

Bliver programmer bedre af at blive evalueret?^{1,2}

Evalueringer griber om sig i den offentlige sektor. Alligevel har vi kun en meget overordnet og begrænset viden om, hvad der kommer ud af at evaluere.



POUL THØIS MADSEN

Forskningscenter for evaluering (FCE)
Institut for Statskundskab,
Aalborg Universitet

I den offentlige sektor anvendes stadigt flere ressourcer på eksterne evalueringer. Men hvad kommer der ud af al den evaluering? Siden tresserne har såvel evaluatorene som de evaluerede stillet sig selv det spørgsmål. Så sent som 2009 betragtede 68 % af evaluatorene i en amerikansk spørgeskemaundersøgelse 'non-use' af evalueringer som et meget stort problem (Fleischer og Christie, 2009, s. 171). Det pessimistiske billede genfinder man også i kvalitative studier: »the informants gave a rather negative overall picture of evaluation use and none of them seemed to be very upset by this... the informants consider it as part of the game, something to be expected ...« (Petersén and Olsson, 2013, s. 11).

Inden for den gren af evalueringsforskningen, som går under betegnelsen 'evalueringsanvendelse', stiller man netop spørgsmål som: Anvendes evalueringens resultater? På hvilken måde? Hvilke barrierer er der for (yderlig) anvendelse? Og hvordan kan anvendelsen fremmes? Især i en international, primært amerikansk, sammenhæng har denne forskning vokset sig stor. Også så stor, at det kan være vanskeligt at bevare overblikket over, hvad vi egentlig 'ved' om evalueringsanvendelse.

Derfor gives der først et overblik over centrale evalueringsanvendelsesbegreber. Siden gives en kritisk analyse af central dansk og international empirisk forskning i evalueringsanvendelse. Som det tredje punkt gør jeg en kort status over, hvad vi så ved og ikke ved om evalueringsanvendelse.

Inden jeg gør det, vil jeg imidlertid forholde mig til to punkter rejst af revieweren. Der er i det følgende ikke en klar definition af, hvad en evaluering er. Her følger jeg Foss Hansen (2003, s. 186) ved at lade de centrale aktører – i mit tilfælde primært evalueringsforskere – bestemme, hvad evaluering er. Det, de betragter som en evaluering, er en evaluering.

Reviewer fremhæver endvidere, at der ikke i mit review indgår evalueringer, der benytter sig af randomiserede kontrollerede forsøg (RCT på engelsk – af eksempelvis den efterfølgende beskæftigelse af ledige med og uden et jobtræningsforløb). Det er imidlertid helt tilfældigt, at de reviewede studier ikke forholder sig til RCT. Samtidig er pointen i artiklen ikke den anvendte evalueringsmetode, men hvorvidt evalueringer anvendes eller ej. Spørgsmålet om RCT, der fylder meget i den generelle evalueringsdebat, er kun vigtigt i denne sammenhæng, hvis man kan påvise, at resultater fra RCT-evalueringer i højere grad finder anvendelse efterfølgende, end tilfældet er for andre typer af evalueringer. Det er der ikke apriori grund til at antage.

4 centrale evalueringsanvendelsesbegreber

På mange måder er store dele af evalueringsforskningen ganske anvendelsesorienteret og dermed også teoretisk lettere tyndbenet. De grundlæggende evalueringsanvendelsesbegreber er imidlertid reelt ikke blevet testet empirisk og er i en vis grad udledt deduktivt. Inden for den danske del af forskningen skelner man mellem hele 7 typer af anvendelse (jf. Dahler-Larsen og Larsen, 2001), medens man inden for den internationale evalueringsforskning oftest opererer med 4 typer

NOTE 1 Tak til den anonyme reviewer og min forskningsgruppe, FCE, for nyttige kommentarer.

NOTE 2 Ved et 'program' forstås evalueringsforskere den aktivitet eller organisation, der evalueres.

af evalueringsanvendelse. Den første og oprindelige forestilling om evalueringsanvendelse er den såkaldte *instrumentelle* anvendelse. Ideen er optimistisk, at man gennem en evaluering af et program udleder nogle anbefalinger, som de evaluerede derefter følger. Principielt er der fire måder, som man kan gøre det på: »Close down or modify ineffective programs, and institutionalize and expand programs that were working well« (Weiss et al, 2005, s. 13).

Den tidlige empiriske anvendelsesforskning fandt imidlertid kun sjældent eksempler på instrumentel anvendelse (jf. også Weiss, 1967). Derfor begyndte evalueringsforskere at lede efter andre typer af anvendelse – også fordi man havde svært ved at forestille sig, at der ikke fandt en eller anden form for anvendelse sted. Herved udvikledes gradvist begrebet *begrebsmæssig* anvendelse. Heri lå primært, at evalueringer ændrede på den måde, hvorpå ansvarlige aktører forstod de evaluerede programmer. Det kunne i princippet føre til hurtige 'nu og her'-ændringer, men det ligger implicit i begrebet, at der først på længere sigt udvikles mere konkrete og synlige forandringer (Petersén og Olsson, 2013, s.3-4).

Den tredje type af anvendelse er den *symbolske* eller *politiske*, hvor det egentlig ikke handler om at evaluere et program med henblik på forandring af det, men om at man gennem selve det at evaluere ønsker at overbevise andre om, at det her område er værd at satse på – også inden evalueringen foreligger. At evaluere kan også være en måde at retfærdiggøre en indsats på: at den så sandelig bliver evalueret. Ved denne type af anvendelse kan der skabes koalitioner, og der kan gennemføres programmer, som det ellers ville have været vanskeligt at opnå politisk accept af. Det kan lidt paradoksalt i princippet have langt større konsekvenser end andre anvendelsesformer.

Gennem og under empiriske studier af anvendelse opdagede man allerede tidligt (eksempelvis Dickey, 1980) en fjerde type af anvendelse, der ikke passede ind i den umiddelbare forestillingsverden, som senere fik navnet *procesanvendelse* af Patton og derfor i høj grad forbindes med hans navn (1997, 1998). Ideen er, at i det øjeblik man begynder at evaluere et program, så påvirkes det. Man kan således empirisk konstatere, at der er i løbet af evalueringsprocessen sker ændringer i adfærd på individ-, program- eller organisationsniveau som følge af evalueringen. Denne vigtige og måske endda i praksis dominerende anvendelsesform ignoreres ofte (eksempelvis i et så nyt studium som Petersén og Olsson, 2013)

I en dansk sammenhæng har anvendelse også blevet tildelt en betydelig opmærksomhed, dog primært teoretisk. Som nævnt blev hele syv anvendelsesformer præsenteret i 2001 af Dahler-Larsen og Larsen: Hhv. instrumentel kontrol og læring, oplysende/debatskabende/værdi-afdækkende, strategisk, taktisk, symbolsk og konstitutiv. Dahler-Larsen har i høj grad markedsført sidstnævnte anvendelsesform – senest også internationalt (2012, 2013).

Frem for også at præsentere denne lange række af begreber (for en uddybning se Dahler-Larsen og Larsen), der i vid udstrækning overlapper og/eller disaggregerer de fire nævnte begreber, så vil jeg kort præsentere endnu en amerikansk begrebsudvikling, som indebærer en problematisering af hele anvendelsesbegrebet. I 2000 foreslog Kirkhart – med en vis gennemslagskraft – at man skiftede ordet 'anvendelse' ud med ordet 'indflydelse'. Ordet anvendelse har en instrumentel og også ret kortsigtet klang og indbygget præcision. Hvorimod indflydelse antyder et noget længere tidsperspektiv og noget mindre præcision. Centrale evalueringsforskere (som Mark og Henry, 2004) fulgte Kirkharts råd, men de gængse begreber er stadig 'use', 'utilization' og på dansk 'anvendelse'.

Normalt introducerer evalueringsanvendelsesartikler ikke flere begreber, men hvad så med følgende helt centrale spørgsmål – også for denne artikel: Hvad kommer der ud af, at evalueringen bliver anvendt? Hvad er de vigtigste positive og negative konsekvenser? Det spørgsmål har evalueringslitteraturen også søgt at belyse. Kirkhart (s. 8) taler om de umiddelbare, 'end-of-cycle' og langsigtede 'resultater'. I andre dele af litteraturen taler man om 'impact' eller 'outcome' (i eksemplvis Patton et al, 1975, hhv. s. 10 og 30), mens man på dansk ofte vil tale om 'virkning'. I det følgende anvender jeg på skift disse 'konsekvensbegreber' alt efter, hvordan det passer ind i sammenhængen.

Et ubesvaret spørgsmål er, hvor de centrale anvendelsesbegreber kommer fra. De sammenfatter ikke kun et empirisk mønster, men er også i høj grad et produkt af en blanding af erfaringsbaseret og deduktiv tænkning. Og så de vækker de tydeligvis genklang – såvel hos dem, der evaluerer, som dem, der bliver evalueret. Men også begreber, der fremstår som selvfølgelige, bør underkastes en empirisk granskning. Det sker i det følgende.

Hvad siger den danske empiri om evalueringsanvendelse og 'outcome'?

Der er en stærkt begrænset dansk empiri om evalueringsanvendelse, og selv dér, hvor man kunne forvente en behandling af spørgsmålet, smutter det mellem hænderne på såvel læser som forsker – til fordel for andre typer af evalueringsspørgsmål. I det følgende analyserer jeg fire empiriske danske evalueringstudier, der også belyser spørgsmålet om anvendelse.

Dahler-Larsen (2002) bygger på en spørgeskemaundersøgelse stilet til 208 socialchefer i amt og kommuner med en svarprocent på henh. 93 og 70 %. Ud fra den seneste evaluering af det område, som de selv er ansvarlige for, bliver de bl.a. spurgt om: 'Hvilke følger havde evalueringsforløbet for den evaluerede indsats?' Hertil svarede de 7 % 'ingen nævneværdig'; 37 %, at evalueringen begrundede en fortsættelse; 45 %, at den førte til mindre justeringer; 23 % større omlægninger, og mindre end 1 % førte til en nedlæggelse af indsatsen (Dahler-Larsen, 2002).

Den anden undersøgelse, som mere gennemført har et anvendelsesorienteret perspektiv, er Moldt-Jørgensens (2007) Ph.d.-afhandling. Her analyseres det, hvordan evaluering af undervisningen påvirker såvel underviseres arbejde som de studerendes

læring. Konklusionen er, at den måde, hvorpå evalueringer anvendes de undersøgte steder, gør underviserne angst, men at det ikke nødvendigvis skyldes evalueringerne, men snarere ledelse eller mangel på samme.

Derudover er det interessant, hvor usynlig evalueringsanvendelsen er i større oversigtsstudier af evalueringsaktiviteter inden for den offentlige sektor i Danmark. I Dahler-Larsens (2003) arbejde for magtudredningen analyseres otte cases (s. 54-77). Det er imidlertid vanskeligt at identificere selve evalueringen i flere af casene. Hvad angår anvendelse, så fremdrages den ikke eksplicit, selv om anvendelsen vies betydelig opmærksomhed teoretisk (hele 17 sider: s. 23-40).

En tæt analyse af casene peger dog på følgende typer af eksempler på anvendelse:

- Instrumentel: a) De internationale læseundersøgelser negative resultater har ført til 'en styrkelse af danskfagets timetal', 'ansvaret for læsning... en integreret del af... flere fag... der sættes penge af til læseforskning (s. 66). b) Evalueringen af håndarbejdsuddannelsen førte direkte til lukning af seminarer (s. 71-74)
- Begrebsmæssig: 'gode diskussioner' (s. 69), 'debatskabende' (s. 71)
- Politisk: 'vi er nødt til at fortælle og dokumentere, hvor dygtige vi er' (s.54)
- Symbolsk: 'det er showpræget, hvad der foregår' (s. 57).

Men bemærk at jeg selv i så relativt detaljerede cases ikke har kunnet finde ud af 'outcome'. Blev eleverne så også bedre til at læse som direkte følge af evalueringen? Hvad kom der ud af lukningen af seminarerne? Var det den rigtige beslutning? Hvad kom der ud af de nævnte gode diskussioner og debatter, som brugere typisk betoner?

Endelig kan det konstateres, at den mest omfattende og meget indsigtsfulde oversigt over danske statslige evalueringer (Foss Hansen, 2003) alene behandler *intenderet* evalueringsanvendelse (undtagen en kort gennemgang af Dahler-Larsens 2002 resultater, jf. Foss Hansen, 2003, s. 65-66).

Den internationale empiri om anvendelse og 'outcome'

Med udgangspunkt i Brandon og Singh (2009) kritiske review af 5 reviewstudier af evalueringsanvendelse har jeg identificeret de 6 første artikler i tabel 1. De forholder sig direkte til spørgsmålet om anvendelse og i en vis udstrækning outcome. Dernæst har jeg systematisk gennemgået centrale evalueringstidsskrifter (*Evaluation Review*, *New Directions for Evaluation*, *Evaluation* og *American Journal of Evaluation*) for at finde nyere bidrag – især med henblik på procesmæssig brug, der som logisk konsekvens af undersøgelsesperioden stort set ikke dækkes af Brandon og Singh. Det førte til en identifikation og analyse af yderligere 7 bidrag (jf. tabel 1), hvoraf 5 primært omhandler den procesmæssige anvendelse af evalueringer. Endelig er de to nyeste artikler, Peck og Gorzalski (2009) og Petersén og Olsson (2013), fundet ved at søge på google.scholar.

Tabel 1 dækker således 15 relevante studier, men giver sig ikke ud for at være udtømmende. Den skulle gerne være selvforklarende på nær kolonnen til højre, der angiver 'type af anvendelse'. Den er udfyldt overvejende ud fra, hvad der er angivet direkte i artiklen, men i flere tilfælde er der tale om min fortolkning. 'Imposed use' er et begreb, der alene behandles i Weiss et al. (2005).

Det kronologisk set første studium er Patton et al (1975), som jeg redegør mere detaljeret for end de øvrige studier, da det i høj grad illustrerer, hvad anvendelsen egentlig er og samtidig slås med de samme metodeproblemer, som vi gør i dag. Med udgangspunkt i 20 evalueringer inden for sundhedsområdet blev såvel evaluatorene som evaluanter interviewet personligt. Evaluatorteamet opgav ganske illustrativt at operere med en forhåndsdefinition på anvendelse, men overlod det til de evaluerede. Det gjorde forskerteamet ved at stille følgende spørgsmål om den indflydelse ('impact') evalueringer havde: 'on program operation, on planning, on funding, on policy, on decisions, on thinking about the program, and so forth ... [men også på den] general thinking on issues that arise from a study or position papers, or legislation' (s. 10).

Med den valgte spørgsmålsformulering spørges der faktisk direkte til instrumentel anvendelse, men alligevel giver informanternes besvarelser snarere mindelser om begrebsmæssig brug (uden at de var introduceret til termen, der ikke fandtes i litteraturen i starten af halvfjerdsere). De typiske konsekvenser af en evaluering viste sig at være, at den 'provided additional pieces of information in the difficult puzzle of program action permitting some reduction in uncertainty' (s. 11), eller at den indgik som 'an additional input into an on-going, evolutionary process of program action' (s. 12). En beslutningstager gav udtryk for, at evalueringen 'resolved a lot of doubt and confusions and misunderstandings' (s. 12). Andre formuleringer er, at den 'confirmed our impressions' (s. 14) eller gav dem en 'new awareness' (s. 14).

I tabel 1 gives der en oversigt over Pattons og fjorten andre empiriske studier af evalueringsanvendelse. Som det fremgår af spalten til højre, så peger de empiriske studier især på, at der finder instrumentel, begrebsmæssig og procesmæssig anvendelse sted. Af tabellen fremgår også to gange politisk/symbolsk anvendelse, men en nærmere analyse viser, at det empiriske belæg er særdeles begrænset.

Hvad udtrykker de femten studier så om de tre dominerende former for anvendelse? At evalueringer har ført til konkrete beslutninger eller en dybere forståelse af de evaluerede programmer – forårsaget af evalueringens resultater og eventuelle anbefalinger og/eller af evalueringsprocessen. Medens det mellem linjerne fremgår, at det forventes, at anvendelsen også fører til ændret outcome.

Et andet vigtigt resultat, der allerede trådte klart frem i Patton m.fl. (1975), er, at slutbrugerne af evalueringer forventer langt mindre af en evaluering, end evaluatorerne gør. Hvis bare

Tabel 1: Oversigt over 15 centrale empiriske anvendelsesstudier i kronologisk rækkefølge fra 1975 til 2013

Kilde	Sektor	Genstandsfelt	Metode	Informant/ respondent	Beskrivelse af anvendelse	Type af anvendelse
Patton et al (1975)	Sundhed	20 evalueringer	Kvalitative interviews	Evaluators og beslutningstagere	(Få) Beslutninger Information	Begrebsmæssig
Alkin/Daillak (1979)	Undervisning	5 evalueringer	Kvalitative interviews	Lokale program- medarbejdere	Gradvis påvirkning af beslutninger	Mellemting mellem instrumentel og begrebsmæssig
Dickey (1980)	Undervisning	47 evalueringer	Spørgeskema Kvalitative interviews	Projektledere	Ændre og bekræfte praksis Evalueringsprocessen	Procesanvendelse
Leviton/Beruch (1983)	Undervisning	21 evalueringer	Kvalitative interviews	Evaluators og opdragsgivere	Lovændringer	Instrumentel
Forss et al (1994)	U-landshjælp	11 (8) evalueringer	Spørgeskema Kvalitative interviews	Key staff	Læring	Begrebsmæssig
Presscill/Cara-celli (1997)	Mange	257 evaluators	Spørgeskema	Evaluators	Forbedre program Information	Instrumentel Begrebsmæssig Procesanvendelse
Forss et al (2002)	U-landshjælp og museer	To cases (evalueringer)	Baseret på egne erfaringer	Egne erfaringer og evalueringer	Nye kurser Programtilpasning Ændrede beslutninger	Procesanvendelse
Presskill et al (2003)	Sundhed	Én case (to evalueringer)	Kvalitative interviews	Advisory board medlemmer	Learning (om evaluering og om programmet)	Procesanvendelse
Ginsburg og Rhett (2003)	Undervisning	Forskellige evalueringer metoder	Review	Forskelligt fra studium til studium	Beslutninger	Instrumentel
Weiss et al (2005)	Narkotikabekæmpelse	4 distrikter/byer i 4 stater	Kvalitative interviews	Skolepersonale, beslutningstagere og politi	Overtalelse Fortsæt/luk Bevidsthedspåvirkning	Instrumentel Begrebsmæssig Politisk 'Imposed use'
Harnar og Preskill (2007)	Mange	1140 evaluators	Spørgeskema	Evaluators	Stakeholder involvement, Processen som outcome	Procesanvendelse
Podems (2007)	U-landshjælp	Én case	Kvalitative interviews	Tre direktører og 7 'beneficiaries'	Fælles forståelse Klargørelse af mål Øget viden om program Empowerment	Procesanvendelse
Amo og Cousin (2007)	Mange	Sammenfatning af 18 studier	Review af primært kvalitative studier	Forskelligt fra studium til studium	Learning Handling Holdninger 'Andet'	Procesanvendelse
Peck/Gorzalski (2009)	Undervisning	16 evalueringer	Kvalitative interviews	16 nøgleinformanter	Begrebsmæssig Gennemført anbefalinger	Begrebsmæssig Instrumentel
Fleischer/Christie (2009)	Mange	1140 evaluators	Spørgeskema	Evaluators	Learning Ændring af strategi	Instrumentel Begrebsmæssig
Petersén og Olsson (2013)	Socialt arbejde	4 Evalueringer	11 kvalitative interviews	Evaluators og opdragsgivere	Rituel kontrolfunktion Anvendes af evaluators	Begrebsmæssig Rituel Symbolsk

evaluater *oplever*, at en evaluering er nyttig, så er de tilfredse, hvorimod evaluators gerne vil se mere synlige resultater, før end de betragter en evaluering som vellykket.

Et tredje vigtigt gennemgående resultat er, at anvendelse på en måde fremtræder som et formål i sig selv (hvad man kan betegne som 'det at anvende'), hvorimod outcome er underbeholdt. Af tabellen fremgår dog følgende eksempler på konsekvenser: Beslutninger, programtilpasning eller forbedring, lovændringer, et nyudviklet kursus, læring, ændring af strategi eller holdninger og endelig gennemførelse af anbefalinger.

dringer, et nyudviklet kursus, læring, ændring af strategi eller holdninger og endelig gennemførelse af anbefalinger.

Sådanne konsekvenser fører givetvis også til ændringer i outcome. Men var der tale om væsentlige ændringer? Kom programmerne også til at fungere bedre – eller værre? Når der eksempelvis tages en beslutning eller gennemføres et lovforslag, så betragtes det implicit som noget virksomt, og som noget, der fører til reelle forandringer. Men det kan man ret

beset ikke vide noget om. Det kendetegner således de femten empiriske studier, at det er aldeles uklart, hvad der i sidste instans egentlig kommer ud af at evaluere. Den generelle konklusion er, at evalueringer anvendes og ofte fører til adfærds- og programændringer, men at vi ikke ved, om det fører til bedre eller værre programmer.

Hvad ved vi så om evalueringsanvendelse?

Hvis man har erfaringer med evalueringsanvendelse, vil man måske studse over artiklens hovedkonklusion. Vi ved da en masse om anvendelse, men afviser denne artikel reelt ikke den viden? Svarer er nej. Der er en klar erkendelse af, at evalueringsanvendelse faktisk finder sted, og at det fører til adfærdsændringer og også ofte programændringer. Kritikken går på, at vi aldrig rigtig får noget at vide om karakteren af og dybden af de programændringer. Fungerer undervisningen nu bedre? Har organisationen fået løst nogen centrale problemer som følge af en evaluering? Hvad har evalueringen betydet for opgaveløsningen? Hvad er outcome?

Ordet 'anvendelse' svarer derfor ganske præcist til det, der rent faktisk måles på i studierne. Begrebet udtrykker, at der er sket eller menes at være sket et eller andet på inputsiden inklusive implementering, men man har højest en helt overordnet viden om, hvad det har ført til på outputsiden.

Hvorfor har outcome ikke været viet større opmærksomhed i evalueringsforskningen? Der er formentlig flere forklaringer. Dels har man har rettet blikket mod, hvor der er lyst, dels er der store metodiske problemer med at afdække hvilke forandringer, der helt eller delvist følger af en evaluering. Den mest sandsynlige forklaring er imidlertid, at evalueringer ofte slet ikke forventes eller har til hensigt at føre til konkrete resultater. De gennemføres primært som led i et ritual, som Dahler-Larsen har fremført det ved flere lejligheder (eksempelvis 1998), og som det også er eftervist empirisk i Petersén og Olsson (2013).

En konsekvens af disse konklusioner er, at vi må forholde os kritisk til den viden, vi mener at have om evalueringsanven-

delse. Som det fremgik ovenfor, så postulerer definitioner af de forskellige typer af anvendelse, at der også er en impact eller outcome af evalueringer, når der sker en anvendelse. Det er der givetvis også ofte – vi ved bare ikke ret meget om det. Vores begreber for evalueringsanvendelse fremhæver, hvad vi gerne vil se, men som vi kun i begrænset omfang har været i stand til at eftervise empirisk.

Konklusion

En væsentlig del af evalueringsforskningen omhandler spørgsmålet om, hvorvidt og hvordan evalueringer anvendes. Den begrebsmæssige og empiriske forskning peger på, at evalueringer:

- anvendes – på en eller anden vis
- ofte fører til ændringer i adfærd

De fakta indikerer også, at evalueringer *kan* bidrage til at skabe væsentlige programforandringer – forbedringer såvel som forværringer. Men hvorvidt det er rigtigt eller ej – dvs. spørgsmålet om 'impact', 'outcome', 'resultater' og 'virkninger' af evalueringer – er stort set ikke undersøgt empirisk.

Mange har peget på, at programforbedringer næppe heller er den reelle hensigt med mange evalueringer. I virkeligheden så handler evalueringer mere om kontrol, politik, symboler og ritualer. En af talsmændene for det synspunkt, Dahler-Larsen (1998), fremfører imidlertid alligevel, at selv 'ritualiserede evalueringer ... kan have meget virkelige effekter' (s. 163). Og det er de effekter, vi ved alt for lidt om og i højere grad bør prøve at afdække i fremtidige empiriske studier. Ikke mindst ud fra et hensyn til de skatteborgere, der bidrager til den millionindustri, som evalueringsbranchen er i dag.

Den viden, der er præsenteret i denne artikel, kan evaluerede anvende (!) til 'empowerment'. De har en klar anledning til at spørge fremtidige evaluatore: »Hvad forventer I, at der kommer ud af at evaluere os? Hvordan vil I sikre, at det også sker i praksis? Er det noget, der efterfølgende vil blive målt på?« Hermed kan de evaluerede blive kritiske brugere af egne evalueringer frem for bare at være passive 'ofre'.

LITTERATURLISTE

- Alkin, M.C., & Daillak, R.H. (1979). A study of evaluation utilization. *Educational Evaluation and Policy Analysis*, 1(4), 41-49.
- Amo, C., & Cousins, J.B. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. *New Directions for Evaluation*, 2007 (116), 5-26.
- Brandon, P.R., & Singh, J.M. (2009). The strength of the methodological warrants for the findings of research on program evaluation use. *American Journal of Evaluation*, 30 (2), 123-157.
- Dahler-Larsen, P. (1998). *Den rituelle refleksion: om evaluering i organisationer*. Odense: Odense Universitetsforlag.
- Dahler-Larsen, P. (2002). *Evaluering kortlagt: en undersøgelse af amters og kommuners evalueringspraksis på det sociale område*. Århus: Systime
- Dahler-Larsen, P. (2004). *Evaluering og Magt*. Magtudredningen. Aarhus: Aarhus Universitetsforlag.
- Dahler-Larsen, P. (2012). Constitutive effects as a social accomplishment: A qualitative study of the political in testing. *Education Inquiry*, 3 (2), juni, 171-186
- Dahler-Larsen, P. (2013). Constitutive Effects of Performance Indicators: Getting Beyond Unintended Consequences. *Public Management Review*, 3 (2), 1-18.
- Dahler-Larsen, P. og F. Larsen (2001). Anvendelse af evaluering – Historien om et begreb, der udvider sig. *Tendenser i evaluering*. Red. P. Dahler-Larsen og H.K. Krogstrup. Odense: Odense Universitetsforlag.
- Dickey, B. (1980). Utilization of evaluations of small-scale innovative educational projects. *Educational Evaluation and Policy Analysis*, 2 (6), 65-77.
- Fleischer, D.N., & Christie, C.A. (2009). Evaluation Use Results From a Survey of US American Evaluation Association Members. *American Journal of Evaluation*, 30 (2), 158-175.
- Forss, K., Cracknell, B., & Samsø, K. (1994). Can evaluation help an organization to learn? *Evaluation review*, 18 (5), 574-591.
- Forss, K., Rebién, C.C., & Carlsson, J. (2002). Process Use of Evaluations Types of Use that Precede Lessons Learned and Feedback. *Evaluation*, 8 (1), 29-45.
- Foss Hansen, H. (2003). *Evaluering i Staten*, Samfundsfagslitteratur.
- Ginsburg, A., & Rhett, N. (2003). Building a better body of evidence: New opportunities to strengthen evaluation utilization. *American Journal of Evaluation*, 24 (4), 489-498.
- Harnar, M.A., & Preskill, H. (2007). Evaluators' descriptions of process use: An exploratory study. *New Directions for Evaluation*, (116), 27-44.
- Kirkhart, K.E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New directions for evaluation*, (88), 5-23.
- Leviton, L. og R. Beruch (1983). Contributions of evaluation to education programs and policy. *Evaluation Review*, 7 (5), 563-598.
- Mark, M.M., & Henry, G.T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10 (1), 35-57.
- Moldt-Jørgensen, C. (2007). *Fra meningsløs til meningsfuld evaluering: Anvendelsen af studentilfredshedsmålinger på de korte og mellem lange videregående uddannelser set fra psykodynamisk systemperspektiv*, Ph.d-serie ved Københavns Handelshøjskole.
- Patton, M.Q. et. al. (1975). *In Search of Impact: An Analysis of the Utilization of Federal Health Evaluation Research*. Minnesota: Minnesota Center for Social Research, University of Minnesota.
- Patton, M.Q. (1997). *Utilization-focused evaluation: The new century text*. Sage.
- Patton, M.Q. (1998). Discovering process use. *Evaluation*, 4 (2), 225-233.
- Peck, L.R., & Gorzalski, L.M. (2009). An evaluation use framework and empirical assessment. *Journal of MultiDisciplinary Evaluation*, 6 (12), 139-156.
- Petersén, A.C., & Olsson, J.I. (2013). An evaluation paradox in social work? An empirical study of evaluation use in connection with temporary programmes in Swedish social work. *European Journal of Social Work*, 1-17.
- Podems, D. (2007). Process use: A case narrative from Southern Africa. *New Directions for Evaluation*, (116), 87-97.
- Preskill, H., & Caracelli, V. (1997). Current and developing conceptions of use: Evaluation use TIG survey results. *American Journal of Evaluation*, 18 (3), 209-225.
- Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation*, 24 (4), 423-442.
- Weiss, Carol (1967). Utilization of evaluation: Toward comparative study. *The use of social research in federal domestic programs, Part III*. Washington DC: Government Printing Office
- Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence how evaluations affect DARE. *American Journal of Evaluation*, 26 (1), 12-30.