at nogle af de kontekstfaktorer som lærere, læsevejledere og ledere på forhånd havde forventet havde indflydelse på om læsevejledning var virkningsfuldt eller ej, ikke blev afspejlet i det materiale der blev indsamlet i forbindelse. Fx viste det sig i de cases vi så på, ikke at have betydning om læsevejlederen var ny i sin funktion.

Det viste virkningsevalueringen

Evalueringen af læsevejlederfunktionen viser at læsevejlederen kan være et effektivt tiltag til at forbedre elevernes læsefærdigheder, men at det afhænger af den måde læsevejlederen varetager sin funktion på. I de cases vi undersøgte, var læsevejledning fx virkningsfuldt når læsevejlederen indtog rollen som medpraktiker. Evalueringen peger også på en række andre kontekstforhold der er vigtige i forhold til hvordan læsevejledning fungerer.

Metodisk har evalueringen også givet en række erfaringer som andre der vil kaste sig ud i virkningsevaluering, forhåbentlig kan få glæde af. Erfaringerne er udkrystalliseret til tre gode råd samlet i boksen nedenfor.

En grundlæggende erfaring vi har gjort os, er at virkningsevaluering fungerer som en velegnet analyseform til at koble en indsats proces og resultater sammen. Men analyseformen leverer ikke hardcore-effekt viden om at læsevejledning i gennemsnit forbedrer elevernes læsekundskaber med et antal procentpoint, målt på en bestemt læsetest. Hvis det er den type viden man efterspørger, må man opsøge andre metoder (se boks 2).

Artiklen bygger på rapporten Viden der forandrer. Virkningsevaluering af læsevejlederen som fagligt fyrtårn der udkommer 21. januar 2009. Rapporten kan downloades fra www.eva.dk.

Boks 1

Tre gode råd til virkningsevaluaterer: Fokuser og simplificer

1. Fokuser problemstillingen

Vi startede evalueringen i Hvidovre Kommune med en ambition om at ville evaluere hele kommunens læseindsats på mellemtrinnet. Det viste sig imidlertid hurtigt som en vanskelig opgave. Dels fordi der var en række forskydninger i implementeringen af kommunens læseindsats, sådan at nogle af tiltagene var velimplementerede og på tidspunktet for evalueringen havde fungeret gennem længere tid, mens andre knap var kommet i gang. Og dels fordi arbejdet med at lave et grafisk overblik over årsags-virkningskæderne i den samlede læseindsats lynhurtigt førte til et uoverskueligt virvar af pile og kasser. Det var derfor et gennembrud for evalueringen da vi formulerede programteorien med een hovedide som omdrejningspunkt – nemlig ideen om faglige fyrtårne som indgang til at udvikle en praksis.

2. Arbejd med programteorier der giver mulighed for simplificering og overblik.

Man kan fremstille programteorien for en indsats på mange forskellige måder – der er ikke én rigtig måde at gøre på. Den programteori der er skildret her i artiklen, har haft mange forskellige skikkelser i løbet af evalueringen. Vi endte med at arbejde med to lag af programteorier. Et lag af programteorien lagde vægt på overblik og simplificering, mens et andet lag lagde vægt på at medtage så mange detaljer som muligt. De to lag af programteorien viste sig sammen som et godt redskab til at strukturere analysen af læsevejledningen.

3. Fokuser på outcomes

Skal man sige noget om effekter og virkninger, er pålidelige data om outcomes alfa og omega. Det er disse data man som virkningsevaluators skal bruge til at afgøre om indsatsen gør en forskel eller ej. Det er også nødvendigt at have et sammenligningsgrundlag for at kunne afgøre om de outcomes man ser, er en succes eller ej. Det kan være et før-efter-perspektiv eller et pålideligt sammenligningsgrundlag. I en del af litteraturen om virkningsevaluering lægges der meget vægt på arbejdet med at opstille virkningskæden. Men en erfaring vi tager med os, er at det at tænke i outcomes, og hvordan man måler dem, bør være et af de allerførste skridt i en virkningsevaluering. Som en bonus er det vores erfaring at data om outcomes også giver input til analyserne af virkningskæderne i indsatsen.

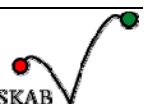
Boks 2

Virkningsevaluering og kvantitativ effektmåling – konkurrerende eller komplementære tilgange?

Debatten om virkningsevaluering har i Danmark været nært knyttet til Pawsons og Tilley's bog om Realistic evaluation fra 1997. Den ide som virkningsevaluering bygger på, at undersøge effekterne af en indsats gennem at opstille en programteori og teste hvordan den lader sig omsætte til virkeligheden, er imidlertid opstået i flere forskellige varianter rundt om i verden på næsten samme tidspunkt. Tilgange der minder meget om, eller næsten er det samme, er fx John Maynes ide om attribution analysis eller Huey-Tsyh Chens ide om Theory-Driven Outcome Evaluation. Det kan være nyttigt at gøre sig klart fordi Pawson og Tilley binder sig meget op på en videnskabsteoretisk retning, den kritiske realisme, der gør at de bruger mange kræfter på at polemisere imod den form for kvantitativ effektevaluering der fx kendes fra SFI Campbells arbejde.

For Mayne og Chen kræver det at arbejde med virkningsevaluering imidlertid ikke at man erklærer sig som kritisk realist, eller at man skal tage afstand fra kvantitativ effektmåling. Tværtimod ser de virkningsevaluering og kvantitativ effektmåling som komplementære tilgange med hver deres fordele og ulemper. De vil også typiske være anvendelige til forskellige typer af programmer. Lodtrækningsforsøg og matchingteknikker egner sig bedst til velafgrænsede, veldefinerede, "modne" programmer, mens virkningsevaluering måske især har sin styrke ved mere "vilde", nye, ikke skarpt afgrænsede programmer. På DES-konferencen 2008 sagde John Mayne i sit oplæg at han ikke så nogen modsætning mellem kvantitativ effektmåling og virkningsevaluering, men at de to tilgange kan supplere hinanden. For en praktisk arbejdende evaluator forekommer det som et konstruktivt synspunkt fremfor at skulle vælge videnskabsteoretisk og metodologisk skyttegrav.

DANIDA's Evalueringskontor er i øvrigt et eksempel på en dansk evaluator der allerede arbejder ud fra en både-og tankegang i forhold til kvantitative effektevaluering og virkningsevaluering, fx i evalueringen af aquakultur-projekter i Bangladesh.



Om Dansk EvalueringsSelskab

DES er åbent for alle med interesse for evaluering. Det kan eksempelvis være personer, som:

1. Forsker og underviser i evaluering
2. Er opdragsgiver til evalueringer
3. Gennemfører interne evalueringer
4. Gennemfører eksterne evalueringer
5. Andre personer med interesse for evaluering, eksempelvis studerende.

Formålet med DES er at være en platform for erfaringsudveksling og diskussion af evaluering-faglige problemstillinger, herunder:

1. Udviklingen inden for evalueringsteori- og metode
2. Nyttiggørelse af evalueringresultater
3. Faglig udvikling
4. Standarder for god evalueringspraksis.

Deltagelse i selskabet rummer følgende fordele

1. Adgang til at diskutere faglige problemstillinger med andre personer, som arbejder med området.
2. Mulighed for at deltage i arbejdsgrupper.
3. Deltagelse i evalueringskonferencer til en reduceret pris.
4. Gratis adgang til gå-hjem-møder. Oplæggene kan komme fra såvel foreningens medlemmer som fra eksterne oplægsholdere.
5. Gratis nyhedsbrev.

Medlemskab

Tegning af medlemskab sker ved at udfylde og indsende en tilmeldingsblanket, som kan findes på www.danskevalueringsselskab.dk. Medlemskabet gælder for resten af kalenderåret.

Medlemskabet kan både tegnes personligt eller som firma/institution. Prisen er kr. 825 pr. år for et personligt medlemskab og kr. 4.125 pr. år for et firma/institutionsmedlemskab. Spørgsmål kan rettes til selskabets sekretariat, Stina Andersen, Stina.Andersen@STAB.RM.DK

Bestyrelsen 2008-2009

<p>Peter Rhode (Formand) Center for Kvalitetsudvikling, Region Midtjylland Regionshuset Århus, Olof Palmes Allé 15 8200 Århus N Tlf.: 87 28 49 10 E-mail: peter.rhode@stab.rm.dk</p>	<p>Tanja Miller (Næstformand) Videncenter for Evaluering i Praksis, CEPRA, University College Nordjylland Sohngårdsholmsvej 51 A, 9000 Aalborg Tlf.: 70 23 50 03 E-mail: TNM@ucn.dk</p>
<p>Müzeyyen Boztoprak (Kasserer) Kontor for effekt og analyse, Arbejdstilsynet Landskronagade 33 2100 København Ø Tlf.: 72 20 87 15 E-mail: mboz@at.dk</p>	<p>Elisabeth Friis Fredskilde Sundhedscentret, Sundhedsområdet, Frederiksberg Kommune Peter Bangs vej 1 opgang G, 2000 Frederiksberg Tlf.: 38 38 45 09 E-mail: elfr01@frederiksberg.dk</p>
<p>Mads Teisen Teknik- og Miljøforvaltningen, Organisations- og Udviklingssekretariatet, Københavns Kommune Njalsgade 13 Postboks 454 1505 København V Tlf.: 33 66 58 60 E-mail: mateis@tmf.kk.dk</p>	<p>Ole Gregersen SFI, Det Nationale Forskningscenter for Velfærd Herluf Trolles Gade 11 1052 København K Tlf.: 33 48 08 70 E-mail: og@sfi.dk</p>
<p>Poul Skov Dahl NIRAS Konsulenterne Buchwaldsgade 35, 3 5000 Odense C Tlf.: 63 12 50 91 E-mail: psd@niraskon.dk</p>	<p>Runa Bjørn Center for Socialfaglig Udvikling Socialforvaltningen i Århus Kommune Værkmestergade 15, 8100 Århus C Tlf.: 89 40 13 31 E-mail: rub@aarhus.dk</p>
<p>Sven-Åge Westphalen Analyse & Data, Servicestyrelsen Skibhusvej 52B, 3. 5000 Odense C Tlf.: 72 42 37 50 E-mail: sw@servicestyrelsen.dk</p>	<p>Tim Jeppesen Det Kommunale og Regionale Evalueringsinstitut (KREVI) Olof Palmes Allé 19, 8200 Århus N Tlf.: 72 26 99 71 E-mail: tj@krevi.dk</p>
<p>Tue Vinter-Jørgensen Danmarks Evalueringsinstitut (EVA) Østbanegade 55, 3. sal 2100 København Ø Tlf.: 35 25 46 59 E-mail: tvi@eva.dk</p>	<p>Sekretariatet Dansk EvalueringsSelskab Att.: Stina Andersen c/o Center for Kvalitetsudvikling Olof Palmes Allé 15 8200 Århus N Tlf.: 87 28 49 11 E-mail: sekretariat@danskevalueringsselskab.dk</p>