

Evaluering i endring

Per Øyvind Bastøe

Direktør, Avdeling for samfunnsøkonomi og offentlig
forvaltning



Jeg vil gjennom dette tilbakeblikket drøfte hvordan omfanget av evalueringsvirksomhet og forståelsen av evaluering har endret seg de siste tiårene – både internasjonalt og nasjonalt her i Norge

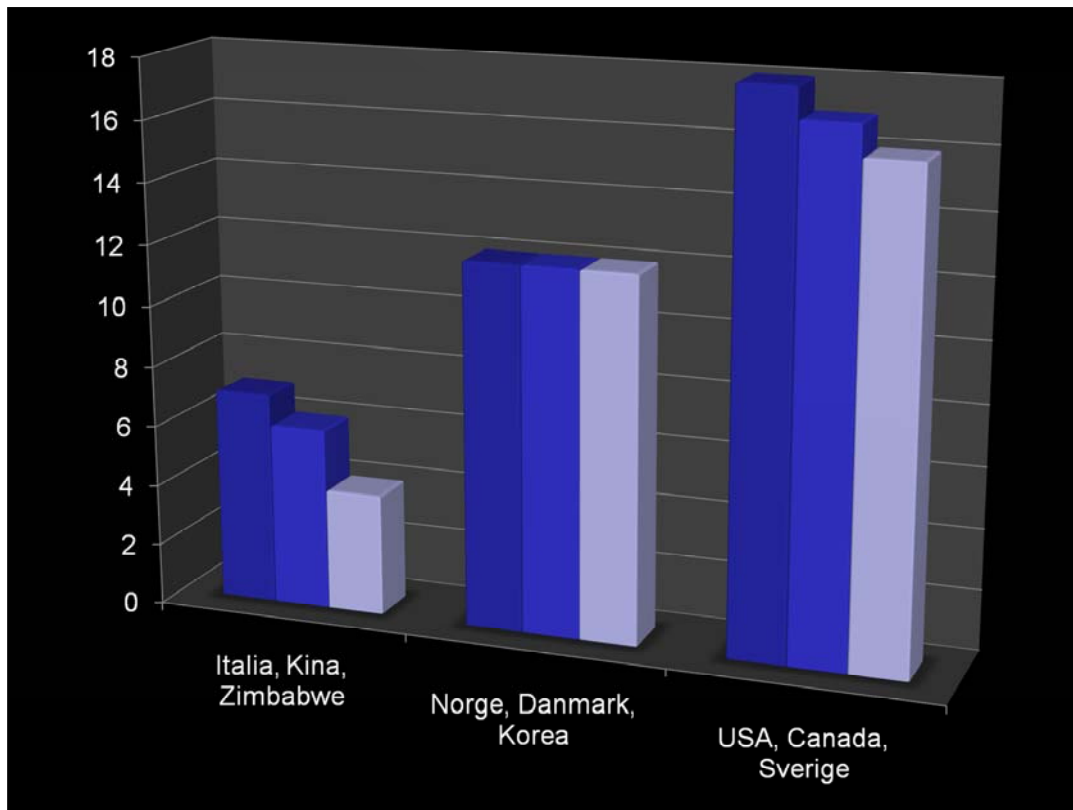


Det er ikke noen dristig påstand å si at det har vært en sterk økning i omfanget av evalueringsvirksomhet. Wikipedia går så langt som å si at: *Within the last three decades there have been tremendous theoretical and methodological developments within the field of evaluation.*

For 20 år siden var "evaluering" et begrep vi her i Norge hadde et lite aktivt forhold til – det ble sjelden brukt og når det først ble brukt var vi ikke helt sikker på hva vi mente med det. Internasjonalt foregikk det meste knyttet til evaluering i USA.

I dag er bildet helt annerledes – både internasjonalt og i Norge – det gjelder omfanget av evalueringsvirksomhet, bruken av evalueringer og forståelsen av hva som er en god evalueringen.

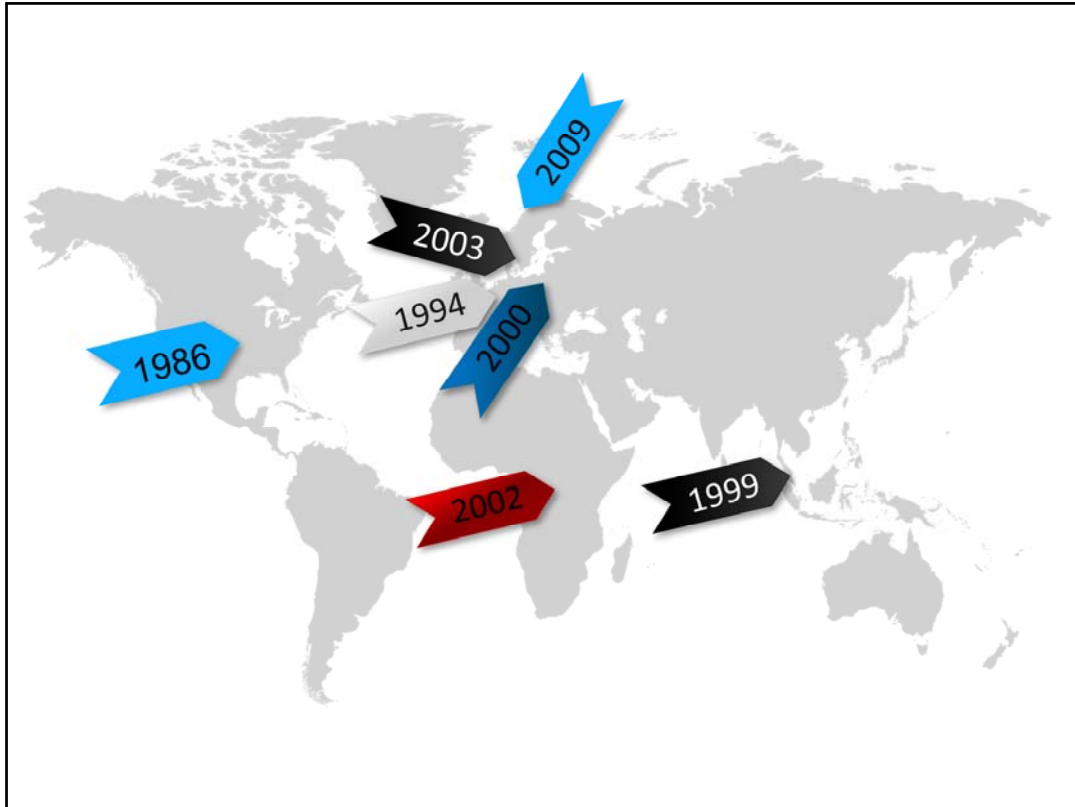
Det er mange ulike forklaringer på denne voldsomme utviklingen de siste tiårene. En av disse er at evalueringsvirksomhet utvikler seg i takt med utviklingen av offentlig sektors tilnærming til planlegging, forvaltning og bruk av kunnskap. I mange land har det de siste tiårene vært en økende oppmerksomhet på brukerorientering, åpenhet, delegering, konkurranse og effektivisering – alt uttrykk for det som har gått under betegnelsen New Public Management. Både revisjon og evaluering er sentrale elementer her. Interessen for New Public Management – i litt ulike avskygninger – har nærmest vært et globalt fenomen. I Europa har særlig EU vært eksponent for denne tenkningen ved å tidlig institusjonaliser evaluering som en integrert del av forvaltning.



The International Evaluation Research Group (der både Peter Dahler-Larsen og jeg er med) gjorde for noen år siden (2002) i det vi kalte "International Atlas of Evaluation" en systematisk gjennomgang av den utviklingen innen evaluering i 21 ulike land.

De 21 landene ble rubrisert ut fra ni ulike kriterier knyttet til graden av institusjonalisering og profesjonalisering av evaluering. Vi fant at noen land var kommet langt (USA, Canada, Australia, Sverige, Nederland og United Kingdom), noen andre var kommet ganske kort (Spania, Italia, Kina m.fl), og noen land (Norge og mange av de andre europeiske landene) var i en mellomkategori i prosessen med å utvikle og institusjonalisere evaluering på nasjonalt nivå.

En av de forklaringene vi trakk fram i boken var graden av eksponering i forhold til amerikansk evaluering. For eksempel forklarte noen av våre svenske kolleger den tidlige utviklingen i Sverige med den sterke koplingen mellom svensk og amerikansk sosiologi på 50- og 60-tallet. En annen forklaring på rask framvekst av evalueringer knytter seg til utviklingen av velferdsstaten og velferdstjenestene. De landene som tidlig fokuserte mye på offentlige utdannings- og sosiale velferdsprogrammer var også tidlig ute med å utvikle evalueringsvirksomheten. Det var nettopp disse områdene den tidligste evalueringsvirksomheten i USA fokuserte på.



Et uttrykk for hvordan framveksten av evalueringsevikrosmhet har eksplodert er etableringen av nasjonale og regionale evalueringeforeninger. På dette kartet har jeg plassert et lite utvalg:

American Evaluation Association kom tidlig på banen – allerede i 1986. European Evaluation Society ble etablert i Haag i 1994. Deretter har foreninger blitt etablert en rekke steder:

- 1999: Malaysian Evaluation Society
- 2000: Polskie Towarzystwo Ewaluacyjne
- 2002: Uganda Evaluation Association
- 2003: Svenska utvärderingsföreningen

Og da til slutt:

- 2009: Norsk Evalueringeforening

Disse foreningene har utvilsomt bidratt til bevisstgjøring, opplæring og utvikling av evalueringsfaget. Jeg har tilsvarende forventninger til hva som vil skje her i Norge.

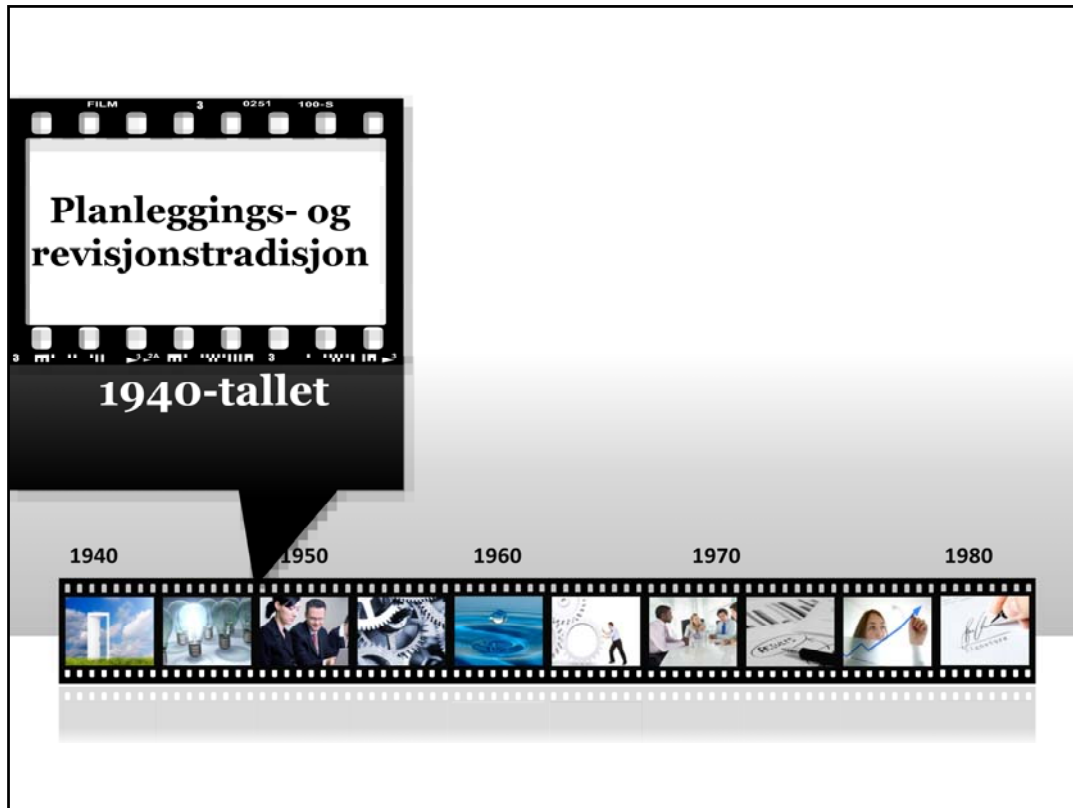
Norge - en sinker?



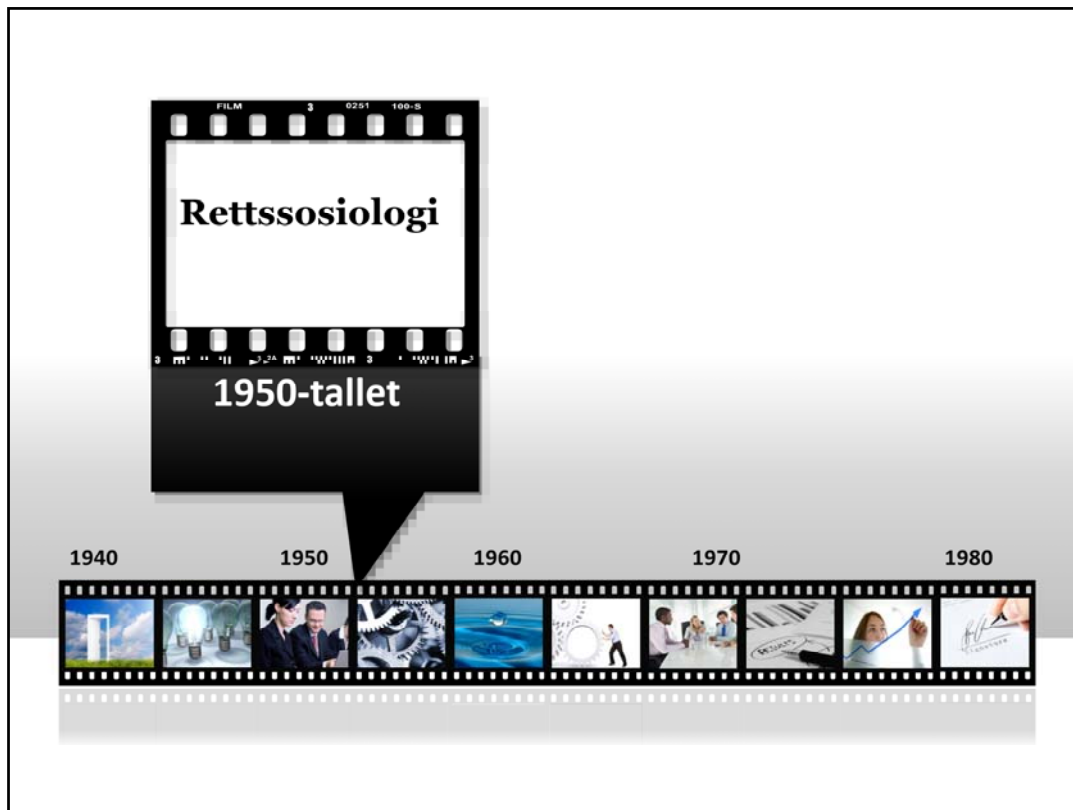
Norge har altså vært sent ut i utviklingen av evaluering som eget felt. I 1990 var "evaluering" på mange måter et fremmedord. Riktignok var det et lite knippe norske medlemmer i American Evaluation Association allerede på slutten av 1980 tallet, og norske evalueringspionerer som Bergjlot Baklien og Karl Erik Brofoss skrev artikler om evalueringsmetoder og evalueringsformer. Men det generelle bildet var at vi hadde problemer med å helt forstå hva evaluering kunne brukes til og hvordan det skilte seg fra andre former for analyser og vurderinger.

I en studie av evalueringsvirksomheten i norske departementene, som Karl Erik Brofoss gjorde i 1997, bekreftes dette uklare bildet. Han fant at alle departementene i perioden 94-96 hadde gjennomført studier og analyser som godt kunne kalles evalueringer (til sammen flere hundre), men at tilnærmingene og bruken av evalueringene varierte mye mellom departementene. Det var altså ingen enhetlig eller sammenhengene forståelse av hva vi forsto med evaluering.

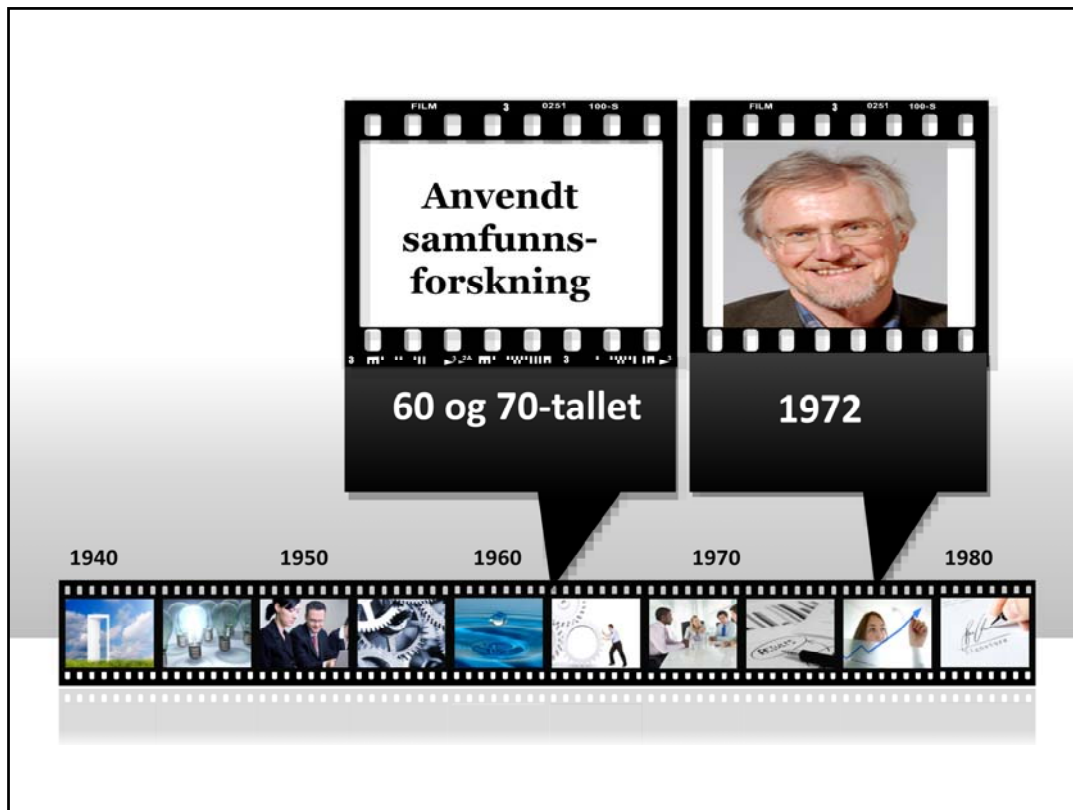
En mulig forklaring som Johan P. Olsen kom med i 1996 var at utviklingen av offentlige reformer i det hele tatt er annerledes i Norge enn i mange andre land – for eksempel USA og Sverige. I følge Johan P. Olsen er offentlige reformer i Norge mer utviklet gjennom små, inkrementelle skritt enn gjennom store, vel planlagte, strukturelle reformer. I motsetning til de fleste andre land i Europa, har vi kommet oss unna behovet for store reformer som konsekvens av økonomiske nedgangstider. Oljeinntektene er selvsagt en viktig grunn til dette. Også vår posisjon på utsiden av EU har gjort at vi har vært annerledes. I stedet for å være opptatt av å kutte i offentlige utgifter, har vår utfordring vært å begrense veksten. Evaluering har ikke nødvendigvis vært betraktet som noe egnet virkemiddel i den sammenheng.



Nå kan man selvsagt argumentere med at vi i Norge i mange år har bedrevet tilsvarende aktivitet under andre benevnelser enn evaluering. For det første har vi hatt en planleggings- og revisjonstradisjon med Finansdepartementet, Norges Bank og Statistisk Sentral Byrå som sentrale aktørene helt siden 1940-tallet. For eksempel etablerte Finansdepartementet sin egen planavdeling allerede i 1963 med ansvar for analyser og langtidsplanlegging.

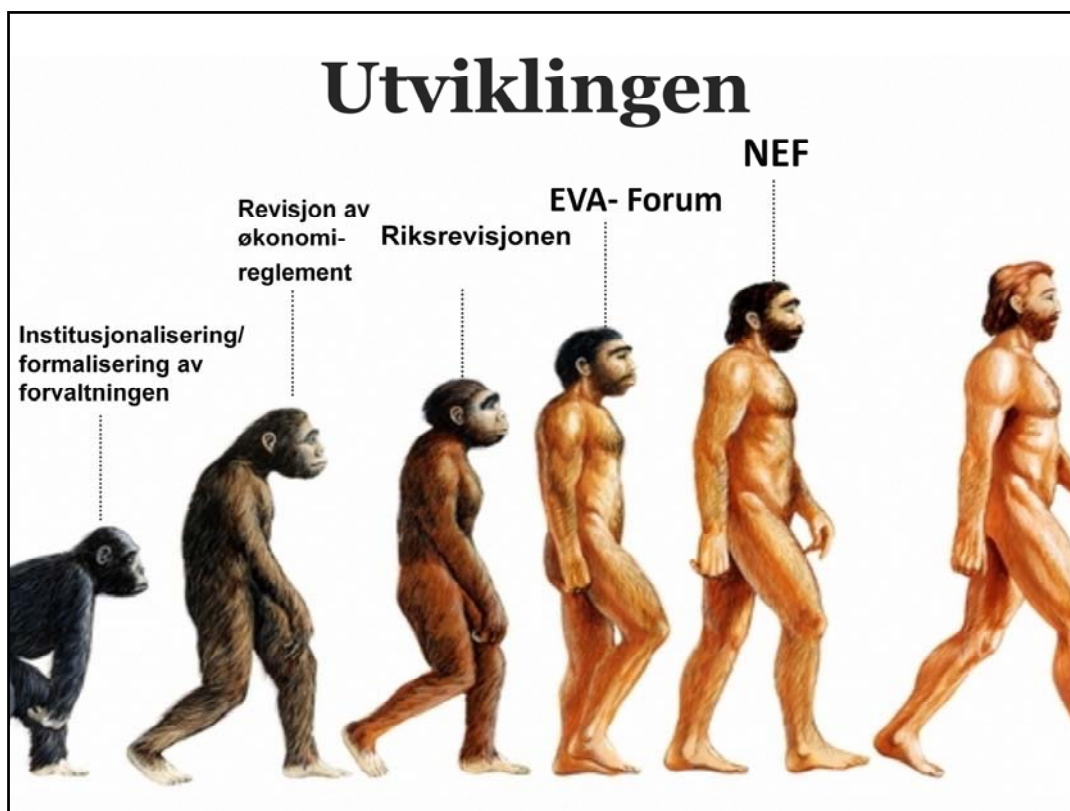


Et annet område som grenser mot evaluering er retts sosiologi. Her ble grunnlaget lagt allerede på 1950-tallet gjennom en studie av ansettelsevilkårene for hushjelper. Et uttrykk for denne tradisjonen er utformingen av den såkalte "utredningsinstruksen" allerede på 70-tallet – med krav om utredning av økonomiske og administrative konsekvenser av nye lovforslag.



En tredje tradisjon som ligger tett opp til det vi kan forstå som evalueringsvirksomhet, knytter seg til anvendt samfunnsforskning, hvor velferdsforskning i ulike varianter har utviklet seg helt fra 60- og 70-tallet. Man kan godt si at Maktutredningen med bl.a. Gudmund Hernes som ble igangsatt i 1972 er et eksempel på denne tradisjonen, og som kanskje gjerne kunne kalles evaluering.

Et unntak fra de generelle utviklingstrekkene i Norge som jeg har beskrevet har vært den omfattende evalueringsvirksomheten på bistandsfeltet. Her har på mange måter evaluering levet sitt eget liv. Allerede i 1978 ble det etablert en egen evalueringsseksjon i Utenriksdepartementet som hadde som ansvar å evaluere bistandsprosjekter og programmer. En håndbok om evaluering og resultatoppfølging ble publisert så tidlig som i 1993. Men man kan vanskelig si at dette i denne tidlige fasen hadde noen særlig innflytelse utenfor bistandsfeltet.



Utviklingen av den eksplisitte evalueringstenkningen i Norge er i stor grad knyttet til institusjonaliseringen og formalisering i den sentrale statsforvaltningen. Forskningsrådet, helsedirektoratet og anvendte forskningsinstitutter innen en rekke sektordepartementsområder har bidratt til systematisk evalueringsarbeid. Sentrale aktører som Bergljot Baklien ved Statens Institutt for rusmiddelforskning og Karl Erik Brofoss ved Norsk institutt for studier av innovasjon, forskning og utdanning, har som nevnt vært nestorer i utvikling av evalueringfaget her i landet. NOKUT – Nasjonalt organ for kvalitet i utdanningen – som ble etablert som et uavhengig kvalitetssikringsorgan i 2003 ved innføringen av kvalitetsreformen i høyere utdanning er også et uttrykk for denne økende formaliseringen.

Et annet viktig skritt i utviklingen av evaluering som egen virksomhet ble tatt med revisjonen av statens økonomireglement på midten av 1990-tallet, der evaluering for første gang ble omtalt som obligatorisk i forhold til offentlige tilskuddsordninger. Etableringen av den første forvaltningsrevisjonsavdelingen i Riksrevisjonen omtrent på samme tid, har helt klart til økt oppmerksomheten på evaluering, og har bidratt til en begynnende forståelse av en mer helhetlig tenkning.

I den sentrale statsforvaltningen har dessuten EVA-forum – som er medarrangør for denne konferansen – spilt en viktig rolle som diskusjonsarea om evaluering. Formålet med dette faglige forumet for personer tilknyttet offentlige virksomheter er å utvikle og utveksle kunnskap om evaluering.

Etableringen av NEF er foreløpig siste trinn i denne utviklingen her i Norge

From Studies to Streams



Parallelt med ekspansjonen i evalueringsaktiviteten har det skjedd en utvikling i forståelsen av hva evaluering omfatter og hva som betraktes som god evaluering. Et uttrykk for en tidlig definisjon finner vi fra 1995 hos en av evalueringens grand old ladies, Elianor Chelimsky: *"A reasonably well accepted definition might be that program evaluation is the application of systematic research methods to the assessment of program design, implementation, and effectiveness."* 10 år senere kommenterte Chelimsky selv at forståelsen av evaluering hadde endret seg: *"we now evaluate both ex ante and ex post."* Evalueringsbegrepet har blitt således bredere og omfatter mer nå enn det gjorde tidligere.

De senere årene har vi også sett en dreining fra fokus på enkeltprosjekter til fokus på større programmer og tverrgående tema. Det fører til at evalueringsoppgavene flyttes fra enkeltstudier til bruk av evalueringkunnskap fra flere ulike kilder – der monitorering og evaluering på sett og vis smelter sammen. Et uttrykk for denne tendensen er økt bruk av begrepet "resultatorientering" og publisering av metaevalueringer og resultatrapporter.

Det at mange av de spørsmål som man søker svar på er sammensatte og komplekse er en av grunnene til at en enkelt evalueringstudie vanskelig gir et tilstrekkelig grunnlag. Dessuten er det selvsagt ikke uvanlig at studier av tilsvarende spørsmål har vært gjort tidligere eller i en annen sammenheng. Fra min verden er det for eksempel opplagt at når man skal finne egnede virkemidler for å bidra til oppbygging av Haiti, så er vi på jakt etter evalueringresultater av tilsvarende gjenoppbyggingsinnsatser i en rekke andre utsatte og sårbare stater. En enkeltstudie kan ikke alene gi de svarene vi er på jakt etter.



En annen diskusjon som stadig har gått igjen de siste tiårene er om en god evaluering skal være uavhengig og objektiv, eller prosessorientert og subjektiv.

En klassisk diskusjon av denne tematikken finnes i essayet "Deltakar eller tilskodar" som filosofen Hans Skjervheim publiserte allerede i 1957. Skjervheims utgangspunkt er at det finnes to helt forskjellige måter å forholde seg til verden på – som aktiv deltaker eller som observatør, altså tilskuer. Hans anliggende var at det å ta et menneske alvorlig, var det samme som å ta den andres meninger opp til ettertanke og diskusjon. Det var å være deltaker. Den objektiviserende innstillingen er en angrepsholdning, hevdet han. Essayet var således en kritikk av vitenskap som maktutøvelse og hersketeknikk – i en tid da vitenskapen gjennomgående ble sett på som sann, objektiv og nesten hellig.

Her illustrert med et bilde fra filmen "Salmer fra kjøkkenet":

Under etterkrigstiden kommer svenske observatører til Norge for å undersøke norske menns kjøkkenbruk, med det formål å effektivisere deres kjøkkenrutiner. Folke Nilsson har fått i oppdrag å studere Isak Bjørviks kjøkkenvaner. Ifølge forsøksinstituttets regler må Folke sitte på en opphøyet stol på Isaks kjøkken og observere ham derfra, men han får aldri lov til å snakke med ham eller gripe inn i Isaks daglige gjøremål. Isak stopper å bruke kjøkkenet og begynner å observere Folke gjennom et hull i gulvet fra soverommet sitt i stedet. Etterhvert begynner de to ensomme mennene å komme overens.



I evalueringsverdenen har denne tematikken særlig vært framme i diskusjonene mellom evalueringstilnærmingene til personer som Michael Skriven – som er opptatt av objektivitet og uavhengighet – og Michael Quinn Pattons nyttefokuserete prosessstilnærming. Denne debatten er på ingen måter ny, men er stadig aktuell. Ikke minst knyttet vurderingene av "evidens based" evalueringer og "impact" evalueringer. Her er det igjen idealene om uavhengighet, objektivitet og vitenskapelighet som understrekes. Er denne type evalueringer bedre og mer høyverdige enn evalueringer som først og fremst søker etter gjennomførbare konklusjoner og anbefalinger?

Dette er også en diskusjon om hva som skaper størst legitimitet og muligheter for endring. Er det uavhengigheten og vitenskapeligheten skal inngir tillit – og dermed fører til at resultatene blir tatt på alvor? Eller er det at evaluatoren er mer en prosesskonsulent som bidrar til at aktørene selv hjelpes til å ta hensiktsmessige og brukbare beslutninger.

Men, man kan spørre om denne bredere forståelsen bidrar til større klarhet, eller til større forvirring. Wikipedia hevder med referanse til Jane Davidson I 2005 at det til tross for sterk utvikling fortsatt er mange fundamentale problemer innenfor evaluering. Ifølge Wikipedia finnes det fortsatt ingen klar definisjon på hva en evaluering omfatter: *"unlike medicine, evaluation is not a discipline that has been developed by practicing professionals over thousands of years, so we are not yet at the stage where we have huge encyclopedias that will walk us through any evaluation step-by-step."*

You get what you measure



"You get what you measure"

Nå skal alt resultatorienteres, monitoreres, kontrolleres og evalueres. Er dette bare et gode, eller er det en fare for at vi går for langt. Det er etter mitt syn nødvendig å ha et realistisk syn på evaluerings rolle og funksjon. Det kan gå for langt – det kan oppstå et resultat paradoks – som gjør at det er det synlige, kortsiktige, det som er lett å telle og måle som får all oppmerksomhet på bekostning av den langsiktige og komplekse.

Er evaluering i ferd med å miste sin kritiske rolle?

Mens evalueringens virksamhet på 60- og 70-tallet var primært rettet mot forbedring av sosialpolitiske programmer, dreiet oppmerksomheten på 80- og 90-tallet seg mot å kutte kostnader og å effektivisere offentlig sektor. Senere har oppmerksomheten vært mer på evaluering som virkemiddel for kontroll og som grunnlag for politiske beslutninger. Budsjettprosessen har blitt en stadig viktigere kanal for bruk og formidling av evalueringens informasjon, og finansdepartementene har blitt de primære etterspørre etter evalueringer. I denne utviklingsprosessen har begreper som nytte, kvalitet og metode fått viktigere plass innen evalueringens diskusjoner. I den sammenheng er det vel også riktig å si at det har utviklet seg en nøktern, og mindre optimistisk syn på evalueringens verdi.

Informasjonsfunksjonen var tydelig i den tidlige fasen der hensikten var å utvikle offentlige utdannings- og velferdsprogrammer. Evalueringens oppgave var da å besvare spørsmål om hva som virker og hva som ikke virker. Senere har evalueringens funksjon i allokering av ressurser vært mer dominerende. Spørsmål om formålseffektivitet og kostnadseffektivitet var på 80- og 90-tallet klart de mest sentrale som skulle besvares gjennom evalueringer. Evalueringens funksjon som legitimering har kommet sterkere til syne det siste tiåret. Det er enklere å bruke evaluering for å bevise at noe fungerer enn det motsatte.

Avslutning

I dette lille tilbakeblikket på utviklingen de siste tiårene ser vi altså at det har vært en enorm utvikling innen evalueringens virksamhet – både internasjonalt og i Norge. Det er spesielt tydelig i den sterkt økende interessen for evalueringer. Et uttrykk for dette er etableringen av Norsk evalueringsselskap. Noen av oss gjorde et mislykket forsøk på å etablere en norsk nasjonal evalueringsselskap tidlig på 90-tallet. Men nå – 15 år senere – har det endelig lyktes.

Vi har valgt å ha "god evaluering" som hovedtema for denne konferansen. Gjennom

Takk for oppmerksomheten!

