

NOTAT 5

ERFARINGER

FRA

EVALUERINGSLITTERATUREN

SET I FORHOLD TIL

MÅL 2-EVALUERINGERNE



INDHOLD

<u>1</u>	<u>INDLEDNING</u>	<u>3</u>
<u>2</u>	<u>MICHEAL QUINN PATTON: TILTÆNKET ANVENDELSE FOR TILTÆNKTE BRUGERE</u>	<u>3</u>
2.1	PATTONS RELEVANS FOR MÅL 2-EVALUERING	4
<u>3</u>	<u>CAROL H. WEISS: OPLYSENDE EVALUERING OG PROGRAMMETS KARAKTER</u>	<u>5</u>
3.1	WEISS' RELEVANS FOR MÅL 2-EVALUERING	5
<u>4</u>	<u>MICHAEL SCRIVEN: EVALUERING SOM VURDERINGENS VIDENSKAB</u>	<u>6</u>
4.1	SCRIVENS RELEVANS FOR MÅL 2-EVALUERING	6
<u>5</u>	<u>VEDUNG: EVALUERING OG DEN PARLAMENTARISKE STYRINGSKÆDE</u>	<u>7</u>
5.1	VEDUNGS RELEVANS I FORHOLD TIL MÅL 2-EVALUERING	7
<u>6</u>	<u>PETER DAHLER-LARSEN: EVALUERINGER OG NY-INSTITUTIONEL TEORI</u>	<u>9</u>
6.1	DAHLER-LARSENS RELEVANS FOR MÅL 2-EVALUERINGERNE	9
<u>7</u>	<u>OPSUMMERING</u>	<u>10</u>

1 Indledning

Ideen med dette notat er at redegøre for 4-5 vigtige strømninger i den internationale evalueringslitteratur og sætte dem i relation til evalueringsdesignet for mål 2-evalueringerne. Notatet ser altså på de principielle muligheder og valg, der ligger ved planlægningen af en evaluering som den for mål 2. Dette er en diskussion, der kun i begrænset omfang har været fremme, da rammerne for mål 2-evalueringerne i store træk har været fastlagt fra EU-kommissionen. Formålet med notatet er altså at belyse og tydeliggøre nogle af de vigtige valg, der foretages, når man bestemmer sig for en bestemt evalueringsmodel.

Det er klart, at dette notat kun kan fremdrage enkelte elementer fra evalueringsteorien, men de er væsentlige både i den forstand, at de har spillet en rolle i udviklingen af evalueringsteorien, og ved at de kan belyse vigtige aspekter ved valg af evalueringsdesign. Derfor vil de også rejse nogle relevante spørgsmål i forhold til den konkrete evaluering af mål 2-programmet. Notatet indeholder ikke en nærmere analyse af evalueringsmodellen for mål 2, men det skitseres, hvilke vigtige valg og fravalg, der er foretaget ved at benytte den form for evaluering, der er praktiseret i forhold til mål 2-programmet.

De vigtige strømninger præsenteres i dette notat ved en af de teoretikere, der har spillet en væsentlig rolle i udviklingen eller udbredelsen af de teoretiske ideer. Det yder naturligvis ikke retfærdighed overfor de udvalgte teoretikers værk at lade dem stå som eksponenter for (bare) et enkelt teoretisk element. På den anden side er en det en instruktiv måde at beskrive teoretiske positioner på. Normalt er evalueringsfeltet domineret af amerikanske forskere. Jeg har ladet mit udvalg få lidt skandinavisk bias ved at inddrage en svensk og en dansk forsker sammen med tre amerikanere, der har haft en dominerende position i evalueringsforskningen de seneste 30 år.

2 Micheal Quinn Patton: Tiltænkt anvendelse for tiltænkte brugere

I sin indflydelsesrige bog "Utilization-Focused Evaluation" giver Patton følgende definition på evaluering:

"Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming. Utilization-focused program evaluation (as opposed to program evaluation in general) is evaluation done for and with specific, intended primary users for specific, intended uses." (Patton, 1997, s.23).

Patton understreger anvendelsen af evalueringer som afgørende. Derfor mener han også, at det er for snævert, hvis man kun definerer evaluering på grundlag af, om et programs målsætninger er blevet mødt. Det vigtige for Patton er, at evalueringer bliver brugt: "Intended use by intended users" (tiltænkt anvendelse af tiltænkte brugere), som det fremhæves adskillige gange (fx ibid. s. 20).

"Utilization-focused evaluation does not advocate any particular evaluation content, model, method, theory or even use. Rather, it is a process for helping primary intended users select the most appropriate content, model, methods, theory, and uses for their particular situation" (Ibid. s. 22).

Selvom om Patton understreger, at evaluator skal basere sig på de professionelle normer for evaluering (systematik, kompetence, ærlighed, respekt mv.), så er det, som det fremgår af citatet ovenfor, anvendelsen af evalueringen, der er helt afgørende. Evaluator skal altså bruge sin ekspertise og kreativitet til at evaluere på en forsvarlig måde, men så anvendelsen er sikret.

Derfor er det også afgørende at afklare, hvad "intended use" er, og hvem de "intended users" er. Evalueringer kan som bekendt have mange anvendelser og funktioner, men Patton fokuserer på den tiltænkte anvendelse og de tiltænkte brugere.

2.1 Pattons relevans for mål 2-evaluering

For alle evalueringer er det fornuftigt at overveje, hvem der skal bruge evalueringen og hvordan. Det er en almindelig bekymring blandt evaluatorene og opdragsgivere, at evalueringer blot skal stå og samle støv. Det, som er kritisk i forhold til andre evalueringstilgange, er, hvordan man definere de "tiltænkte brugere" og den "tiltænkte anvendelse".

For et program som mål 2-programmet, der i sig selv er komplekst og indgår i et komplekst politisk beslutningssystem, er det spørgsmålet, i hvilken grad det er muligt (praktisk og politisk) at afgrænse både den tiltænkte anvendelse og de tiltænkte brugere af evalueringen.

Man kan enten vælge at sige, at det reelt kun er en begrænset gruppe af policyansvarlige inden for strukturfondspolitikken, der interesserer sig for implementeringen af mål 2-programmet. Deres informationsbehov er derfor tæt knyttet til implementeringen af programmet og deres tiltænkte instrumentelle brug af evalueringen forstået sådan, at evalueringen skal hjælpe til at forbedre indsatsen og (forhåbentlig) legitimere den hidtidige indsats. Da denne gruppe er relativt homogen i interesser og forståelse af programmets struktur og funktion, vil det være muligt at lave en skræddersyet evaluering til målgruppens behov og deres tiltænkte brug.

Da målgruppen primært består af embedsmænd involveret på forskellige niveauer i strukturfondspolitikken, er spørgsmålet for det første, om det er politisk holdbart at lave en evaluering med så snæver en målgruppe og begrænset anvendelsesområde. Hvis målgruppen udvides til organisationernes repræsentanter og politikerne i de forskellige beslutningsorganer inden for strukturfondspolitikken, kan det hævdes, at alle de officielt udpegede interesser indgår i evalueringernes målgruppe. Men bliver målgruppen hermed udvidet med lægmænd i forhold til strukturpolitikens administration, kan der desuden stilles spørgsmål ved, om alle relevante interesser er repræsenteret i de forskellige beslutningsorganer inden for strukturfondspolitikken. Det første problem betyder, at evalueringen for at tilfredsstille hele målgruppen må gøres mindre instrumentel i forhold til den tekniske implementering af programmet og mere betone de politiske og strategiske aspekter ved gennemførelsen af programmet. Det andet problem er, at programmets kompleksitet og indgriben i andre politikområder gør, at andre, end de der er repræsenteret i beslutningssystemet, har en legitim interesse i evalueringen. Disse problemer sættes på spidsen, når det gælder regional strategi, hvor der er en bred interesse for udviklingsinitiativer.

Opsummerende kan det siges, at Patton sætter fingeren på et vigtigt (ømt) punkt med hensyn til evalueringer: Bliver de brugt til det, de er tiltænkt, af dem de er tiltænkt? Det er i høj grad et relevant spørgsmål for mål 2-evalueringerne. Samtidig rejser det første spørgsmål så automatisk spørgsmålet, hvem der er de tiltænkte brugere af evalueringen, og hvad der er den tiltænkte brug af evalueringen? Det er mindst ligeså relevant at afklare i forbindelse med evalueringen af mål 2, som skal have en stærk regional forankring.

3 Carol H. Weiss: Oplysende evaluering og programmets karakter

To væsentlige synspunkter fra Carol Weiss' forfatterskab vil blive fremhævet i denne sammenhæng: **For det første** at evalueringens anvendelse skal ses i et længere og bredere perspektiv end den umiddelbare anvendelse, som Patton repræsenterer. Den viden evalueringer tilvejebringer, vil ifølge Weiss have en oplysende effekt på længere sigt og sjældent en umiddelbar instrumentel anvendelse. **For det andet** fremhæver Weiss, at programmer fungerer efter andet end rationelle præmisser. Man må derfor inddrage organisations- og beslutningsteoriens indsigter for at kunne forstå, på hvilke præmisser komplicerede programmer skal evalueres.

Den holdning og indsigt, der lægger vægt på evaluering som oplysning (enlightment) er primært baseret på erfaringer fra den føderale regeringsadministration i USA. Weiss fremhæver en række grunde til, at evalueringer får indirekte indflydelse frem for at blive anvendt direkte. Blandt andet at evalueringer kan true vigtige interesser, og politiske aktører og embedsmænd bruger information til mange formål og de er underlagt en række bindinger (Shadish et al., 1991, s.54f).

Weiss er naturligvis ikke den eneste, der har fremhævet, at det giver store problemer at anskue programmer ud fra en antagelse om at de fungerer rationelt. Det der gør hendes bidrag særligt relevant i denne sammenhæng, er at hun har udviklet sine bidrag til evalueringsteorien gennem sit arbejde i komplekse policysystemer.

3.1 Weiss' relevans for mål 2-evaluering

Selvom der ikke direkte kan trækkes paralleller fra Carol Weiss' oprindelige erfaringsgrundlag til mål 2-programmet, så er hendes erkendelse af, at mange evalueringer primært har en oplysende effekt, en påmindelse om, at evalueringer af komplekse programmer som fx mål 2-evalueringerne skal skrives med henblik på også at have en langsigtet effekt og ikke kun umiddelbar anvendelse. Det udspringer altså af en beskrivelse af, hvordan mange evalueringer får en effekt, men det er nærliggende at stramme indsigten til en normativ anbefaling, der går på, at så skal en evaluering i hvert fald i det omfang og betydning som mål 2-evalueringen, indeholde analyser, der giver mulighed for dybere forståelse af, hvordan programmet virker.

Når man ser på mål 2-programmet, virker det oplagt, at det må være en væsentlig faktor i en evaluering at overveje, hvordan man kan tage højde for erkendelsen af, at programmer af den art kun kan forstås på baggrund af en grundlæggende analyse af programmets organisering, den politiske kontekst og det brede spektrum af intentioner og motiver, der præger programimplementering. I praksis betyder det, at man fx bør se på de forskellige interesser, der er involveret i implementeringen af mål 2-programmet. I forhold til dette projekts problemstilling om regional

strategi vil det være oplagt at se nærmere på samspelet mellem aktørerne i den regionale strategi udvikling og deres interesser set i forhold til planerne for programmet.

Opsummerende må det konkluderes, at de to elementer i Weiss' bidrag til evalueringsteorien umiddelbart har relevans for evalueringen af mål 2-programmet: Evalueringerne af mål 2-programmet har potentiale for at afdække grundlæggende vilkår for regional udviklingspolitik, og det er afgørende i analysen af mål 2-programmets funktion, at de mange aktørers interesser og motiver inddrages i evalueringerne.

4 Michael Scriven: Evaluering som vurderingens videnskab

Scrivens definition på evaluering er følgende: "*Evaluation determines the merit, worth, or value of things. The evaluation process identifies relevant values or standards to what is being evaluated, performs empirical investigation using techniques from social sciences, and then integrates conclusions with the standards into an overall evaluation or set of evaluations*" (Scriven, 1991).

Det element, som skal fremhæves her, og som Scriven står som den fremmeste eksponent for i evalueringsteorien, er vægtningen af de værdier og standarder, der skal ligge til grund for en evaluering. Scriven slår fast, at evaluering netop adskiller sig fra samfundsvidenskab ved at skulle vurdere, om en given ting (fx et program) lever op til de værdier og standarder, som er fastlagt. Den skarpe fokus på vurdering af evalueringsobjektet (evaluanden), understøtter han også i ganske klart sprog: "*Bad is bad and good is good and it is the job of evaluators to decide which is which*" (citeret fra Shadish et al., 1991). Det er ikke et trivielt synspunkt, da mange evalueringer af forskellige grunde (fx i håb om at sikre anvendelse) bliver draget ind i implementeringen af fx et program og dermed ikke kan eller ønsker at drage konklusioner om programmets værdi. I forlængelse heraf understreger Scriven også, at hvordan evalueringen skal bruges ikke er et spørgsmål evaluatør skal beskæftige sig med. anbefalinger om fremtidig politik er derfor ikke noget evaluatør skal blande sig i, hvis man følger Scrivens logik. Her er altså fremhævet to elementer fra Scriven:

1. Værdier, standarder og vurdering er afgørende i evaluering.
2. Anvendelse af evaluering, fx i form af anbefalinger, skal ikke være en del af evaluering.

4.1 Scrivens relevans for mål 2-evaluering

I forhold til mål 2-evaluering henleder de to fremhævede elementer fra Scriven opmærksomheden på, hvilke standarder og værdier evalueringerne baserer sig på, og hvorvidt det er hensigtsmæssigt, at evaluatør kommer med anbefalinger i evalueringen. Svaret til det sidste må umiddelbart være nej, hvis man følger Scrivens argumentation, men det er tæt knyttet til første spørgsmål på den måde, at hvis evaluatør skal eller ønsker at integrere anbefalinger i evalueringen, så er det i hvert fald helt afgørende, hvilke værdier og standarder evalueringens vurderinger og efterfølgende anbefalinger bygger på.

Som det tidligere er blevet nævnt, er der mange interesser repræsenteret inden for regional udvikling, og det er derfor vigtigt, hvordan værdierne for evalueringernes vurderinger fastlægges. Selvom interesse modsætninger og måske mere grundlæggende forskelle i opfattelsen af, hvordan regional politik kan/skal virke, ikke er tilstede i det besluttende forum for evalueringerne, så må

der alligevel tages højde for, at de interessemodsætninger kan eksistere i det omgivende samfund.

Omvendt kan evalueringer af områder som regionaludvikling, hvor vidensgrundlaget hele tiden udvikler sig, være med til empirisk at afprøve mere teoretiske antagelser på området.

Opsummerende kan det siges, at de to teoretiske pointer fra Scriven stiller nogle afgørende spørgsmål til, hvordan evalueringer som fx af mål 2-programmet gribes an, men hvordan der findes en balance i forhold til andre hensyn som anvendelse mv. er et kompliceret spørgsmål. Men kompleksitet kan ikke undskylde, at man går på kompromis med de krav, som Scrivens tilgang stiller evaluators overfor.

5 Vedung: Evaluering og den parlamentariske styringskæde

Professor Evert Vedung fra Uppsala Universitet er en dominerende skikkelse i den skandinaviske evalueringsforskning, og hans bog "*Utvärdering i politik og förvaltning*" er en af de mest benyttede lærebøger i Norden om evaluering.

Det element i Vedungs tænkning, der skal fremhæves her, er evalueringens grundlæggende forankring i det repræsentative demokrati og den offentlige forvaltning. Det betyder, at evalueringer indgår i den parlamentariske styringskæde, og at de derfor henter deres legitimitet fra det repræsentative demokrati: "*Utvärderingens reson är att i efterhand systematisk klarlägga och bedöma utfall, slutprestationer eller förvaltning af insatser i offentliga sektorn för att åstadkomma mer rationella beslut*" (Vedung, 1998, s. 35), og senere hedder det "*Politik kan inte ersättas med vetenskapelig utvärdering. Det är viktigt att utvärderingens rätta roll i offentlig politik ständigt diskuteras. Att den bör spela en roll är uppenbart. Men en aldrig så väl underbyggd utvärderingsfunktion kan inte ersätta politisk debatt och beslutsfattande*" (Vedung, 1998, s.207).

Ud fra dette synspunkt er evalueringens funktion i demokratiet afgørende, og de skal se på mere end økonomisk produktivitet og effektivitet, fx spørgsmål som retssikkerhed, demokrati mv. (Ibid. s, 204f).

5.1 Vedungs relevans i forhold til mål 2-evaluering

Det er en forudsætning for strukturfondspolitikken i Danmark og dermed også mål 2-programmet, at det er indlejret i den komplicerede, men demokratiske styringskæde, der lidt forenklet går fra EU's ministerråd over Folketing og regering ned til de nationale myndigheder, der implementerer programmerne.

Men kompleksiteten af myndigheder, der henter deres legitimitet fra forskellige demokratiske organer, gør, at evalueringer på et område som strukturpolitikken stilles overfor spørgsmål, om hvilke aktører og myndigheder der har legitime forventninger til evalueringens gennemførelse og resultater. Samtidig er strukturpolitikken et så institutionaliseret område, der på sin vis er isoleret fra den direkte politiske opmærksomhed, som er mange andre politikområder til del. Derfor vil de institutionelle interesser let kunne komme til at dominere dagsordenen også for evalueringerne.

Her kan Vedungs tilgang til evaluering være med til at huske på, at evalueringerne i sidste ende henter deres legitimitet fra de relevante politiske beslutningstagere. At det kan være kompliceret at udrede, hvordan forholdene er mellem de involverede aktører, er ingen undskyldning for ikke at forholde sig til problemet. En konkret problemstilling i forhold til denne undersøgelses emner er, hvordan de regionale udviklingsstrategier og interesser afvejes i forhold til nationale strategier og til dels overordnede EU-strategier. Der er et spillerum ved implementeringen af mål 2-programmet og det er derfor interessant og vigtigt, hvordan det rum udnyttes, hvilket evalueringer vil være et oplagt redskab til.

6 Peter Dahler-Larsen: Evalueringer og ny-institutionel teori

Heller ikke professor Peter Dahler-Larsens spændvidde ydes nogen retfærdighed i denne sammenhæng. Det element som her skal fremhæves ligger i forlængelse af Carol Weiss' fokus på programmets eller evaluandens karakter, men hvor Weiss' i første række ser dette ud fra sine praktiske erfaringer og traditionelle politologiske teorier om interesser og magt, angriber Dahler-Larsen i sin bog *"Den rituelle refleksion"* fra 1998 denne problematik med sociologisk ny-institutionel teori. Bogen gennemgår grundigt de teoretiske forudsætninger for denne teoridannelse, men her vil kun blive fremdraget et enkelt element, der er væsentligt i forhold til mål 2-evalueringerne.

Det element, som skal fremhæves her er, at institutioner, herunder organisationer, *"ikke kan reduceres til eller forklares ud fra organisationernes formelle formålstjenlighed eller for den sags skyld aktørgruppers snævre egeninteresser"* (Ibid. s. 46). *"Aktører" og "interesser" er ikke givet forud for institutionsdannelsen, men etableres i institutioner"* (Ibid. s. 48). Videre er det interessant, at *"Når institutioner virker under normale omstændigheder, så viser de sig netop ved ikke at kunne ses: man tager for givet, hvem der skal handle, og hvad handlingerne går ud på"* (Ibid. s.49).

Denne synsvinkel har en række konsekvenser, ikke bare for den måde programmet opfattes på, men også for evalueringens anvendelse, værdier og vidensgrundlaget (herunder metoden) for evalueringen. Her vil blive fokuseret på vidensgrundlaget for evalueringen, da fx metodevalget er tæt forbundet med den virkelighedsopfattelse, der ligger i den institutionaliserede strukturpolitik. Samtidig er evalueringens vidensgrundlag, fx udtrykt i metodevalg, med til at understøtte eller bekræfte det institutionaliserede politikområde, som mål 2-programmet er en del af.

6.1 Dahler-Larsens relevans for mål 2-evalueringerne

For en pragmatisk anlagt evaluering, hvor der skal tages hensyn til et stort antal krav fra EU-Kommissionen kan en institutionel tankegang som den netop præsenterede virke teoretisk og vanskelig at integrere i evalueringspraksis. Det kan sandsynligvis også give evaluator visse problemer, hvis han ruller et sådant begrebsapparat frem, men på den anden side rejser tilgangen mange vigtige spørgsmål, som netop en evaluering kan give mulighed for at tage op, hvor den "daglige" implementering af programmet ikke indbyder til den form for refleksion. Tilgangen kan bruges pragmatisk, så man ikke behøver rejse grundlæggende ontologiske spørgsmål, men diskutere hvordan fx metodiske valg er med til at bestyrke opfattelsen af programmet som fx et rationelt beslutningssystem, hvis succes kan måles på nogle objektive kriterier. Ved at åbne mulighed for at andre kræfter, også nogle mindre rationelle, er på spil i programmet, kan man måske nå frem til en bedre forståelse af, hvordan programmet fungerer, og hvorfor man måske ikke får de resultater, man forventer.

Dette er en omfattende diskussion, som ikke skal uddybes nærmere her, men det skal som ved de andre teorier anføres, at denne tilgang kan tilføje mål 2-evalueringerne en væsentlig dimension, hvis den bliver anvendt af evaluator.

7 Opsummering

Denne skitseagtige gennemgang af særlige aspekter hos fem fremtrædende teoretikere inden for evalueringsteorien har resulteret i følgende punkter, som det kunne være nyttigt at tage højde for i mål 2-evalueringer:

1. Bliver evalueringerne brugt til det, de er tiltænkt, af dem de er tiltænkt?
2. Hvem der er de tiltænkte brugere af evalueringen, og hvad er den tiltænkte brug af evalueringen?
3. Evalueringerne har potentiale for at afdække grundlæggende vilkår for regional udviklingspolitik.
4. Det er afgørende i analysen af mål 2-programmets funktion, at de mange aktørers interesser og motiver inddrages i evalueringerne.
3. Værdier, standarder og vurdering er afgørende i evaluering.
4. Anvendelse af evaluering, fx i form af anbefalinger, skal ikke være en del af evaluering.
5. Evalueringerne henter deres legitimitet fra de relevante politiske beslutningsorganer.
6. Ved at åbne mulighed for at andre faktorer end de rationelle er på spil i programmet, kan man bedre forstå, hvordan programmet fungerer, og hvorfor man måske ikke får de resultater, man forventer.

Teoretikerne deler ikke teoretisk udgangspunkt, men det vil alligevel være muligt at bruge de enkelte tilgange inden for samme evaluering, da de enten retter sig mod forskellige dimensioner af evaluering eller er komplementære.