

# EVALUERING – FALSKMØNTNERI, LEJEMORD ELLER ÆRLIGT HÅNDVÆRK

*Benny Karpatschof, lektor, dr. phil  
Institut for Psykologi, Københavns Universitet*

Der har i de sidste årtier været et boom i evaluering af psykosocial behandling. Der foretages utallige evalueringer, som omfatter projekter, organisationer, behandlingsmetoder, ja sågar hele skoler eller retninger, så som kognitiv terapi. Det kan man som professionel inden for branchen jo nok være tilbøjelig til at se på med en vis velvilje, men ind imellem melder der sig en bekymring over, hvad denne formidable bølge af undersøgelser nu egentlig dækker over. Der er nemlig tale betydelige kvalitetsforskelle, evalueringsbølgen slæber med sig, der er både revl og krat.

Jeg skal her opridse nogle hovedproblemer i denne genre, dels ud fra mine egne erfaringer som hovedevaluator og som konsulent<sup>1</sup>, dels ud fra en systematisk læsning af nogle mere principielle artikler i tidsskriftet *Evaluation Review*.

## **Evaluering som videnskabelig disciplin**

Evaluering fik først en videnskabelig basis, da Campbell og Stanley (1966) udgav den banebrydende bog *Experimental and Quasi-experimental designs for research*. Med dens klare begrænsninger, der bl.a. består i en kvantitativ orientering, en gruppemæssig designopfattelse og et ateoretisk effektbegreb, har bogen på godt og ondt sat den scene, de fleste evalueringer har optrådt i.

Jeg skal senere vende tilbage til disse begrænsninger, der stadig præger en meget stor del af nutidens evalueringer.

## **Evaluering som praktisk styringsredskab**

De fleste evalueringer har slet ikke karakter af egentlig forskning, idet de mangler det almene perspektiv, som må karakterisere en videnskabelig undersøgelse. I stedet har de et rent praktisk formål, hvilket indebærer, at deres perspektiv er snævert knyttet til et afgrænset praksisfelt - med de særlige betingelser, arbejds måder og målsætninger der karakteriserer dette. Eksempler kan være:

- en socialpædagogisk institution vil gerne vide, om deres psykosociale behandling har nogen effekt på ”utilpassede indvandrere-unge”

---

<sup>1</sup> Se (Karpatschof, 2001a, 2001b , 2003a, 2003b), Karpatschof & -Rue 2003.

- et behandlingscenter for ludomane skal dokumentere sin eksistensberettigelse
- et forsøg med beboerråd på et plejehjem skal evalueres.

Her er der sjældent nogen større interesse i at almengøre erfaringerne fra det specifikke praksisfelt til en mere generel sektor. Altså i de nævnte tilfælde til, hvad der generelt synes at bide på utilpassede indvandrere-unge, hvad der får ludomane til at holde på deres mønter, og hvad der bidrager til at aktivere plejehjemsbeboere. Endnu sjældnere at løfte blikket ligefrem til et helt alment teoretisk plan, der jo kunne handle om, hvordan mennesker i det hele taget bliver animeret til at ændre livsførelse. Og det er måske meget godt, at sådanne evalueringer beskedent holder sig deres snævre praktiske sigte, for det vil næsten altid være ret umuligt at foretage en overbevisende almengørelse.

Det hænger sammen med de generelle forskelle mellem praktisk og forskningsmæssig orienteret virksomhed. Praksisfeltet er afgrænset og specifikt, men underkastet nogle krav om netop praktisk effektivitet. Forskningsfeltet forsøger (mere eller mindre dygtigt) at overskride sine egne begrænsninger med henblik på at opnå en erkendelse af forhold, der ikke er lokaliseret snævert i tid og rum. Hvor praksisfeltet spørger om dens erfaringer er praktisk anvendelige, interesserer forskningsfeltet sig for om dens erfaringer er videnskabeligt gyldige.

## **Evaluering som strategisk våben**

Hvor de to første perspektiver, det forsknings- og det praksisrettede, således er meget forskellige, men ud fra deres formål lige agtværdige, kommer vi nu til den problematiske, for ikke at sige betændte side af evaluering. Når evaluering omfatter en aktivitet i et bestemt praksisfelt, er den nemlig ikke altid bare underkastet nogle praktiske betingelser af indlysende og åbenbar karakter, men aktiviteten kan meget vel være genstand for en politisk kamp mellem modstridende målsætninger og interesser. Dette er der naturligvis ikke i sig selv noget forkert ved, men i denne sammenhæng melder problemerne sig, når konflikterne ikke udspiller sig i en åben politisk kamp, men foregår underjordisk, og når konfliktens parter, ligeledes underjordisk, søger at spænde en evaluering for deres respektive vogne.

Nu er det imidlertid naivt principielt at afvise en evaluering i et politisk kontroversielt område. For så ville der ikke være mange felter tilbage at evaluere, og det ville i øvrigt være nogle af de mest interessante felter, der blev elimineret. Spørgsmålet er blot, hvordan man kan indgå som evaluator i et politiseret felt på en måde, der hverken kompromitteres af naivitet eller korruption.

Politiseringen kan her være lokaliseret til det undersøgte praksisfelt, fx en organisation eller et projekt. Naiviteten består da i, at evaluatoren som en pantsat bondedreng uafvidende lader sig bruge til et bestemt på forhånd givet, men skjult formål. Korruptionen ligger i med åbne øjne at lade sig misbruge. Den korrupte evaluator bliver falskmøntner, når hun eller han mod bedre vidende blåstempler en organisation eller projekt. Og lige så korrupt, men endnu mere skummelt kan evaluator optræde i rollen som lejemor, når opdragsgiveren ønsker at nedlægge en organisation eller stoppe et projekt, evt. at finde en syndebuk for et forløb der løbet af sporet.

Politiseringen kan også være knyttet til et mere generelt praksisfelt, af psykosocial karakter, fx integration af indvandrere eller behandling af bestemt afvigerkategori. Hvis dette gene-

relle praksisfelt er nationalt politiseret, så er evaluator underkastet et pres, der kan være endnu mere massivt end i det første tilfælde med lokaliseret praksis. En evaluering af modersmålsundervisning for indvandrerbørn vil være øretævernes holdeplads, uanset om konklusionen er positiv eller negativ. Og en afprøvning af behandling af ungdomskriminelle vil, næsten uanset udfald, vække modstand fra enten den konservative eller liberale fløj inden for kriminalpolitikken.

Jeg skal vende tilbage til, hvordan man som evaluator kan tackle politiseringens problem.

## **Evalueringens metodeproblemer (evaluering som disciplin)**

Selv om en evaluering indgår som styringsredskab i et afgrænset praksisfelt, og selv om den ikke har noget særligt alment perspektiv, så har den, i vor af videnskabelig tænkning prægede kultur, i reglen brug for at påberåbe sig en undersøgelsesmetode af en vis videnskabelig karakter. Her er problemet imidlertid at der faktisk har udviklet sig to meget forskellige evalueringsskoler. Den kvantitative og den kvalitative.

Den kvantitative retning inden for evaluering blev som nævnt lanceret med fejende kompetence af Campbell og Stanley. Inden jeg nu går i gang med at påpege denne retnings indbyggede begrænsninger, må det understreges, at der ikke er tale om en afstandtagende kritik, og at disse begrænsninger ikke er at betragte som fejl, men som karakteristika, der lige så vel kan anskues som dyder. Det forhold gælder jo også inden for kunstens verden, hvor maleriet almindeligvis mangler en dimension, men til gengæld kan trylle med farven, og hvor skulpturen i det store og hele har det lige omvendt. Campbell-Stanley-traditionen er præget af en stringent videnskabelig tænkning, ofte præget af en sofistikeret brug af statistisk forsøgsplanlægning og analyse, men den har tre klare begrænsninger:

- Den er baseret på kvantitative data, dvs. entydige oplysninger, typisk kategori-svar fra et spørgeskema, eller en score i en test.
- Den holder sig til gruppemæssige designs
- Dens effektbegreb er ateoretisk

Fordelen ved *kvantitative* data er måske ikke så meget deres pålidelighed som deres metapålidelighed, det er nemlig forholdsvis let at måle deres pålidelighed, uanset om denne nu er lav eller høj. Til gengæld står det ofte skralt til med gyldigheden. Kommunikationen i en spørgeskemaundersøgelse er meget ofte både snæver og usikker. Forstår forsøgspersonen spørgsmålet? Fatter undersøgeren, hvori svaret består? På den anden side er rent kvalitative undersøgelser, hvilket i realiteten vil sige de semistrukturerede interviews, ofte noget blafrende i selve interviewmetoden og uanset kodningsteknik hyppigt vilkårlig i analysen. Her har jeg selv haft gode erfaringer med en særlig kombination af kvalitativ og kvantitativ metode, hvor indledningsvis et hold af særligt trænedede interviewere foretager semistrukturerede interviews, som så efterfølgende bliver kodet dels kvantitativt af interviewerholdet selv efter nogle nøje definerede og instruerede kodningsanvisninger, dels kvalitativt af evaluator med et teksanalysesystem. (Karpatschhof, 2001a, 2003a).

De *gruppdesign*, som Campbell og Stanley beskrev, afslører hurtigt deres oprindelse inden for pædagogikkens felt. De er knyttet til forsøg, hvis genstand er grupper, fx skoleklasser,

og ikke som mange psykologer ville foretrække til enkeltindivider. Her må det understreges, at det eksperimentbegreb, som Campbell og Stanley instituerede, slet ikke er en naturlig del af den psykologiske disciplin, men snarere har oprindelse i dels pædagogikken, dels lægevidenskaben med det kliniske randomiserede forsøg rettet mod en enkelt, eller nogle få kontrollerede variable. Psykologien har derimod en anden tradition fra sin eksperimentelle oprindelse, en tradition som dels er knyttet til individer snarere end til grupper, og som dels anvender en anden type af eksperimentel design, nemlig det komplette multifaktorielle eksperiment, hvor en række kontrollerede variable kombineres systematisk. Et mere psykologirelevant alternativ til gruppedesignet er således efter min opfattelse en individorienteret forløbsundersøgelse, hvor man har righoldige informationer dels vedrørende de undersøgte personer, dels af forløbet. Der er desuden ved psykologisk behandling ofte forskningsetiske argumenter mod gruppedesigns af forsøgs- kontrolgruppe-kategorien. Psykologer vil generelt være mere afvisende end læger over for systematisk at unddrage mennesker en psykologisk hjælp, der er af vital betydning. Her er læger øjensynligt mere hårdhodede.

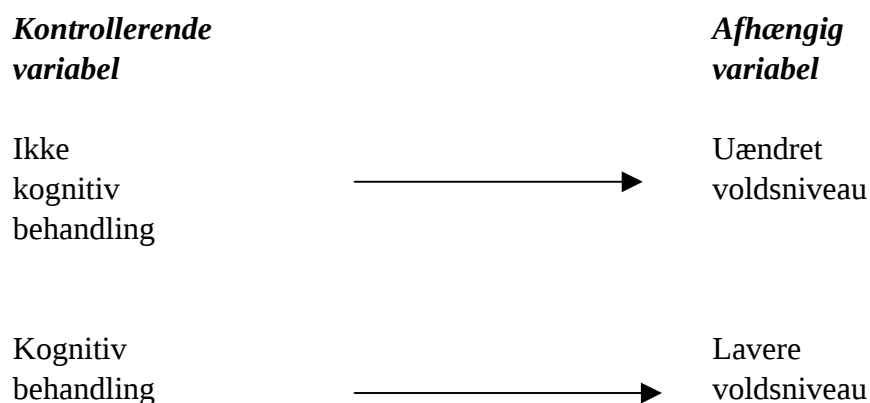
Endelig er Campbell og Stanley-traditionen knyttet til ateoretisk *effektbegreb*, hvis positive side er muligheden for at have et simpelt og veloperationaliseret effekt-begreb, mens den negative side er, at vi - når vi nu har fået at vide, at faktor X er virksom over for adfærden Y - så lades i stikken med hensyn til en viden om, *hvornår*, sådan mere præcist, denne virkning fremkommer, og *hvordan* og *hvorfor* den gør det.

Hvor de to første begrænsninger er af mere principiel karakter, og nok kun kan mildnes ved, at det kvantitative gruppedesign kombineres med en kvalitativ undersøgelsesdel, da er der faktisk nogle muligheder for at overskrive det ateoretiske ved effektbegrebet. Jeg kan her hen-vise til to artikler i Evaluation Review.

Anthony Petrosino (2000) kritiserer det nøgne binære effekt-begreb, hvor vi blot konstaterer, at interventionen X har en virkning på adfærden Y, og han foreslår en udvidelse, der for så vidt stadig ligger inden for den oprindelige positivistiske ramme. Han peger nemlig på to andre typer af variable ud over den kontrollerede variabel X (interventionsvariablen) og den afhængige variabel Y (virkningsvariablen). Den første type er mediatorer, det vil sige nødvendige mellemliggende variable mellem intervention og virkning. Lad os sige, at vi evaluerer et forsøg med kognitiv behandling af unge voldskriminelle. Man træner dem i at overveje deres handlinger inden de bringes til udførsel.

Lad os først se på den simple, binære effekt-model:

## Den simple (binære) effektmodel

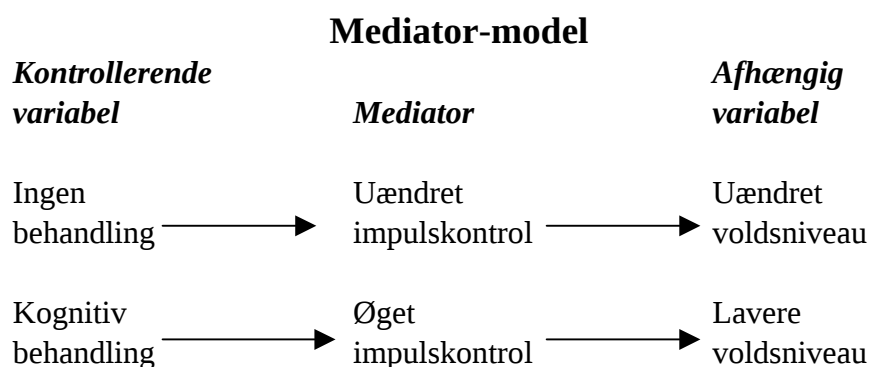


Figur 1

Lad os nu antage, at der mellem den eksperimentelle og den afhængige variabel er en mellemliggende variabel, som fx kunne være impuls kontrol.

Det er altså her modellen, at den kognitive behandling ikke kan opfattes som en automatisk og direkte årsag til mindre vold, men at den virker gennem forøget impuls kontrol, og kun under denne forudsætning.

At denne udvidelse af det binære effektbegreb er empirisk og ikke bare teoretisk kan vi se, hvis vi inddrager den som observationsvariabel. Vi kan da teste modellen på følgende måde:



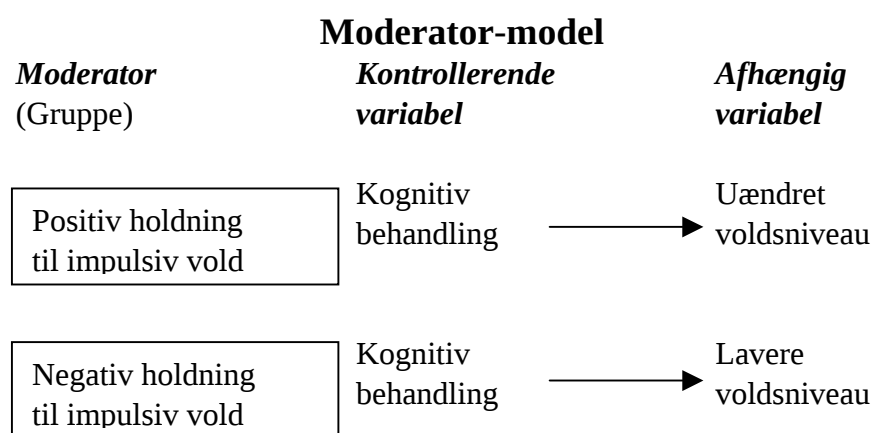
Figur 2

Ved at inddrage den formodede mediator som observationsvariabel kan vi altså her vise, at kognitiv behandling kun influerer på voldsniveauet, hvis det sker gennem ændret impuls kontrol. Dette har i så fald ikke blot en praktisk relevans, nemlig at rette den kognitive behandling mod forøget impuls kontrol, men det kan måske samtidig pege på metodens begrænsninger,

nemlig at den kun påvirker voldsniveauet for dem, der ad kognitiv vej lader sig animere til at forøge deres impuls kontrol.

Og hermed er vi ved at bevæge os ind på det andet begreb i artiklen, moderator-variable.

En moderator er en variabel, der virkningsmæssigt indgår i sammenhæng med den kontrollerede variable. Ud fra den her opstillede model kunne vi tænke os, at vi kan operere med to grupper af de unge voldsmænd. Den ene opfatter deres lave impuls kontrol som problematisk, ja måske ligefrem som uheldig. Den anden ser, måske ud fra en macho-præget subkultur, voldelig impulsivitet som en positiv egenskab. Hvis vi inddrager denne variabel, som altså ikke er en mellemliggende, men en forklaringsvariabel, ja måske ligefrem en selektionsvariabel, ligestillet med den kontrollerede, så har vi følgende observationsskema:



**Figur 3**

Da vi her har to uafhængige variable, såvel behandlingsvariablen som moderator-variablen, bør vi have ikke to, men 4 grupper, hvor de to variable er kombineret.

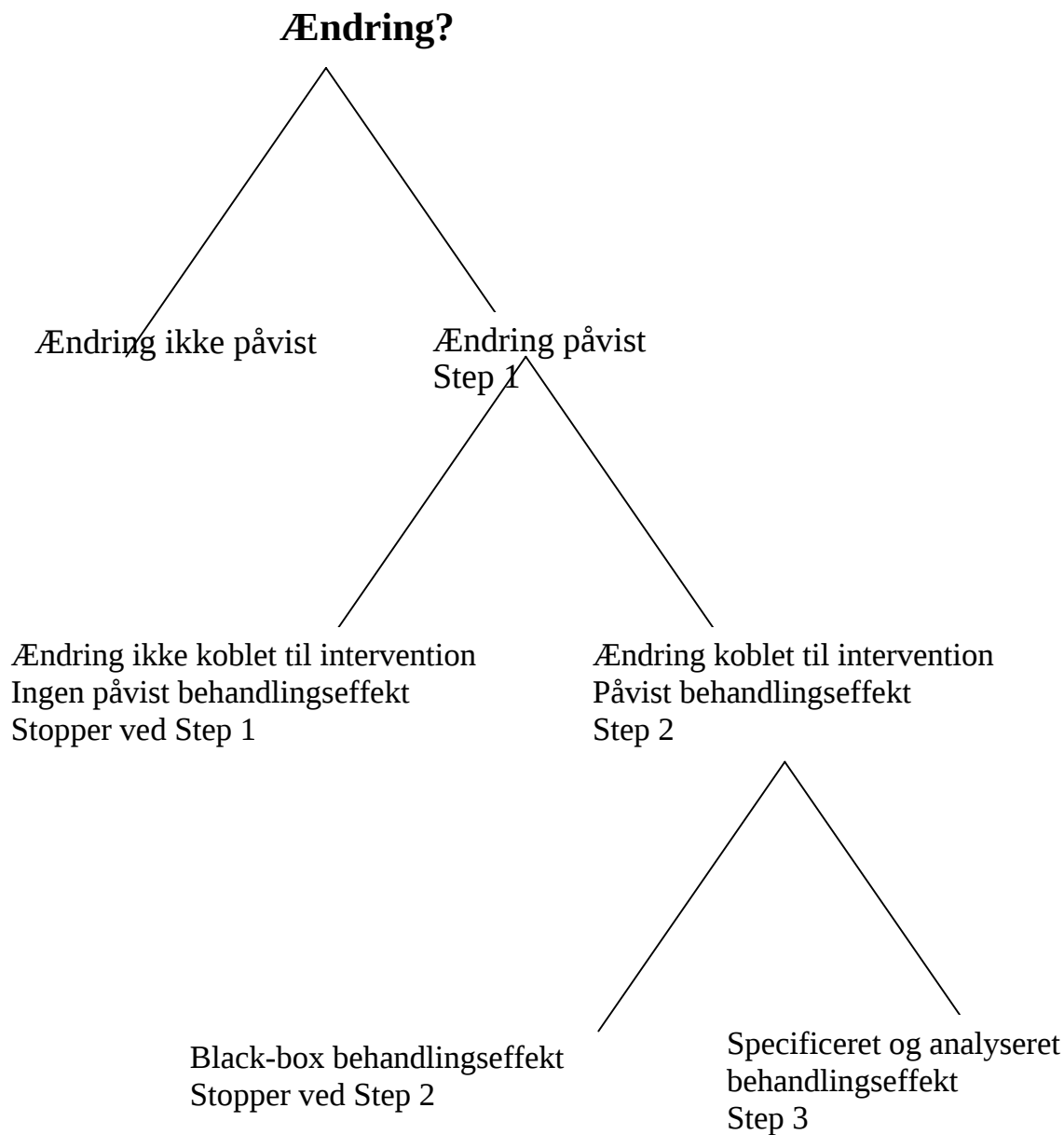
Der er naturligvis intet i vejen for at kombinere mediator- og moderator-modellen, således at man indfører en moderator som ekstra uafhængig variabel og samtidig observerer evt. medierende processer.

Endnu længere end Petrosino går Birckmaer & Weiss (2000), der diskuterer begrebet teori-baseret evaluering. De gennemgår i artiklen en række undersøgelser af denne art og peger på, at teoribaseret evaluering kan føre til bedre evaluering, og at det i nogle tilfælde endda kan føre til en bedre teori, fordi empirien fra evalueringen nødvendiggør en modifikation af teorien.

Perspektivet i såvel mediator-moderator-udvidelsen som i teoribaseret evaluering er altså for det første en berigelse af den nøgne binære effekt-undersøgelse, som kun kan give os en black-box dokumentation af, at der er en størrelse X, der indvirker på en anden størrelse Y. Berigelsen kan pege på, *ad hvilke veje* indvirkningen finder sted, og måske ligefrem give os en forståelse af, *hvornår* interventionen virker, *hvornår* den *ikke* gør det, og *hvorfor*. For det andet giver den mere teoriorienterede evaluering en mulighed for at overskride den erkendelsesmæssige begrænsning, som ellers karakteriserer evalueringspraksis.

En mulighed for et samspil mellem teori og praksis.

Hvis vi skal sammenfatte de forskellige former for effektundersøgelser kan vi opstille nedenstående træstruktur:



Figur 4

Ved step 1 har vi kun påvist *ændring*, men vi har endnu ikke vist, at denne ændring er en *effekt*, dvs. at den er kausalt koblet til interventionen.

Ved en black-box dokumentation, step 2, har vi vist, at behandlingen *virker*, at effekten altså ikke er resultatet af fremmede faktorer, men vi har alene påvist en black-box-effekt. En

effekt, som vi ikke ved så meget om. Først i næste skridt, i step 3, får vi kastet lys over de afgørende spørgsmål:

- a. *Under hvilke betingelser* (moderatorer) har vi effekten
- b. *For hvilke personer* (moderatorer)
- c. *Ad hvilke veje* (mediatorer) virker effekten
- d. Hvordan vi kan *forstå* behandlingens virkning

Med al respekt for de forbedringer, som de førnævnte mediator- og moderator-modeller kan indebære, er det efter mine erfaringer først gennem en kvalitativ analyse, at disse spørgsmål, og her især de to sidste, kan besvares. Disse erfaringer med en overskridelse af black-box evaluering ved en kombination af kvantitativ og kvalitative metoder er der redegjort for i evalueringsrapporterne (Karpatschof 2001a, 2003a).

## **Evalueringens praktiske problemer (evaluering som praktisk styringsredskab)**

Som flere gange fremhævet har evaluering ikke alene et konkret stykke praksis som sin genstand, men den bliver selv en del af denne praksis. Evaluator må altså, selv om hun eller han kommer fra forskningens elfenbenstårn, indstille sig på at træde ind i det pulserende liv, på at vise sin duelighed mere som praktiker end som forsker. Dueligheden som praktisk evaluator er snævert knyttet til de etiske problemer, vi skal behandle i næste afsnit, men her skal vi se på den mere håndværksmæssige side af evaluering. Selv om en forskningsmæssigt *placeret* eller i alt fald *uddannet* evaluator naturligvis har udviklet sin basale kompetence inden for en videnskabelig disciplin, så får hun eller han nu brug for evner af en helt anden karakter. Det er i høj grad de samme evner, som efterspørges i en moderne erhvervsvirksomhed. Evner til teamwork fx, til at aflægge eventuelle unoder i retning af akademisk individualisme, til at indgå i en specifik organisationskultur. Og ikke mindst en orientering mod tempo og effektivitet, snarere end mod forskningens uendelige sandhedssøgen. For så vidt som der er tale om behandlingsinstitution eller for den sags skyld en erhvervsvirksomhed, med et bestemt klientel, og det er arbejdet med dette klientel der skal evalueres, er opgaven ikke blot at få et tilstrækkeligt kendskab til organisationen, men også til klientellet.

I mit eget arbejde som evaluator har jeg derfor set mit initiale arbejde som to forskellige feltstudier. Det ene felt er organisationen, det andet er klientellet. Ja i reglen er der faktisk ikke 2, men mindst 3 felter at opnå kendskab til, nemlig opdragsgiverens hovedorganisation, den datterorganisation som skal evalueres, og endelig klientellet.

I min evaluering af støtte-kontaktpersonordningen i Københavns Kommune var opdragsgiver-organisationen Kommunens Familie- og Arbejdsmarkedsforvaltning, den evaluerede organisation var gruppen af støtte-kontaktpersoner, fordelt på et centralt center og en snes lokale voksenteams, mens klientellet var gruppen af de sindslidende, som er brugere af ordningen.

Samarbejdet med opdragsgiveren blev sikret gennem mit arbejde i en lille arbejdsgruppe og med en større følgegruppe. Kontakten med den evaluerede organisation og med klientellet

blev opnået gennem adskillige møder, hvor jeg ud fra min initiale uvidenhed blev ved med at stille spørgsmål, indtil et billede af de respektive felter udkrystalliserede sig.

## Evaluerings etiske problemer (evaluering som strategisk våben)

En evaluerings placering i praksis betyder, at den ikke bare er rettet mod at opnå viden om et bestemt felt, men at føre til konkrete afgørelser om et projekt eller en organisations fremtid. Dette indebærer et moralsk ansvar for evaluator. En del af professionalismismen i denne branche er derfor efter min opfattelse ikke at fralægge sig etiske fordringer i professionalismismens navn, men derimod at have gjort sig nogle etiske overvejelser om, hvad man kan, bør og vil foretage sig som evaluator.<sup>2</sup>

Faktisk kan man have et større eller mindre ansvar over for såvel opdragsgiver, den evaluerede datterorganisation, klientellet og offentligheden. Disse forskellige former for ansvar er igen knyttet til evaluators multiple status som fagperson, som lønnet konsulent og som samfundsborger.

Som *fagperson* har man et ansvar over for sit fag, over for den disciplin, man tilhører, ja over for kravet til videnskabelig hæderlighed og ubestikkelighed.

Som *lønnet konsulent* har man et undertiden kontraktligt og altid moralsk ansvar over for opdragsgiver, et ansvar for at være sin hyre værd. Desuden vil en evaluering medføre et ansvar over for såvel den evaluerede datterorganisation som for dets klientel.

Som *samfundsborger* har man et ansvar over for offentligheden i de samfundsmæssige anliggender, som evalueringen berører eller influerer på.

Disse tre former for ansvar kan imidlertid meget let komme i indbyrdes konflikt. Lige som den hårdt prøvede tjener i Goldonis komedie, *En tjener - to herrer* kan man være usikker på, hvilken herre det er, man skal tjene. Er det hensynet til sandheden, eller til opdragsgiver, til den undersøgte datterorganisation, til klientellet, eller de overordnede samfundsmæssige anliggender?

Hvad gør man, hvis ens konklusion går mod opdragsgivers ønsker eller opfattelse?

Hvordan forholder man sig til at skulle give en negativ bedømmelse af den evaluerede gruppe medarbejdere, eller til en klientgruppe, der måske kommer til at miste et tilbud efter evalueringen?

Hvorledes forener man sin professionelle status med ens forpligtelse som samfundsborger, og hvad stiller man op med sine principielle politiske holdninger til områdets problemer?

Der er mange etiske dilemmaer i denne konfliktzone, og det er slet ikke muligt at give et patentsvar. Med hensyn til det overordnede politiske plan kan det på den ene side være uacceptabelt at optræde upolitisk og fastholde, at ens forpligtelse er professionel og ikke samfundsmæssig. Men modsætningen til en sådan påstået neutralitet med dens afpolitiserede aura af saglighed kan være en privat politisering, hvor man på ensidig og usaglig måde inddrager sine egne holdninger af privat karakter og dermed misbruger sin status som evaluator til en rollesammenblanding.

---

<sup>2</sup> I Morris & Jacobs (2000) beskrives en spørgeskemaundersøgelse blandt medlemmer af det amerikanske evalueringsselskab om deres forhold til etiske problemer, og det viser sig, at disse ikke er et fokusområde for ret mange af de adspurgte.

Mine erfaringer peger derfor hverken på et klart og ensidigt valg mellem bestemte ansvarsaspekter eller på vattede forsøg på kompromis'er. Derimod vil jeg pege på den form for strategisk fantasi, hvor man søger at finde en praktisk løsning på problemet. En løsning, der måske ikke er dadelfri i alle henseender, men som man har fundet frem til som den bedst mulige.

Af diskretionsgrunde skal jeg her afholde mig fra at gå for langt ind i mine egne erfaringer, men vil igen pege på en artikel fra Evaluation Review, hvor de politiske og etiske valg nok ikke er ganske sammenfaldende med mine egne synspunkter, men hvor der er tale om nogle tankevækkende og intelligente overvejelser.

I Bomberg & Waldo (2002) fortæller nogle medarbejdere ved *Juvenile Justice Educational Enhancement Program, JJEED*, Floridas organisation for evaluering af pædagogiske tiltag for ungdomskriminelle, om de problemer, denne evalueringsorganisation har haft med sammenstødet mellem sine teoretisk og empirisk baserede erfaringer og synspunkter på den ene side og holdningerne hos de lovgivere, der har magten over kriminalpolitikken og styringen af dens institutioner.

Umiddelbart er konflikten antagonistisk og ubodelig. De professionelle inden for det særlige pædagogiske felt vedrørende ungdomskriminelle har entydige resultater, der peger på:

at *pædagogiske* initiativer er bedre end straf  
at *små* institutioner er bedre end store  
at *offentlige* institutioner er de bedste, almennyttige NGO-institutioner de næstbedste, og dårligst er de private kommercielle virksomheder, som endnu er ukendte hos os, men i hastig vækst i USA.

Og lovgiverne. Ja blandt dem råder den stik modsatte opfattelse.

Her synes der kun at være to mulige løsninger. På den ene side en Don Quichote-agtig dødsforagt, hvor man utrættelig fremsender rapporter til lovgiverne, hvori man dokumenterer, hvor tåbelige lovgivernes synspunkter er. Og på den anden side en kujonagtig opportunisme, der giver kejseren, in casu de folkevalgte, hvad de nu gør krav på.

*JJEED* afviser den første strategi som spild af tid. Den fører kun til, at de videnskabeligt baserede undersøgelser bliver afvist som akademisk pjank, der ikke er denne verden (og den slags forekommer jo også, sandt at sige). Men de afviser også, rimeligt nok, det sidste alternativ som en videnskabelig kapitulation.

I stedet har de gjort noget helt tredje. De har forsøgt en dialog med de implicerede parter, både opdragsgiver, her lovgiverne, og de evaluerede organisationer, her behandlingsstederne. Således har de i starten sønderlemmende evalueringer af kæmpestore kommercielle ungdomsfængsler med en meget ringe pædagogisk standard været ledsaget af konkrete anbefalinger af at gøre institutionen mindre og undervisningen bedre. Disse anbefalinger er meget ofte blevet fulgt, ikke mindst fordi borgerretsorganisationer i flere tilfælde med held har sagsøgt enkeltstater for at forsømme lovgivningens undervisningspligt. Og så snart en sådan silo for utilpassede unge har forbedret sig (og dermed de unge), så følger naturligvis en ny og mere positiv evaluering.

Jeg har fremhævet denne artikel, uden at jeg dermed vil påstå, at tilsvarende konflikter mellem vore hjemlige "kriminologiske smagsdommere" og deres populistiske kritikere kan klares på tilsvarende vis, men den er et eksempel på den evalueringsmæssige fantasi, som ved siden af solid metodisk kompetence og konkret viden om de evaluerede genstande er en afgørende kvalitet for den, der hverken vil være falskmøntner eller lejemorder, men helst en ærlig håndværker.

## Referencer:

- Birckmaer, J. & Weiss, C. H. (2000). Theory-based evaluation in practice: What do we learn? *Evaluation Review* (24:4), 407-341.
- Bomberg, T. E. & Waldo, G. P. (2002). Evaluation research, policy and politics. *Evaluation Review* (26:3), 340-351.
- Campbell, D.T. & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Karpatschof, B. (2001a). *En evaluering af brugernes oplevelse af SKP-ordningen i Københavns kommune* (evalueringsrapport til Københavns kommunes Familie- og arbejdsmarkedsforvaltning, okt. 2001).
- Karpatschof, B. (2001b). At blive behandlet ordentligt - eller bare blive behandlet -om brugerstyret kontakt i socialpsykologien og om psykopatologisk reduktionisme *Psyke & Logos*, 22(2): 695-716
- Karpatschof, B & Rue, K. (2003). En støtte der er brug for - Brugernes oplevelse af støttekontakt-person-ordningen i Københavns Kommune. I: Steen Bengtsson (Ed.). *Handicap, kvalitetsudvikling og brugerinddragelse*, AKF forlaget.
- Karpatschof, B. (2003a). *Evaluering af Vil:kan – Egmont Fondens forsøgs- og udviklingsprojekt for børn der har mistet forældre eller som har forældre med en livstruende sygdom*. København: Egmont Fonden.
- Karpatschof, B. (2003b). Når sorgens fugle flyver bort – Evalueringen af Egmont Fondens forsøgs- og udviklingsprojekt med hjælp til børn i sorg. I: *Egmont Fondens Årsskrift 2002*. København: Egmont Fonden.
- Morris, M. & Jacobs, L. R. (2000). You got a problem with that? Exploring evaluators' disagreement about ethics. *Evaluation Review* (24:4), 384-406
- Petrosino, A. (2000). Mediators and Moderators in the Evaluation of Programs for Children – Current Practice and Agenda for Improvement. *Evaluation Review* (24:1), 47-72.