

Kvalitativ evaluering

- en anmeldelse af Ian F. Shaw (1999): **Qualitative Evaluation**. Sage.

Anmeldt af Ph.d. stipendiat *Tanja Miller*, DIG, SDU

Bogen rejser væsentlige problemstillinger om evaluering på den ene side og forskning på den anden side. Bogens genstandsfelt er ikke begrænset til uddannelsesforskning og uddannelsesevaluering, men har som intention at beskæftige sig tværfagligt og overordnet med de spørgsmål, der rejses.

En gennemgående præmis i værket er opsplitningen af videnskaben i grundforskning, strategisk og anvendt forskning som det traditionelle forsøg på at skelne mellem en uafhængig og mere bundet forskning. En skelnen, der får konsekvenser for evalueringer, hvis evaluering er forskning? - her underforstået at vores horisont er videre end evalueringsforskningen selv. – Dette spørgsmål præciseres yderligere ved kun at have relevans for den kvalitative evaluering. Hermed lægges afstand (via videnskabelig kritik og mindre videnskabelige undladelser) til de store kvantitative systemevalueringer.

Bogens mission er således at udrede, hvordan kvalitativ evaluering adskiller sig fra kvalitativ forskning ved at have forskellige forehavender, i sine teoretiske ophavshistorier og konstruktioner samt ved sine anvendelsespotentialer og praksisser.

Bogen er struktureret efter teori/praksis matrixen. Første del beskæftiger sig således med udredninger af teorierne og deres indbyrdes relationer, mens anden del indeholder kritiske analyser og diskussioner af (kvalitativ) evaluering i praksis.

Rejsning af problemstillingerne

Indledningskapitlet starter med et 30-årigt citat af Lee Cronbach: "What the evaluation field need is a good social anthropology" – og anvendes i konteksten som tidsangiver på, hvor lang tid "kampen" inden for evalueringsforskningen har fundet sted. Aktuelt er området oversvømmet med inspiration fra antropologien – alle evalueringer med respekt for sig selv inddrager og baseres på vidnesbyrd fra feltet, der evalueres. Fra fængselsvæsnet til politik og social velfærd. Kombinationer af kontrollerede evalueringer af social intervention holdt op over for adfærdsmæssigt udbytte verificerer, hvad der dur og ikke dur inden for de forskellige områder af samfundslivet. Og i tilfælde af at brugerne, professionelle eller politikere ikke kan genkende sig selv, institutionen, professionen eller samfundet i resultaterne forklares det som kommunikati-

onsproblemer eller som modstand mod forandringer. Altså ved psykologiske 'brist' ved evaluanten.

"that the relationship between evaluative purpose and qualitative social science is a – possible *the fulcrum* for generating good evaluation inquiry" (s. 1)

Så det er ikke inspirationen fra antropologien/etnografien, Shaw opponerer imod, men måden hvorpå det sker. Han skelner mellem en bredt orienteret og en mere snævert orienteret gøren brug af inspirationerne. Den bredt orienterede er overvejende praksisbåret og inkluderer ansvarlighed, tætte arbejdsrelationer mellem forskeren og praktikerne, en generel forpligtigelse til at fæstne vurderinger til empiriske facts og en eksplicit trækken på forskningsresultaterne som støtte i beslutningsprocesser. Ligeledes er der et pres på forskningen til, at resultaterne (her inkluderes processerne) er gennemførlige og afstedkommer en professionel udvikling.

Den mere snævre retning er akademisk båret. Denne retning anvender termene mere restriktivt og har mere tillid til målinger, instrumentalitet, objektivitet af 'evidence' og adfærdsmæssige indicier på god professionel praksis, quasi eksperimenter og er lejlighedsvis negative over for metodologien i kvalitative evalueringer.

Shaws forbehold over for begge udgave, men specielt den snævre er, at alt for meget er formodet. Hvad betyder, at noget er baseret på noget andet, hvad betyder 'i praksis' og endelig, hvad betyder 'evidence'?. Som han siger om disse retninger

"The confidence in measurement within evidence-based practice betrays a naïve view of how easy it is to grasp an understanding of reality. The failure to problematize the evidence/practice relationship is the Achilles heel of evidence-based practice". (s. 3)

Shaw lister en række aktuelle problemer i evalueringsforskningen: amerikansk og professionel etnocentrisme, at evaluering udelukkende opfattes som programevaluering, at evaluering af f.eks. uddannelser eller andre områder er separeret fra personlig evaluering og endelig diskussioner af rækkevidden af evalueringsresultater i en bevægelse fra det rent lokale til det helt generelle.

Kvalitativ evaluering

Men hvad er så kvalitativ evaluering, og hvordan er forholdet til kvalitativ forskning?

Det første spørgsmål er naturligt nok; hvad er evaluering? Shaw lægger ud med Chelimskys tredeling af evaluerings formål: evaluering til måling af resultater, effekter og effektivitet, evaluering for udvikling eller som styrkelse af f.eks. en institution og endelig evaluering som vidensform – for at opnå en dybere forståelse af specielle områder. Og løfter hermed en vigtig detalje i evalueringsforskningen ud i lyset, nemlig at sondringer ofte foretages ved hjælp af formål og et efterfølgende valg af metoder.

Når han så spørger til forskellene på forskning og evaluering, opløser han svarrunden i tre:

- Kan forskning og evaluering være det samme/forskelligt?
- Er forskning forskellig fra evaluering?
- Skulle forskning og evaluering være det samme eller forskelligt?

Pointen med opløsningen er at vise, at det første spørgsmål er af teoretisk karakter, det næste af empirisk og det sidste af moralsk karakter. Han hævder videre, at mange problemer stammer fra en sammenblanding af niveauerne i diskussionerne – som for det meste fokuserer på det empiriske.

Efter illustrationen af denne svaghed går han videre med at adskille forskningen og evalueringen – dette ved at inddrage de store evalueringsforskere. Denne argumentation vil jeg ikke udrede her, da den dukker op igen i gennemgangen af ”kongerækken” .

Kap. 2: Evalueringsteoretikere og kvalitativ evaluering

Shaw indleder med at understrege, at han ikke agter at evaluere evalueringsteoretikerne, men blot udlede opkomsten, holdninger og forslag til definitioner på kvalitativ evaluering.

Jeg går ikke i dybden med alle¹, men har valgt statements, der støtter Shaw’s argumentation.

Campbell inddrages med den væsentlige elaborering af validitetsspørgsmålet. Han skabte begrebet ’intern validitet’, der refererer til slutninger om forholdet mellem afhængige og uafhængige variabler. Ellers er Campbell mest kendt for eksperimentelle og quasi-eksperimentelle forskningsdesigns. Han betragtede kvalitativ viden som både tests på og byggeklodser i kvantitativ viden: ”this is not to say that such common-sense naturalistic observation is objective, dependable or unbiased. But it is all that we have. It is the only route to knowledge – noisy, fallible and biased though it².

Michael Scriven, præsenteres som den store nyskaber af sproget³ via nye begreber inden for evalueringsforskningen og skal her bare citeres for sit grundsynspunkt; at evaluering undersøger og bestemmer værdi og merit. Han vendte spørgsmålet om mål og effekter på hovedet.

”Rather than ask, ’there the objectives. Have they been achieved’?, Scriven would prefer ’Here is the programme. What’s the effects?’” Han åbner altså blikket for, at effekter kan være

¹ Donald Campbell, Michael Scriven, Lee Cronbach, Robert Stake, Carol Weiss, Ernest House, William Reid, Elliot Eisner, Egon Guba og Yvonna Lincoln.

² Campbell, 1979:50 – her s. 22

³ F.eks. målfrie evaluering, summativ og formativ, meta-evaluering osv

både tilsigtede og ikke-tilsigtede . Merit og værdibestemmelser af et program kan således ikke vurderes/gøres op kun i forhold til de foreliggende mål med programmet.

Lee Cronbach, som af Shaw beskrives som undervurderet af etablissementet, vægter den formative evaluering højt. Han påpeger, at evaluering ikke kan betragtes som en før- og efter situation, men som mens, mens, mens situationer og siger: "A theory of evaluation must be as much a theory of political interaction as it is a theory of how knowledge is constructed".

Robert Stake præsenteres som den ledende inden for kvalitativ evaluering med vægt på udvikling af case-study metoder. Han skaber en model – responsiv evaluering – der fokuserer på programmets aktiviteter fremfor programmets formål, som indretter sig på lokale stakeholderes ønsker om information, som udreder værdier ved at rapportere både succeser og fiaskoer og endelig ved at være deltagende observatører.

Hovedargumentationen for synsvinklen formulerer Stakes således: "tenacious expectation that knowledge can make a sick education well", as in assumption that research leads to knowledge which leads to improved practice".⁴

Stakes fortjeneste er først og fremmest, at blikket rettes mod deltagelse/deltagerne og case-studier som metode.

Carol Weiss' forskning i evalueringsforskningens tilknytning til politisk forskning ændrede synet på evalueringer for stedse. Hun påpegede den legitimerende funktion, som alle evalueringer har. En dobbelt legitimation, idet programmernes præmisser ikke problematiseres og dernæst ved, at rapporterne/resultaterne legitimerer enten status quo eller ændringer. Det var i øvrigt Carol Weiss, der gennem sine empiriske studier gjorde det bemærkelsesværdige fund, at evalueringsresultater kun i meget sjældne tilfælde blev anvendt ifølge sit formål (med at lave evalueringen).

'Instrumental use seems in fact to be rare, particularly when the issues are complex, the consequences are uncertain, and a multitude of actors are engaged in the decision making process, i.e., in the making of policy".⁵

Dette fund, sammen med moderniseringens stigende kompleksitet, har afstedkommet bl.a. bestemmelse af evalueringer snarere som ritualer og organisationskulturelle udfordringer end som vurderinger.

Carol Weiss er endvidere kendt for begrebet "oplysende" evaluering i betydningen, at evalueringer kan kaste lys over noget⁶.

⁴ s. 27.

⁵ Weiss 1980:397, her s. 29.

⁶ En af Peter- Dahler-Larsens syv evalueringsanvendelser. Se "Tendenser i evaluering" (2001), s. 215.

Ernest House tager afstand fra de foregående ud fra en fællespræmis (for dem), nemlig at øget viden gør mennesker lykkeligere eller i hvert fald mere tilfredse. I stedet betragter han evaluering som noget meget mindre rationelt og meget mere hvilende på argumenter og overbevisninger. Evaluering kan aldrig hævdes at være sikker viden og har indbygget en normativ fortolkning af tilværelsen.

”In essence, the pluralist model confuses issues of interests with conflicts of power. It can balance only those interests that are represented – typically those of the powerful”.⁷

House argumenterer for evalueringsmodeller, der kan fange variationer og irregulariteter ved bred inddragelse af deltagerne og ved, at forskeren (den der evaluerer) anvender kvalitative metoder.

Ja. Så er historien snart fortalt. De to næste - William Reid og Elliot Eisner - har hver for sig stået for nyskabelserne: evaluering som udvikling af praksis og endelig reflektiv evaluering. Med de to sidstnævnte er sammenhængen med koblinger mellem læreprocesser og evaluering blevet tydelig. Og til sidst skal nævnes Guba og Lincoln, som fortæller evalueringsforskningens historie med konstruktivistiske briller og dermed forbinder evaluering med alle andre fænomener og forskningstraditioner. Deres eget bidrag til evalueringsforskningen kan kaldes fortolkende eller narrativt, og vi nærmer os f.eks. etnografer og litteraturfortolkere.

I resten af bogen (6 kapitler) behandles kvalitativ evaluering fra forskellige synsvinkler. Nogle praktiske i den forstand, at allerede udførte kvalitative evalueringer vurderes teoretisk og metodologisk. Kapitel tre er viet til udredningen af paradigme-krigen inden for evalueringsforskningen. Paradigmerne benævnes postpositivisme, kritisk konstruktivisme og pragmatisme. Et meget velskrevet og spændende kapitel - men tungt.

De sidste er mere som en værktøjskasse og besvarer spørgsmål om, hvordan der kan bedrives feltarbejde med kvalitative evalueringsmetoder. Det drejer sig f.eks. om analyser af observationer, tekst og dokumenter, livshistorier, fokusgrupper, interviews og simulationer.

Afslutning

Bogen henvender sig til alle med interesse i evaluering – både som forskning og i praksis, og den bidrager til at kaste lys over og skabe orden i et meget uoverskueligt (og gigantisk stort) område. Bogen giver grundige og velbegrundede udredninger af forhold mellem f.eks. teori og praksis, epistemologi og ontologi og viden og magt. Den behandler ydermere etiske og politiske implikationer af den ekspanderende evaluerings ”industri” og maner til eftertanke og kritiske refleksioner over – ikke så meget, hvad evalueringer bliver brugt til – som til, hvordan evalueringer er med til at skabe vores virkelighed og opfattelser af den. Konkluderende om

⁷ House, 1991:240, her s. 31.

bogens stærke sider vil jeg fremhæve den del af evalueringsforskningen, der drejer sig om vidensproduktion og konsumption.

At diskussionen også er aktuel i Danmark, kan bl.a. dokumenteres i pjecen ”Pædagogisk forskning, udvikling og uddannelse” som ”Statens Humanistiske Forskningsråd” udgav her først på året. Her er der en spændende artikel af Kirsten Weber, RUC, hvori hun bl.a. siger:

”... Det hænger sammen med, at uddannelsesforskningen (til forskel fra den traditionelle pædagogiske forskning) gennem de seneste årtier netop har udviklet sig teoretisk i forsøg på at begribe netop dén samfundsmæssige udvikling, der – også – har sat evaluering på den politiske dagsorden” (Pædagogisk forskning, udvikling og uddannelse, s. 19).

Evaluerer er også i Danmark historisk knyttet til den forvaltningmæssige diskurs med sin effektmåling og rationaliseringsinteresse, men vi diskuterer evaluering i dag også som organisationskulturel udfordring og som en dimension i mentalitetsændringer. Og en evalueringsforskning, der tilbyder sig med bestemte og vigtige former for vidensproduktion.