

ACTA UNIVERSITATIS UPSALIENSIS
Uppsala Housing and Urban Studies
2

Samtal om idéer, utvärdering och samhällsförändring

En vänbok till Evert Vedung

Redaktörer
Marianne Danielsson & Nils Hertting



UPPSALA
UNIVERSITET

2024

© Författarna 2024
Distribueras under licensen CC BY 4.0

ISBN 978-91-513-2008-3

Acta Universitatis Upsaliensis
Distributör: Uppsala universitetsbibliotek, Box 510, 751 20 Uppsala, Sverige
www.uu.se/bibliotek, uup@uu.se
<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-520419>

Tryck: Stibo Complete, Horsens, Danmark
Omslagsfoto: David Naylor
Porträttfoto: Mikael Wallerstedt
Sättning och omslag: Ulrika Wahlberg

INNEHÅLL

FÖRORD

| | |
|--|---|
| En vänbok till en vän av det goda samtalet <i>Marianne Danielsson & Nils Hertting</i> | 9 |
|--|---|

OM ANALYS AV IDÉER I PRINCIP OCH PRAKTIK

KAPITEL 1

| | |
|---|----|
| Aversionen mot demokratisk bearbetning <i>Sverker Gustavsson</i> | 17 |
|---|----|

KAPITEL 2

| | |
|---|----|
| Evert, Skytteanum och jag: Ett personligt möte och en personlig utveckling <i>Mats Lindberg</i> | 29 |
|---|----|

KAPITEL 3

| | |
|--|----|
| Taxonomies in evaluation research and practice <i>Kim Forss</i> | 41 |
|--|----|

KAPITEL 4

| | |
|--|----|
| Forskningens relevans, användbarhet, användning och nytta – 30 år senare <i>Verner Denvall</i> | 51 |
|--|----|

KAPITEL 5

| | |
|--|----|
| Idéanalys och utvärdering <i>Leif Lewin</i> | 65 |
|--|----|

OM UTVÄRDERINGENS ANSPRÅK, ANSATSER OCH ANVÄNDNING

KAPITEL 6

- Behöver evaluering ikke længere demokratiet? 73
Peter Dahler-Larsen

KAPITEL 7

- Utvärdering och annan kunskap 87
Jan-Eric Furubo

KAPITEL 8

- An evaluation manager perspective – inspiration from Evert 99
Per Øyvind Bastøe

KAPITEL 9

- Theory-Based Stakeholder Evaluation Revisited: 107
An Exploratory Review
Morten Balle Hansen

KAPITEL 10

- Evalueringsbølgers dynamik 119
Hanne Foss Hansen

KAPITEL 11

- De andra diffusionsvågorna: Utvärdering i förändring 133
Wolfgang Meyer & Reinhard Stockmann

KAPITEL 12

- Att använda eller inte använda – det är frågan 147
Per Mickwitz

KAPITEL 13

- Taking side-effects seriously 159
Nicoletta Stame

KAPITEL 14

| | |
|--|-----|
| I huvudet på en utvärderare <i>Rolf Sandahl</i> | 169 |
|--|-----|

OM SAMHÄLLSSTYRNING OCH SAMHÄLLSFÖRÄNDRING

KAPITEL 15

| | |
|---|-----|
| Politiska intentioner och utvärdering i kohandelns tid: En pågående dialog med Evert Vedung <i>Bo Bengtsson</i> | 181 |
|---|-----|

KAPITEL 16

| | |
|--|-----|
| Typologies of policy instruments: 25 years of developments and challenges <i>Frans L Leeuw</i> | 193 |
|--|-----|

KAPITEL 17

| | |
|---|-----|
| Mångtydiga idéer och den haltande implementeringens rationalitet – fallet deltagandebaserad samhällsplanering <i>Marianne Danielsson & Nils Hertting</i> | 207 |
|---|-----|

KAPITEL 18

| | |
|--|-----|
| Rationalitet, energipolitik och policyprocess <i>Ylva Norén Bretzer</i> | 219 |
|--|-----|

KAPITEL 19

| | |
|--|-----|
| I framtidens skugga – om tiden i implementeringsprocesser <i>Stefan Sjöblom & Sebastian Godenhjelm</i> | 231 |
|--|-----|

KAPITEL 20

- Lämna mig inte! Några reflexioner kring
internationella skilsmässor 243
Li Bennich-Björkman

ETT YRKESLIV OCH EN YRKESBANA

KAPITEL 21

- Evert Vedung – en forskarbana med statskunskap,
utvärdering och oväntade händelser 253
Joakim Tranquist

KAPITEL 22

- Om serendipiteter och 269
Svenska utvärderingsföreningen
Sanja Magdalenic

KAPITEL 23

- Intervention theory and outcomes of 283
an international teaching collaboration:
A personal reflection from Brazil
Luiz Pedone

KAPITEL 24

- Evert – et eksempel på den gode mentor i evaluering 295
Olaf Rieper

KAPITEL 25

- Serendipity in place 301
Terry Hartig

- OM FÖRFATTARNA 307

FÖRORD

En vänbok till en vän av det goda samtalet*Marianne Danielsson & Nils Hertting*

Evert Vedung är en samtalets man. I hans närhet är konversationen alltid prioriterad – som nöje och som metod. Det som nyss bara var en passage och en halvslapp formulering kan plötsligt fokuseras och redas ut. "Vi kan väl prata om det där en stund", säger Evert, och vi blir påmindna om att i detta sällskap antas den trivsamma pratstunden alltid ha vetenskaplig potential. Byggstenarna finner vi i Everts logiska skolning, breda beläsenhet, historiska intresse och sinne för detaljer. För oss som fått möjlighet att vara med är det ett privilegium – ett slags samtalets fest.

Det var ett tag sedan Evert Vedung formellt sett gick i pension – faktiskt närmare två än ett decennium sedan. Men det hindrar inte Evert från att vara fortsatt aktiv som skribent och forskare, och som iakttagare och tolkare. Vi som tagit initiativ till denna bok menar att det nu är hög tid för en festskrift som uppmärksammar Evert och Everts långa forskargärning. Inte för att forskargärningen är slut, det ser vi inga indikationer på, utan för att vi så länge önskat tacka och hylla hans insatser, både som forskare, kollega och som mentor, på vitt skilda forskningsområden och i vitt skilda delar av världen. Vi har tänkt att det bästa sättet att göra det på är att erbjuda ett smörgåsbord av inlägg i diskussioner som Evert på ett eller annat sätt är inblandad i, och kan tänkas vilja fortsätta.

I boken samlas bidrag som knyter an till Everts forskning på områden såsom utvärdering och utvärderingsmetod, styrning och implementering, idéanalys, samhällsplanering och energipolitik, unionsupplösningar och secessioner, samt mer personliga reflektioner om personen, kollegan och mentorn Evert Vedung. Vi har därför kallat boken *Samtal om idéer, utvärdering och samhällsförändring: En vänbok till Evert Vedung*.

Volymens första del – *Om analys av idéer i princip och praktik* – inleds med *Sverker Gustavssons* kapitel om vår tids populistiska tendenser. I kontrast till populismens *Aversion mot demokratisk beredning* (bidragets titel) påtalar han vikten av vad han kallar demokratisk bearbetning innan ett beslut går till votering, och efterlyser en idéanalys som inte bara tar sakfrågan utan också andra positioner i det politiska landskapet i beaktande. I nästa kapitel knyter *Mats Lindberg* under titeln *Evert, Skytteanum och jag* direkt an till Everts betydelse för Lindbergs egna intellektuella utveckling, samtidigt som texten också utgör en analys av och positionering i ett bredare samtal om den idéorienterade samhällsvetenskapens förutsättningar. I de två nästkommande bidragen kopplar texterna till Everts mer konkreta insatser för att bringa begreppslig reda på specifika områden. *Kim Forss* tar sig an frågan om taxonomier i sitt bidrag *Taxonomies in Evaluation Research and Practice*: Vad är en bra taxonomi, och hur de taxonomier som görs bli bättre? *Verner Denvall* återvänder i *Forskningens relevans, användbarhet, användning och nytta* till Everts undersökning av vad det betyder att forskning ska vara "relevant". I kapitlet problematiseras fokuset på användbarhet och det avslutas med en beskrivning av "relevansparadoxen" – en situation där mer evidens och kunskap än någonsin lever sida vid sida med "fake news" och kunskapsförakt. Under den lakoniska titeln *Idéanalys och utvärdering* avslutar *Leif Lewin* den första delen av boken genom att lyfta fram ett släktskap mellan Everts tankar om idéanalys genom det rationella samtalets lins och hans bidrag inom utvärderingsfältet.

I bokens andra del hittar vi nio bidrag om *utvärderingens anspråk, ansatser och användning*. Först ut är *Peter Dahler-Larsen* som i sitt bidrag ställer den finurliga frågan om utvärdering inte längre

behöver demokrati – *Behøver evaluering ikke længre demokrati?* Mer precist tar sig Dahler-Larsen an de normativa och empiriska grundvalarna och den empiriska realismen i Everts, mer eller mindre explicita, koppling mellan utvärdering och den representativa demokratins styrningskedja. Därefter skriver *Jan-Eric Furubo* om utvärderingens särskilda karaktär som kunskapsform. I sitt kapitel *Utvärdering och annan kunskap* knyter han särskilt an till en mekanismcentrerad syn på (kausal) kunskap och lyfter dess implikationer för utvärderingars användning. *Per Øyvind Bastøe* bidrar med en reflektion om utvärderingsanvändning med ett managementperspektiv – *An evaluation manager perspective – inspiration from Evert* – som samtidigt ger en inblick i norsk diskussion om utvärdering.

I sitt kapitel *Theory-Based Stakeholder Evaluation Revisited* återvänder *Morten Balle Hansen* till hans och Everts artikel om ”teori-baserad intressentutvärdering” och undersöker hur deras modell har kommit till användning. En slutsats är den använts på olika sätt och på vitt skilda politikområden men förefaller vara som mest fruktbar när den offentliga insatsen bygger på *co-creation* på tvärs professioner och organisationer. *Hanne Foss Hansens* kapitel, *Evalueringens dynamik*, knyter an till Everts idé om hur utvärderingstrender sköljt över oss i mer eller mindre distinkta vågor. Med hjälp av *Kjell Arne Røviks* teori om idédiffusion kompletteras Everts begreppsapparat och ansatsernas fruktbarhet undersöks och illustreras med utvecklingen av centrala undersökningar av elevers färdigheter i den danska folkskolan. Också *Wolfgang Meyer* och *Reinhard Stockmann* knyter an till Everts böljemetamorfi. I sitt – på svenska egenskrivna – kapitel *De andra diffusionsvågorna* följer de upp Everts beskrivning och frågar sig vad som egentligen hänt därefter.

I kapitlet *Att använda eller inte använda – det är frågan* sätter *Per Mickwitz* fokus på utvärderingens användning genom tre korta fallstudier. Användning av utvärdering, föreslår *Mickwitz*, främjas bland annat av att framtida användare dras in redan i utvärderingsprocessen. I anknytning till Everts resonemang och teoretiserande om olika typer av utvärdering och utvärderingsmodeller

lyfter *Nicoletta Stame* i sitt bidrag *Taking side-effects seriously* fram vikten av att uppmärksamma bieffekter i utvärdering. Genom att låta Everts begreppsutredningar kring olika utvärderingsmodeller möta andras, utreds och utvecklas i kapitlet viktiga distinktioner kring effekter i och av offentlig politik. *Rolf Sandahl* avslutar denna del av boken med att reflektera över vad som egentligen är så särskilt och säreget med utvärdering. I sitt bidrag *I huvudet på en utvärderare* antyds att svaret inte är uppenbart och att det har blivit allt svårare att skilja utvärdering från andra utvärderande verksamheter på ett strikt vis.

Del tre, *Om samhällsstyrning och samhällsförändring*, inleds med *Bo Bengtssons* kapitel *Politiska intentioner och utvärdering i kohandelns tid*. I sitt bidrag problematiserar Bengtsson en kritik av måltvärdering som inte minst Evert utvecklat, och lanserar samtidigt sin egen modell för utvärdering utifrån institutionaliserade intentioner. Kapitlet är ett bidrag till utvärderingslitteraturen som särskilt ägnar uppmärksamhet åt de utvärderingsmetodologiska implikationerna av en samhällsstyrning präglad av uppgörelser och kohandel illustrerade utifrån samtida hyrespolitik. I sitt kapitel tar sig *Frans L. Leeuw* an ett av Everts mer uppmärksammade bidrag – typologin över styrmedel med ”piskan”, ”moroten” och ”predikan”. I kapitlet *Typologies of policy instruments: 25 years of developments and challenges* diskuterar Leeuw begränsningar och utmaningar med typologier generellt och typologier över styrmedel i synnerhet. Särskilt argumenterar han för värdet av genomtänka typologier när de kombineras med (förklarande) teorier. *Ylva Norén Bretzer* ger i sitt bidrag *Rationalitet, energipolitik och policyprocess* en noggrann översikt av Everts bidrag till svensk energiforskning. På köpet blir kapitlet en slags argumentation för vikten av, och fruktbarheten i, statsvetenskapliga perspektiv på energifrågan och på policy och policyprocesser mer generellt.

I *Stefan Sjöbloms* och *Sebastian Goldenhjems* kapitel, *I framtidens skugga – om tiden i implementeringsprocesser*, får vi ytterligare ett kapitel som knyter väl an till Everts forskningsagenda. I kapitlet ställs särskilt tidsaspektens betydelse för implementering i fokus. I anslutning till Everts egna bidrag betonas vikten av anteciper

och fenomenet "anteciperande styrning" preciseras och utreds. I bidraget *Mångtydiga idéer och den haltande implementeringens rationalitet* närmar sig också *Marianne Danielsson* och *Nils Hertting* genomförandefrågan. Närmare bestämt tar kapitlet sig an att formulera ett, i den vedungska idéanalysen förankrat, förslag till varför idéen om deltagandebaserad samhällsplanering trots sin popularitet tycks notoriskt svår att implementera. I det sista kapitlet i denna sektion – *Lämna mig inte! Några reflektioner kring internationella skilsmässor* – tar sig *Li Bennich-Björkman* an de riktigt stora samhällsförändringarna. I kapitlet drar hon utifrån Everts doktorsavhandling om unionsupplösningen mellan Norge och Sverige från 1972 slutsatser om Ukrainas och Rysslands förhållande efter Sovjetunionens fall, såsom den ter sig med dagens krig som sorglig relief. Sättet att skriva om och samtala med varandra i ett "rationellt politiskt samtal" lyfts som en särskiljande och kritisk faktor.

I bokens sista del möter vi ett antal betraktelser som från olika perspektiv kastar ljus över *Everts yrkesliv och yrkesbana*, i flera fall med personlig vinkel. *Joakim Tranquists* bidrag – *Evert Vedung – en forskarbana med statskunskap, utvärdering och oväntade händelser* – bygger på en intervju med Evert om hur hans intresse för utvärdering föddes och utvecklats. I kapitlet hörs Everts egen stämma tydligt. *Luiz Pedones* beskrivning av hur han och Evert kom att arbeta tillsammans i Brasilien och Sverige ger bilden av en nyfiken forskare som tog amerikanska idéer hem till Sverige. Pedones kapitel *Intervention theory and outcomes of an international teaching collaboration* framhåller hur olika forsknings- och utbildningssamarbeten har lett Evert ut i världen, och inte bara längst de mest upplöjda stigarna.

Olaf Rieper ger i sitt kapitel *Evert – et eksempel på den gode mentor* en personligt hållen beskrivning av hur viktigt det kan vara att ha en ledsagare och mentor, samtidigt som vi får en inblick i den danska utvärderingsmiljön och svensk-danska utvärderingskopplingar. *Sanja Magdalenic* har i sitt bidrag – *Om serendipiteter och Svenska utvärderingsföreningen* – grävt i arkiven för att kunna ge en bild av hur det gick till när Svenska utvärderingsföreningen bildades, och Everts roll i det.

Terry Hartig, slutligen, bidrar i *Serendipity in place* med en betraktelse från livet som Everts långvariga kollega på Institutet för bostads- och urbanforskning (IBF). Det är en personlig betraktelse av kollegan Evert, som gärna låter den kollegiala samvaron och det personliga mötet gå hand i hand, på och utanför arbetsplatsen. Det är en bild som också säger något mer allmängiltigt om personen Evert. Kanske kan man därtill säga att vi, med alla bidrag lagda intill varandra, får lite akademisk samtidshistoria. Den understryker i så fall att den svenska universitetsvärlden inte bara skapas av inåtvända tänkare och drillade elitister. Den kan också vara en fruktsam bas för en driftig, samhällstillvänd entreprenör i idéer.

Såväl Evert själv som flera av de medverkande i denna volym lyfter gärna fram tillfälligheternas betydelse för hur det blev som det blev. Och det är en viktig poäng. Men det faller sig minst lika naturligt att beskriva Everts gärning i termer av medvetna val, med ganska tydlig och egen kompassriktning. Han har ofta varit före sin tid. Han var en aktiv nätverkare långt innan det var vanligt. Han har knutit kontakter med stora internationella stjärnor under ett långt yrkesliv men lika mycket vinnlagt sig om att knyta och hålla kontakter med studenter och stretande yngre kollegor, och delat med sig av både kontakter och möjligheter. Vi är många som har Evert att tacka för att ett första utvärderingsuppdrag kom i vår väg.

Akademikern Evert hade också tidigt goda och täta kontakter med förvaltning och praktiker, långt innan det sågs som meriterande i den akademiska miljön. Evert bidrog till att studenter som läste hans kurser fick en idé om hur yrkesarbete med utvärdering och analys av styrning och implementering kunde te sig. Evert var också tidigt en aktiv kommunikatör av sina forskningsresultat, inte bara inom akademien utan också utanför den. Ett sätt att uttrycka detta är att Evert konsekvent och nyfiket riktat sig utåt, mot sin omgivning i vid mening, och på detta sätt skapat osedvanligt många personligt färgade och vänskapliga professionella kontakter. Vi som medverkar i denna bok utgör, vågar vi lova, blott en mindre del av detta vittförgrenade nät.

Med ett kontaktnät som Everts följer det närmast per definition att alla som kanske skulle vilja vara med inte har fått frågan – det beklagar vi! Andra är med i tanken, fast möjligheten att skriva inte funnits. En given medverkande i en festskrift till Evert Vedung är Everts vän och kollega Gustav Petersson. Gustav var också initialt en av redaktörerna. Fast han efter uppstartsfasen varit tvungen att avstå från att medverka i arbetet har Gustavs initiala insatser för att entusiasmera medverkande, och hans engagemang för att äntligen ge Everts gärning en välförtjänt hyllning, varit viktiga för färdigställandet av boken. Vi vill även särskilt tacka Ulrika Wahlberg, IBF, för praktiskt stöd och engagemang i detta arbete.

Fler skulle kunna nämnas, men vi tors inte börja tänka på hur lång den listan skulle kunna bli. Vi vågar dock påstå att Evert med de flesta måttstockar nått ovanligt många, i ovanligt många olika sammanhang, på ovanligt många ställen i världen.

Avslutningsvis vill vi därför bara säga, tack Evert!

Samtalet fortsätter.

Uppsala, januari 2024

Marianne Danielsson & Nils Hertting

KAPITEL 1

Aversionen mot demokratisk bearbetning*Sverker Gustavsson*

När världsläget var som mörkast hösten 1940 dristade sig ett antal demokratiskt sinnade svenska universitetslärare, under ledning av Gunnar Aspelin och Göte Turesson, att publicera en brett upplagd samlingsvolym. Kapitlen behandlade den innehållsliga problematiken inom snart sagt alla akademiska ämnen. Bidraget om statskunskap var skrivet av Herbert Tingsten. Dennes centrala påpekande var att grundfrågor, som länge hade behandlats i en atmosfär av sunt förnuft och kortfristiga synpunkter, på nytt hade framträtt som levande och väsentliga.

Vad detta innebar hade Tingsten året innan, i förordet till sin bok om konservatismens idéer, beskrivit som att uppgiften var att genom kritik höja nivån på debatten. Även i fråga om politiskt omstridda ståndpunkter borde en rationalisering av meningsutbytet vara syftet. Det fanns ingen anledning, hade han framhållit, varför sådana utsagor skulle vara befriade från kritisk prövning bara därför att de var kontroversiella.

Till detta argument återknöt Evert Vedung, då han 1977 publicerade sin bok om det rationella politiska samtalet. Den fråga som han ställde sig var hur politiska budskap mest fruktbart kan tolkas, ordnas och prövas. Sedan dess har kollegerna Björn Badersten,

Ludvig Beckman, Stefan Björklund och Mats Lindberg bidragit till ytterligare tankereda och uttrycksskärpa om vad som bör menas med kritik och rationalitet. Kraven på stringens och precision har med tiden blivit allt bättre tillgodosedda (Vedung, 2018).

Dagens politiskt mörknande perspektiv är inte utan likheter med situationen 1940. Vad som fortsatt frapperar är hur föga mönstergillt de globalt, europeiskt och nationellt stora frågorna hanteras. Det vi ser är inte bara en teknokratiskt utan också en populistiskt motiverad *aversion* mot demokratisk bearbetning. Kring denna dubbla motvilja finns det skäl att närmare reflektera. Jag redovisar här tre försöksvis formulerade tankar om varför partier, organisationer och redigerade media trängs tillbaka från uppgiften att främja växelverkan mellan stat och samhälle.

Först uppmärksammar jag den motvilja mot demokratisk bearbetning, som tillfördes svensk politik under senare delen av nittonhundratalet. Parlamentariskt sammansatta utredningar blev då alltmera sällsynta företeelser. På motsvarande sätt inskränktes möjligheten att i riksdagen förhandla om regeringens förslag till statsbudget. Min andra reflexion avser populism och teknokrati som principiella alternativ till den ”normalt” fungerande demokratin. Min tredje undran gäller vad som skiljer en demokratisk bearbetning från en uteslutande sakfrågestyrd idéanalys. Vilken är den springande punkten?

FRÅN BEARBETANDE TILL VOTERANDE PARLAMENTARISM

Något som kan förbrylla en internationell betraktare är varför det svenska statskicket har undgått att bryta samman. Av historiska skäl harmonierar inte sättet att välja riksdag med sättet att bilda regering. Regeringssättet gjordes år 1917 parlamentariskt, trots att valsättet tio år tidigare hade gjorts proportionellt. En politisk ordning med flera i stället för två partier blev följd. Nästan utan undantag har landet som följd av detta systemfel under mer än hundra år kommit att styras av regeringar utan egen majoritet (Lewin, 2020).

Ett nära till hands liggande svar på frågan om det uteblivna sammanbrottet hänvisar till den breda uppslutningen. Under mellan-

krigstiden, andra världskriget, det kalla kriget och i efterdyningarna av nittiotalskrisen, pandemin och angreppet på Ukraina råder nationell samling. Genom att vädja till det upplysta egenintresset har regeringar utan majoritet kunnat vinna tillräckligt stöd för utrikespolitiken, välfärdsstaten, unionsmedlemskapet och partsautonomin på arbetsmarknaden. Först sedan grundstämningen blivit aggressiv har förutsättningarna för brett samförstånd minskat.

En sådan historieskrivning förbiser hur motsättningar faktiskt hanteras. Uppslutningen som sådan är inte förklaringen. Frågan är vad aktörerna gör för att främja uppslutningen. Sättet att överbrygga mål- och intressekonflikter är det väsentliga. Slående med tanke på detta är hur parlamentarismen med tiden övergår från att vara i första hand *bearbetande* till att vara företrädesvis *voterande*. Förloppet under 1900-talet var i korthet följande:

För att medverka till rösträttsutvidgningen 1907 ställde högern som villkor att mandaten skulle fördelas proportionellt och inte som tidigare efter majoritetsval i enmansvalkretsar. Högern ville på det sättet säkerställa att Sverige inte blev ett tvåpartisystem dominerat av liberaler och socialdemokrater. Detta förbehåll gjorde att Sverige blev ett land med fler än två partier och där det blev svårt att bilda regering.

Tio år senare tvingades Gustaf V acceptera principen om parlamentarism, det vill säga att regeringen inte skulle förutsätta kungens utan riksdagens förtroende. Ministärer utan majoritet blev därigenom regel snarare än undantag. Lärobokslösningen på problemet att frigöra sig ur detta dilemma är majoritetsval. En sådan omläggning går praktiskt inte att åstadkomma så länge det anses böra råda partipolitisk enighet om sättet att välja riksdag. De mindre partierna låter sig inte godvilligt utplånas.

Bristen på samstämmighet mellan sättet att välja riksdag och sättet att utse regering doldes länge av två för svensk politik under större delen av nittonhundratalet karaktäristiska företeelser. Jag syftar på förekomsten av parlamentariskt sammansatta utredningar och att regeringens förslag till statsbudget realbehandlades av partierna i riksdagen.

Stora frågor efter 1917 blev utrikespolitiken, författningen, försvaret, arbetslösheten, bostäderna, skatterna, skolan, pensionerna, socialförsäkringarna och regionalpolitiken. Målkonflikter och intressekonflikter inom dessa områden stöttes och blöttes i mångårigt verksam utredningar med företrädare för samtliga partier jämte kommuner, regioner och organisationer.

De partier som inte satt i regeringen fick härigenom ett inflytande, som svarade mot deras andel av valmanskåren. Lika viktigt var att företrädare för skilda idéer och intressen blev förtrogna med hur problemen på de olika områdena sakligt och politiskt var beskaffade. Det framgick att politik inte bara handlar om maktinnehav utan också om sakliga skillnader.

Samma inverkan hade att regeringens förslag till budget fram till 1997 avhandlades som en öppen fråga av partierna i riksdagen. Bearbetningen kunde modifiera vad regeringen föreslog. Utöver den effekt som uppnåddes genom tvärpolitiska utredningar var hanteringen av budgeten ägnad att höja och hålla uppe den intellektuella nivån.

Det omständliga arbetet med att genom utredningar och budgetförhandlingar förankra och nysansera bidrog till att i partier, kommuner, regioner och organisationer bredda intresset för möjliga framkomstvägar. Båda dessa korrektiv – parlamentariska utredningar och realbehandling av budgeten – kan antas ha räddat det historiska försöket att förena proportionalism med parlamentarism. Spänningen mellan de båda grundprinciperna reducerades därigenom.

Poängen som jag här vill göra är att betydelsen av dessa båda mildrande korrektiv från och med 1980-talet kom att kraftigt reduceras. Idén att tvärpolitiskt utreda fortlever i dag bara i fråga om författnings- och säkerhetspolitiken. Vare sig bostadsbristen, skatterna, sjukvården, arbetslösheten, socialförsäkringarna, skolan eller migrationen bearbetas längre inom ramen utredningar av detta slag.

Också riksdagens realbehandling av budgeten har avvecklats. Budgetlagen från 1997 föreskriver att varje riksmöte inleds med en röstning om saldot. Behandlingen blir därmed inte längre

tvärpolitiskt bearbetande. Meningsskiljaktigheter kan inte längre preciseras och resultera i differentierade kompromisser med inflytande också för oppositionen. Allt ställs på sin spets i en avgörande votering.

Eliminerandet av de båda korrektiven förtjänar att närmare beskrivas och förklaras. Räcker det att hänvisa till den tidsanda som började göra sig gällande under nittonhundratalets sista tredjedel? Eller berodde det framför allt på att idén om demokratisk bearbetning kom att bristfälligt försvaras? Mest fruktbart är att betona den senare synpunkten.

Det nyliberala tänkande som bröt igenom kring 1980 rekommenderade inte bara ett skifte från Keynes till Friedman som vägvisare för den ekonomiska politiken. Lika viktigt var att ta fasta på vad Friedrich Hayek redan 1944 hade framhållit om den bearbetande välfärdsstaten som "vägen till trälldom".

Blandekonomin var enligt Hayek inte någon symmetrisk kompromiss mellan marknad och demokrati utan ett sluttande plan. Likt den trojanska hästen tjänade den allmänna rösträtten som inkörsport för en alltför stor offentlig sektor. Följaktligen gällde det att inte praktisera en tvärpolitiskt upplagd bearbetning. Att förhandla om vilken som var bäst avvägning mellan demokrati och marknad var missriktat. Bara genom restriktivitet på den punkten kunde systemet skaka av sig betydelsen av särintressen. Vad som var allmänintresset borde vara en fråga för experter att avgöra.

Historiskt viktigare än inverkan på ekonomin blev inverkan på politiken. De nyliberala föregångsmännen stack aldrig under stol med vad de ansåg om blandekonomin (Burgin, 2012, Jones, 2012, Gerstle, 2023). Breda uppgörelser borde undvikas. Den utbredda acceptansen av denna rekommendation kom att underminera den kritiska infrastruktur (Müller, 2021, s. 90 ff), som den politiska liberalismen i form av yttrandefrihet, föreningsfrihet, rättsstat och principen om legitim opposition möjliggör och förutsätter.

TEKNOKRATI OCH POPULISM

Problematiken kring vad som i våra dagar omtalas som teknokrati och populism grundläggs under årtiondena efter 1848. Alltsedan

dess har frågan varit hur de folkliga kraven på breddat ansvarstagande skall kunna tillgodoses utan att detta får någon nämnvärd saklig inverkan.

Oron för det vanstyre som kunde tänkas följa av ett växande folkligt deltagande troddes efter nedkämpandet av 1848 års revolutionära tendenser kunna stillas. Det hade nämligen visat sig att inte bara liberaler och socialister var att räkna med. Med Napoleon III i spetsen visade sig även konservativa grupper kunna utnyttja principen om majoritetsstyre. Frågan blev hur nationalismen och kritiken mot industrialismen skulle kunna inordnas i ett system präglad av rösträttsutvidgningar (Tingsten, 1933, Sternhell, 1996, Berman, 2019).

Mot denna bakgrund försöker socialliberalt, kristdemokratiskt och socialdemokratiskt sinnade politiker, opinionsbildare och intresseföreträdare på ett "mittenorienterat" sätt under de följande etthundrasjuttiofem åren undgå såväl teknokrati som populism. Demokratisk bearbetning blev den huvudsakliga metod, varmed målkonflikter och intressemotsättningar hanteras. Särskilt under mellankrigstiden och de trettio första åren efter andra världskriget handlade debatt och politik i Frankrike, Italien och Tyskland om hur "den breda mitten" mest effektivt kunde slå vakt om demokratin som en förening av majoritetsstyre och politisk liberalism (Müller, 2011, Hacke, 2018, Biebricher, 2023).

Den statsvetare som i våra dagar mest uppfordrande beskriver hur strukturen av partier, organisationer och redigerade media pressas tillbaka från uppgiften att vara en bearbetande mellaninstans, är den i Leiden i Holland och vid det europeiska universitetet i Florens verksamme irländaren Peter Mair. Det sker i hans 2013 postumt utgivna bok om hur växelverkan mellan stat och samhälle utarmas i takt med att inslaget av pluralistisk bearbetning bortfaller. Därigenom öppnas ett vakuum, som låter sig fyllas av populistiska och teknokratiska förhållningssätt.

Utrymmet för växelverkan mellan stat och samhälle är enligt Mair på väg att fyllas av populism och teknokrati som varandra ömsesidigt förstärkande krafter. Populisten säger att partier, organisationer, domstolar, förvaltningar och fritt arbetande medier

hindrar den verkliga folkmeningen från att sätta sin prägel på lagstiftning och budgetering. Den som invänder är inte lojal mot folket i singularis. Teknokraten säger att inslaget av populistisk bearbetning hindrar expertisen från att få sista ordet. Den som invänder är inte lojal mot idén om vetenskap och beprövad erfarenhet. Båda menar sig företräda allmänintresset i kamp med olika särintressen.

Sedan 2013 har den av Mair hävdade synpunkten kommit att ytterligare underbyggas. Det sker i två steg. Först framställs populism och teknokrati som *konträrt* motsatta begrepp (Urbinati, 2014, Caramani, 2017). Enligt detta första steg befinner sig bearbetningen under dubbel attack i vad som liknar ett tvåfronts-krig. Det följer av att populism och teknokrati framställs som varandras motsatser. Båda anses ha spetsen riktad mot partier, organisationer, förvaltningsapparater och redigerade media. Men det rör sig ännu så länge om motstående krafter. Populisten vill rädda demokratin från expertvälde. Teknokraten ser som sin uppgift att rycka undan mattan för populismen.

Det andra steget togs några år senare. I stället för att tolka populism och teknokrati som konträrt motstående krafter bör vi ta fasta på vad dessa gemensamt angriper. Det är vad de senast anlända bidragsgivarna hävdar. Tendenserna till teknokrati och populism är inte motsatta utan *komplementära*. Båda är uttryck för samma aversion. Båda är inriktade på att undanröja den sakliga och politiska betydelsen av mellankommande instanser.

Dagens brittiska, franska och italienska politiker håller på att etablera en ny form för politisk bearbetning. Den är vad engelsmannen Christopher Bickerton och italienaren Carlo Invernizzi Accetti kallar "teknopopulism". Det sker i en bok från 2021 med den titeln. Denna nya metod präglas av att populism och teknokrati inte bara förstärker varandra ömsesidigt utan *förenas* i sin motvilja mot mellaninstanser. Följden blir att betydelsen av bearbetande partier, organisationer och media förflyktigas. Tomrummet fylls av opinionsundersökningar, direktkontakt med väljarna, anlitan av PR-byråer och ett ostentativt hänvisande till specialiserad expertis. Att partier, organisationer, förvaltningar

och opinionsbildare i största möjliga mån hålls borta från bearbetningen är det centrala för politiker av det slag som i dag dominerar i brittisk, fransk och italiensk politik (Bickerton & Accetti, 2021).

Enligt Bickerton och Accetti är en total sammansmältning av populism och teknokrati nära förestående. En sådan fusion ter sig både lovande och skrämmande. Lovande för den som tror och hoppas, att en fördjupad medvetenhet kan hejda en ytterligare minskning av utrymmet för bearbetning. Skrämmande för den som tror och hoppas att idén om att kunna gjuta samman majoritetsstyre och politisk liberalism inte skall visa sig ha varit förgäves.

DEN SPRINGANDE PUNKTEN

Min tredje undran tar sikte på vad som skiljer bearbetning från idéanalys. Innebörden av detta förstnämnda begrepp tydliggjordes under mellankrigstiden av den österrikiske juristen och statsvetaren Hans Kelsen. Det skedde med polemiken riktad mot dennes tyska kollega Carl Schmitt. För den senare var den avgörande poängen att folket bara finns i singularis. Enligt denna renodling var kompromisser vare sig möjliga eller önskvärda. Varje idé om gemensamma fördelar var illusorisk. Vän stod alltid mot fiende (Gustavsson, 2022a).

Mot idén om demokratin som renodlat majoritetsstyre – utan vare sig yttrandefrihet, föreningsfrihet, rättsstat eller legitim opposition – vände sig Kelsen. Genom sin bok om demokratins väsen och värde blev han den ledande företrädaren för idén om folkstyret som en oupplöslig förening av politisk liberalism och majoritetsstyre. Hans kritik riktade sig inte bara mot diktaturen utan också mot populismen i den av Schmitt företrädade meningen. Att demokratin förutsätter politisk liberalism var enligt denne en historisk villfarelse.

Vad som i våra dagar leder till en *ökad* aversion mot bearbetning är lätt att föreställa sig. Det är att medvetet höja konfliktnivån genom att aktivt avhålla sig från tolerans och måttfullhet som "skyddsräcken" (Levitsky & Ziblatt, 2018, 97 ff) för möjligheten av fortsatta samtal. Det sättet att förhålla sig rycker undan grunden för ömsesidig respekt i politiken, förvaltningen, rättsväsendet, den

allmänna debatten och det specialiserade yrkeslivet (Gustavsson, 2020).

Svårare är att vända på resonemanget och söka svar på frågan om vad som leder till en *minskad* aversion mot bearbetande mellaninstanser. Vad finns att göra för att Kelsen mer än Schmitt skall komma att sätta sin prägel på det framtida tänkandet och handlandet? Mitt svar är att vi – till att börja med – behöver göra oss klara över vad som utmärker en demokratisk och därmed principiellt tvärpolitisk bearbetning av en problematik jämförd med en analys som bortser från att flera aktörer och andra sakfrågor också finns med i bilden.

Enligt Vedung, Badersten, Beckman, Björklund och Lindberg börjar en idéanalys med att sakfrågan tolkas och preciseras. Det sker med stöd av ett för undersökningsuppgiften fruktbart avpassat analyschema. Exempel på sådana scheman är – enligt vad jag själv som föreläsare brukar föreslå – skuggteorin hos Robert Dahl, trilemmat hos Jon Elster och fyrgruppsteorin hos Don K. Price. Den sortens scheman klargör problem, ståndpunkter och argument om såväl verklighetsomdömen som värderingar och rekommendationer (Gustavsson, 2022b).

Skuggteorin hos Dahl tydliggör, att statskunskapen inte har något vetenskapligt grundat svar på frågan om vilka som är folket, vilken som är rätt avvägning mellan förtroendevalda och anställda samt inte heller om vad som är syftet med sammanslutningen. Demokratins styrka – jämfört med diktaturen och populismen – är att den inte erbjuder några värdeneutrala svar (Dahl, 1989, s. 3 ff; Gustavsson, 2018). Bara genom bearbetning kan den ta upp kampen med kraven från skuggteorin om klara besked.

Likadant förhåller det sig med vad Elster säger om det trilemma – majoritetsstyre, rättsstat och effektivitet – som politiska aktörer har att hantera. Inget av kraven på folkligt stöd, rättsstatlighet och handlingskraft trumfar något av de båda övriga. Konflikter mellan dessa tre ambitioner antas fortlöpande behöva bearbetas i syfte att kunna uppnå provisoriskt hållbara lösningar (Elster, 1988, s. 1 ff).

På motsvarande sätt förhåller det sig med det antagande om fyra informellt sammanlänkade grupper som möter hos Price.

Problemet att komma fram till en praktiskt ändamålsenlig avvägning mellan förtroendevalda, tjänstemän, yrkesverksamma och forskare har inte heller någon statsvetenskapligt självklar lösning. Konstitutionell relativitet gäller enligt honom som grundprincip för all politisk organisation (Price 1965, s. 163 ff).

Det som gör statskunskapen speciell är att den i motsats till nationalekonomin inte syftar till någon i strikt mening systematisk och heltäckande lärobyggnad. Dahl, Elster och Price låter sina analyser av olika empiriska och normativa problem disciplineras av sina schematiska antaganden. Med deras hjälp tydliggör de vilka intressekonflikter och målkonflikter som finns att hantera. Därigenom kompenserar de för det faktum att statskunskapen inte är någon vetenskap i vare sig naturvetenskaplig eller normativt systembyggande mening. Sättet att arbeta hör hemma i kategorin humaniora. Förutsättningen är ett vetenskapsbegrepp, som rymmer alla tre sätten att förhålla sig.

Minsta gemensamma nämnaren för alla tre arbetssätten är *fallibilism*. Intagna ståndpunkter är provisoriska och förtjänar att upprätthållas till dess att meningsutbytet har givit stöd för att hävda någon annan uppfattning. Kraven på rationalitet och intersubjektivitet är desamma. Det gäller bara att det finns bättre skäl än de dittills anförda. Så upprätthålls skillnaden mellan att vara auktoritativ och att auktoritärt slå fast vad som gäller (Gustavsson, 2023, s. 346 ff).

Därmed har jag antytt vad som skiljer en demokratisk bearbetning från en uteslutande sakfrågestyrd idéanalys. Pudeln kärna och den springande punkten är att sättet att ställa problem, inta ståndpunkter och hävda argument inte bara gäller sakfrågan. Bearbetaren beaktar därutöver sin egen och andra aktörers förmåga att strategiskt (Lewin 2017, s. 28 ff) sätta makt bakom orden. Bearbetningen tar sikte på vad som går att åstadkomma och inte "bara" vad som från saklig synpunkt är den bästa lösningen.

Av denna insikt följer endera av två slutsatser. Den ena är att maktfördelningen sist och slutligen avgör. För en nyliberal eller marxistiskt inspirerad betraktare är "ren" idéanalys förspild

möda. Ty han eller hon "vet" redan att materiella faktorer sist och slutligen faller utslaget.

Enligt den andra slutsatsen – hävdad av Tingsten liksom av Vedung, Badersten, Beckman, Björklund och Lindberg – finns ingen annan sista instans än argumentets primat. Öppenhet om hur det faktiskt förhåller sig, vad som är eftersträvansvärt och vad som bör rekommenderas gäller såväl sakfrågan som strategin. En grundläggande aversion mot demokratisk bearbetning går då inte att försvara.

Referenser

- Badersten, B. (2006). *Normativ metod – att studera det önskvärda*. Lund: Studentlitteratur.
- Beckman, L. (2005). *Grundbok i idéanalys – det kritiska studiet av politiska texter och idéer*. Stockholm: Santérus.
- Berman, S. (2019). *Democracy and dictatorship – from the Ancien Régime to the present day*. Oxford University Press.
- Bickerton, C.J. & C. I. Accetti. (2021). *Technopopulism – the new logic of democratic politics*. Oxford University Press.
- Biebricher, T. (2023). *Mitte/Rechts – die internationale Krise des Konservatismus*. Berlin: Suhrkamp.
- Björklund, S. (1977). *Den uppenbara lösningen – om möjligheten till en objektiv ståndpunkt i värdefrågor*. Stockholm: Aldus.
- Burgin, A. (2012). *The great persuasion – reinventing free markets since the depression*. Harvard University Press.
- Caramani, D. (2017). Will vs. reason – populist and technocratic forms of political representation and their critique to party government. *American Political Science Review*, 111 (1): 54-67.
- Dahl, R.A. (1989). *Democracy and its critics*. Yale University Press.
- Elster, J. (1988). Introduction. I J. Elster & R. Slagstad (Red.), *Constitutionalism and democracy* (s. 1-18). Cambridge University Press.
- Gerstle, G. (2023). *The rise and fall of the neoliberal order – America and the world in the free market era*. Oxford University Press.
- Gustavsson, S. (2018). Skuggteorin tydliggör problemet. I S. Gustavsson, C.-M. Jonsson & I. Lindberg, *Vad krävs för att rädda demokratin?* (s. 161-202). Stockholm: Premiss förlag.
- Gustavsson, S. (2020). Hur demokratin överlever. I J. Heuman, J. Sandström & S. Widmalm (Red.), *Oumbärliga samtal – humaniora och samhällsvetenskap i svensk offentlig debatt* (s. 65-91). Göteborg: Daidalos.
- Gustavsson, S. (2022a). Demokrati, diktatur och populism. I K. Abrahamsson, P.-O. Jansson & T. Åkesson (Red.), *Demokratin som bildningsväg* (s. 79-91). Stockholm: Carlssons.
- Gustavsson, S. (2022b). Fyrgruppsteorin. I T. Bull, I. Mattson & J. Teorell (Red.), *Makt, politik, opinion – en vänskrift till Olof Petersson* (s. 71-87). Stockholm: Dialogos.

- Gustavsson, S. (2023). Öppenhet i stället för alternativlöshet. I L. Bennich-Björkman, S. Gustavsson & M. Lindberg (Red.), *Hålla huvudet kallt – om distanserat engagemang i en uppjagad tid* (s. 331-351). Göteborg: Daidalos.
- Hacke, J. (2018). *Existenzkrise der Demokratie – zur politischen Theorie des Liberalismus in der Zwischenkriegszeit*. Berlin: Suhrkamp.
- Hayek, F. A. (1944). *The road to serfdom*. London: Routledge & Kegan Paul.
- Jones, D.S. (2012). *Masters of the universe – Hayek, Friedman and the birth of neoliberal politics*. Princeton University Press.
- Kelsen, H. (1929). *Vom Wesen und Wert der Demokratie. Zweite umgearbeitete Auflage*. Tübingen: Mohr.
- Levitsky, S. & D. Ziblatt (2018). *How democracies die – what history reveals about our future*. London: Penguin.
- Lewin, L. (2017). *Ideologi och strategi – svensk politik under 130 år. Sjätte omarbetade och utvidgade upplagan*. Stockholm: Carlssons.
- Lewin, L. (2020). *Konsten att bilda regering när ingen har majoritet*. Göteborg: Daidalos.
- Lindberg, M. (2018). The VDP-triad in ideational analysis, I-II. *Statsvetenskaplig tidskrift*, 120 (2): 277-359 och (4): 435-554.
- Lindberg, M. (2023). Vetenskaplighet och normativ relevans – om distanserat engagemang som grundhållning. I L. Bennich-Björkman, S. Gustavsson & M. Lindberg (Red.), *Hålla huvudet kallt – om distanserat engagemang i en uppjagad tid* (s. 191-245). Göteborg: Daidalos.
- Mair, P. (2013). *Ruling the void – the hollowing of Western democracy*. London: Verso.
- Müller, J.-W. (2011). *Contesting democracy – political ideas in twentieth century Europe*. Yale University Press.
- Müller, J.-W. (2021). *Democracy rules*. London: Allen Lane.
- Schmitt, C. (1923). *Die geistesgeschichtliche Lage des heutigen Parlamentarismus*. München & Leipzig: Duncker & Humblot.
- Sternhell, Z., ed., (1996). *The intellectual revolt against liberal democracy 1870-1945. International conference in memory of Jacob L. Talmon*. Jerusalem: The Israel Academy of Sciences and Humanities.
- Tingsten, H. (1933). *Demokratiens seger och kris – den författningspolitiska utvecklingen 1880-1930*. Stockholm: Bonniers.
- Tingsten, H. (1939). *De konservativa idéerna*. Stockholm: Bonniers.
- Tingsten, H. (1940). Statsvetenskapliga arbetsuppgifter. I G. Aspelin & G. Turesson (Red.), *Vetenskap av i dag* (s. 194-215). Stockholm: Hugo Gebers förlag.
- Tingsten, H. (1941). *Idékritik*. Stockholm: Bonniers.
- Urbinati, N. (2014). *Democracy disfigured – opinion, truth and the people*. Harvard University Press.
- Vedung, E. (1977). *Det rationella politiska samtalet – hur politiska budskap tolkas, ordnas och prövas*. Stockholm: Aldus.
- Vedung, E. (2018). Fyra typer av statsvetenskaplig idéanalys. *Statsvetenskaplig tidskrift*, 120 (2): 197-235.

KAPITEL 2

Evert, Skytteanum och jag: Ett personligt möte och en personlig utveckling

Mats Lindberg

STUDENTÅREN – OCH EVERT

Jag minns Evert från min tid som student vid Skytteanum runt 1970. Jag höll just på med min C/D-uppsats "Staten, socialdemokratin och socialismen", den som senare kom att publiceras som bok (Dahlkvist, 1975). Evert hade den säregna förmågan att med ständig nyfikenhet fråga vad vi studenter hade för tankar och åsikter, och jag fick flera gånger frågan vad jag egentligen menade med min strukturmarxistiska begreppsapparat som allmänt uppfattades som snårig. Han hade ett särskilt tonfall och glittrande intresserade ögon när han undrade: "Men vad är själva kärnan?" Med stolthet blev jag och ett par andra C-studenter inbjudna till en informell studiecirkel som han gav för doktorander där man läste och diskuterade den episke Arnold Brechts livsverk *Political Theory. The Foundations of Twentieth Century Political Thought* (1959).

Att Arnold Brecht hade varit högt uppsatt, offentligrättslig jurist i Weimar-republiken, och en av få höga ämbetsmän som öppet motsatte sig Hitlers maktövertagande 1933, hade jag ju ingen aning om då. Ännu mindre visste jag att han hade avskedats, emigrerat till USA och mottagits väl av amerikanska akademiker.

Sedan blev han utrikespolitisk rådgivare åt USA:s regering och återvände till Tyskland efter 1945 där han deltog i arbetet med att utarbeta Förbundsrepublikens författning (!). Som vänsterinriktad student med pågående intellektuell utveckling i raketfart var jag för självupptagen för att inse det historiska och politiska drama som Brechts liv och verk innehöll. Tillräckligt spännande framstod emellertid hans politiska filosofi om värde-relationism, att inget politiskt värde (som naturen, nationen, livet, friheten, rättvisan, rasen, klassen, Guds lag med mera) har ett inneboende *intrinsic value* i sig självt (essentialism, fundamentalism) utan bara i relation till något eller någon. Det förment värdefulla måste värderas av någon (subjektivt, argumentativt). Kritiken av den gamla naturrätten, kristen eller liberal, stod därmed i centrum, liksom kritik av all politisk-filosofisk essentialism, såsom den tyska rasens överlägsenhet eller arbetarklassens historiska uppgift. Det senare var litet hårdsmält för en nyss övertygad marxist, men samtidigt var Brechts anti-essentialistiska position intellektuellt lockande. Dessutom hade Karl Popper framlagt en liknande anti-essentialism i sin bok *The Open Society* (1945, kap. 11 och 25).

Att Brechts politiska teori i sin tur var en systematiserande vidareutveckling av en ny-kantiansk tankelinje från Weber, Radbruch och Kelsen – där man också finner den svenske Axel Hägerström – var ett utmanande idéhistoriskt perspektiv. Den intellektuella utmaningen var: Kunde man vara marxist och ändå acceptera Brechts anti-essentialism och värderelationism? När vi C-studenter senare blivit doktorander försökte vi oss på detta konststycke! Det underlättades av Althusserskolans scientistiska strukturmarxism, att marxistisk samhällsteori enbart skulle ses som vetenskap och inte som normativ politisk filosofi eller metafysisk historieteori. Vi kom följdriktigt därför också att – ironiskt och pejorativt – etiketteras som ”ny-kantianer” av mer trosvissa, maoistiska marxist-leninister i Uppsala, till exempel professorn i ekonomisk historia, Bo Gustafsson. Jag blev därför knappast imponerad några decennier senare när socialkonstruktivismen och poststrukturalismen med dessas kritik av *intrinsic values* bredde ut sig som en ”nyhet”. Jag menar, denna kritik var ju hundra år

gammal, men i händerna på Weber, Popper och Brecht ändå med bibehållen, sanningssökande rationalism (se Lindberg, 2023).

Skytteanum var en fantastisk miljö att komma till som student runt 1969–1970. Kurslitteraturen var klokt sammansatt. Det märkte jag när jag senare i livet försökte utforma den perfekta statsvetenskapliga utbildningen! Jag gissar att det var den insiktsfulle, avgående professorn Carl Arvid Hessler som var arkitekten. Där fanns till exempel Herbert Tingstens arbeten med hans förmåga att lyfta fram idéstriden och idédebatten i den empiriska politiken. Där fanns också Leif Lewins *Planhushållningsdebatten* (1967) med samma förmåga. På B-nivån fanns också Karl Popper (1945) som trots sin kritik av marxistisk metafysik lärde oss studenter allt vi behövde veta om Marx! Dessa författare fick mig att tänka politik på ett idéanalytiskt sätt.

Våra lärare på A-nivån var utmärkta med Mats Bäck i svensk och jämförande politik, Donald Söderlind i förvaltning och den imponerande Stefan Björklund i politisk teori. Den senare hade just skrivit boken *Politisk teori* (1968) där han i utgångsläget beskrev allt politiskt tänkande, och jämförde de politiska ideologierna, utifrån Herbert Tingstens tre basala kategorier: värderingar, beskrivningar och handlingsnormer. Han gick i senare upplagor också i vidareförande debatt rörande Brechts värderationism. Björklund försvarar tanken om basala grundvärden, Björklund är naturrättsteoretiker. Men samtidigt ser han inte politisk-filosofiska grundvärden som "objektiva" i metafysisk eller ontologisk mening. Han betraktar naturrättsliga grundidéer som människors lika värde, jämlikhet och rättvisa som något subjektivt argumenterat och politiskt konstruerat, inte som ontologiska sanningar. Björklunds filosofiska grundhållning är Poppers kritiska rationalism och den senare John Rawls' icke-essentialistiska politiska liberalism. Den kritiska diskussionen – "argumentets primat" – ser han som kunskapssökandets alfa och omega, även i normativ politisk filosofi. Björklund undviker därmed att hamna i en postmodern sanningsrelativism, men samtidigt undviker han att hamna i essentialistisk metafysik. Detta synsätt förde Björklund sedan vidare i flera böcker med Rawls, Hume, Adam Smith och

Karl Popper som hjältar och med Carl Schmitt och Heidegger som skurkar.

DOKTORANDÅREN – OCH EVERT

Stefan Björklund och Sverker Gustavsson var lärare i den ny-tänkta doktorandkursen statsvetenskapliga analysinriktningar; de blev senare mina första handledare och livslånga forskarvänner. Sverker Gustavsson såg till att vi 1972 fick en eminent kurs i vetenskapsteori som gavs av den spännande Håkan Törnebohm, med Thomas Kuhn, Karl Popper, Imre Lakatos och Mario Bunge som centralgestalter. Denna kurs satte djupa spår i min vetenskapsuppfattning. Stefan hade därtill handledaransvaret för en hel grupp Althusserianska och Brechtianska värderelationer (!) där jag ingick. Med tålmodigt intresse iscensatte han dynamiska gruppseminarier i sin fjällstuga i Sälen, baserade på långa (och ibland inte så välskrivna) avhandlingsutkast. Stefan framförde också 1973 en principiell kritik av Jörgen Westerståhls "positivism" och hävdade att statsvetenskaplig forskning alltid måste ha värderingar som utgångspunkt, till exempel i val av problem och analysbegrepp.

Evert var ansvarig för doktorandkursen i kvalitativ idéanalys. Han var inspirerande och dialogisk. Hans vetenskapliga grundhållning sammanfattade han gärna med ett citat från Brecht: "The mere selection of a problem, i.e., of a question for worthwhile inquiry, is as a rule no mechanical act but an act of the creative mind..." (Brecht, 1959, s. 30f). Kreativitet och nyfikenhet skulle sättas i centrum, problemet och grundfrågan var det avgörande, teori och metod fick sedan anpassas till det problem som skulle lösas. Detta var ju också Sverker Gustavssons och Stefan Björklunds ständiga budskap.

Evert var intresserad av språkfilosofi. Vi mötte i kursen den litteratur och de teman som han sedan sammanfattade i *Det rationella politiska samtalet* (1977) om analys av politiska budskap (eng. 1982). Jag hade läst (analytisk) teoretisk filosofi uppe på Villavägen 5 och var intresserad av logik, semantik och Wittgenstein, men här blev saken plötsligt politiskt konkret. Det började med Vedungs distinktion mellan innehållslig och funktionell idéanalys, att det är

en avgörande skillnad att å ena sidan försöka bringa klarhet i det idémässiga innehållet i politiska budskap, och att å andra sidan studera (de klargjorda) idéernas och begreppens ursprung eller politiska funktion.

Kursen fortsatte med argumentationsanalys. Här mötte vi som kurslitteratur Arne Naess (1970) och Jan Evers (1970). Vi skulle bibringas kunskaper om vad som är relevanta och irrelevanta argument för och emot i en sakfråga, samt tydliga normer för saklig respektive osaklig argumentation. Arne Naess' bok avslutades med reflexioner om skillnaden mellan saklig debatt, propaganda och förkunnelse; kunskaper som för övrigt verkar vara ännu mer viktiga i dagens politiska debattklimat.

Evert problematiserade därtill begrepp som "mening" och "tolkning" utifrån Naess och andra samtida. Viktiga blev de enkla distinktionerna mellan "lexikalisk tolkning" och "kontextuell tolkning", liksom mellan "systematisk tolkning" och "rekonstruktiv tolkning". Här fick vi inblick i den lingvistiska filosofins debatter, till exempel den sene Wittgensteins kontextuella språkfilosofi eller John Austins teori om "performativa satser". Dessa verkade i sin tur ha haft mäktiga föregångare i Ogden och Richards grundläggande verk *The Meaning of meaning* (1923) som för mig varit en ständig inspirationskälla. Där föregrips den sene Wittgenstein och tas avstånd från den franske lingvisten de Saussure och dennes grammatiska strukturalism. Evert var djupt intresserad av termers mångtydighet och sociala kontextbundenhet, vilket också är kärnan i Ogden och Richards diskussion, liksom hur mångtydighet och vaghet, samt irrelevant argumentation, kan användas retoriskt och propagandistiskt i politiken. Sådana frågor hade för övrigt just aktualiserats i svensk debatt genom filosofen Gunnar Fredrikssons mycket lästa bok *Det politiska språket* (1962).

Denna duvning i det politiska språkets semantik har följt mig genom hela livet och var en stor tillgång när jag skulle ta mig an mitt avhandlingsprojekt om Karl Marx' *Kapitalet*, med dess erkänt snåriga terminologi, analyserad i ljuset av tre dominerande marxistiska filosofier med *deras* specifika terminologier och språkvärldar. Jag använde också Everts föreslagna tolkningsmetoder "syste-

matisk” respektive ”rekonstruktiv” tolkning (Dahlkvist, 1978/ Lindberg, 2013, s. 459 not** samt s. 567ff och 616ff). Jag är oändligt tacksam för Everts språkfilosofiska inspel i mitt liv. De olika tankelinjer och problem som han visade på – och några till – har följt mig sedan dess.

FORSKARÅREN – OCH EVERT

1993 tog jag, tillsammans med mina gamla vänner Jörgen Hermansson (Uppsala) och Bo Lindensjö (Stockholm) initiativet till att vitalisera arbetsgruppen för politisk teori inom Statsvetenskapliga förbundet. Den fick senare namnet *Swedish Network in Political Theory*. I våra diskussioner började det fram emot 2010 växa fram en tanke att Herbert Tingsten borde lyftas fram som förebildlig för vår disciplin. Vi bildade en arbetsgrupp för en nytugivning av hans skrifter. Där ingick förutom Jörgen och jag själv, Ludvig Beckman, Ulf Mörkenstam och Björn Badersten. Vi fick också stöd från nyblivna professor Skytteanus, Li Bennich-Björkman. Det hela utvecklades till slut till ett projekt att digitalisera hela Tingstens vetenskapliga och journalistiska produktion (cirka 3 500 poster) finansierat av RJ och lett av Björn Badersten, samt med placering vid Lunds universitetsbibliotek under väktarskap av anrika Statsvetenskaplig Tidskrift. Jag stötte vid ett tillfälle ihop med Evert. Vi kom som hastigast att prata om digitaliseringen av Tingsten och de viktiga distinktionerna mellan värdeomdömen, verklighetsbeskrivningar och handlingsrekommendationer. Detta möte skulle få konsekvenser.

Jag hade runt 2012 bestämt mig för att börja skriva om Tingsten och dessa tre analytiska grundbegrepp för statsvetenskaplig analys. När Evert sedermera fick uppdraget att för *Statsvetenskaplig Tidskrifts* räkning göra ett temanummer om idéanalys anmälde jag mitt intresse och skisserade en artikel om Tingstens tre grundbegrepp; värdeomdömen, verklighetsbeskrivningar och handlingsrekommendationer. Evert blev högeligen entusiastisk och ett intensivt samarbete inleddes. (Se Vedung, 2018.) Särskilt uppskattade han min systematisering av detta tänkande som han menade genomsyrade hela hans bok från 1977 men som han nu ansåg

att han inte hade lyft fram tillräckligt, det var ju så självklart då. Vi beslöt också att jag på engelska skulle använda termerna *values*, *descriptions* och *prescriptions*. Evert var särskilt glad för att jag sammanställde dessa analytiska grundbegrepp till en argumentationsteoretisk analysmodell som gavs namnet *the VDP-triad*. Denna var dock inget mer än ett framlyftande av Stefan Björklunds gamla modell från boken *Politisk teori* (1968; 1970). Jag var dock (som vanligt) överdrivet omständlig och hade mycket att reda ut. Jag ville nämligen förtydliga det Tingstenska och Skytteanska tänkandets inplacering i nutida samhällsvetenskap och ideologiteori i bredare mening; det vill säga dess relationer till allmän samhällsteori (Parsons, Lasswell & Kaplan), marxistisk ideologiteori (Althusser, Pécheux, Bourdieu) samt lingvistiska diskursstudier (van Dijk, Johnstone), kritisk diskursanalys (Fairclough, Wodak) samt den samhällsfilosofiska, poststrukturalistiska så kallade "diskursteorin" (Laclau-Mouffe, Howarth, Torfing) (se Lindberg, 2018b, s. 437–534).

Jag kom att urskilja tre ideologibegrepp (Lindberg 2018b, s. 437). För det första det "pejorativa" marxistiska ideologibegreppet (Seliger 1976; Seliger 1977), där "ideologi" hos den härskande kapitalistklassen bara skenbart handlar om sanning och rätt. I stället uttrycker ideologin dess maktposition och maktperspektiv, medan "ideologi" hos den underordnade arbetarklassen ger denna ett "falskt medvetande" som osynliggör dess underordnade position. För det andra den allmänna samhällsteorins sociologiska ideologibegrepp, där "ideologi" ses som det konsensusbaserade tänkande, den "kultur" som håller samman en grupp eller ett samhälle och legitimerar dess struktur och institutioner. För det tredje statsvetenskapens politiska ideologibegrepp, där "ideologier" ses som de idéer som motiverar och vägleder aktörers politiska handlande i konflikt och idéstrid om hur samhället ska utvecklas och förändras, respektive försvaras och bevaras.

Jag kom också att urskilja tre diskursbegrepp som jag pedagogiskt försökte sammanfatta (Lindberg 2018b, s. 467). För det första det disciplinärt lingvistiska, Discourse (L), som ser diskurs som konversation eller kommunikation, där diskursstudier är allmänna

studier av hur språk och kommunikation används; näraliggande det lingvistiska allmänbegreppet språkpragmatism. För det andra det foucauldianska, Discourse (F), som härrör från det marxistiska, pejorativa ideologibegreppet med dess maktperspektiv, och använder ordet "diskurs" i ungefär samma betydelse som marxismens "ideologi", det vill säga som benämning på ett idéinnehåll i en språkvärld som inrättar en maktsituation, som till exempel globaliseringsdiskursen. Samt slutligen det konventionella, Discourse (C), som används sedan gammalt i traditionell filosofi och konventionell bildning som betyder ordnat samtal eller principiell diskussion, på franska även föreläsning eller föredrag.

Jag menar med detta att Michel Foucault själv självklart och bildat – paradoxalt nog kanske – använder det konventionella diskursbegreppet Discourse (C) i betydelsen "konversation" eller "föredrag", till exempel i sin programförklaring i installationsföreläsningen "*L'ordre du discours*" (Foucault, 1971). Detta problematiseras sällan av hans efterföljare som har bestämt sig för ett marxiserande, "foucauldianskt" diskursbegrepp enligt ovan, där ordet 'diskurs' betyder ungefär detsamma som marxismens 'ideologi'. Foucauldianer läser därför oftast denna rubrik som att det är termen "discours" som är huvudsaken, och som här programmatiskt förs in som något nytt i den samhällsvetenskapliga teoribildningen. Men enligt min mening är det viktiga och nya – för Foucault och hans åhörare – det som signaleras av termen "*l'ordre*".

Mitt arbete med detta verkade aldrig bli färdig. Språkligt präglades artiklarna också av ofärdighet, pendlande mellan *context of discovery* och *context of justification*. Samtidigt hade jag fått det hedersamma uppdraget att skriva kapitlet om kvalitativ idéanalys i Kristina Boréus och Göran Bergströms engelska översättning av metodboken *Textens mening och makt*. Mitt bidrag fick namnet "*Qualitative analysis of Ideas and Ideological content*" (Boréus & Bergström, 2017). Inte förrän 2018 kom min dubbelartikel i tryck, i Everts temanummer om idéanalys, "*The VDP-triad in ideational analysis, I-II*". Jag hade hållit på med detta sedan 2014 och redaktörernas tålamod överträffades bara av Everts eget.

I någon mening var min dubbelartikel på sammanlagt 200 sidor en sorts maratonlopp, med sin kritiska genomlysning av nästan alla framträdande teorier om ideologi och ideologiskt tänkande, där VDP-teorin framställdes som *"the strong alternative"*. Jag hade inte orkat genomföra detta omfattande arbete utan Everts entusiasm, tålamod och ständigt uppmuntrande stöd.

SKYTTEANUM SOM TANKEKOLLEKTIV OCH TANKESTIL – EVERTS BETYDELSE

När jag började segla i egen sjö, och började ge egna doktorandkurser i Göteborg från 1980 (senare i Karlstad och Örebro) förvånades jag över att det som jag ville förmedla inte alls var mina specialkunskaper om Marx och marxistisk filosofi utan den intellektuella grundhållning och syn på statsvetenskapen som jag hade mött vid Skytteanum. Det är ju när man står i katedern som ett sanningens ögonblick infinner sig! De stackars doktoranderna fick läsa både Brecht och Vedung! Jag började med viss förvåning – och växande stolthet – mer och mer tänka på mig själv som skyttean. Tingstenprojektet och min dubbelartikel om VDP-teorin blev min återbetalning.

Betraktat i efterhand tänker jag att det fruktbara och vibrerande under mina år som student och doktorand var att ett "tankekollektiv" vid Skytteanum under åren 1967–1977 höll på att etablera och institutionalisera en speciell och mycket fruktbar statsvetenskaplig "tankestil". Det jag mötte var inte enbart en etablerad vetenskapsuppfattning utan en som fortfarande var under utveckling. Jag insåg inte förrän senare hur viktigt detta möte var och vad det bar på. I min dubbelartikel om VDP-teorin djärvdes jag sammanfatta denna tankestil och ge den ett namn: "uppsalaskolan i idéanalytisk statsvetenskap" (Lindberg, 2018a, s. 279–280 samt 2018b, 450). Denna benämning har naturligtvis baktanken att knyta an till den mer välkända uppsalaskolan i filosofi och Axel Hägerström.

Som jag ser det var det en utveckling i flera på varandra följande steg. Grundläggande var Herbert Tingstens basala distinktioner mellan värdeomdömen, verklighetsomdömen och politiska handlingsrekommendationer, samt hans rationalistiska kritik av

de politiska ideologiernas metafysiska eller försanthållna antaganden. Ett viktigt vidareförande steg togs sedan med Leif Lewins *Planhushållningsdebatten* (1967, s. 77). Utgående från Tingstens analys av den tidiga socialdemokratins ”stora” och ”lilla”, marxistiska perspektiv införde Lewin en distinktion mellan politiska värderingar och ideal ”högst upp i idévärlden” och det politiska tänkandet och handlingsförslagen i den vardagliga politiken. Det var en distinktion som jag senare i VDP-teorin anknöt till och förde vidare, med stöd av internationella forskare, såsom ”fundamental” respektive ”operativ” nivå i det ideologiska tänkandet (Lindberg, 2018b, s. 519–524).

Tingstens distinktioner blev tydligt införda och så att säga institutionaliserade som analytiska grundbegrepp i Evert Vedungs och Sverker Gustavssons doktorsavhandlingar från 1971 (Vedung, 1971, s. 48–51; Gustavsson, 1971, s. 16–19, 30). Tillsammans med Björklund (1968, s. 16; 1970, s. 28–31) hade alltså en metametodisk grundsyn förtydligats, där den språkfilosofiska distinktionen mellan värderingar, beskrivningar och handlingsrekommendationer lades som grund.

Ytterligare ett steg togs när Evert i *Det rationella politiska samtalet* (1977), förutom den metodiska skillnaden mellan innehållslig och funktionell idéanalys, presenterade de nämnda språkfilosofiska aspekterna av argument, mening och tolkning, samt hur empirisk respektive normativ utvärdering av politiska budskap bör gå till. Han menar där också att även normativa politiska budskap kan utsättas för vetenskaplig prövning, och att statsvetenskapen därmed också kan uttala sig vetenskapligt om vad som ”bör” vara. Där drog Evert en för dåtiden djärv slutsats. Statskunskapen behöver inte begränsa sig till att enbart vara en vetenskap *om politik* utan kan också vara en vetenskap *i politik*; men med bibehållen analytisk distans och kunskapsprövande attityd. Därmed hade den statsvetenskapliga uppsalaskolans tankestil fått en sammanfattande språkfilosofisk och metametodisk inramning. I detta ligger Everts stora bidrag.

SLUTORD

I en festskrift kan man få vara både festlig och personlig. Att ha haft Evert som inspiratör, kollega och vän – under åren som student och doktorand samt senare som forskare och universitetslärare – är helt enkelt en ynnest och en gåva. Tack Evert!

Referenser

- Audi, R. (2001). *The Architecture of Reason*. Oxford University Press.
- Audi, R. (2023). *Of Moral Conduct*. Cambridge University Press.
- Björklund, S. (1968). *Politisk teori*. Aldus (2. uppl. 1970, 4. uppl. 1976).
- Björklund, S. (1991). *Forskningsanknytning genom disputation*. Acta Universitatis Upsaliensis.
- Boréus, K & Bergström, G. (Eds.) (2017). *Analyzing Text and Discourse. Eight Approaches for the Social Sciences*. Sage.
- Brecht, A. (1959). *Political theory. Twentieth Century Political Thought*. Princeton University Press.
- Dahlkvist, M. (1975). *Staten, socialdemokratin och socialismen*. Prisma.
- Dahlkvist, M. (1978). *Att studera Kapitalet. Första boken*. Cavefors (ny upplaga som Lindberg 2013, Arkiv förlag).
- Eliaeson, S., Mindus, P. & Turner, S. P. (Eds.) (2014). *Axel Hägerström and Modern Social Thought*. The Bardwell Press.
- Evers, J. (1970). *Argumentationsanalys för jurister*. Gleerups.
- Fleck, L. (1935/1997). *Uppkomsten och utvecklingen av ett vetenskapligt faktum, Inledning till läran om tankestil och tankekollektiv*. Brutus Östling.
- Foucault, M. (1971) *L'ordre du discours*. Presse Universitaire Francais.
- Foucault, M. (1993) *Diskursens ordning*. Brutus Östling.
- Fredriksson, G. (1962) *Det politiska språket*. Prisma/Cavefors (många senare upplagor).
- Gustavsson, S. (1971). *Debatten om forskningen och samhället*. Almqvist & Wiksell.
- Lewin, L. (1967). *Planhushållningsdebatten*. Almqvist & Wiksell.
- Lindberg, M. (2018a) The VDP-triad in ideational analysis, part I. *Statsvetenskaplig tidskrift*, 2018(2), 277–359
- Lindberg, M. (2018b) The VDP-triad in ideational analysis, part II. *Statsvetenskaplig tidskrift*, 2018(3–4), 435–554.
- Lindberg, M. (2023). Vetenskaplighet och normativ relevans. Distanserat engagemang som grundhållning. I Bennich-Björkman, L., Gustavsson, S. & Lindberg, M. (Red.), *Hålla huvudet kallt – om distanserat engagemang i en uppjagad tid* (s. 191–245). Daidalos.
- Naess, A. (1970). *Empirisk semantik*. Norstedts.
- Ogden, C.K. & Richards, I.A. (1923). *The Meaning of Meaning. A Study of the Influence of Language upon Thought*. Harcourt, Brace & Co (ny upplaga 2013, Mertino Publishing).
- Popper, K. (1945). *The Open Society and its Enemies, Vol I–II*. Routledge and Kegan Paul (många senare upplagor).
- Seliger, M. (1976). *Ideology and Politics*. George Allen & Unwin.

Seliger, M. (1977) *The Marxist Conception of Ideology*. Cambridge University Press.

Vedung, E. (1971). *Unionsdebatten 1905*. Almqvist & Wiksell

Vedung, E. (1977). *Det rationella politiska samtalet. Hur politiska budskap ordnas och prövas*. Aldus.

Vedung, E. (1982). *Political Reasoning*. Sage.

Vedung E. (2018). Fyra typer av statsvetenskaplig idéanalys. *Statsvetenskaplig tidskrift*, 2018/2, 197–235.

KAPITEL 3

Taxonomies in Evaluation Research and Practice

Kim Forss

STARTING POINTS – JUXTAPOSING TWO TAXONOMIES

One of the many things that are common to both research and evaluation is a fondness for classification. But while there is an extensive and lively discussion of taxonomies in science, there is so far little on the subject in the evaluation discourse. With these reflections I hope to increase the attention to taxonomies among those who study and practice evaluation.

For many, the city and university of Uppsala is connected to the creation of taxonomies. I'm thinking of the binomial naming system with its nested categories: species, genus, family, order, and so on, which has prevailed in zoology and biology. The rank-based method of classifying living organisms was originally popularized by Linnaeus in Uppsala. The greatest innovation, and still the most important aspect of this system, is the general use of binomial nomenclature, the combination of a genus name and a second term, which together uniquely identify each species of an organism. The Linnaean system has proven robust, and it remains the only working classification system at present that enjoys universal scientific acceptance. Still, even this seemingly perfect system has its flaws,

for example although the number of ranks is unlimited, in practice any classification becomes more cumbersome the more ranks are added.

The Argentinian writer Jorge Luis Borges¹ had many interests, one of which was classification and taxonomies. In an essay on John Wilkins he made reference to an old Chinese encyclopaedia called *The Celestial Emporium of Benevolent Knowledge*. Borges meant to lay out the challenges of human attempts to classify the world and to open our eyes to the ‘ambiguities, redundancies and deficiencies’ of classification schemes. According to this ancient Chinese encyclopaedia, animals can be classified into: a) those belonging to the emperor, b) those that are embalmed, c) tame or trained ones, d) suckling pigs, e) mermaids and sirens, f) those that are fabulous, g) stray dogs, h) those included in the present classification, i) frenzied ones, j) innumerable ones, k) those drawn with a very fine camelhair brush, l) other ones, m) those that have recently broken a water pitcher, n) those that from a long way off look like flies.

Not quite the Linnean approach to classification! I can hear Borges laughing when he created this taxonomy (there is hardly any doubt that such an encyclopaedia ever existed outside his mind). Nevertheless, it is a thought-provoking comment on the problems of any complex classification system as much as it is an entertaining and ironical comment on our urge to classify. Practically speaking, the examples of Linnaeus and Borges suggest that it is possible to look for the qualities of taxonomies; what makes for a good taxonomy and which pitfalls to look out for when creating and using taxonomies?

PROBLEMS AND PITFALLS IN TAXONOMIES

Many evaluations start off by creating taxonomies to analyse the phenomena being evaluated. In one example of a policy

1) Borges, Jorge Luis (1999), “John Wilkins’ Analytical Language”, in Weinberger, Eliot (ed.), *Selected nonfictions*, Eliot Weinberger, transl., Penguin Books, p. 231, ISBN 0-14-029011-7. The essay was originally published as “El idioma analítico de John Wilkins”, *La Nación* (in Spanish), Argentina, 8 February 1942, and republished in *Otras inquisiciones*.

evaluation², the authors developed a model to analyse the implementation of policies, distinguishing between four ‘methods’ that were identified during the study. They argued that policies can be implemented through 1) funding, 2) training and capacity building, 3) facts and information, and 4) normative arguments and political dialogue. Having recently read *Carrots, Sticks and Sermons: Policy Instruments and their Evaluation* (Bemelmans-Videc, Rist and Vedung, 1998) I was struck by some features in their approach to classifications of policy instruments/methods.

Even though the four terms used in the evaluation can be empirically verified it turned out to be quite difficult to use them for analytical purposes, both for the authors of the evaluation and for me as a reader. In practice, training and capacity building are usually activities that require funding. Facts and information are hard to distinguish from arguments and political dialogue, and the former would naturally develop into the latter. Training would contain facts and information, and perhaps also arguments. Without going into further details, it is easily seen how the taxonomy in this evaluation could not really be used to assess methods of policy implementation. The categories were not distinct from each other and tended to overlap, making it unclear what different methods of policy implementation there were to evaluate. The three policy instruments of sticks, carrots and sermons, on the other hand, may seem more difficult to use as the terms themselves are not used by policy makers. Neither sticks, nor carrots, nor sermons are immediately visible in the world of policy implementation. However, the terms have a common-sense appeal, and they are easily understandable. They refer to processes that include both funding, structures and content of funding, and purpose. If the evaluation I read had been based on that taxonomy (sticks, carrots and sermons), it might well have analysed which policy instruments were used, how they were used, and what the results

2) I have chosen to keep the reference anonymous. The evaluation I refer to has strengths and weaknesses, and my point is not to undertake a comprehensive critique of it. I use it to illustrate an aspect of taxonomies, and I think that if I disclosed the evaluation, I would need to use more space here to give proper account of its qualities.

were, and it could have assessed the balance between instruments. A better taxonomy would have led to a better evaluation.

The second example comes from attempts to classify the uses of evaluation. Once we tried to penetrate deeper into the mysteries of what “process use” of evaluation really is (Forss, Rebien and Carlsson, 2002)³. We developed a method for empirical analysis of evaluation processes and constructed a taxonomy of five different types of process use⁴. It was possible to compare them, point to advantages and disadvantages, and not least how to improve different kinds of use through appropriate choices of method. However, when I compare our taxonomy of process use to other taxonomies of use, not least the ones proposed by Vedung (1997, 2021), I’m now struck by our bias; our method, and perhaps our background, lead us to look for the positive aspects of the use of processes for evaluation. The taxonomies of evaluation use developed by Vedung and others rightly point to legitimizing use, camouflage, postponing and other types of use that could also be described as ‘misuse’. In retrospect, I’m inclined to believe that process use also has several ‘non-legitimate’, or problematic uses in organisational politics, stakeholder competition and power games. Our approach was inductive, based on a few case studies. Perhaps it would have been necessary to add deductive methods to detect the less obvious, hidden and undesirable types of use.

These two examples show that the construction of taxonomies in evaluation is a critical task, and mistakes can lead to missed opportunities in knowledge building, and to downright faulty and eschewed conclusions.

3) Initially the analysis of how evaluations were used took its starting point in a completed evaluation report. In different studies (DsUD 1990:63), we often heard that ‘the report itself was not that useful, we knew more or less what it would contain and what to do. However, it was useful to be through the process’. Were the interview respondents merely polite, or what exactly did they mean when they said the process was useful? That’s what we wanted to explore. ‘Use’ can start well before a report is written, indeed already when the evaluation is conceptualized and then throughout the process of inquiry up to the publication of a report.

4) For example; to strengthen the programme itself – irrespective of the evaluation, to build networks among stakeholders, to develop alternative approaches to problems, to generate commitment and enthusiasm, etc. as elaborated in the article in *Evaluation*.

DEFINING TAXONOMIES

A reader might object that the examples I point to are something quite different than the Linnean system of binomial naming and the playful ontological inquiry of Borges essay. However, in essence, they represent similar processes in trying to understand the world by defining objects and how they relate to each other. The origin of the word taxonomy is Greek, from *taxis*, meaning arrangement or order, and *nomos*, meaning law or science. (<https://en.wikipedia.org/wiki/Taxonomy>). A taxonomy is: “the result of naming and classifying items into groups within a larger system according to their similarities and differences. The resulting term’s structure, which can be either flat or hierarchical in nature, is used to provide a conceptual framework for discussion, analysis or information retrieval” (ISO, 2011).

This may seem simple enough, but there are still many meanings of the word taxonomy. The original meaning, the study of classification, is not much used and instead the term is mostly used to designate all kinds of classification. It has crossed disciplinary boundaries; from its origins in natural sciences, it has spread to the social sciences. The term taxonomy can be used both in a narrow sense, to mean a hierarchical classification or categorization system, and in the broad sense, in reference to any means of organizing concepts of knowledge (Hedden, 2016). In this sense, a taxonomy is an example of knowledge organization or knowledge organization structure. It is in this broader sense that a discussion of taxonomy can be useful in evaluation research.

CREATING TAXONOMIES

Where do taxonomies come from, or rather, how are they created? The space here is too short to allow for any extensive account⁵, but let us look at some of the choices on approach and method:

1. *Inductive or deductive approaches*. The two examples of taxonomies mentioned above were generated by the authors,

5) A comprehensive account of the steps in a process of creating taxonomies is found in Lambe (2007).

based on empirical studies – interviews with policy makers in the one, and interviews during case studies in the other. They were purely inductive and made little use of current research. There is no indication that the taxonomies were verified by practitioners or tested in other contexts. In retrospect an element of deductive analysis might have improved the quality by revealing the weaknesses mentioned above. On the other hand, a purely deductive approach might be less useful in an evaluation context. Stakeholders in an evaluation need to recognize and relate to classification, otherwise the risk is that the findings cannot be used. A balance between deductive and inductive approaches would be advisable.

2. *Analytical or synthetic approaches.* Evaluation is often seen as an analytical process and the creation of a taxonomy is also a process of analysis. However, in the same manner as that process seems to require a balance between inductive and deductive approaches, it requires a balance between analysis and synthesis. Analysis means to break down complex ideas into smaller fragmented concepts to come up with an improved understanding. Synthesis, on the other hand, means to go beyond the concepts being studied, to seek understanding in the way the studied object relates to its environment. The intellectual process of analysis stays within the system being studied, the process of synthesis goes beyond that.

3. *Didactic or participatory approaches.* Many of the methodological guidelines on taxonomies emphasize a participatory approach, that is, involving various stakeholders in the design of taxonomies. I would think that Linnaeus worked on his own as well as with his disciples, and the value of participation would certainly depend on context and purpose. Like other approaches, there are advantages and disadvantages.

4. *Qualitative or quantitative approaches.* Numbers and quantitative analysis can play an important role in the design of taxonomies, not least when it comes to understanding the overall structure of the taxonomy – it's hierarchical levels, the branching of a decision tree, the relations between terms. Nevertheless, the overarching understanding and organizing

principles are likely to build on qualitative insights, leaps of the imagination – and a certain amount of brainstorming (in one or many brains).

These four points are hardly even a skeleton framework, so how does an evaluator learn to design taxonomies? There are handbooks, guidelines and more general introductions to the subject (Bailey, 1994; Doty et al, 1994; Hedden, 2016). Google has several suggestions on training opportunities. Many courses are offered within university programmes on knowledge management, information systems, data management and similar subjects (Hedden, 2016). There are shorter components of larger programmes and there are stand-alone courses by faculty members and consulting firms.

These training opportunities are made for people in information sciences writ large. Within the field of evaluation there is not much to be found. I have not seen any courses in evaluation that contain a section on taxonomies. Textbooks that are used in evaluation capacity building seldom elaborate on design, use and analysis of taxonomies. The competency frameworks that intend to show what an evaluator must know do not mention anything about skills in taxonomy⁶. It is thus not surprising that at least some taxonomies used in evaluation are less useful if not downright misleading. Designing taxonomies is an art that needs to be developed, practiced and fine-tuned. And taxonomies need to be evaluated⁷!

QUALITIES OF TAXONOMIES

It is a daunting task to suggest a taxonomy of the qualities of taxonomies, but clearly taxonomies can be of different qualities – they can be good or bad. Here might be some qualities to look for.

6) See for example the EES Evaluation Capabilities Framework www.europeanevaluation.org or the UNEG Evaluation Competences Framework www.unevaluation.org

7) A pioneering effort was presented in Uppsala four years ago: Szopinski, D., Schoormann, T., & Kundisch, D. (2019): Because Your Taxonomy is Worth It: Towards a Framework for Taxonomy Evaluation. In: *Proceedings of the European Conference on Information Systems (ECIS)*, Stockholm-Uppsala, Sweden.

First, a good taxonomy should be comprehensive, meaning that it should cover all relevant categories and subcategories without significant gaps. To take an example, early attempts to distinguish different types of use of evaluation mentioned instrumental and conceptual use; soon process use was added; subsequently legitimizing use, and so on. If a taxonomy is not complete, if it does not cover the full spectrum which it purports to cover, it is less useful and potentially misleading.

Second, a good taxonomy should be constructed with mutually exclusive categories and subcategories. It should be clear why an item falls into one category and not in another. It is sometimes said that items observed should fit into one category only, as otherwise there would be overlap and confusion. Returning to the Chinese encyclopaedia, an animal could both belong to the emperor and look like a fly at a long distance – and hence the categories overlap which is one of the characteristics that we find amusing when reading the list.

Third, a good taxonomy should be scalable and adaptable, allowing for the inclusion of new items and categories as needed. The first descriptions of how evaluations were used stem from the 1960s, and some 60 years later the discipline has grown immensely, and it is reasonable to expect that use has found new forms and has become more varied. The Linnaean classification system has been able to expand and accommodate large numbers of living organisms that were not known in the 18th century. The Linnaean system has a clear and logical structure that accommodates change over time.

Are there other qualities we should be aware of? Perhaps a good taxonomy should be intuitive and easy to grasp? It is appropriate to quote Albert Einstein: *“Everything should be made as simple as possible, but no simpler.”* Some things are more difficult to grasp than others, and some taxonomies hence less intuitive than others. I would think that ‘intuitiveness’ is a quality to keep in mind, rather than to apply automatically.

Another quality could relate to transparency, that is, it should be possible for others to understand how the taxonomy was created,

to replicate the process, or possibly to criticize it because of how it was created. Much like transparency is a quality of evaluation – and of any research process, so it could also be an important aspect of a taxonomy. Few taxonomists in the past fulfilled such a standard, but perhaps the times are changing.

Could usefulness be a quality criterium? Sometimes yes and sometimes no; most of us agree that evaluations should be useful, but evaluations are goal oriented and have a purpose. The exercise of creating taxonomies is not necessarily teleological. I guess it depends, if a taxonomy is developed in order to serve a purpose, then its usefulness in doing so is a quality criterium.

WHERE TO FROM NOW

In this essay I have argued that taxonomy is a neglected topic in the discourse on evaluation methods, in understanding evaluation, and in the practice of doing evaluations. Designing taxonomies is an art and a skill which needs to be developed, practiced and refined.

I can imagine many ways in which the topic of taxonomies could be developed. For example, the curriculum of training programmes in evaluation can be developed to include sections on taxonomy, and similarly course material, introductory texts on evaluation methods, etcetera could be made to include the topic of taxonomy. The subject might have its place parallel to chapters on causal analysis, methods for data collection, and other equally necessary topics. I cannot remember that taxonomies have ever featured in any of the professional meetings organised by the European Evaluation Society or by national societies (but I could certainly be proven wrong on this).

Often a good way to learn about a subject is to look for examples of good practice. There is a literature on taxonomies outside the evaluation discipline, and within the discipline there are interesting examples to study and learn from, for example: Bemelmans-Vidéc, Rist and Vedung 1998; Vedung, 1998; Vedung, 2010; Vedung, 2021.

Evaluation is a constantly developing field of practice and over the past decade it has grown increasingly complex. Evaluation

is called upon to play a role in moving towards sustainable futures, it is asked to address non-linear phenomena and multiple causality, long-term perspectives, and it must be transparent, useful and effective. With such challenges, evaluation needs tools in the inquiring process – and a good taxonomy would be one of the best companions to have.

References

- Bailey, K. D. (1994). *Typologies and Taxonomies: An Introduction to Classification Techniques*. London: Sage Publications.
- Bemelmans-Vidéc, M.-L., Rist, R.C. & Vedung, E. (1998). *Carrots, Sticks and Sermons: Policy Instruments and Their Evaluation*. New Brunswick: Transaction.
- Doty, D. H. & Glick, W. H. (1994). Typologies as a Unique Form of Theory Building: Toward Improved Understanding and Modeling. *The Academy of Management Review*, 19(2), 230–251.
- DsUD 1990:63. *Bra Beslut: Om effektivitet och utvärdering i biståndet*. Stockholm: Regeringskansliet.
- Forss, K., Rebien, C. & Carlsson, J. (2002). Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback. *Evaluation*, Vol 8 (1) 29 – 45.
- Hedden, H. (2016). *The Accidental Taxonomist (2nd Ed.)*. Medford, NJ: Information Today.
- International Organization for Standardization (2011). *Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval* (ISO Standard No. 25964-1:2011)
- Lambe, P. (2007). Preparing for a Taxonomy Project. In *Organising Knowledge: Taxonomies, Knowledge, and Organisational Effectiveness* (s. 153–183). Oxford: Chandos Publishing.
- Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Because Your Taxonomy is Worth It: Towards a Framework for Taxonomy Evaluation. In: *Proceedings of the European Conference on Information Systems (ECIS)*. Stockholm-Uppsala, Sweden.
- Vedung, E. (2010). Four Waves of Evaluation Diffusion. *Evaluation*, 16(3), 263–277.
- Vedung, E. (1998). *Public Policy and Programme Evaluation*. New Brunswick: Transaction.
- Vedung, E. (2021). The Twelve-Part Typology of Evaluation Uses. *Zeitschrift für Evaluation* 20. 2021 (1), p. 101-130.
- Whittaker, M. & Breinger, K. (2008). Taxonomy Development for Knowledge Management. In *World Library and Information Congress: 74th Ifla General Conference and Council*, 10-14 August 2008, Québec, Canada.

KAPITEL 4

Forskningens relevans, användbarhet, användning och nytta – 30 år senare*Verner Denvall*

I december 1993 avslutade dåvarande docenten vid statsvetenskapliga institutionen vid Uppsala universitet Evert Vedung en skrift med titeln *Forskningens relevans, användbarhet, användning och nytta*. Det är en delrapport från ett treårigt FoU-projekt finansierat av Byggforskningsrådet (BFR) med syfte att utveckla en modell som skulle underlätta bedömning av sektoriell FoU-kunskap. Ambitionen var att skapa checklistor som skulle kunna rutinisera ansökningar, beredning och utvärdering av forskning och därmed avgöra vad som var samhällsrelevant.

Evert tog på sig uppgiften att på 59 sidor utreda hur forskningsresultat kan mätas med utgångspunkt att forskning måste vara användbar även utanför forskarsamhället (1994). Everts framställning koncentreras kring vad som menas med det lilla begreppet "relevans".

I detta bidrag återvänder jag till Everts skrift och undersöker hur hans argument om vad som gör forskning relevant står sig trettio år senare. Framställningen präglas av dagens diskussioner om forskningens användbarhet med begrepp som kunskaps- och faktaresistens.

VAD KOM EVERT FRAM TILL?

Skriften ingick i "NOGA-projektet" som syftade till att utveckla en modell för att underlätta bedömning av sektoriell FOU-kunskap: vad var användbart i denna form av forskning? Ambitionen var att utveckla checklistor vid bedömning av vad som var relevant forskning. Modellarbetet presenteras i skriftens förord författat av projektledaren Ola Nyquist. Everts skrift är nr 1 av sexton publicerade rapporter 1994 – 1996.¹

I sitt förord – eller företal enligt författaren själv – skriver han att syftet med NOGA-projektet är utvärderingsinriktat. Det ska utforma metoder för utvärdering av sektorsforskningen. Evert tackar ett tiotal personer från såväl NOGA-projektet som andra forskare för kommentarer och läsning av tidigare versioner av texten. Skriften får väl närmast betraktas som en slags vetenskaplig essä, dvs en text i en friare form än den traditionellt uppbyggda med syfte, metod, teori, resultat och diskussion. Evert har utnyttjat sex riksdagstryck mellan 1979–1993 och presenterar ett fyrtiotal övriga referenser av blandad kompost. Allt från en odaterad stencil till läroböcker, SoU-rapporter och internationellt publicerade vetenskapliga artiklar och böcker, av exempelvis Michael Scriven, Carol Weiss och Vedung själv.

RELEVANS ÄR ETT OKLART BEGREPP

Vi forskare startar ofta med att utreda centrala begrepp. Det kan tyckas enkelt att redovisa en lexikal begreppsdefinition av relevans, men Evert stöter snart på problem. Till att börja med måste man fråga sig för vem forskning är relevant skriver han. Det visar sig nämligen att sektorsforskning inte är ett enhetligt begrepp utan vem som har behov av forskning och vilken slags forskning som är betydelsefull innehåller flera spänningsfält. Ett sådant spänningsfält är synen på användbarhet mellan långsiktig grundforskning och mera kortsiktig tillämpad forskning. Evert går till argumentationsanalys för att söka vägledning kring hur man ska förstå relevansbegreppet och konstaterar att något måste höras till saken för

1) Se <http://libris.kb.se/bib/1854533>

att kunna betraktas som relevant – annars är det irrelevant. Han slutar med att relevans måste ses i termer av att forskning måste ”omfatta intressant för *vem* i vilket *hänseende* och vid vilken *tid*” (1994, s. 20). Någon större hjälp av lexikal förklaring tycks inte Evert ha fått.

RELEVANS FÖR VEM?

Han konstaterar att relevans i forskning inte bara kan handla om dess resultat utan föreslår ett bredare synsätt där relevans måste kopplas till inomvetenskapliga processer och ämnen, för människor och för miljöer. Evert kommer därigenom in på den spännande frågan om för vem som forskning kan ha relevans för och därigenom blir användbar. Evert tar nu hjälp av kritisk samhällsvetenskaplig forskning och konstaterar att forskning inte bara används för att lösa samhällsproblem utan även används som ett maktmedel. Eftersom forskning kan användas destruktivt är därför inte användbarhet en särskilt bra indikator på forskningens samhällsrelevans.

RELEVANS OCH NYTTA

Bortom begreppet användning finns dessutom frågan om nyttan. Forskning ska användas för nyttiga ändamål varför man bör skilja på god användning och dålig användning. Här kan man göra kostnads- och intäktsanalys och väga kostnader för forskning mot dess nytta och samhällsrelevans. Men konstaterar han – den forskning som är nyttig för en person kan vara skadlig för annan. Kanske skall man utgå från att den forskning som faktiskt implementeras också är särskilt användbar? Nja, konstaterar Evert, ”forskningens relevans står för så mycket mer än att höra till saken” (1994, p. 37). Så här långt i framställningen står det klart att författaren snarare upptäckt nya problem än lyckats lösa relevansfrågan i forskning. Han konstaterar nämligen att för det första kan forskning ha betydelse för samhällsproblem men ändå vara oanvändbar för aktörer som inte förstår eller inte sympatiserar med dess resultat. För det andra behöver forskningsresultat bedömas ur en nyttighetsaspekt, dess resultat ska vara gott.

RELEVANS OCH UTVÄRDERING

Evert sammanfattar genom att ta hjälp av vad utvärderingsforskningen har att säga om användning. Nu kan vi börja ta del Everts framtida analyser av utvärderingsanvändning presenterat i ett flertal artiklar och böcker under kommande decennier (Vedung, 1995, 2011, 2015, 2021). I skriften presenteras fem modeller. Till att börja med ingenjörsmodellen som instrumentell och rationell användning av forskningsresultat. Vidare tar Evert upp Carol Weiss banbrytande arbete där hon identifierat en upplysningsmodell (även konceptuell eller begreppslig användning). Den innebär att polletten trillar ner efter ett tag, det vill säga resultat är kanske inte omedelbart och instrumentellt användbara utan först efter ett tag. Vidare presenterar Evert den interaktiva användningen av forskning då forskningsresultat inte används linjärt utan snarare organiskt, som en sorts information tillsammans med allt möjligt annat. Avslutningsvis berör Evert legitimerande och taktisk användning av forskning genom att forskning även utnyttjas som ammunition i konflikter eller rent av taktiskt för att vinna tid eller hamstra fördelar; forskning pågår, stör ej!

RELEVANSENS POLITIK

Evert avslutar i god utvärderingsanda med att även säga något om hur forskningsanvändning bör mätas. Jag tar inte upp denna del här då den leder oss bort från skriftens huvudtematik om relevans. I stället avslutar jag med att redovisa de hypoteser som Evert för fram i skriftens avslutning. Han konstaterar något uppgivet att taktisk användning inte är något som det talas högt om. Inte desto mindre tycks den spela en viktig roll i forskningspolitiken. Han tar då upp det svenska energiforskningsprogrammet från 1973 och framåt som hade ”drag av symbolisk politik över sig. Den väldiga satsningen på forskning tjänade som en täckmantel för att legitimera fortsatt utbyggnad av kärnkraften. Åtskilliga tvivlade på att forskningen skulle kunna åstadkomma något i sak” (1994, s. 52). Nog skulle det även kunna sägas om kärnkraften år 2023?

FORSKNINGENS RELEVANS – 30 ÅR SENARE

Det är lättare att bedöma dagsaktuella frågor än att sia om framtiden. Evert kunde knappast föreställa sig den starka utvecklingen för forskning som ägt rum under de trettio år som gått sedan han författade sin skrift. Han var i alla händelser helt korrekt då han menade att forskningsrelevans rimligen bör kopplas till forskningens förmåga att hantera samhällsproblem.

Dagens narrativ talar om samhällsutmaningar. För varje mandatperiod sedan 1980-talet har tillträdande regeringar lämnat nya forskningspolitiska propositioner till riksdagen. Dessa propositioner speglar den sittande regeringens syn på forskning, definierar angelägna samhällsproblem och pekar ut riktningen för kommande satsningar på forskning. Låt mig illustrera med den senaste forskningspolitiska propositionen (2020/21:60).

- Forskningspolitikens mål anges som att Sverige ska vara ett av världens främsta forsknings- och innovationsländer.
- Sveriges FoU-investeringar uppgår till 3,4 procent av BNP år 2019 och Sverige är det land inom EU som satsar mest på forskning och utveckling i relation till ekonomins storlek.
- Sverige har i flera år legat på första plats på EU-kommissionens rankinglista som mäter medlemsländernas innovationsförmåga.
- Sveriges universitet och högskolor är den största statliga sektorn räknat till antalet anställda.
- Svensk forskning av hög kvalitet anses enligt propositionen bidra till "vetenskapliga genombrott, hantering av samhällsutmaningar, utbildning av hög kvalitet, kompetensförsörjning och näringslivets och samhällets utveckling.
- Forskning och innovation anses nödvändigt för att skaffa den kunskap och de lösningar som behövs för att skapa ett hållbart samhälle.

- Forskningen är numera starkt internationaliserad och Sverige deltar i ett stort antal projekt och samarbeten inom EU.
- Anslagen till universitet, högskolor och forskningsråd kommer att öka.

Forskning ska leda till samhällspåverkan och propositionen pekar också ut ett antal angelägna samhällsutmaningar som forskningen behöver satsa på. I propositionen 2020/21 preciseras de som fem områden: klimat och miljö, hälsa och välfärd, digitalisering, kompetensförsörjning och arbetsliv samt ett demokratiskt och starkt samhälle. I forskningspolitiken beskrivs nu forskningens relevans i termer av hur den bidrar till att lösa några aktuella samhällsproblem. Identifieringen av just dessa samhällsutmaningar som forskningsproblem demonstrerar forskningsanvändningens nuvarande politisering.

Internationellt ökar mängden forskning kraftigt mätt i den globala artikelpresentationen. Forskningen levererar alltså. Tillväxttakten har varit ungefär fem procent årligen enligt Vetenskaprådet (2018). Inom flera områden är tillväxttakten oerhört omfattande och dessutom snabb, en "exponentiell utvecklingstakt" enligt framtidsforskaren Joakim Skog.²

För att sammanfatta har det skett en omfattande utveckling av svensk och internationell forskning, forskning är numera ett hög-prioriterat politiskt mål och högskoleutbildning och forskning är statens mest omfattande verksamhet mätt i antalet anställda. Dessutom sker en omfattande FoU-verksamhet inom näringslivet. Med Everts språkbruk tycks det som om forskningen ska betraktas som relevant. Att så mycket forskning produceras måste väl betyda att den är användbar och gör nytta?

FORSKNINGENS IRRELEVANS – 30 ÅR SENARE

Jag ska avsluta genom att ta upp mörka moln i den annars så ljusa uppfattningen om forskningens höga relevans. De var hart när omöjliga att förutspå för trettio år sedan, även om Evert berörde

2) <https://www.sydsvenskan.se/2023-08-24/kan-framtidstankande-bli-den-nya-svenska-framgangssagan>

några när han pekade på hur forskningsresultat kan användas för att legitimera politik och utnyttjas av taktiska skäl. Evert nämnde också, närmast i förbifarten, att viss forskning kan vara irrelevant om den inte hör till saken. Han gör emellertid inte själv någon större poäng av denna iakttagelse.

Dock har tiden nu sprungit ifrån Evert. Vi befinner oss i de alternativa sanningarnas tidsålder. Konspirationsteorier, filterbubblor, fakta- och kunskapsresistens och postdemokrati är några av dessa mörka moln. Det mobiliseras dock till försvar för fri forskning och dess betydelse i ett liberalt samhälle. Vetenskapsjournalisten Emma Frans har i flera böcker granskat forskningens relevans och sökt informera om hur forskningsresultat bör tolkas för att kunna uppfattas som trovärdiga (Frans, 2023). Propositionen 2020/21 lyfter särskilt fram vikten av att forskning ska tillåtas vara fri och att ett oberoende sökande av kunskap måste främjas.

Men de mörka molnen spökar. Forskningsanvändning är en starkt politiserad verksamhet. Jag kommer att peka på ifrågasättandet av demokratiska institutioner och hur värdet av kunskap blivit relativiserat. Jag kommer särskilt att utgå från arbetet i tanke-smedjan *Participation* vid Lunds universitet (Denvall & Iwarsson, 2022) och Lundasociologen Mikael Klintmans bok *Knowledge resistance* (Klintman, 2019).³ Som forskare upplever vi ofta att våra forskningsresultat marginaliseras, och det är inte alltid möjligt att effektivt förmedla deras betydelse även när de grundas på solida och pålitliga vetenskapliga undersökningar. Även om universitet och andra institutioner söker öka forskningens genomslagskraft, är utmaningarna att nå ut till en bredare publik och övertyga om våra slutsatser högst påtagliga.

Det är särskilt problematiskt när politiska ledare framträder med en övertygad ton i frågor där deras kunskapsunderlag är begränsat, även om omfattande vetenskaplig forskning finns tillgänglig. Denna diskrepans mellan politisk retorik och vetenskaplig evidens är en bekymmersam trend. Offentlig debatt präglas av en tendens att förkasta vetenskaplig evidens till förmån för starka åsikter.

3) Det är naturligtvis en ren tillfällighet att det råkar vara arbeten från Lunds universitet som varit mina huvudsakliga inspirationskällor och att Uppsala universitet blivit utelämnat.

Det verkar som om styrkan i ett påstående korrelerar med dess sanningshalt, och detta fenomen är en utmaning för vetenskapens och rationellt tänkandes legitimitet. Ju mer aggressivt och övertygande ett påstående framförs, desto större sanningsvärde tillskrivs det, oavsett det vetenskapliga underlagets styrka. Rationellt tänkande är nödvändigt för att samhället ska kunna utvecklas i demokratisk riktning hävdar Steven Pinker (2021). Men vi människor är nu inte alltid så rationella utan psykologiska, kognitiva och affektiva aspekter påverkar och styr.

ÖPPEN FORSKNING?

I antologin *Participation* pekar vi på hur dagens förändrade samhällsförutsättningar har skapat en miljö där nya aktörer har möjlighet att aktivt delta inom områden som forskning, innovation och planering. Dessa förändringar är drivna av flera samhällstrender, inklusive globalisering, etablerandet av sociala medieplattformar, ökad betydelse av nätverkssamhällen och framväxten av nya aktörer inom kunskapsproduktionen. Kunskap skapas inte bara inom universitet och högskolor. Dessa dynamiker har medfört ökad komplexitet i samhället och utmanat de traditionellt etablerade institutionernas roll och inflytande, inklusive massmedia, myndigheter och oss akademiska forskare.

Vi tar även upp digitaliseringens framsteg. Genom nya kanaler och plattformar för kommunikation och deltagande kan individer, brukare och medborgare uttrycka sina åsikter, vilket skapar synlighet. De kan nu organisera sig kring frågor som engagerar dem, något som öppnat upp för identitetspolitik. Idealt framstår samarbete och nätverksbyggande som nödvändiga förutsättningar för att påverka och påverkas inom dessa komplexa samhällsstrukturer.

Men denna utveckling har en dubbel karaktär. Å ena sidan har den öppnat dörren för antidemokratiska och auktoritära krafter att aktivt främja sina agendor, inklusive rasistiska och nationalistiska kampanjer. Det har gett upphov till nya politiska rörelser som utmanar de demokratiska normerna och spelreglerna. En studie utförd av V-Dem, ett institut vid Göteborgs universitet, visar att regeringspartier i många demokratiska stater under de senaste

årtiondena har antagit illiberala tendenser (Lührmann m.fl., 2020). Dessa inkluderar minskat engagemang för politisk pluralism, öppen demonisering av politiska motståndare, ökad brist på respekt för grundläggande minoritetsrättigheter samt uppmuntran av politiskt våld.

Statsvetaren Pippa Norris har visat hur förändringarna i en amerikansk politisk kontext bäddat för konspirationsteorier, urholkat tilliten till demokratiska institutioner och ökat samhällets polarisering (2022). Hon kopplar detta till kunskapsproduktionens digitalisering. I den digitala tidsåldern är information riklig men ofta opålitlig. Alternativa sanningar, "fake news" är vardag och gör det allt svårare att skilja på fakta och lögn. Hon diskuterar hur sociala medieplattformer och internet har förändrat informationslandskapet, vilket gör det lättare för desinformation att spridas snabbt.

Dessutom – enligt Nicholas Carr – förändrar vår exponering för internet och oavbrutna tillgång till smartphones vårt sätt att söka kunskap (Carr, 2020). När all information finns ett knapptryck bort behöver vi inte längre memorera utan det räcker att ställa frågor. Då blir det allt viktigare hur och till vilken sökmotor vi kopplar upp oss till. Hans förutsägelse är dystopisk; våra hjärnor förändras. Eftersom informationsmängden ständigt ökar blir vi alltmer impulsiva istället för reflexiva: "Far from enhancing human intelligence, I argue, the Internet degrades it" (2020, s. 1).

För att fortsätta den dystopiska vägen så varnade fysikern Max Tegmark i sitt sommarprogram 2023 för att världen riskerar att utplånas om inte samhället tar kontroll över en skenande utveckling av AI-forskning.⁴ Den hypersnabba utvecklingstakten i forskningen, otaliga kommersiella satsningar på generativ AI och riskerna med att illvilliga krafter använder AI destruktivt oroar många forskare. Man skulle kunna säga att forskningsrelevansen är så pass hög att den till och med skapar nya problem. Med Everts språkbruk är dess nytta tvivelaktig.

Å andra sidan, och i motsättning till de ovan nämnda utmaningarna, innebär den nuvarande tidsåldern också en rad initiativ som

4) <https://sverigesradio.se/avsnitt/max-tegmark-sommarpratara-2023>

syftar till att stärka individers engagemang i vardagsfrågor och att stödja deras möjligheter att delta i det offentliga samtalet. Forskningspropositionen betonar vikten av att medborgare får tillgång till forskningsinformation och forskningsråd som FORTE kräver att forskare ska redovisa hur de samarbetat med medborgare i sin forskning. Dessa initiativ strävar efter att ge röst åt medborgarnas angelägenheter och skapa möjlighet för öppna och konstruktiva dialoger.

FAKTA- OCH KUNSKAPSRESISTENS

Faktaresistens och kunskapsresistens utgör två aktuella och komplexa fenomen som Mikael Klintman har analyserat (Klintman, 2019). Faktaresistens kan definieras som en medveten tendens att undvika tillgänglig information och ny kunskap, trots att den är tillgänglig. Kunskapsresistens har en djupare dimension då den nya kunskapen utmanar individers etablerade världsbild och ideologiska övertygelser. Detta fenomen kan observeras i olika sammanhang där lojalitet gentemot djupt rotade religiösa och politiska övertygelser ger en särskild motståndskraft mot alternativa uppfattningar. Reportageboken *Det svenska missnöjet* är ett intresseväckande exempel (Lindell & Pelling, 2021). Två forskare försökte här förstå hur missnöjda låginkomsttagare resonerade. De samtalade med över 300 personer i sex orter som lastade problem med otrygghet och dålig ekonomi på invandrare snarare än att reflektera över varför den förda ekonomiska politiken drabbade dem själva så hårt.

En central aspekt inom fenomenet kunskapsresistens är att den åsikt som budbäraren förmedlar anses lika betydelsefull som själva informationen. Detta perspektiv på kunskapsförmedling kan förstås med hjälp av begreppet "Moral Tribes", och hur kunskap fungerar som en social sammanhållande kraft inom grupper. När aldrig så valid och säkerställd kunskap kommer i konflikt med våra grundläggande värderingar, sociala behov och grupptillhörigheter, uppstår en konflikt som kan ge upphov till kunskapsresistens.

Långt före termen "kunskapsresistens", utförde psykologen Leon Festinger och hans kolleger en studie om hur individer förnekar

faktiska omständigheter som går emot deras tidigare förväntningar (1956). De undersökte hur en grupp människor som förutspått jordens undergång reagerade när deras förutsägelse inte förverkligades, och jorden fortfarande existerade. De ändrade inte sin ursprungliga tro, trots det uppenbara motsägelsefulla beviset. I stället formulerade de en förklaring som rättfärdigade deras tidigare trosföreställningar. Deras starka kollektiva övertygelse om jordens undergång hade, enligt deras tro, på något sätt förhindrat katastrofen.

Festinger (1962) fördjupade sedan sin forskning och utvecklade teorin om kognitiv dissonans. Den belyser de mentala och emotionella konflikter som människor upplever när de konfronteras med motsägelsefull information eller upplevelser. I fallet med teorin om jordens undergång skulle människorna som trodde att jorden skulle gå under och ändå fann den oförändrad uppleva en kognitiv dissonans. För att lindra denna obehagskänsla hade de olika strategier till sitt förfogande. De kunde antingen ändra sin ursprungliga övertygelse och acceptera de faktiska omständigheterna eller, som här, förneka fakta och konstruera en alternativ tolkning som på något sätt bevarade deras kollektiva sammanhållning. Därigenom kunde obehagskänslan mildras utan att grundläggande övertygelser behöver ändras.

Kunskap är alltså i många avseenden en social konstruktion. Vi människor har en tendens att tro på information som delas av våra sociala grupper eller likasinnade. När en övertygelse delas av andra med liknande åsikter, kan det förstärka individens vilja att hålla fast vid den, särskilt om den upplevs som betydelsefull eller central för självbild eller grupptillhörighet. Detta understryker vikten av att förstå de psykologiska mekanismer som ligger bakom kunskapsresistens och de komplexa sätt på vilka vi människor hanterar och tolkar information som strider mot befintliga uppfattningar. Forskningen tycks har svårare att hävda sin relevans idag jämfört med då Evert skrev sin rapport.

RELEVANSPARADOXEN

Ja, du Evert, forskningens infrastruktur har då sannerligen förändrats under de senaste decennierna. Vi befinner oss i den digitala

revolutionen med supersnabb förändringstakt. Aldrig förut har världen haft en så omfattande och viktig kunskapsmassa till förfogande som samhällets forskning har skapat. Den har fört fram inte bara vaccin under pandemin utan nya former av energi och AI, något som förändrar världen. Relevansparadoxen innebär att å andra sidan utmanas forskning av "fake news" och desinformation och processer där människor får information som åsiktsbekräftelse. Om maskinerna, algoritmerna och AI tar över undergrävs inte bara demokratin utan människans förmåga att förstå på ett djupare plan. Till och med människans existens tycks vara hotad. Hur ska detta sluta?

Vetenskaplig kunskap, även om den anses trovärdig och reflexiv, är påverkad av sin sociala och historiska kontext. Det finns kanske vissa likheter mellan oss forskare och präster när vi förmedlar våra budskap och förväntar oss att omgivningen ska respektera dem. Emellertid tenderar moraliska och politiska aspekter att överskugga vetenskaplig kunskap, vilket leder till att budskapet inte alltid uppfattas som neutrala och därmed saknar relevans.

Både budskap och kunskap har en tendens att stärka sammanhållningen inom sociala grupper. Det är möjligt för varje grupp att bära på sin egen sanning. Det är energibesparande och skapar trygghet inom de grupper man tillhör. För mycket kunskap kan hota denna stabilitet, och det är utmanande att ta till sig information som kan kräva handling och förändring.

Oaktat dessa dystra iakttagelser måste ändå forskningens potential i utformning av ett demokratiskt och starkt samhälle lyftas fram. Kritiska och välinformerade medborgare är mer benägna att hålla politiska ledare ansvariga och att kräva insyn och ansvar från offentliga institutioner. Forskning har relevans genom att bidra till att motverka spridning av desinformation och polarisering av samhället. Välinformerade medborgare bör vara mindre benägna att styras av propaganda eller auktoritära ideologier, vilket gör oss alla mera motståndskraftiga mot manipulation.

Referenser

- Carr, N. (2020). *The shallows: What the Internet is doing to our brains*. WW Norton & Company.
- Denvall, V., & Iwarsson, S. (Red.) (2022). *Participation. Vad - När - Hur*. Studentlitteratur.
- Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford University Press.
- Festinger, L., Riecken, H., & Schachter, S. (1956). *When prophecy fails: A social and psychological study of a modern group that predicted the destruction of the world*. University of Minnesota Press.
- Forskningsproposition 2020/21:60 (2021). *Forskning, frihet, framtid – kunskap och innovation för Sverige*.
- Frans, E. (2023). *Expertparadoxen. Ska jag alltid lyssna på forskarna?* Bonnier fakta.
- Klintman, M. (2019). *Knowledge resistance. How we avoid insight from others*. Manchester University Press.
- Lindell, J., & Pelling, L. (2021). *Det svenska missnöjet*. Atlas.
- Lührmann, A., Medzihorsky, J., Hindle, G., & Lindberg, S. I. (2020). *New Global Data on Political Parties: V-Party*. Hämtad: https://www.v-dem.net/documents/8/vparty_briefing.pdf
- Norris, P. (2022). *In praise of skepticism: Trust but verify*. Oxford University Press.
- Pinker, S. (2021). *Rationalitet. Varför vi behöver det och hur vi får mer*. Fri tanke.
- Vedung, E. (1994). *Forskningens relevans, användbarhet, användning och nytta* (Vol. 17). Bygghälsningsrådet.
- Vedung, E. (1995). Utvärdering och de sex användningarna. I B. Rombach & K. Sahlin-Andersson (Red.), *Från sanningssägande till styrmedel: Moderna utvärderingar i offentlig sektor*. Nerenius & Santerus.
- Vedung, E. (2011). Spridning, användning och implementering av utvärdering. I B. Blom, L. Nygren, & S. Morén (Red.), *Utvärdering i socialt arbete: utgångspunkter, modeller och användning*. Natur & Kultur.
- Vedung, E. (2015). Six Uses of Evaluation. I V. Hennefeld, W. Meyer, & S. Silvestrini (Red.), *Nachhaltige Evaluation? Auftragsforschung Zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann* (s. 187-210). Waxman.
- Vedung, E. (2021). Twelve uses of evaluation. *Zeitschrift für Evaluation*, 20(1), 101-130.
- Vetenskapsrådet. (2018). *Vetenskaplig produktion. Analys av det vetenskapliga forskningssystemet*. Hämtad: <https://www.vr.se/analys/rapporter/vara-rapporter/2018-06-01-vetenskaplig-produktion---analys-av-det-vetenskapliga-forskningsystemet.html>

KAPITEL 5

Idéanalys och utvärdering*Leif Lewin*

Idéanalys och utvärdering – sammanställningen av dessa båda begrepp kan synas överraskande. Idéanalysen är en förnämlig gammal konst med doft av bibliotek och arkiv, utvärderingen en modern teknik hemmahörande i dagens luftkonditionerade kontorslandskap. Men det finns ett inre samband mellan de båda verksamheterna. Inspirerade av Jürgen Habermas hävdade vi som var unga doktorander på sextiotalet och skrev idéanalytiska avhandlingar, att genom att pröva ideologiernas hållbarhet och rensa ut det som var osammanhängande eller saknade empiriskt stöd bidrog vi till att sanera den offentliga debatten och göra den transparent och rationell (Vedung, 1971, s. 36; Lewin, 1972b, s. 14). Det är samma syfte, kan man säga, som vägleder moderna utvärderingar. Var den politiska intervention som är föremål för granskning välöverlagd? Uppnåddes det avsedda syftet? Eller fick reformerna andra effekter än de avsedda? När man sätter sig in i Evert Vedungs statsvetenskapliga gärning, kan man se hur det ena ger det andra och hur Evert i skrift efter skrift slår en bro mellan idékritik och policyanalys.

På sextiotalet kände vi emellertid att vi arbetade i motvind. Hindren restes inte av vår professor; Carl Arvid Hessler var tvärtom vår förebild och själva ikonerna för gammaldags idéhistoria

med egna, stilistiskt utsökta arbeten om den svenska religionsfrihetens problem. Men från nästan alla andra var motståndet massivt. "Behaviorism" hette den nya modeläran som ville ersätta idékritik med kvantitativ analys och, om politiska idéer alls skulle studeras, förordade undersökningar inte av idéernas hållbarhet utan av den roll de politiska programmen spelade i det politiska systemet (se till exempel Sainsbury 1968). Mot detta protesterade vi både genom att framhärda med en traditionell utformning av våra avhandlingar för att visa att idékritiken i allra högsta grad behöll sin relevans för en statsvetenskaplig förståelse av den politiska verkligheten och genom att i särskilda metoduppsatser argumentera för det berättigade i att studera "innehållet" och inte bara "funktionen" hos de politiska idéerna (Lewin, 1972a).

Den innehållsliga idéanalysen uppträder i olika former. En är den "klargörande" analysen, doktrinhistoriskt förknippad med just Habermas. Genom att korrigera de många avsiktliga och oavsiktliga missuppfattningar och förvrängningar som präglar den politiska debatten kan den ömsesidiga förståelsen förbättras och människor i många fall komma överens. En annan metod är den "giltighetsprövande" ansatsen i Poppers anda. Verklighetspåståendena i de politiska ideologierna kan undersökas precis som vetenskapliga utsagor och felaktiga eller förvirrade uppfattningar elimineras. Slutligen finns den "normativa" idéanalysen i skuggan av Rawls, som går över gränsen mellan vara och böra – inte så att de normativa utgångspunkterna kan "bevisas" men väl så att man utifrån några få grundprinciper, om vilka de flesta är överens, logiskt kan resonera sig fram till vissa allmänna slutsatser om det goda samhället. Därtill kommer, som sagt, den "funktionella" idéanalysen, som inte primärt syftar till att belysa idéernas hållbarhet utan det politiska systemets sätt att fungera (Lewin, 2019).

För dessa metodfunderingar visade dock Carl Arvid Hessler föga intresse. Han var en intuitiv humanist utan håg för vidlyftig teori, Habermas hade han aldrig hört talas om och, som Evert själv berättar, ogillade han Everts metodkapitel i avhandlingen (Vedung, 2018, s. 160). Jag förde fram våra tankar om idéanalysens mission att förädla det politiska samtalet också som fakultetsopponent på

Everts avhandling och under den efterföljande, skriftliga, ömsesidigt respektfulla replikskiftningen skulle man kunna säga att Evert tog det avgörande steget från idéhistoria till utvärdering, som han senare förklarar i förordet till sitt mästerverk *Det rationella politiska samtalet* (1977, s. 9):

Att ägna sig åt något så "kvalitativt" och "traditionellt" [som idéanalys] betraktas nog av en del som en mindervärdig och perifer sysselsättning. Detta kan bero på att man helt felaktigt likställer innehållslig idéanalys med politisk idéhistoria. Kanske tror man också att den innehållsliga idéanalysens metoder blott kan tillämpas på ideologier och mer storslagna tankesystem med liten praktisk relevans. Men det nya som skett sedan Herbert Tingsten utsatte de politiska ideologierna för sin idékritik är att samma principiella tillvägagångssätt börjat tillämpas på uttalanden fällda i ett politiskt vardags-sammanhang.

Nu var det väl framför allt Evert själv som börjat tillämpa idékritiken på vardagspolitiken – men han skulle få många efterföljare. Jag hade startat ett större forskningsprogram "Politik som rationellt handlande" (i vår forskarverkstad handlade det som synes mycket om "rationalitet" vid denna tid), på engelska *Politics As Rational Action (PARA)*; arbetet inleddes med ett internationellt symposium inom ramen för universitetets 500-årsjubileum och resulterade i en volym med samma namn som forskningsprogrammet och som Evert och jag redigerade tillsammans (Lewin & Vedung 1980). Jag bjöd in medarbetarna vid institutionen att ansluta sig till programmet och bad var och en att välja en samhällssektor för tillämpning av den modell för rationellt handlande som höll samman programmet. Evert fastnade för energisektorn. Det visade sig vara ett lyckokast. Kärnkraftsfrågan hade gjort energipolitiken till dagens hetaste politiska stridsfråga. Delegationen för energiforskning i Stockholm fattade intresse för Everts rationalistiska analys av energipolitiken och utöver bidragen till PARA:s publikationslista kom Evert med åren att genomföra en rad energipolitiska utvärderingar.

Som så ofta är fallet för vår generations statsvetare gav sedan kontakterna med amerikanska kolleger en skjuts åt forskningen. I början av åttiotalet vistades Evert vid *Lyndon Johnson School of Public Affairs* vid Texasuniversitetet i Austin. Där kom han i kontakt med *American Evaluation Society* och i denna inspirerande miljö fick han möjlighet att i ett flertal skrifter förfina de statsvetenskapliga utvärderingsmetoderna. Om Everts professionella utveckling har det för övrigt skrivits en ytterst läsvärd uppsats (Tranquist, 2015, se även Tranquist denna volym).

På samma sätt som är fallet med idéanalysen tar sig utvärderingar många olika uttryck och litteraturen på området tillväxer explosionsartat. Därvidlag excellerar Evert själv med många utarbetade klassificeringar av olika ansatser. Jag kan här inte redogöra för alla dessa distinktioner och komplikationer utan får nöja mig med några påpekanden.

Syftet med utvärderingar över huvud taget är alltså att ta reda på konsekvenserna av en viss intervention. En möjlighet är naturligtvis att resultatet blev det avsedda, att politiken lyckades. Men ända sedan Robert Mertons berömda uppsats från mitten av trettio-talet har statsvetarna mera intresserat sig för de fall där resultatet blev ett annat än det förväntade och oförutsedda omständigheter i stället avgörande för utfallet (Merton, 1936).

Frågan om politikerna har kontroll över händelseutvecklingen är central för det politiska ansvarsutkrävandet och därmed för den representativa demokratins möjlighet att fungera som det var tänkt. När något går fel, är det en vanlig strategi för politikerna att söka fly från sitt ansvar genom att skylla ifrån sig och säga att över detta hade jag ingen kontroll; det är andras fel att det blev som det blev: världsmarknaden, byråkraterna, koalitionspartnern eller rent av Historien själv som är förutbestämd att gå i viss riktning oberoende av enskilda aktörers handlande. Har jag inte något val, har jag heller inte något ansvar (Weaver, 1986; Lewin, 2007).

Att politikerna har "kontroll" är sålunda viktigt för att väljarna ska kunna hålla politikerna ansvariga. En utvärdering kan hjälpa politikerna att uppnå sådan kontroll. Ibland är syftet med en utvärdering i stället "kunskapsutveckling". Utvärderingen bidrar till

förståelse av grundläggande händelser och processer. Ett tredje syfte kallas "främjande". Utvärderingen går ut på att förbättra det granskade programmet, så att det blir mera effektivt. Slutligen talar utvärderingsforskarna också självkritiskt om utvärdering som "Potemkinkuliss". Huvudmannen använder utvärderingen för att ge sken av att allt står väl till. En utredning tillsätts för att vinna tid. En utvärderare tillkallas för att visa upp en front av kompetens eller med uppdrag att ensidigt leta upp positiva inslag. För sådana försök betackar sig naturligtvis den sanne utvärderaren.

Hur avgör man om en viss politik lyckats eller inte? Ett sätt är att jämföra med aktörernas intentioner. Men medborgarna och skattebetalarna ställer ofta en annan fråga: helt enkelt blev politiken "bra"? Liksom idéanalytikern får då utvärderaren axla den normativa manteln och jämföra resultatet med kriterier som hämtas från annat håll än utredningsdirektiv eller propositionstexter, kanske från den politiska filosofin med dess principer om rättvisa, jämlikhet eller legitimitet. Normativ utvärdering är en grannliga uppgift, som måste utföras med omdöme för att rapporten inte ska klassas som subjektiv och politisk. Kriterierna måste, precis som för den normativa idéanalysen, just vara sådana allmänt omfattade grundprinciper som jag nämnde ovan.

Evert återkommer ofta till bränsledeklarationen. Med det exemplet vill han visa hur intrikat ett utvärderingsuppdrag kan vara. Huvudpoängen med en bränsledeklaration är inte att förse bilköparna med korrekt information om bränsleåtgång. Syftet är att påverka dem att köpa mera bensinsnåla bilar och därmed uppnå lägre energiförbrukning i samhället. Men detta i sin tur är inte entydigt. Man kan antingen tolka målet som en absolut nedgång i personbilssektorn eller som en relativ minskning per körd mil. För att avgöra om bränsledeklarationen är en framgång eller inte måste vi med andra ord ta reda på både om svenskarna köper mera energisnåla bilar och om detta leder till minskad bränsleåtgång per körd mil. Evert framhåller ytterligare komplikationer i detta exempel men jag skyndar här till hans allmänna slutsats att vi kanske borde förskjuta intresset från sakpolitik till procedurer för utvärdering, det Evert kallar institutionslinjen. Vi ska kanske

prioritera byggandet av oberoende institutioner, som kan utföra utvärderingar på ett professionellt sätt.

Politiskt handlande leder ibland till överraskningar, det är utvärderingsforskningens huvudbudskap. En speciell form av överraskning kallas serendipitet, ännu ett ord med en alldeles speciell doft, i det här fallet av 1700-talets skräckromantik efter beteckningens upphovsman den skönlitteräre författaren Horace Walpole. Innebörden är att om något oväntat inträffar – en olycka, en tillfällighet, en katastrof – kan det ibland vändas till något gott. Så upptäcktes penicillinet, när Alexander Fleming återkom till sitt laboratorium efter semestern och hittade en odiskad bakteriekultur som det börjat växa bakteriedödande mögel på. Så upptäcktes röntgenstrålarna, när Wilhelm Röntgen fann ett märkligt ljussken från några papper en bit från det glasrör täckt av svart papper han arbetade med för att studera elektriska urladdningar. Så upptäcktes Amerika, när Christoffer Columbus sökte finna sjövägen till Indien.

Sextiotalets underlägsenhetskänsla bland dem som arbetade med kvalitativa metoder har förbytt i självförtroende och stolthet. I referenslistan nedan kan läsaren se att Evert i en uppsatsrubrik numera rent av placerar idéanalysen ”i universums mitt”. Och visst har pendeln svängt och intresset för kvalitativ analys återkommit! Det hindrar inte att utvärderingar idag utsätts för kritik. Bakgrunden är genombrottet för den nya förvaltningspolitik som kallas *New Public Management (NPM)*. Under åttiotalet började den starka välfärdsstaten att upplevas som alltför stark, alltför centraliserad, alltför stelbent. Förändringens vind började blåsa med marknaden som förebild och decentralisering, privatisering och konkurrens som metoder både för att effektivisera välfärdspolitiken och för att göra medborgarna mera delaktiga. Men staten får inte abdikera (Lewin, 2014). Staten måste i efterhand komma och se efter om målen uppnåddes – i skolan, i vården, i rättssamhället, i forskningen. Med decentraliseringen följde kravet på utvärdering och en formidabel utvärderingsindustri tog sin början. Den allmänna uppfattningen idag torde vara att denna utvärdering gått för långt och blivit alltför tidskrävande. Den fäster också

uppmärksamheten vid fel saker. NPM hämtar sina ideal från redovisningsbranschen med dess trånga synsätt, som det kan vara frestande att kopiera men som svarar illa mot förhållandena i många offentliga verksamhetsgrenar. Redovisningsbranschen "koloniserar" det område som ska undersökas med sina högst speciella, kvantitativa och ekonomiska mätningar och tvingar fram en omfattande dokumentation, som ska ligga till grund för utvärderingarna. På så sätt har ett "utvärderingsmonster" sett dagens ljus som tvingar medarbetarna att sitta framför ofta dåligt fungerande datorprogram och drar bort dem från deras egentliga arbetsuppgifter (Power, 1997; Lindgren 2006). För Evert är denna sorts utvärderingsverksamhet främmande. För Evert är de lättköpta indikatorerna och mätningarna otänkbara. Hela hans utvärderingsforskning går ut på att fånga upp kärnverksamheten i den offentliga politiken.

I inledningen till den flitigt använda *Utvärdering i politik och förvaltning* sätter Evert in utvärderingsforskningen i sitt idéhistoriska sammanhang. Han ser den som en del av en rationalistisk tradition, som accelererat under senare årtionden. Vissa forskare vill dra perspektivet ännu längre tillbaka, ända till jordens skapelse, kallar Gud den förste utvärderaren och citerar Första Mosebok: "Och Gud såg på allt som han hade gjort, och se, det var mycket gott." Att använda stringenta begrepp, att resonera sammanhängande, att belägga sina teser bör vara målet för såväl vardagspråk som utvärderingar. Klarhet, logik, evidens, intersubjektivitet är förnuft förädlad till vetenskap.

Den moderna utvärderingsforskningen är otvivelaktigt en del av en rationalistisk strömning i politik och förvaltning som vuxit fram i moderna västerländska stater sedan andra världskriget och särskilt efter 1960. Tanken har varit att den offentliga politiken ska rationaliseras och bli mera vetenskaplig ... Att utvärdera är att blicka bakåt i syfte att förbättra manövreringen framåt. Utvärderingens roll är att systematiskt ta fram och bedöma information om den förda politikens genomförande och resultat för att ge underlag för mer rationella och långsiktiga beslut (Vedung, 1991, s. 15 och 18).

Så framstår Evert som en nyttoinriktad upplysningsfilosof, som genom kreativitet och målmedvetenhet lyckats förnya den metod han en gång skolades in i genom att oförväget tillämpa den på ett nytt område.

Referenser

- Lewin, Leif (1972a). Om studiet av de politiska ideologiernas innehåll och funktion, *Statsvetenskaplig tidskrift*.
- Lewin, Leif (1972b). *Statskunskapen, ideologierna och den politiska verkligheten*. Installationsföreläsning. Rabén & Sjögren.
- Lewin, Leif (2007). *Democratic Accountability. Why Choice in Politics is Both Possible and Necessary*. Harvard University Press.
- Lewin, Leif (2014). *Staten får inte abdikera. Om kommunaliseringen av den svenska skolan*. SOU 2014:5. Fritzes.
- Lewin, Leif (2019). Politisk idéanalys enligt Habermas, Popper och Rawls, *Statsvetenskaplig tidskrift*.
- Lewin, Leif & Vedung, Evert eds. (1980). *Politics As Rational Action. Essays in Public Choice and Policy Analysis*. D. Reidel.
- Lindgren, Lena (2006). *Nya utvärderingsmonstret. Om kvalitetsmätning i den offentliga sektorn*. Studentlitteratur.
- Merton, Robert K. (1936). The Unanticipated Consequences of Purposive Social Action, *American Sociological Review*.
- Power, Michael (1997). *The Audit Society: Rituals and Verifications*. Oxford University Press.
- Sainsbury, Diane (1968). A Critique of Leif Lewin's Planhushållningsdebatten, *Statsvetenskaplig tidskrift*.
- Tranquist, Joakim (2015). The Oral History of Evaluation: the Professional Development of Evert Vedung, *American Journal of Evaluation*.
- Vedung, Evert (1971). *Unionsdebatten 1905: en jämförelse mellan argumenteringen i Sverige och Norge*. Almqvist & Wiksell.
- Vedung, Evert (1977). *Det rationella politiska samtalet. Hur politiska budskap tolkas, ordnas och prövas*. Aldus.
- Vedung, Evert (1991). *Utvärdering i politik och förvaltning*. Studentlitteratur.
- Vedung, Evert (2018). Idéanalys från ankdamm till universums mitt, *Statsvetenskaplig tidskrift*.
- Weaver, R. Kent (1986). The Politics of Blame Avoidance, *Journal of Public Policy*.

KAPITEL 6

Behøver evaluering ikke længere demokratiet?

Peter Dahler-Larsen

INDLEDNING

Hvordan finder en evaluator ud af, hvilke evalueringskriterier, der er legitime? Hos Evert Vedung er svaret klart. Legitime evalueringskriterier er rodfæstet i det repræsentative demokratis parlamentariske styrekæde (Vedung, 2009, s. 94). Hvis man for eksempel evaluerer ud fra mål, er målene ikke selvopfundne eller vilkårligt sat af en given opdragsgiver eller evaluator. De mål, der tæller, er kun dem, som er institutionelt anerkendte, ved at de er sanktioneret i styrekæden. Det repræsentative demokratis styrekæde har normativt primat.

En nysgerrig sjæl, som ser sig rundt i det aktuelle samfunds evalueringslandskab må imidlertid undre sig. I praksis har det repræsentative demokratis styrekæde en begrænset rolle at spille. Det vil jeg underbygge ved at argumentere for

- at evaluatore og udbydere af evalueringsinformation begrunder evaluering på mange andre måder eller endda undlader en sådan begrundelse uden at bekymre sig om nogen parlamentarisk styrekæde.

- at en hel del evaluering finder sted uden at være institutionelt ansvarlig i forhold til den parlamentariske styrekæde
- og, at en hel del evaluering, selv om den står i forbindelse med den parlamentariske styrekæde, samtidig gør dette på en måde, hvor hverken styrekæden eller dens primat står klart.

Tilsammen betyder dette, at det empiriske grundlag for det repræsentative demokratis styrekæde som primært legitimeringsprincip for evalueringskriterier står svagt og måske stadig svagere.

Så spørgsmålet er: Kan et normativt princip, så som princippet om den parlamentariske styrekædes normative primat i evaluering, være værdifuldt, selv om dets empiriske fundament er ved at falde sammen?

Besvarelsen af spørgsmålet skrider logisk frem. Først redegøres for den parlamentariske styrekæde i et repræsentativt demokrati som normativt princip i evaluering. Så argumenteres der for – ud fra ovenstående tre teser – at det empiriske grundlag for dette princip er svækket. Til sidst diskuteres meningsfuldheden i at holde fast i dette princip, for eksempel i undervisningen af nye evaluatore.

DET REPRÆSENTATIVE DEMOKRATIS STYREKÆDE

Det kendetegner legitime evalueringskriterier, at de er sanktioneret i den parlamentariske styrekæde i et repræsentativt demokrati. Dette forklares måske allermest tydeligt af Evert Vedung i forbindelse med målopfyldelsesmodellen (Vedung, 2009, s. 94-95). I et demokrati udgår al magt fra folket. Herfra forsyner hvert led i kæden det følgende led med et mandat til at handle på det foregående leds vegne ud fra princippet om delegering. De overordnede har imidlertid ikke selv tid til at udføre alle de vedtagne opgaver. De delegerer opgaveløsningen nedad i styrekæden. Opgaveløsningen vurderes ud fra demokratisk fastsatte mål. Det er målopfyldelsesmodellens store fortjeneste, at den er så tydeligt forankret i en demokratisk statsform.

På dette grundlag kan bestemmes fire legitime kontrollensyn i et repræsentativt demokrati (Vedung, 2009, s. 167-69). Via evaluer-

ing kan folket kontrollere de folkevalgte. De folkevalgte kan kontrollere embedsmænd. Embedsmænd kan kontrollere leverandører af ydelser og services. Og folket kan direkte kontrollere serviceproduktionen. Evaluering *sikrer* ikke, at styrekæden virker. Evaluering undersøger *om* den virker. Evaluering findes, fordi styrekæden i praksis er fejlbarlig. Det er styrekædens normative princip imidlertid ikke. For at finde ud af, om en kæde er bristet, bruger man kædens ide som normativt grundlag for bedømmelsen. Dette grundlag er objektivt, altså adskilt fra hvad du og jeg i øjeblikket synes om de valgte mål og evalueringskriterier. Derfor giver evalueringens forankring i styrekæden et objektivt svar på, hvorfor jeg skal respektere din evaluering. Evaluatorene er børn af det repræsentative demokrati.

Sådan lyder min tolkning af Vedungs forankring af evaluering i den parlamentariske styrekæde. Styrekæde er primær, det vil sige "øverst" i et værdihierarki og tidsligt "først" i et ræsonnement om evalueringens tilrettelæggelse.

Prioriteringen skinner igennem i Vedungs eget arbejde. Hos ham sker inddragelse af brugere i evaluering for eksempel inden for rammerne af det repræsentative demokratis styrekæde som den primære institutionelle ramme (Dahlberg & Vedung, 2001).

En interessant og kompleksitetsskabende variant af denne tankegang ser man hos Hertting og Vedung (2009). Her er ideen, at en evalueringens genstand kan og bør vurderes ud fra den officielt vedtagne styringsdoktrin, den hviler på. Kompleksiteten kommer af, at et repræsentativt demokrati kan legitimere flere samtidige styringsdoktriner (så som f.eks. målstyring, netværksstyring og brugerstyring) og at en given evaluering også kan have til opgave at korrigere en given styringsdoktrin (Hertting & Vedung, s. 2009: 83).

Den underliggende ide er imidlertid, at en sådan ny, heroisk opgave for evaluering fortsat er defineret på legitim måde i det repræsentative demokratis styrekæde.

ET VAKLENDE EMPIRISK FUNDAMENT

Er alle evaluatorene og evalueringssudbydere fortalere for princippet om forankring af evalueringens normative grundlag i den parlamen-

tariske styrekæde? Nej, tydeligvis ikke. Lad mig begrunde svaret under tre teser.

Tese 1. Evaluering legitimeres uden henvisning til den parlamentariske styringskæde

Vi har i de seneste årtier set en vækst i brugen af forskellige nøgleord til at begrunde evaluering. Læring har stået centralt. Læring kan naturligvis være knyttet til politik ført inden for det repræsentative demokratis rammer, men læring kan også benyttes bredt og uspecifikt. Alle kan lære noget i forskellige retninger. Evaluering har også annammet termer som "transformativ", "emancipatorisk", "deliberativ", "participatorisk", "empowerment", "responsiv", "innovativ", "systemisk". Hver af disse kan givetvis åbne for en mangfoldighed af interessante perspektiver, men de fleste bibetydninger vil enten være uafklarede i forhold til det repræsentative demokrati eller endda stå i et eller andet konkurrenceforhold hertil.

Allermest interessant er det måske, at der også tilbydes evalueringsinformation, som tilsyneladende ikke suppleres af nogen begrundelse eller legitimering.

Et eksempel er Google Scholar, som oplyser om citationstal for forskere og deres publikationer. En række evalueringsfaglige spørgsmål fremstår uklare. Det er uoplyst, hvilket begreb, antal citationer er en operationel indikator for. Der findes ikke – mig bekendt – en offentligt tilgængelig beskrivelse af, hvordan Google Scholar teknisk og operationelt definerer en "citation". Det kan være problematisk, hvordan indikatoren i givet fald bedst fortolkes i lyset af disciplin, subdisciplin, forskerens køn og alder med mere. Ligeledes er sproget af betydning for tolkningen af citationsdata, fordi engelsk som publikationssprog både kan repræsentere øget kvalitet og øget cirkulation. Blandt to sproglige varianter af samme tekst har den engelske version ofte højere citationstal end en tilsvarende tekst udgivet i et mindre sprogområde. (For eksempel har "Public Policy and Program Evaluation" af Evert Vedung væsentligt flere Google Scholar citationer end "Utværdering i Politik och Forvaltning" af samme forfatter, selv om sidstnævnte værk hverken står tilbage i originalitet eller kvalitet.)

Det er uklart, hvilke evaluative afgørelser, Google Scholar ønsker at understøtte og i hvilket omfang citationer udgør et grundlag for officielle evaluative afgørelser. I det daglige kan hensynet til citationer have konstitutive virkninger på forskeres arbejde, som i omfang væsentligt overstiger hensyn til vedtagne nationale forskningspolitikker. Google Scholar giver i løbet af sekunder et overblik over forskelle mellem forskere uden at man behøver læse deres publikationer. Det får mange forskere til at kere sig en hel del om deres citationstal (Fochler og de Rijcke, 2017). Det sker også før, at Google Scholar citationer når frem til at blive brugt officielt i officiel evaluering. Men de kan øve betragtelig indflydelse på officielle evalueringsbeslutninger for eksempel ved at rangere forskere og etablere præmisser for bedømmelser.

En del anerkendte forskere opretter deres personlige side i Google Scholar. De har mulighed for at rense ud i dobbeltregistreringer eller fjerne uvedkommende tekster med fejlagtige navnesammenfald. Den enkelte forsker hjælper dermed Google Scholar med at varetage udbudsfunktionen.

Vi står med et evaluativt materiale, som både er udbudt og anvendt (også af kendte og dygtige evalueringsforskere), uden at der foreligger nogen officiel og aftalt normativ begrundelse. Det er ”op til den enkelte”, hvad disse evaluative data skal bruges til. Tilsyneladende behøver udbyderen ikke at redegøre for hensigten eller legitimiteten heri. Ikke blot undlader udbyderen at begrunde sig i forhold til nogen parlamentarisk styrekæde. Den undlader helt at begrunde sig overhovedet. Evaluering ser ud til at leve fint uden demokratiet.

Tese 2. Evaluering finder sted uden institutionel ansvarlighed over for den parlamentariske styrekæde

Store tekniske og institutionelle kræfter frisættes i evalueringslandskabet. De er knyttet til internationale organisationer, til moderne ideer, omfattende infrastrukturer til skabelse og formidling af data og til omsætning af evalueringsinformation til praktiske handlinger og beslutninger. Det sker via socio-materielle og digitale netværk helt uden om landenes parlamentariske styringskæder (Dahler-Larsen, 2019).

Ingen regering har bedt Google om at udarbejde og offentliggøre Google Scholar data. Hvorfor er man overhovedet kommet på ideen? Sandsynligvis fordi det er billigt for Google at træne en maskine til at identificere noget, der meget ofte ligner en videnskabelig publikation med en forfatter. Ved at offentliggøre maskinens fund (og bede forskere opstille deres egne sider) knytter Google både forskere og personer interesseret i publikationer til virksomheden. Den enkelte forsker kan endda abonnere på nyheder om citationer af sit eget værk. Det styrker selvtilliden og bygger forbindelser mellem forskningsmiljøer. Måske formidles endda blandt forskere – en interessant og vidende gruppe af interessenter – et positivt billede af Google. Det er antageligvis af nytte for virksomhedens omdømme og dermed kommercielle interesser på længere sigt, også selv om Google Scholar citationer stilles gratis til rådighed. Men alt dette er gætterier. Der foreligger som sagt ingen formel redegørelse og begrundelse.

Tese 3. Selv når evaluering står i forbindelse med den parlamentariske styrekæde, gør den det på en uklar måde

Betragt som eksempel de internationale PISA-undersøgelser af skolen. PISA er "Program for International Student Assessment." De nationale regeringer står naturligvis bag beslutningen om de enkelte landes deltagelse i PISA. Regeringerne åbner for datafangst blandt skoler og elever.

Men styrekæden er ikke tydeliggjort, hverken af regeringerne eller folkene bag PISA. De færreste danske lærere ved, hvem der laver PISA-undersøgelser og hvor de skal henvende sig, hvis de vil klage over dem. OECD har ikke etableret et midlertidigt telt på Rådhuspladsen i København, hvor aktører i den danske skole kan tale med ministeren og ophavsmændene bag undersøgelsen. Hvis der var en styrekæde, var den aldrig eksplicit, synlig og kontaktbar.

Selve formen i en international sammenligning betinger, at der ikke kan tages udgangspunkt i hvert enkelt lands curriculum. I stedet stipulerer en række eksperter nogle almene kompetencer, som antages at være relevante i fremtidens konkurrencesamfund.

Der er naturligvis teknisk set tale om en "delegering" fra den danske regering til organisationen bag PISA, nemlig OECD, som er en økonomisk samarbejdsorganisation. Organisationens overordnede idealer handler om økonomi og konkurrence (hvor de f.eks. kunne have været et ideal om international vidensdeling og dannelse). Efterfølgende kan man blive forvirret som aktør i den nationale skolepolitik. Gælder det om at realisere allerede fastsatte nationale uddannelsesmål eller om at score højt i PISA? Når forskere efterfølgende undersøger effekten af PISA på de nationale skolepolitikker, ser det i deres analyser ud til, at PISA har stor indflydelse på skolepolitikken (Grek, 2009).

Lad mig dermed udfordre en ide om, at den parlamentariske styrekæde har primat i betydningen "ligger først i en tidlig rækkefølge." I eksemplet med PISA er det som sagt korrekt, at beslutningen om at deltage i PISA er taget "først" på en legitim måde i den parlamentariske styringskæde. Men tilbagevirkninger på uddannelsespolitikken i anden omgang synes at være større og mere omfattende end regeringens indflydelse på PISA i første omgang.

Yderligere et eksempel: I Danmark har man vedtaget, at nationale test skal være digitalt selvværdende og højst må tage 45 minutter. Det har store konsekvenser i faget dansk. Man har helt opgivet at bede elever om at formulere noget selv. Et kreativt produkt kan ikke vurderes digitalt. I stedet er testen i dansk reduceret til en test i læsning. De tekniske begrænsninger i evalueringsformen ender med at indsnævre evalueringskriterierne på en måde, som man måske ikke helt havde gennemskuet, da man tog de første beslutninger om testenes design.

Effekten – at evaluering virker tilbage på politikken – gentager sig i mere komplekse former. På det seneste har man i et kompromispræget politisk forlig vedtaget frivillige test i faget historie. En faglig ekspertgruppe har peget på det problem, at der er en sådan frihed i tilrettelæggelsen af historiefaget i skolen, at man en given frivillig test i en given skole kan kommet til at teste eleverne i stof, som de ikke har gennemgået. En løsning kunne være deklarerer varianter af de frivillige test i historie med hver sit emne, så den enkelte

lærer kunne vælge en test i et emne, der var gennemgået. Det ville formodentlig medføre, at undervisningen på forhånd blev rettet mod de opstillede test-emner. Fænomenet ville være en variant af den velkendte "teaching-to-the-test." I et forsøg på at undgå den effekt foreslår en politiker derfor at vedtage obligatoriske læseplaner i historie. På den måde vil undervisningen i kraft af en officiel beslutning blive bragt i overensstemmelse med testens krav. Forslaget består med andre ord i at *tilpasse de officielle mål til testen, ikke i tilpasse testen* til de officielle mål.

Lovgivningen er naturligtvis i dette eksempel ikke sat ud af kraft. Men den spiller andenviolin. Evalueringsmaskineriet har nået en sådan grad af selvstændiggørelse og kraft, at det i eksemplet stiller krav til det politiske system eller i hvert fald bliver brugt som en bekvem begrundelse for et forslag om at ændre lovgivningen. At styrekæden som sådan stod bag den oprindelige beslutning om test er interessant nok. Men nu bliver styrekæden ramt af konsekvenserne af sin beslutning, som den tydeligtvis ikke havde tænkt igennem. Det er svært i eksemplet at se, at styrekæden som sådan skulle have primat.

Der er også eksempler på, at evalueringskriterier bliver selvbekræftende, efter at de måske på et tidspunkt (for længe siden, under tidligere regeringer), blev vedtaget legitimt. Som eksempel kan tjene det danske BFI-system. BFI var et system for optælling af værdien af forskningsmæssige publikationer, hvor det blev givet point til artikler og bøger alt efter tidsskrifternes og forlagernes prestige. Efterfølgende blev en lille del af de offentlige midler til universiteterne fordelt i forhold til disse point. Ved en evaluering af BFI-systemet kunne man så udtale sig positivt om hele systemet efter at have registreret en stigning i BFI-point i den undersøgte periode. En positiv udvikling i systemets eget evalueringskriterium tages til indtægt for hele systemets succes. Det stod mindre klart, om der var sket en indfrielse af de politiske målsætninger, som måtte have været begrundelsen for systemets indførelse oprindeligt.

Ultimativt blev hele BFI-systemet nedlagt politisk. Så den parlamentariske styrekæde kan godt vise sine muskler. Men mine

eksempler viser, at styrekæden – i bedste fald – er en spiller blandt mange.

HVORFOR ER DET BLEVET SÅDAN?

Udbredelse af digital teknologi har gjort det muligt for en meget bred vifte af aktører at indsamle og formidle evalueringssinformation. Konsekvenserne af evalueringssinformationer drages uden om de nationale regeringer via hjemmesider og netværker. Organisatoriske felter med fælles værdier og normer findes på tværs af landegrænser. Internationale organisationer spiller en betydelig rolle som normsættere også for national politik.

Evalueringssinformation kan også produceres af konsulenter eller kommercielle virksomheder uden for politiske systemer for eksempel som forbrugerinformation eller "gratis goder", der leveres som "tilgift" til andre ydelser i en markedssammenhæng.

Udbydere af evalueringssinformation så som internationale organisationer, sociale mediehuse, tænketanke og konsulenthuse er kun i meget begrænset omfang genstand for demokratisk kontrol og feedback-mekanismer, der kan gøre dem følsomme for input om værdikonflikter og konstitutive virkninger. Disse forskellige institutionelle sammenhænge giver evalueringssystemer hver deres særskilte kriterier. Endda ser man, at brugen af evalueringssinformation bliver et argument i sig selv, også blandt seriøse evalueringssforskere (Esposito og Stark, 2019).

Evaluering som moderne fænomen bidrager til denne fragmentering ved at splitte aktører op i forskellige roller: Eksperter, statistikere, psykometrikere, evalueringssrådgivere, konsulenter, beslutningstagere, managers, brugere. Det er kun en begrænset delmængde af disse roller, der er integrerede i den parlamentariske styrekæde. Det er kun ved opbydelsen af den yderste abstraktions-evne, man bag ved disse roller kan fornemme et "res publica", et fælles offentligt anliggende.

Tilbagevirkningerne af evalueringssystemer på politiske sager og sociale relationer er til gengæld så omfattende, at de ikke lader sig begribe med gængse samfundsvidenskabelige midler. Udviklingerne i teknik og organisation går så stærkt, at der ikke er tid nok i

en repræsentativ styrekæde til at indsamle information og træffe bindende beslutninger om mål, værdier, teknologier og evaluering (Rosa, 2005). De nationale regeringer er ikke længere styrende. De er reaktive. Og selv evalueringssubjekter.

Hvor et grundlæggende problem tidligere var, at de politiske ledere ikke selv havde tid til at udføre de politikker og opgaver, som de selv vedtog, så er et stort problem idag, at de ikke har tid til at tage sig af konsekvenserne af evaluering, heller ikke i tilfælde af evalueringssystemer, de selv har vedtaget. Men hvor det første problem havde en institutionel løsning, nemlig delegering og etablering af en styrekæde, så mangler det andet problem fortsat en løsning.

TILBAGE TIL SPØRGSMÅLET

Kan et normativt princip være værdifuldt, selv om dets empiriske fundament er ved at falde sammen? Svaret er ja.

Der er, måske paradoksalt, fortsat grund til at være opmærksom på Evert Vedungs optagethed af, hvordan legitime evalueringsskriterier begrundes, også selv om det i svindende grad empirisk foregår sådan, som Vedungs normative ideal foreskriver.

Med Vedung in mente er vi i god ret til at spørge kritisk til enhver evaluering: Hvordan er værdigrundlaget begrundet? Dette spørgsmål udgør en vaccine mod skødesløshed og ligegyldighed i valget af evalueringsskriterier. Mange aktuelle evalueringsskemaer fremhæver ikke betydningen af *res publica* for os. De lader i stedet som om, de ikke har brug for demokratiet. De finder begrundelser (eksplicit eller ofte implicit) andetsteds. Det kan være i, at de er metodisk raffinerede: Evaluering fra den såkaldte "øverste hylde" rent metodisk. Det kan være i, at de angiveligt gør én parat til en uafvendelig international konkurrence. Det kan være, de understøtter "læring" uden dog at forklare, om læring i en retning er bedre end en anden. Yderligere medvirker opsplitningen af talrige roller i forhold til evaluering (specialist, rådgiver, objekt, manager, udbyder, projektleder, intern evaluator, ekstern evaluator osv) yderligere til en fragmentering af forestillinger om, hvad man er fælles om på tværs af rollerne.

Hvis nogen borger puster røg fra cigaretter ud, er der andre borgere, som ånder den ind. Hvis et land går i krig, er konsekvenserne ikke kun individuelle. Vi er, hvad enten vi bryder os om det eller ej, lænket til et fælles anliggende, et "res publica." Evaluering kan hjælpe os til at opdage og genopdage "res publica".

VEJE FREM

Hvis samfundet er noget vi lærer at kende ved at opdage nogle fælles erfaringer (Dewey, 2017), kan vi så bruge evaluering til at skabe sådanne erfaringer og genopdage, at vi hænger sammen i et samfund? Hvis vi kan svare på det, kan vi samtidig finde grunde til, at jeg overhovedet skal respektere din evaluering. Og på, hvordan vi kan reetablere forbindelserne mellem evaluering og demokrati.

Ifølge Rosanvallon (2011) har vore hidtidige forestillinger om demokratisk legitimitet har været baseret på fiktioner for eksempel om "Folket." Folket har mistet sit enhedspræg og det er ikke længere oplagt, at et flertal repræsenterer en totalitet. En lang række forskellige mindretal gør også krav på forskellige former for legitimitet. Sidst, men absolut ikke mindst, kan vi ikke længere antage, at det legitime valg af en regering, der tager beslutninger på legitime måder, også forsyner indholdet i en vedtagen politik med stabil legitimitet. Samfundet er de facto på vej væk fra en monistisk ide om legitimitet.

I Rosanvallons perspektiv udvikler demokratiet sig imidlertid med historien. Der kan være frem, vi endnu ikke har lært (Rosanvallon, 2009). Måske kan vi etablere nye forbindelser mellem evaluering og demokrati. Han antyder, at det kan ske gennem principper om upartiskhed, refleksivitet og åbenhed for dialog. Men hans tanker adskiller sig fra den klassiske forestilling om det repræsentative demokrati ved at legitimitet i Rosanvallons øjne mangfoldig, flydende og ustabil.

Vi kan undervise kommende evaluatore i et bredt samfundsmæssigt perspektiv på evalueringskultur, evaluativ information og dens mange samfundsmæssige konsekvenser og ikke at forveksle evaluerings selvforståelse med virkeligheden.

Vi kan øge opmærksomheden på konstitutive virkninger af evaluering ved at forske i dem, undersøge dem og gøre dem til genstand for offentlig debat (Espeland og Sauder, 2007).

Vi kan anerkende, at vi i egenskab af brugere, professionelle, ledere og borgere medvirker til at fuldbyrde virkninger af evaluering, når vi bruger og genbruger evalueringsinformation, uden at det er resultat af nogen officiel beslutning. Vi har altså et medvirkende ansvar, også uden for styrekæden.

Vi kan opfinde institutionelle mekanismer, således at nationale og internationale udbydere af evalueringsinformation responderer mere direkte på input fra de mennesker, der omfattes af evalueringernes konsekvenser.

Vi kan eksperimentere med demokratiske formater for deliberation om evalueringers og evalueringssystemers virkninger.

Vi kan optælle og påpege de mange og stigende omkostninger ved evalueringssystemer og dermed bane vejen for, at evaluering ikke altid er den bedste løsning på givne problemer. Nogle politiske problemer kræver politiske løsninger. Eksistentielle problemer kræver eksistentielle løsninger.

Det har altid været evalueringens opgave at gribe fat i et misforhold mellem ideal og realitet. Nu er det tiden at tilse misforholdet mellem ideal og realitet i evalueringens felt selv. Derfor er det vigtigt at blive ved med at stille spørgsmål om evalueringens værdimæssige begrundelse, som Evert Vedung gør, også selv om fremtidige generationer må finde nye sammenhænge mellem evaluering og demokrati.

Referenser

- Dahlberg, M. og Vedung, E. (2001). *Demokrati och Brukarutvärdering*. Studentlitteratur.
- Dahler-Larsen, P. (2019). *Quality. From Plato to Performance*. Palgrave.
- Dewey, J. (2017). *Offentligheden og dens problemer [The Public and its Problems]*. Hans Reitzel.
- Espeland, W. og Sauder M. (2007). Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology* 113(1), 1-40.
- Esposito, E., & Stark, D. (2019). What's Observed in a Rating? Rankings as Orientation in the Face of Uncertainty. *Theory, Culture & Society*, 36(4), 3–26
- Fochler, M. og de Rijcke, S. (2017). Implicated in the indicator game? An experimental debate. *Engaging Science, Technology, and Society* 3, 21-40.
- Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, 24(1), 23-37
- Hertting, N. & Vedung (2009). *Den utvärderingstäta politiken*. Studentlitteratur.
- Rosa, H. (2005). The Speed of Global Flows and the Pace of Democratic Politics. *New Political Science* 27(4), 445-459.
- Rosanvallon, P. (2009). *Demokratin som Problem [Democracy as the Problem]*. TankeKraft Förlag.
- Rosanvallon, P. (2011). The Metamorphoses of Democratic Legitimacy: Impartiality, Reflexivity, Proximity. *Constellations* 18(2), 114-123.
- Vedung, E. (2009). *Utvärdering i politik och förvaltning*. Studentlitteratur.

KAPITEL 7

Utvärdering och annan kunskap

Jan-Eric Furubo

Våren 1979 springer Evert och jag ihop på Drottninggatan i Stockholm. Båda är vi ju uppsaliensare, och känner igen varandra på det diffusa sätt som gör det motiverat att hälsa, men också att klargöra frågor kring när och var. När detta är gjort frågar Evert, som den mer seniora av oss, om vad jag sysslar med. Jag svarar att jag arbetar på Industridepartementets energisparkkommitté och mera specifikt med en, av regeringen inte helt kristallklart formulerad uppgift, utvärdering av energisparkkampanjer. Evert utbrister då något i stil med att "det är ju ungefär det jag också gör, så vi måste träffas och prata". Efter att Evert något mera beskrivit det uppdrag han fått av den då existerande energiforskningsnämnden bestämmer vi oss för att träffas. Enligt min kalender gjorde vi också det på förmiddagen den 17 maj i energisparkkommitténs lokaler på Kungsgatan 56. Detta första samtal ledde fram till en rad andra och efter att ha lämnat institutionen för dagen stannade Evert inte sällan till för diskussioner i min lägenhet på Sankt Larsgatan några minuter från Skytteanum. Evert berättar senare att "We sat there discussing problems of Evaluation while enjoying a few cups of coffee. We formed an informal discussion circle on evaluation" (Tranquist, 2015, s. 573, se även Tranquist denna volym).

När jag ser tillbaka på dessa allra första kontakter med Evert våren 1979 är det tydligt att många av de drag hos vännen Evert som jag kommit att uppskatta, var tydliga redan vid dessa första möten. Evert är en generös person, glad att få dela med sig av kunskaper och erfarenheter. Evert är en nyfiken person, han ville få reda på hur saker fungerar – inte bara i stort utan ofta också in i minsta detalj. Nyfikenheten leder ofrånkomligen till samtal och ligger därmed nära hans vilja och förmåga till kontaktbyggande. Evert är också en optimistisk person. Detta har även präglat hans syn på utvärdering. Det gäller utvärderingens förmåga att med samhällsvetenskapliga metoder besvara frågor om genomförandet och effekterna av olika statliga interventioner, liksom hur denna kunskap kan användas för att förbättra dessa och utforma nya. Denna tilltro till utvärdering har också varit avgörande för den roll Evert fått i svensk utvärdering, det ligger nära till hands att tala om ett entreprenörskap. Till en del har detta säkert bedrivits inom akademien, men om detta kan andra vittna bättre, klart är dock att det också i hög grad varit riktad mot förvaltningen. Ett tidigt uttryck för detta var att Evert, tillsammans med bland andra Göran Hermerén och Tore Nilstun 1985 arrangerade en något mer avancerad tvåveckorskurs i utvärdering, finansierad av Nordiska forskarkurser. Tack vare Evert kom genom denna kurs också de första kontakterna att byggas mellan flera av de ledande amerikanska utvärderingsforskarna och nordisk utvärdering.

När det nu gått ett antal decennier efter Everts och mina första diskussioner om utvärdering som väl för oss båda var präglade av optimism och tilltro till utvärderingens värde ter det sig inte orimligt att ställa mera utvärderande frågor om utvärdering själv. En sådan fråga gäller i vilka situationer utvärdering är särskild användbar och därmed också i vilka situationer den är mindre användbar.

Diskussionen av denna fråga kräver dock en precisering av den företeelse som frågan gäller; vad är egentligen utvärdering.

EVALUATION FORMAS I USA

När Evert och jag i slutet av 1970-talet dryftade frågor om utvärdering så var det förvisso inte en okänd företeelse i Sverige.

Svenska forskare, inte minst pedagoger hade vid vistelser i USA kommit i kontakt med och skrivit om *evaluation*. Även inom offentlig förvaltning talades i olika sammanhang om evaluation eller utvärdering som denna nya verksamhet kom att benämnas i Sverige.

Att beskriva evaluation som en ny verksamhet är inte självklart. Men det är, om än en viss försiktighet bör iaktas vid mera precisa tidsangivelser, i 1960-talets USA evaluation formas till en särskild praktik som senare kommer att bli ett centralt inslag i många länders politik och förvaltning.¹ När Derlien beskriver framväxten och spridningen av evaluation och konstaterar att länder som Sverige, Canada och Förbundsrepubliken Tyskland följde "the U.S. efforts relatively early" så är det just denna praktik han beskriver spridningen av (Derlien, 1990, s. 148).

Det som sker i USA under 1960-talet föregås, förstås, av olika skeenden. Beskrivningen av dessa ligger utanför denna framställning. Summariskt är det dock möjligt att peka på tre faktorer som under 1960-talet i USA konvergerar på ett sätt som gör att de kan ses som förutsättningar för en takeoff för evaluation. Den första är expansionen av *social sciences* under den första hälften av 1900-talet. Nya discipliner, och omorientering av gamla discipliner resulterade i metoder och teorier som kunde användas för att bygga kunskap om politiskt relevanta fenomen (Prewitt et al, 2012). Den andra förutsättningen har att göra med forskarnas ökade intresse att hjälpa till att finna lösningar på samhällets problem (Prewitt et al. 2012). Den tredje förutsättningen är framväxten av en mera teknisk eller vetenskaplig syn, forskning borde i ökade grad användas vid konstruktionen av olika program. John F. Kennedy framhåller vid ett tillfälle att landets problem kunde ses som tekniska snarare än politiska frågor (Fischer, 2003, s. 6).

De amerikanska forskare som under 1960-talet och 1970-talet börja skriva om evaluation, till exempel Suchman och Weiss, ser också evaluation som en ny praktik som ska förbättra program och policier genom att säkra ett inflöde av kunskap om hur dessa fungerar. Även om det förekom hänvisningar bakåt som när

1) En något mera omfattande beskrivning av framväxten av evaluation i USA och den därpå följande importen till Sverige finns bland annat i Furubo 2021.

Suchman skriver om ett nyväckt intresse för "long-neglected aspect of social research – the evaluation study" så är det uppenbart att evaluation är en praktik som just håller på att utvecklas (Suchman, 1967, s.1f). House skriver också många år senare om "the time that the contemporary practice of evaluation developed" (House, 1993, s.138).

I de länder som därefter importerar evaluation beskrivs det också som en, låt oss säga, amerikansk nymodighet. Det är notabelt att man i de länder som under 1970-talet introducerar evaluation inte knyter an till tidigare traditioner. Det fanns ju förvisso i ett land som Sverige också långt tidigare ett intresse för att vetenskapligt studera statliga insatser, till exempel effekterna av regleringar, något som är väl dokumenterat genom bland annat det offentliga utredningsväsendet och i inlägg av enskilda forskare. Att detta inte var unikt för Sverige framgår av Nicoletta Stames konstaterande att det som skett tidigare i olika länder kom att överskuggas av nymodigheten från USA. I den europeiska utvärderingsdiskussionen var "*The relationship with U.S evaluation...everywhere a point of reference*" (Stame, 2013, s. 367).

SYFTAR TILL ATT FÖRBÄTTRA

I sökandet av en definition av evaluation refereras många gånger till Scrivens definition som också finns i åtskilliga läroböcker. Scriven definierar evaluation som "the process of determining, merit, worth or value of something, or the product of that process" (Scriven, 1991, s. 93). I motsats till Scriven har dock de flesta utvärderingsforskare, inklusive Vedung, valt att definiera evaluation som en form av praktik syftande till att förbättra offentliga verksamhet. En representant för dessa utvärderingsforskare är Weiss som framhåller att "The basic rationale for evaluation is that it provides information for action" (1972a, s. 318.) och att det handlar om att förbättra ett program eller en policy (1998, s.4). Vedung är redan i sin första mer omfattande skrift om utvärdering *Energipolitiska utvärderingar 1973-1981* tydlig med att han med utvärdering menar något som är en del av beslutsprocesserna inom politik och förvaltning (1982, 7ff). Detta är

också något som han slår fast i de olika upplagorna av *Utvärdering i politik och förvaltning*.

Det finns ett värderande element i utvärdering, men det innebär förstås inte att varje värderande påstående om till exempel en förd politik är en utvärdering. Allmänna påståenden om att politiken på ett område varit framgångsrik eller för den delen misslyckad uppfyller varken Weiss krav på *systematisk assessment* (1998, s. 4) eller det noggrannhetskrav Vedung betonar (2009, s.22). En evaluation måste också vara förklarande, det gäller att klargöra vad i en utveckling som orsakats av det som är utvärderingens objekt. En jämförelse av utvecklingen på ett område i förhållande till olika politiska mål kan därmed inte ses som en utvärdering (även om gränsdragningarna ibland kan vara svåra).

Det senare har betydelse för den fråga som ska diskuteras i det följande. När evaluation formas som en praktik i 1960-talets USA är det tydligt att evaluation inte är detsamma som forskning som allmänt sett kan komma till nytta vid utformningen eller förbättrandet av ett program eller en policy. Objektet i en evaluation är något som görs, en "intentional action designed to influence people or change a material situation" som von Riecken formulerar det redan under 1950-talet (von Riecken, 1972, s. 86). Kortare uttryckt handlar det om en "purposive social action", för att låna en formulering av Merton (1936, s.896). I den fortsatta diskussionen kommer jag oftast att använda det ord som Vedung själv ofta använder, *intervention* (t.ex. 2009, s. 201f).

Avgränsningen innebär att utvärdering, och den kunskap utvärdering ger, har avgränsats från mycket annan kunskap som är viktig i offentligt beslutsfattande. Och detta ger upphov till frågor om relationen mellan utvärdering och annan kunskap. Det ska understrykas att det fortsatta resonemanget och slutsatserna just handlar om utvärdering som avses att ge kunskap som kan användas för att förbättra en intervention. En annan diskussion skulle kunna föras kring utvärdering som mer uteslutande är inriktad på att utröna hur personer eller organisationer levt upp till sitt ansvar i olika situationer. Men den diskussionen förs inte här.

UTVÄRDERING MÅSTE SAMPELA MED ANNAN KUNSKAP

Oavsett hur beslutsfattarna kommit fram till att det finns ett problem eller en framtida utveckling som man vill påverka genom ett beslut om en viss intervention förutsätts någon form av antaganden om vad som förbinder denna intervention med den önskade utvecklingen. Det handlar med andra ord, för att knyta an till Everts och mina diskussioner på 1970-talet, om vad som till exempel ökar hushållens benägenhet att genomföra energihushållningsåtgärder. Är det kunskap? Och i så fall är det kunskaper om hur energikonsumtion påverkar miljön eller är det kunskap om den privatekonomiska nyttan av att genomföra vissa åtgärder?

Tanken att offentliga interventioner på något sätt förutsätter antaganden om kausala mekanismer gör att Pawson and Tilley har beskrivit program "som theories incarnate" som börjar "in the heads of policy architects", som därefter läggs i verkställarnas händer för att ibland föras över till "the hearts and minds of programme subjects." (Pawson & Tilley, 2004, s. 3).

Att det är svårt att tänka sig att det fattas beslut om olika interventioner utan att det finns någon form av antaganden om hur interventionen kan skapa förlopp som leder till det som anses önskvärt, betyder förstås inte att dessa antaganden behöver vara riktiga, grundade i forskning eller ens explicit formulerade. Det betyder dock att beslutsfattare när de fattar beslut kan vara i stort behov av annan kunskap än den som utvärdering ger.

Ett exempel får illustrera detta. En forskare kan studera hur sociala, biologiska och psykologiska mekanismer påverkar hur barn lär sig läsa. Oavsett om forskaren från början motiverats av att resultaten av forskningen kan vara användbara vid utformningen av olika politiska insatser, kan vi förstås lätt tänka oss att resultaten kan användas i konstruktionen av åtgärder, som till exempel hur lärarna ska utbildas eller förändringar i läroplaner. Att en studie av något fenomen och olika påverkande mekanismer, till exempel kring hur lärarnas förväntningar påverkar skolresultat, frekvensen av energihushållningsåtgärder eller utveckling av handel mellan länder, kan användas vid utformningen av politiska

åtgärder betyder inte att dessa studier kan kallas för utvärderingar. Om vi skulle kalla alla studier som ger kunskap som kan vara av nytta vid till exempel beslut och genomförande av statliga verksamheter för utvärderingar – ja, då blir utvärdering liktydigt med näst intill all kunskapsproduktion.

Om en studie ska kallas en utvärdering beror alltså på vad som är studiens objekt. Många studier som ger kunskap som kan användas vid utformningen av politiska interventioner handlar om biologiska, psykologiska, ekonomiska och andra mekanismer som gör att individer eller organisationer handlar på ett visst sätt. En utvärderings objekt är däremot inte dessa mekanismer per se, utan någonting – en intervention – som genomförs i syfte att sätta dessa mekanismer i spel. Det bör dock påpekas att utvärdering ibland kan bidra till att bygga kunskap om dessa mekanismer, något som Weiss tidigt pekade på då hon framhåller att evaluation under vissa omständigheter kan bidra till att utveckla det som hon beskriver som *basic knowledge*. När utvärderaren undersöker hur framgångsrikt ett program varit och förklarar detta kan det leda till ”discoveries about basic concepts of human behaviour and social structures” (Weiss, 1972b, s. viii).

NYTTAN AV UTVÄRDERING KAN VARA STÖRRE ELLER MINDRE

Det förhållandet att utvärdering handlar om interventioner som existerar och att utvärderingen gäller något som pågår eller avslutats gör att frågan om relationen mellan evaluation och annan kunskap blir särskilt betydelsefull i situationer som kräver mera kraftfulla förändringar.² Det handlar om situationer som innebär att beslutsfattarna tvingas lämna den välkända terräng som utgörs av tidigare fattade beslut. Det kan vara ändrade synsätt om vad som är problem, om betydelsefulla förändringar i samhället, en

2) Vedung har alltid sett utvärdering som något som sker i efterhand, ex post. Det som utvärderas är något som pågår eller har genomförts. Det är också, vågar jag påstå, den vanligaste uppfattningen inom området. Det ska dock noteras att det finns ett skiftande språkbruk och också något olika meningar gällande detta. Det resonemang som förs här, och som utgår från att utvärderings objekt är en existerande eller avslutad intervention, torde dock vara giltigt även om den specifika interventionen endast existerar på ritbordet.

ackumulation av inkrementella förändringar som nått vägs ände eller att beslutsfattarna tvingas hantera omskakande kriser.

En konsekvens av att utvärderingens objekt är en beslutad intervention var att utvärdering både byggde på och präglades av tanken att de verksamheter som utvärderades förändrades stegvis, inkrementellt. De amerikanska utvärderingsteoretikernas texter är präglade av en stark övertygelse om att den arsenal av program som satts in gradvis skulle leda till ett bättre samhälle. De betyder naturligtvis inte att de inte tänkte sig att utvecklingen kunde brytas, för det fanns förvisso hot. Men om dessa kunde avvärjas skulle den gradvisa utvecklingen fortsätta.

Utvecklingen av program, eller mer allmänt interventioner, sågs därför som en sekventiell process, inte sällan illustrerad med en cirkel med förutsebara moment eller stationer där en station var utvärdering, något som också Vedung framhåller i sin bok 1982 (s. 8). I denna process skulle utvärdering ge underlag för att undan för undan förbättra olika verksamheter. Eftersom resonemanget handlar om justeringar av existerande interventioner, som syftade till att hantera definierade problem, var det naturligt att se den kunskap som utvärderingar gav som betydelsefull. Utvärderingarna handlade ju just om dessa interventioner.

För att leverera information till denna ständigt pågående process av justeringar och förbättringar byggdes också olika utvärderingssystem upp. Inte sällan handlade det då om att ta fram information som byggde på samma antaganden som den program- eller interventionsteorin, som låg till grund för själva interventionen (Leeuw & Furubo, 2008).

Tanken att beslutsfattande i politik och förvaltning handlar om att successivt förbättra det som redan görs är högst rimlig. Mycket pekar på att offentlig verksamhet sällan omprövas från grunden. En anledning till detta har tagits upp av många författare: Att förändra är kostsamt. Det kan ofta kosta mer att förändra en policy än att fortsätta på inslagna vägar. En förändring skapar osäkerhet, förändrar relationer och konsekvenserna är ofta oklara. Pierson pekar på att "the *relative* benefits of the current activity compared with once-possible options increase over time" (Pierson 2004, s. 21).

Samtidigt når inkrementalistiska skeenden ibland sin slutpunkt. Det blir då uppenbart att det inte är möjligt att försätta med marginella förändringar och mindre avvikelser från den tidigare inslagna kursen. Baumgartner och Jones tar upp detta i sin diskussion av *punctuated equilibrium*. De pekar på att ackumulation av inkrementella förändringar kan vara viktig, men att de stabila utvecklingsförloppen ibland bryts på ett sätt som innebär krav på en dramatisk nyorientering (Baumgartner & Jones, 1993). Tanken att stabila historiska faser avbryts av perioder i vilka människorna söker eller tvingas nya framgångsvägar är dock inte en produkt av de senaste decenniernas samhällsforskning. Tanken uttrycks av Abraham Lincoln, en politiker vars gärning mer än många andras kan placeras i ett icke-inkrementellt skeende. I sitt budskap till kongressen 1862 skriver Lincoln: "The dogmas of the quiet past, are inadequate to the stormy present ... As our case is new, so we must think anew, and act anew" (Lincoln, 1862).

När beslutsfattare funderar på åtgärder för att förbättra inom befintliga policykonstruktioner och institutioner behöver de kunskap om en rad frågor som gäller just de befintliga interventionerna, alltså sådan kunskap som utvärdering kan ge. I diskussionen om nya prioriteter, nya aktioner, ter det sig dock troligt att sådan kunskap, vars objekt är befintliga interventioner och existerande organisatoriska strukturer, är mindre relevant. Det som behövs är i stället kunskap som kan vara till hjälp för att välja mellan olika färdriktningar och strategier för att hantera nya problem och/eller värdeförskjutningar.

EN FÖRÄNDRAD SYN PÅ ANVÄNDNING AV UTVÄRDERING?

Det är förstas inte möjligt att veta om svensk politik i framtiden kommer att vara mindre inkrementell än den varit tidigare. De argument som talar för detta, och som kan te sig närmast triviala, har ofta att göra med tid och avstånd. Under senare decennier har vi sett att vad som händer i ett land, på till exempel på olika marknader men också vad gäller spridning av farsoter, mer eller mindre omedelbart påverkar länder i andra delar av världen, stora förändringar utlöses nu inte bara av skeenden i det egna landet. Till detta

kan läggas att spridningen av nya idéer, inklusive sådana som ifrågasätter grundläggande samhälleliga strukturer, nu sker snabbare än för några decennier sedan.

Och det är inte heller svårt att föreställa sig att beslutsfattare i växande grad kan tvingas konfrontera kriser, som gör att de tvingas hantera extrema osäkerheter under extrem tidspress. I sådana lägen måste diskussioner och analyser som rör värdemässiga utgångspunkter, påverkande faktorer, alternativa lösningar, bieffekter och så vidare som "normalt" kan ägnas åt pressas in under veckor eller kanske till och med dagar.

Om politikerna i framtiden kommer att ställas inför nya problem och därmed tvingas agera i mindre känd terräng påverkar det utvärderingens roll och användning. Rimligen borde i sådana lägen kunskap om vilka mekanismer med vars hjälp ett skeende kan påverkas efterfrågas, och oavsett om denna kunskap har producerats genom utvärderingar av tidigare gjorda interventioner eller på något helt annat sätt. I stället för frågan "vilka utvärderingar finns som kan vara användbara?" kommer politikerna kanske att ställa frågan "vad säger forskningen om vad vi kan göra nu, vilken kunskap är relevant och var finns den, vilka är de ledande experterna och hur snabbt kan de komma hit?"

Detta har direkt bäring på den kanske mest diskuterade frågan inom utvärderingsområdet, användning av utvärdering. Att frågan är central har förstås att göra med att utvärderingens grundidé är att den ska användas för att förbättra interventioner. Redan tidigt framhölls att denna process inte behövde vara direkt eller instrumentell, tvärtom framhölls processens komplexitet och många både avsedda och icke avsedda former av användning, något som inte minst Vedung framhållit (2009, s. 255ff). I grunden, menar jag, har dock denna diskussion handlat om hur de utvärderingar som görs ska nå beslutsfattare. Det är dessa, politiker eller ämbetsmän, som har setts som målgruppen.

Det resonemang som här förts innebär att beslutsfattare när de tvingas att "think anew, and act anew" inte i först hand frågar, eller bör fråga, efter utvärderingar av vad som gjorts tidigare. Beslutsfattarna frågar i stället, eller bör fråga, efter kunskap som gör det

möjligt att bedöma antaganden om de kausala mekanismer som är avgörande för synen på värdet av olika handlingsalternativ. Precis som Weiss påpekade kan utvärdering bidra till att bygga denna kunskap. Detta leder till en förskjutning i synen på vad som är målgruppen för utvärdering; det blir viktigare att kanalisera den kunskap som utvärderingar ger till dem som beslutfattare riktar sina frågor än till beslutsfattarna själva. Genom att denna målgrupp i högre grad är verksamma inom discipliner och kunskapsfält kommer resultaten från utvärderingar också i allt högre grad prövas i förhållande till tidigare teori och empiri. Utvärdering blir därmed också en del i byggandet av den kunskap som är nödvändig i kriser eller nyorienteringar av politiken inom olika samhällsområden.

Referenser

- Baumgartner, F.R. & Jones, B.D. (1993). *Agendas and Instability in American Politics*. Chicago, IL: The University of Chicago Press.
- Derlien, H-U. (1990). Genesis and Structure of Evaluation Efforts in Comparative Perspective. I Rist, R.C. (Red.) *Program Evaluation and the Management of Government*. New Brunswick, NJ: Transaction Publishers.
- Fischer, F. (2003). *Reframing Public Policy*. Oxford: Oxford University Press.
- Furubo, J.E. (2021). *Utvärdering och kontroll. Tilltro och ansvarsutkrävande i demokratin*. Lund: Studentlitteratur.
- House, E. (1993). *Professional Evaluation*. Newbury Park, CA: Sage.
- Leeuw, F-L. & Furubo, J.E. (2008). Evaluation systems. What They are and Why Study them. *Evaluation*, Vol.14, No.2. 157-169.
- Lincoln, A. (1862). *Annual Message to Congress*.
<http://www.abrahamlincolnonline.org/lincoln/speeches/congress.htm> (hämtad 2020-05).
- Merton, R.K. (1936). The Unanticipated Consequences of Purposive Social Action. *American Sociological Review* 1(6).
- Pawson, R. & Tilley, N. (2004). *Realist Evaluation*.
www.urban-response.org/system/files/content/resource/files/main/pawson---tilley-%282004%29-realist-evaluation.pdf (hämtad januari 2024)
- Pierson, P. (2004). *Politics in Time. History, Institutions and Social Analysts*. Princeton, NJ: Princeton University Press.
- Prewitt, K., Schwandt, T.A. & Straf, M.L. (2012). *Using Science as Evidence in Public Policy*. National Research Council. Washington, DC: The National Academy Press.
- Scriven, M. (1991). *Evaluation Thesaurus*. Newbury Park, CA: Sage
- Stame, N. (2013). An European Evaluation Theory Tree. I Alkin, M.C. (Red.), *Evaluation Roots. A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks, CA: Sage.
- Suchman, E.A. (1967). *Evaluative Research. Principles and Practice in Public Service and Social Action Programs*. New York: Russell Sage Foundation.

- Tranquist, J. (2015). The Oral History of Evaluation: The Professional Development of Evert Vedung. *American Journal of Evaluation*, 36(4), 570–583.
- Vedung, E. (1982). *Energi politiska utvärderingar 1973-81*. DFE rapport NR 52. Delegationen för Energiforskning.
- Vedung, E. (2009). *Utvärdering i politik och förvaltning*. 3 uppl. Lund: Studentlitteratur.
- von Riecken, H.W. (1972). Memorandum on Program Evaluation. I Weiss, C.H. (Red.) *Evaluating Actions Program. Readings in Social Science and Education*. Boston, MA: Allyn & Bacon.
- Weiss, C.H. (1972a). Utilization of Evaluation: Towards Comparative Study. I Weiss, C.H. (Red.) *Evaluating Actions Program. Readings in Social Science and Education*. Boston, MA: Allyn & Bacon.
- Weiss (1972b). *Evaluation Research. Methods of Assessing Program Effectiveness*. Englewood Cliffs, NJ: Prentice Hall, Inc.
- Weiss, C.H. (1998). *Evaluation. Methods for studying Programs and Policies*. Second edition. Upper Saddle River, NJ: Prentice Hall Inc.

KAPITEL 8

An evaluation manager perspective – inspiration from Evert

Per Øyvind Bastøe

I will present some reflections based on more than 30 years of experience as manager and performer of evaluations in public administration in Norway and internationally. It is interesting to note that Evert Vedung has been a source of inspiration throughout this period.

As a manager, I have been particularly concerned that evaluations are useful for improving the business they relate to. There are many, perhaps too many, instruments available for governance and learning in public administration. Thus, it is not necessarily the case that many evaluations are better than few evaluations. Any evaluation should have a clear purpose and be carried out in a way that is beneficial. There are two aspects I would like to highlight here where the inspiration from Evert Vedung has been particularly strong. The first concerns the understanding of what evaluation is, and what it is not. The second concerns what the use of evaluations means.

When I received my first evaluation assignment from the Research Council of Norway at the end of the 1980s, the use of the term evaluation was almost completely absent in Norway. As documented in a study of the development of evaluation activities

in 21 countries in 2002 (Furubo, Rist & Sandahl, 2002), Norway was assessed as not being at the forefront of the development of evaluation. The study was based on nine different criteria related to the scope of evaluation activities, the availability of evaluation competence, public discussions, the prevalence of evaluation associations and various aspects of the institutionalization of evaluation activities in the public sector. The United States, Canada, Sweden and Germany were considered to constitute the first phase in the development of evaluation activities. Although progress has been made since then, there is still no general concern with and uniform understanding of evaluation in the Norwegian public administration.

UNDERSTANDING WHAT EVALUATION IS AND IS NOT CAN AFFECT QUALITY

Therefore, in the late 80s, in my infancy as an evaluator, I had to rely on international literature to understand what evaluation entailed. At that time, there was a strong discussion among the leading academics in the United States (Michael Q. Patton, Carol Weiss, Michael Scriven, and others) about the purpose and nature of evaluation, especially whether the evaluator should be understood as an external and objective valuer, or play a participatory, process-oriented role. I obviously had most sympathy for the latter, not least from my sociology education where I was taught that as a researcher one can never be isolated from one's object of study, never neutral and never objective. Your own values and perceptions will influence what you observe. This position was partly influenced by the Norwegian philosopher Hans Skjervheim, who in the 50s wrote the essay "Participant or observer" (in Norwegian: "Deltakar og tilskodar") (Skjervheim, 1957), in which he argued that sociologists who objectify often fail to take seriously the content of what is said. He believed that when someone stands outside, when someone deprives others of subject status and obscures their own commitment, one should react.

No wonder, then, that I felt at home in what Evert described about the understanding of what evaluation is (Vedung, 1991),

and which he elaborated on later. His description (Vedung, 1997) of evaluation as “the careful assessment *ex post* of the merit of the content, administration, output, outcome effects, and organization of public sector interventions, which is intended to be useful in practical action situations,” clarified my understanding of the term.

In particular, the emphasis on evaluation taking place *ex post* I found to be a helpful delimitation. It is notably different from other commonly used definitions of evaluation. For example, in its guide, the Norwegian Ministry of Finance defined evaluation as systematic data collection, analysis and assessment of not only something that has happened or is ongoing, but also of something that is planned (Ministry of Finance, 2005), i.e. also including what others refer to as “*ex-ante*” evaluations. An *ex-ante* evaluation is done before an intervention or initiative has started. That is, it evaluates the program design and the program for implementation. An *ex-ante* evaluation has many similarities to a risk assessment or impact analysis.

A broad definition of evaluation that also includes *ex ante* evaluations is, in my view, more disruptive than clarifying. Or as Evert (Vedung, 2009) puts it: “Careful assessment of planned public interventions is not evaluation” (In Swedish: “Grundlig bedömning av påtänkta offentliga åtgärder är däremot inte utvärdering.”). Worth noting is that Evert emphasizes that his definition of evaluation does not mean that evaluation is only about completed interventions. Evaluation is about terminated but also ongoing interventions, but not about interventions contemplated and planned only. Evaluation is not about interventions *ex ante*, i.e. before they are adopted; it is about adopted interventions and terminated interventions. *Ex post* in the case means *after* inauguration, *after* decision to adopt, *not after termination*.

The Norwegian professor Kristin Tornes (Tornes, 2012) touches on something similar when she describes some characteristics of evaluation that make this tool different from other approaches to systematizing knowledge and experience. Among other things, she points out that evaluation involves describing and assessing. Furthermore, that evaluation is action-relevant and intended to

be useful. She also points out that evaluations are historical in the sense that they concern something that has been completed or put into practice.

An understanding of evaluation in line with what emerges in the Ministry of Finance's guide, may blur the difference between evaluation and other forms of analysis and risk assessments. Evaluation is not "informed speculation" about the future, but rather careful assessment of real experiences –past or ongoing. An understanding in line with what Evert expressed, by reserving the concept of evaluation to apply to the assessment of either something that is ongoing, or something that has been completed may contribute to understand that evaluation has distinct features different other management tools.

The implication of an unclear understanding is well recognized. For example, a study of the use of evaluation in the Norwegian public administration in 2016 (DFØ 2016) showed great variation in understanding and approach to evaluations. At the same time the study found weak linkages between evaluations, management and learning. An OECD study of policy evaluations in 2019 (OECD, 2019) found something similar in most of the 42 countries included, despite a formal commitment to conduct evaluations. It showed that most countries faced major challenges in using evaluations as a basis for political decisions due to absent strategies for conducting evaluations and inadequate human resources in terms of competence and capacity.

RECOGNIZING MORE THAN ONE EVALUATION MODEL – AVOIDING A STRAITJACKET APPROACH

The close link to performance management in the Norwegian government guidance material may furthermore contribute to evaluation being perceived primarily as a tool for analyzing the extent to which the formulated objectives of an intervention have been achieved. If evaluation is understood in this way, it may limit the use of evaluations and the understanding of how evaluations are conducted. Evaluation is not constrained to only assessing causes and effect of an intervention. Evaluation can also take place

as process evaluation or follow-up evaluation; That is, an evaluation while something is being implemented, for example by evaluating the program or change theory, or by evaluating various aspects of the actual implementation of the intervention.

Consequently, there are different ways to conduct an evaluation, many different models. Evert has a useful breakdown of models into five different families: goal-result models, goal-process models, relevance models, actor models and economic models (2009). This categorization is a good contribution to understanding the similarities and differences between the different ways of conducting an evaluation.

Decisive for the choice of model is, of course, the context within which the evaluation is to take place, what is to be evaluated, what is the purpose of the evaluation, what questions the evaluation should answer, what the results of the evaluation should be used for, and what resources and how much time is available to carry out the evaluation.

My experience is that many evaluations are limited either by only analyzing achievement of intended objectives, or by following a straitjacket approach not tailored to the purpose, the context, and the intended use of a specific evaluation (for instance by always including all the international evaluation criteria, see OECD, 2021). Such evaluations most often end up less useful.

UNDERSTANDING THAT EVALUATIONS ARE MEANT TO BE USEFUL

The whole point of an evaluation is that the results should be used by someone other than those conducting the evaluation, for learning, improvement, management or as a basis for decision-making. One aspect of this, which has been frequently discussed over the years, is regarding the extent of the involvement and participation of those affected is crucial for use of evaluations. Evert is one of those who has actively participated in this debate (Vedung, 2017). Like Patton, he argues that the purpose of evaluation is to be useful for practical action, ref. the definition “intended to be useful in practical action situations.” He adds that evaluation is meant to

create a basis for continued practical considerations, i.e. it is something that might be used in public governance.

Another aspect is how use is perceived. It is quite common to assess use only on the basis of whether the recommendations are implemented. Evert has helped expand the understanding. His distinction between the use of the product and the use of the process I have found to be particularly helpful (Vedung, 2017). Product use concerns the use of the final evaluation report, both in writing and in the dissemination and communication of the report's findings and recommendations. Process use is about use in terms of the activities that the evaluation involves, from announcement to planning and execution of the evaluation. His six categories of use (instrumental use, conceptual use, legitimizing use, tactical use, ritual use, and constitutive use) contribute to understand that use is much more than the simple notion of implementing recommendations.

Instrumental use entails a direct use of findings and recommendations as a basis for decisions. This means that decisions are influenced by the evaluation. They would have been different without the evaluation. It's worth noting that Evert also links instrumental use to findings that emerge through the evaluation process. In contrast to direct instrumental use, the conceptual use means that the evaluation is interpreted and translated in the context in which the users find themselves. In conceptual use, the evaluation and evaluation process affect the thinking of the potential users. This may mean that users refer to multiple evaluations and also that they combine the evaluations with other knowledge and insights. Completely different is the legitimizing use, which means that the evaluation is used to legitimize a decision or action that users have already decided. The evaluation is used as a justification for what one should do. This implies that the evaluation does not lead to a change of perception or decision but could mean that one becomes more certain of what one already intends to do. Use may also depend on tactics and rituals. For instance, by delaying a decision until the evaluation is available.

CONCLUDING REMARKS

Evaluations are often controversial. They can reveal something that doesn't work and may not be pleasant to bring to light. In my career as evaluation manager, I have often been met with various forms of resistance, at least in the short term. Interestingly, however, evaluations have often ended up being used despite initial resistance. Not necessarily in line with the recommendations. But rather by creating new insights, placing something on the agenda, or contributing to improvements in the long term.

The insights that Evert Vedung have presented and shared through a long life as evaluation researcher, academic and adviser has been an inspiration far beyond Sweden. It has certainly helped to give me a sense of assurance as evaluator and evaluation manager. A confidence that is based on understanding the strengths and weaknesses of evaluation. Moreover, to understand that use is far more nuanced than one immediately thinks.

References

- Bastøe, PØ (2020). *Fem forutsetninger for gode evalueringer*. (Five prerequisites for good evaluations.) Oslo. Gyldendal Norsk Forlag.
- Ministry of Finance (2005). *Veileder til gjennomføring av evalueringer*. (Guide to conducting evaluations.)
- Furubo, J.-E., R.C. Rist & R. Sandahl (2002). *International Atlas of Evaluation*. Transaction Publishers.
- OECD (2019). *Institutionalisation, Quality and use of Policy Evaluation. Governance lessons from countries experience*. GOV/PGC(2019)51/ANN
- OECD (2021). *Applying Evaluation Criteria Thoughtfully*. OECD Publishing, Paris, <https://doi.org/10.1787/543e84ed-en>
- Skjervheim, H. (1957). *Deltakar og tilskodar*. (Participant or observer.) Essay. Stencil series. Department of Sociology, University of Oslo.
- The Norwegian Agency for Public and Financial Management (DFØ) (2016). *Bruk av evaluering i statlig styring*. (Use of evaluation in state governance.)
- Tornes, K. (2012). *Evaluering i teori og praksis*. (Evaluation in theory and practice.) Trondheim: Akademika forlag.
- Vedung, E. (1991). *Utvärdering i politik och förvaltning*. (Evaluation in policy and management.) Lund: Studentlitteratur.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. Transaction Publisher. New Brunswick. New Jersey.
- Vedung, E. (2017). *Six-times-two Uses of Evaluation*. Uppsala: Evert Vedung's Personal Archive.

KAPITEL 9

Theory-Based Stakeholder Evaluation Revisited: An Exploratory Review

Morten Balle Hansen

INTRODUCTION

This chapter provides an exploratory review of the literature related to the theory-based stakeholder evaluation model (TSE-model). In 2010 Hansen and Vedung introduced a novel approach to program theory evaluation that they called theory-based stakeholder evaluation or the TSE model for short (Hansen & Vedung, 2010). The article was based on earlier empirical research in Danish municipal eldercare (Hansen & Vedung, 2005).

The TSE model is embedded in a much broader trend within the evaluation research community to get into “the black box” and clarify or reconstruct the intervention theory (also called program, implementation or change theory) behind a specific intervention (a program, a project, a reform). What kind of problems is intended to be solved in the initial situation? How is the intervention supposed to work to create change? Why is the situation aimed for preferable to a situation without the intervention?

This clarification of the theory behind the intervention can be useful for evaluators when they empirically examine how the intervention works but can also be useful for other analytical and

practical purposes. For instance, leaders may use the invention theory as a unifying narrative concerning the meaning of the intervention.

Important contributions within the theory-based tradition of evaluation research have been made by seminal evaluation scholars such as Huey-Tsyh Chen (1990), Carol Weiss (1997, 2007), Nicoletta Stame (2004) and more recently John Mayne (2015). In this review however, the focus is on the TSE-citing literature.

The basic notion of the TSE approach is to introduce a political dimension to theory-based evaluation designs. Many theory-based evaluation designs tend to assume that a relatively unproblematic agreement about how and why a specific program/intervention should work and produce specific outcomes could be accomplished. Hansen and Vedung argue that in some types of evaluation, most likely in large multisite, multilevel, controversial interventions, this assumption is unfortunate since there may be tensions and conflicts concerning the program/intervention and the outcomes it is supposed to enhance. When trying to construct one unitary intervention theory, an unfortunate “designed blindness” to tension and conflict may be the outcome. It is for such contexts that the TSE-model has been elaborated.

For a context characterized by more or less open tensions and conflict between different intervention theories the TSE model combines four basic recommendations:

- Tripartite intervention theory scheme: That it is useful to distinguish between three elements of an intervention theory: (a) a situation theory concerning the context in which the intervention takes place. (b) a causal theory concerning how an intervention is supposed to make an impact. (c) a normative theory concerning why the intervention is preferable to other or no interventions (Hansen & Vedung, 2010, pp. 300–301)
- Principle of charity (or reason): That pertinent actors are presumed to have reasonable intervention theories and the evaluator should try to elicit these from the available evidence (Hansen & Vedung, 2010, pp. 300–302).

- Extended system model with pertinent collective actors: That for comparison intervention theories can usefully be packed into an input-output system scheme – that is an input-conversion-output-outcome model – and extended by pertinent collective actors (Hansen & Vedung, 2010, pp. 300 and 304).
- Five-Step Model of TSE: That a TSE model could be usefully conducted in five steps: (a) Reconstruction of the embedded intervention theory. (b) Search for pertinent stakeholders and relevant theories. (c) Selection of pertinent stakeholders and relevant previous research. (d) Reconstruction of two or more pertinent stakeholder theories. (e) Comparison of stakeholder theories and findings from relevant previous research (Hansen & Vedung, 2010 Figure 1, page 302)

The outcome of the TSE analysis is the elicited stakeholder interventions theories and the systematic comparison of their similarities and differences.

It provides important insights in the dynamics of the intervention and can be used to design an evaluation of the intervention in various ways. One possibility is to frame the intervention theories as competing and to test to what extent the empirical evidence supports one of the intervention theories more than the others. Another possibility is to use the elicited stakeholder theories to enhance the evaluators sensitivity to the presence of competing causal logics embedded in an intervention. As we shall see this latter use of the TSE model seems to be an important merit of the model.

But the TSE analysis can also be useful in other kinds of policy analysis than evaluations and this paper provides an exploratory review of the literature and critically examine how the TSE-model has been used.

The following five research questions will be examined:

- RQ1: What are the basic quantitative characteristics of the TSE-citing literature?
- RQ2: How has the TSE-model been used – empirically and theoretically?
- RQ3: To what extent where the four basic recommendations used?
- RQ4: What modifications of the model have been suggested?
- RQ5: What seems to be the most promising use and novelties associated with the model?

METHODOLOGY

Since 2010 the TSE model has generated some interest among evaluators, evaluation scholars and social scientists. From December 2016 to September 2023, the TSE article has been viewed or downloaded 2482 times from the AJE homepage¹. It has 3634 reads and 89 citations at ResearchGate, has been cited 158 times in Google Scholar, 55 times in Scopus and 45 times in Web of Science (17 September 2023).

The 55 SCOPUS indexed publications were selected for further analysis. One duplicate where excluded, reducing the number to 54. The quantitative characteristics of the selected publications were analyzed in the SCOPUS database using their metrics.

47 publications where then downloaded, examined, and coded (see next quantitative characteristics section). A distinction between empirical and conceptual papers were used to code the papers. Within this typology themes and issues were qualitatively assessed in each group (see empirical use section and conceptual issue section).

1) <https://journals.sagepub.com/doi/abs/10.1177/1098214010366174>

BASIC QUANTITATIVE CHARACTERISTICS OF THE TSE-CITING LITERATURE

The 54 publications (see appendix) citing the TSE article and indexed in SCOPUS were published between 2012 and 2023. The highest number of citations were in 2017 (10 citations) and the lowest in 2020 (1 citation).

Most of the publications are articles (43 publications), followed by book chapters (9 publications) and books (2 publications).

The sources publishing the TSE-citing literature the most are not surprisingly evaluation journals such as *Evaluation* (5 publications), *Canadian Journal of Program Evaluation* (4 publications), *American Journal of Evaluation* (3 publications) and *African Evaluation Journal* (2 publications) but the sources are otherwise spread among a variety of journals within social science, business and management, environmental science, psychology, medicine, economics, energy, and engineering.

The TSE-citing literature has primarily been published by researchers from Europe and Northern America, with notable exceptions. Researchers from Denmark (13 publications) and Sweden (10 publications) have published the most followed by researchers from United Kingdom (9 publications), Canada (7 publications), Netherlands (6 publications), Belgium, Finland, Spain and the United States (all 5 publications), Norway and South Africa (3 publications).

The by far most TSE-citing (and also most cited) author is the Canadian scholar John Mayne (6 publications), followed by N. Clyck, M.B. Hansen, C. Timmermann, R. Van Caudenberg, and L. Van Praag (3 publications).

The two most cited articles in the TSE-citing literature are written by John Mayne (2012, 2015) and focus mainly on conceptual issues and challenges related to the construction of a unitary intervention theory.

Besides the SCOPUS metrics the 54 TSE-citing articles were coded according to their analytical approach (see table 1).

| ANALYTICAL APPROACH | NUMBER OF ARTICLES |
|----------------------------|--------------------|
| 1. Empirical quantitative | 1 |
| 2. Empirical qualitative | 10 |
| 3. Empirical mixed methods | 13 |
| 4. Conceptual | 21 |
| 5. Review | 2 |
| 6. Other/not known | 7 |
| Total number of articles | 54 |

Table 1: Analytical approach of included TSE-citing articles indexed in SCOPUS

The 47 downloaded TSE-citing publications were also coded according to, to what extent the TSE-model where of low (13 publications), medium (12 publications) or high (21 publications) importance in the publication. In the low importance category, the TSE model where mentioned, but not really used. In the high importance category, the TSE model where really used and/or discussed. In what follows a focus is given to the 21 high importance publications.

EMPIRICAL USE OF THE TSE APPROACH

– SELECTED EXAMPLES

16 of the 21 empirical publications really used the TSE model in their empirical analysis.

- three publications concerning *public health issues*,
- five publications concerning *environmental issues*,
- four publications concerning *education issues*,
- four articles concerning *social issues* and
- one article concerning *other issues*.

In what follows four articles – one from each of the four policy areas mentioned above – are selected for further analysis. RQ 2-5 are in focus in the analysis:

- RQ2: How has the TSE-model been used – empirically and theoretically?
- RQ3: To what extent where the four basic recommendations used?
- RQ4: What modifications of the model have been suggested?
- RQ5: What seems to be the most promising use and novelties associated with the model?

Public health policies – Interprofessional collaboration (Vestergaard & Norgaard, 2018): This research project analyzed an interprofessional collaboration (IPC) intervention in a Danish hospital from 2012 to 2015 by means of the TSE model combined with other evaluation approaches. The intervention involved five departments and eight professions and aimed at developing coherent practices across health professions by optimizing patient pathways with the establishment of interprofessional teams as a core element.

Hospitals delivery of patient health care tend to be hampered by poor coordination across departments, insufficient collaboration, and little shared work experience among health professionals. In response, interprofessional collaboration (IPC) has gained popularity worldwide as an effective method to enhance collaborative practices across health care professions (Vestergaard & Norgaard, 2018, p.185). The qualitative methods used were document analysis and focus group interviews. From the TSE model especially the tripartite scheme of situation-, normative and causal theory as well as the distinction between five stakeholder groups (pertinent actors) are used in the article in the three tables presenting the empirical findings. Such analysis of co-creation across different professions seems to be a promising area in which to use the TSE-model.

Environmental policies – the Lima-Paris Action Agenda (LPAA) (Widerberg, 2017): This article examines the Lima-Paris Action Agenda (LPAA) by applying a theory-based evaluation approach. It uses recommendations from the TSE-model, but only constructs one intervention theory by means of documents and interviews. In that sense it uses a theory-based approach, but not a stakeholder approach. From the TSE-model it uses the definition of intervention theory, the tripartite scheme of situation-, causal- and normative theory and the extended system model with pertinent collective actors (2017, p. 729). This is really the intervention theory embedded in the LPAA intervention and thus step one in the TSE-model, and all the recommendations for step one is used. Widerberg's analysis is useful and illuminating since it shows many of the necessary processes to make the LPAA intervention work. It could have been extended to become a TSE project and include the intervention theories of the pertinent actors, who is envisioned to deliver in crucial parts of the intervention. But that would have been a larger and more resource demanding research project. Widerbergs analysis however indicate that Environmental and climate policies is an area where the TSE model can prove very useful as an analytical tool.

Education policies – involving Roma parents (Wauters et al., 2017): In this case-study of the involvement of Roma parents in a primary school in Ghent in Belgium is analyzed. A case-study approach fits nicely with the TSE-model and other theory-based approaches because it allows a comprehensive exploration of the processes of the intervention.

The study indicates that there were few disagreements about the intervention though some tension between team members working on the 'care' aspect and teachers working solely within the classroom where uncovered. It is the general theory-driven notion of uncovering the causal links of the intervention that seems to have been useful in the project, which is an important accomplishment. The TSE-models designed sensitivity to conflict however, helped uncover the tension between care and classroom. Thus, this study indicates that even in low-conflict interventions

the TSE-model may be of use in order to uncover the minor conflicts that may cause an inferior impact of a program.

Social policies – Data Recording in Performance Management (Groth Andersson & Denvall, 2017): This well conducted case-study show, through theoretical perspectives as well as a qualitative empirical analysis of a social intervention in Denmark, three different and somewhat conflictual, but equally convincing recording logics for reporting data in performance management systems in organizations (Groth Andersson & Denvall, 2017, p. 197). The empirically analyzed intervention is a measurement tool inspired by the “Outcomes Star” from the UK, but substantially translated and adapted to the Danish context and to different types of social clients. The social intervention is a performance measurement tool that measures social client’s ability to master their life on 10 dimensions by means of a 10 point Likert scale (2017, pp. 193–194).

Initially the notion of competing logics and the tripartite model for theories of change from the TSE model, where used to conduct the analysis. “The result was two overarching logic models/theories of change: one formulated by leaders and managers and the other formulated by frontline staff, both representing understanding and usage of the performance measurement tool and underlying rationales. These two logic models served as the basis for additional interviews that focused more specifically on the informants’ individual actions and use of the measurement tool. The interviewees were asked if they recognized the overall logic model and to elaborate on and discuss the actions and assumptions it described. These interviews were processed using a more open approach, and emerging themes were coded. From this analysis, the initial logic models were further developed and nuanced, drawing the picture of coexisting and shifting logic models among the frontline staff (2017, p. 195)”.

The resulting three recording logics by frontline staff (2017 Table 1, p. 197) found in the study – tactical, client-centered and “By the book” – seems convincing and likely to be found in most organizations who makes extensive use of performance

measurement. The study show how the TSE-model can be used to uncover hierarchical conflicts related to an intervention.

CONCEPTUAL ISSUES RELATED TO THE TSE APPROACH

Four of the 21 conceptual publications discussed the TSE-model in some depth, while most conceptual publications merely mention the TSE-model as an approach to be consulted in the complex interventions where tensions, conflicts and politics indicate that more than one intervention theory is needed.

While recognizing that competing intervention theories may be present in some cases, most of the theory-based evaluation tradition focus on the challenging task of formulating one unitary intervention theory.

The TSE-model has furthermore been incorporated extensively in some recent reviews of the theory-based evaluation approach in relation to complex interventions such as international development programs and Marine Spatial Planning (Nkwake, 2013; Varjopuro, 2019).

The TSE-model, complexity and ambiguity (Dahler-Larsen, 2018; Mbava & Dahler-Larsen, 2019): One of the conceptual and practical challenges related to the TSE-model is how the relation between the intervention theories and the pertinent stakeholder groups should be understood. The raison d'être of the TSE-model is the recognition of the possible presence of multiple different and partly competing intervention theories. This opens the theory-based evaluation approach to tensions and conflict. The TSE-model propose to construct two or more different, but internally unitary stakeholder theories and then compare their similarities and differences. Such stakeholder unity should however not be overstated. The social policy case-study of performance management discussed above for instance, indicate that a closer look uncovered three rather than two intervention theories and that individual actors within one stakeholder group could represent different intervention theories over time.

This is highlighted in recent publications by Peter Dahler-Larsen and Nombeko P. Mbava (2018; 2019), who suggest

ambiguity rather than conflict and politics as the more promising approach. This approach is a novel and interesting addition to the theory-based evaluation toolbox and should be further tested in future evaluations. Both approaches however are attempts to make theory-based evaluation useful in more complex interventions characterized by tension, ambiguity, and conflict.

CONCLUDING DISCUSSION

Usually theory-based evaluations construct one intervention theory and then test its feasibility or use it as an organizing framework for process-oriented evaluations. In the case of more complex interventions sometimes involving several groups of stakeholders in different organizations, at different hierarchical levels perhaps involving both local, regional, national and international levels, the assumption of one unitary intervention theory may be counterproductive. It tends to enforce designed blindness to tensions, conflicts, disagreements, and ambiguity. This may even be the case in single organizations with a single professional group disagreeing about the possible impact of a novel type of interventions.

The Theory-based Stakeholder Evaluation model (the TSE-model) was formulated to use the theory-based evaluation approach in such complex interventions. The exploratory review indicate that the TSE-model has been useful in several evaluations in all the major welfare areas of the modern welfare state: public health, environmental policy, education, and social policy. The explicit recognition of the problems related to the construction of one unitary intervention theory has perhaps been the most important contribution of the TSE-model. The four recommendations of the TSE model have all been used to varying extent. Perhaps more as a conceptual toolbox to pick ideas from rather than a take it or leave it model.

The TSE-model has primarily been used in Northern Europe and Northern America. It is primarily used in research designs using qualitative or mixed methods designs, while its utility in quantitative studies remains to be seen.

The TSE-model has been useful in major policy areas like public health, education, climate policy and social policy. Across these policy areas the TSE model has been especially useful in interventions requiring cocreation among several professional groups and in interventions requiring hierarchical coordination across several organizations.

Note: The chapter has benefited substantially by insightful comments from senior researcher Kristian Høyer Toft, UCL University College, Denmark.

References

- Chen, H. T. (1990). *Theory-Driven Evaluations* (3103). SAGE.
- Dahler-Larsen, P. (2018). Theory-Based Evaluation Meets Ambiguity: The Role of Janus Variables. *American Journal of Evaluation*, 39(1), 6–23.
- Groth Andersson, S., & Denvall, V. (2017). Data Recording in Performance Management: Trouble With the Logics. *American Journal of Evaluation*, 38(2), 190–204. Scopus.
- Hansen, M. B., & Vedung, E. (2005). *Common Language in the Organization of Elder Care. Evaluation of a Standardized Classification System* [In Danish: Fælles sprog i ældreplejens organisering. Evaluering af et standardiseret kategorisystem]. Syddansk Universitetsforlag.
- Hansen, M. B., & Vedung, E. (2010). Theory-Based Stakeholder Evaluation. *American Journal of Evaluation*, 31(3), 295–313.
- Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, 18(3), 270–280.
- Mayne, J. (2015). Useful Theory of Change Models. *Canadian Journal of Program Evaluation*, 30(2), 119–142.
- Mbava, N. P., & Dahler-Larsen, P. (2019). Evaluation in African contexts: The promises of participatory approaches in theory-based evaluations. *African Evaluation Journal*, 7(1), 1–9. Scopus.
- Nkwake, A. M. (2013). Working with assumptions in international development program evaluation: With a foreword by Michael Bamberger. In *Working with Assumptions in International Development Program Evaluation: With a Foreword by Michael Bamberger* (p. 184). Springer New York; Scopus.
- Stame, N. (2004). Theory-based evaluation and varieties of complexity. *Evaluation*, 10, 58–76.
- Varjopuro, R. (2019). Evaluation of marine spatial planning: Valuing the process, knowing the impacts. In Zaucha, J., Gee, K. (Eds.), *Maritime spatial planning: past, present, future* (pp. 417–440). Springer International Publishing; Scopus.
- Vestergaard, E. & Norgaard, B. (2018). Interprofessional collaboration: An exploration of possible prerequisites for successful implementation. *Journal of Interprofessional Care*, 32(2), 185–195.
- Wauters, J., Van Mol, C., Clycq, N., Michielsen, J., & Timmerman, C. (2017). Involving Roma parents: Analysing the good practice of a primary school in Ghent. *British Journal of Sociology of Education*, 38(3), 287–306.
- Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, 21(4), 501–524.
- Weiss, C. H. (2007). Theory-Based Evaluation: Past, Present, and Future. *New Directions for Evaluation*, 114(Summer), 68–81.
- Widerberg, O. (2017). The “Black Box” problem of orchestration: How to evaluate the performance of the Lima-Paris Action Agenda. *Environmental Politics*, 26(4), 715–737.

KAPITEL 10

Evalueringsbølgers dynamik

Hanne Foss Hansen

INDLEDNING¹

Teorien om og analysen af evalueringsbølger er et af Evert Vedungs mange interessante og inspirerende forskningsmæssige bidrag. Det er et vigtigt bidrag, hvor Evert viser, hvordan forskellige former for evalueringstænkning – konkret den videnskabelige bølge, den dialogorienterede bølge, New Public Management (NPM)-bølgen og evidensbølgen – siden 1960'erne er rullet ind over den vestlige verden. Forandringerne minder om modens skift. Nye evalueringsformer udvikles og udbredes men afsvækkes igen efter nogen tid. Ældre evalueringsformer erstattes imidlertid ikke af nye former. Bølgerne aflejrer sedimenter på deres vej og ideerne overlever som særlige kulturlagre, der fortsat lever og kan reaktiveres og kombineres.

Bølgemetaforen er inspirerende, men de aktører, der er udviklere og bærere af evalueringstænkning og derved får tænkningen til at rulle fra land til land og sektor til sektor er anonyme. Med afsæt i Everts begreber og tænkning vil mit bidrag på den baggrund reflektere over følgende spørgsmål: Hvordan kan den

1) En stor tak til Olaf Rieper for frugtbare kommentarer til et tidligere udkast til denne tekst.

empiriske analyse af evalueringssideernes udvikling og rejse over lande-, sektor- og organisatoriske grænser videreudvikles? Hvordan kan vi empirisk forklare evalueringssølgernes dynamik og skiftende evalueringssideers implementering? Begrebsmæssigt og teoretisk vil mit bidrag hente inspiration i nyere teoridannelser om ide- og vidensoverførsel særligt i den norske forsker Kjell Arne Røvik's univers, virus- og oversættelsesteorien.

I afsnit 2 nedenfor præsenteres Evert's tænkning om evalueringssølgere nærmere. Derefter præsenteres i afsnit 3 Kjell Arne Røvik's begrebs- og teoriunivers. Afsnit 4 diskuterer et illustrativt eksempel på en evalueringsside, der rejser ind i Danmark – konkret anvendelse af test i folkeskolen. Eksemplet tolkes i lyset af Evert's sølgemetaphor kombineret med Kjell Arne Røvik's virus og oversættelsesteori.

EVALUERINGSBØLGER

Evert bruger begrebet evalueringssølgere til at beskrive, hvordan evaluering har udviklet sig som en offentlig styringsaktivitet fra 1960'erne og fremefter (Vedung, 2012 & 2016). Han beskriver, hvordan følgende fire sølgere siden 1960 har rullet ind over den vestlige verden: 1) Den videnskabelige sølger, 2) Den dialogorienterede sølger, 3) New Public Management (NPM)-sølgere og 4) Evidenssølgere. Forskellige tilgange til evaluering har i hver af disse sølgere været forbundet med mere generelle styringsdoktriner. Alle fire sølgere har aflejret sediment, der i dag eksisterer i de offentlige sektorer. Tilsammen har disse aflejringer dannet tidens ideer og praksisser på evalueringssølgere.

Tabel 1 sammenfatter de væsentligste karakteristika ved de fire evalueringssølgere, deres start- og udviklingstidspunkt, deres syn på evaluering samt karakteristikaene ved den styringsdoktrin og det grundsyn/den filosofi, de er indlejret i.

DEL 2: OM UTVÆRDERINGENS ANSPRÅK, ANSATSER OCH ANVÆNDNING

| KARAKTERIS- TIKA: EVALUERINGS- BØLGE: | START- OG UDVELINGS- TIDSPUNKT | EVALUERINGS- SYN | STYRINGS- DOKTRIN | GRUNDSYN/ FILOSOFI |
|--|---|--|--|---|
| Den videnskabelige bølge | Start i 1950'erne, konsolidering fra midten af 1960'erne | Videnskabelig evaluering og test af offentlige inter- ventioner før fuld implementering | Ingeniørmodel for offentlig planlægning, policyudvikling og -implementering | Radikal rational- isme, målmiddel rationalitet, konsekvenslogik, kausalitet- stænkning |
| Den dialog- orienterede bølge | Midten af 1970'erne | Interessent- evaluering, senere brugerevaluering | Interessent invol- vering med sigte på læring, kon- sensuddannelse, politisk accept og demokratisk legitimitet | Samtalerationalitet, kommunikativ rationalitet |
| NPM-bølgen | Slutningen af 1970'erne, konsolidering i 1980'erne | Kvalitetssikring, benchmarking, resultatmåling, KPI'er, mål- opfyldelseevaluering, omkostnings- effektivitet | Mål- og resultatstyring | Nyliberalisme, principalagent tænkning, tro på lederskab, kunde- fokusering |
| Evidensbølgen | Midt 1990'erne | Systematiske reviews (forsk- ningsoversigter) vedrørende resul- tater af offentlige interventioner udarbejdet med afsæt i metoder placeret højest muligt i evidens- hierarkiet. Randomiserede, kontrollerede forsøg (RCT) er idealet | Evidensbaseret politik | Kausalitets- tænkning |

Tabel 1: Evalueringsbølger

Ifølge den videnskabelige bølge, der udvikledes fra start 1950'erne, burde offentlige interventioner testes med afsæt i en eksperimentel metodik, før de blevet udrullet fuldt ud. Fra midten af 1970'erne udvikledes den dialogorienterede bølge. Her var interessentevaluering og brugerevaluering i fokus. Inddragelse skulle sikre læring og give offentlige interventioner demokratisk legitimitet. Fra slut 1970'erne kom så NPM-bølgen. Fokus var på målopfyldelsesevaluering, resultatmåling og omkostnings-effektivitet. Da Evert publicerede sine vigtige bidrag om evalueringsbølgerne i 2010'erne var evidensbølgen det sidste skud på stammen. Denne bølge var til dels en reaktivering af den videnskabelige bølge, men evidensbevægelsens pointering af vigtigheden af udarbejdelse af forskningsoversigter, de såkaldte systematiske reviews, ikke blot på det medicinske område men på alle offentlige politikområder var ny (Hansen & Rieper, 2009). Tiltroen til mulighederne for at generalisere eksisterende solid viden var forøget.

Evert rejser selv (Vedung, 2016) spørgsmålet om, hvilke nye bølger der kan være på vej. En kandidat, som Evert nævner, er løbende evaluering (i EU terminologi "ongoing evaluation"), en evalueringsform der foregår og afrapporteres kontinuerligt parallelt med at en intervention implementeres. Sigtet er at facilitere læring og sikre, at evalueringsresultaterne bringes til anvendelse. "Regulatory impact assessment", en evalueringsform, der sigter mod at forbedre lov-kvalitet blandt andet ved at minimere uønskede side-effekter, kan nævnes som en anden mulig kandidat.

Evert's tænkning om evalueringsbølger er et særdeles frugtbart og vigtigt bidrag på to dimensioner. For det første, viser Evert os, hvordan evalueringsfeltets faglige udvikling ikke er foregået i et autonomt rum af evalueringseksperter, men derimod i tæt interaktion med udviklingen i forvaltningspolitiske styringsdoktriner. Denne kontekst har evalueringslitteraturen, der fremfor alt har haft fokus på udvikling af evalueringsdesigns og -metoder, kun sjældent haft blik for (se dog Hansen, 2005, for en diskussion af sammenhængen mellem statsmodeller og evalueringsmodeller). For det andet, viser Evert os, hvordan udviklingen af nye styringsdoktriner har inklu-

deret udvikling af nye former for evalueringstænkning, et perspektiv litteraturen om offentlig forvaltning og offentlig sektor styringsudvikling kun sjældent har haft blik for. Everts store styrke, der danner konteksten for udviklingen af tænkningen om evalueringsbølger er hans historiske forvaltnings- og evalueringsfaglige overblik og indsigt.

Bølgemetaforen er yderst inspirerende. Bølger ruller ind og ud, skifter størrelse, karakter og retning. Bølger kan for en tid lægge sig som smult vand men vender på et tidspunkt tilbage. Bølger er alt-overvejende vind- og strømbestemt. Aktører spiller en marginal rolle.

Evert er bevidst omkring, at evalueringsbølgerne ikke kan forventes at slå igennem på samme tid og med samme intensitet i alle lande og sektorer, og opfordrer derfor til, at kolleger i forskellige lande lader sig inspirere til, at udarbejde deres egne analyser og fortællinger om ligheder og forskelle i landes og sektorerers evalueringshistorie. Dette er en god ambition men spørgsmålet er, om der ikke vil være behov for et mere rigt og systematisk begrebsapparat, hvis mere udførlige komparative analyser skal udvikles. Jeg vil i det følgende argumentere for, at der med fordel kan hentes inspiration i litteraturen om idé- og vidensoverførsel specielt i den norske forsker Kjell Arnes Røvik's rige begrebsverden, som vil blive præsenteret i næste afsnit.

IDÉ- OG VIDENSOVERFØRSEL

Den norske organisationsforsker Kjell Arne Røvik (2007, 2011, 2016 & 2023) har udviklet en begrebsrig teori om, hvordan ideer og viden overføres mellem organisationer, sektorer og lande. I forlængelse af hans tidlige arbejder bruger jeg her betegnelsen virus- og oversættelsesteorien. Virus skal i denne sammenhæng ikke opfattes som et problem, der skal reguleres via regler, normer og samfundssind. Virus skal derimod opfattes som en metafor, der er lærerig i relation til de dynamikker der udspiller sig, når ideer og viden rejser, parallelt til Evert's tænkning om nye evalueringsbølgers indtog som modeskift. Selv har Kjell Arne Røvik i tiden efter corona droslet brugen af virusmetaforen ned, men jeg synes fortsat den er en inspirerende metafor. Virus- og oversættelsesteoriens kerne og begrebsmæssige rigdom er sammenfattet i tabel 2.

| | |
|---|--|
| KERNEBEGREBER I VIRUS- OG OVERSÆTTELSESTEORIEN | Begreber applikeret på evalueringsideer |
| VIRUS: MUTATION, INAKTIVERING, REAKTIVERING | Evalueringsideer rejser over lande-, sektor og organisatoriske grænser båret af aktører, for eksempel forskere, konsulenter og offentlige myndigheder. En evalueringside kan mutere under rejsen og den kan inaktiveres i en periode for måske at reaktiveres igen senere. |
| MODTAGER: IMMUNITET, INKUBATIONSTID, GLEMSEL | Rejsen kan foregå hurtigt eller langsomt afhængig af karakteristika ved modtageren. |
| DEKONTEKSTUALISERING | En evalueringside's rejse påbegyndes, når en aktør oversætter en evalueringspraksis til en idemæssig beskrivelse – fra praksis til tale. |
| KONTEKSTUALISERING | En ide's indrejse sker ved at en aktør oversætter den idemæssige beskrivelse til en praksis i en ny kontekst – fra tale til praksis. |
| OVERSÆTBARHED | En praksis/ide's oversætbarhed kan variere afhængig af dens eksplicitet, kompleksitet og indvævedhed i andre praksisser. |
| OVERSÆTTELSESTYPER (MODUS) | Oversættelse af en evalueringside kan ske via reproduktion i form af kopiering, modifikation i form af at noget lægges til eller trækkes fra eller som en mere radikal forvandling. |
| OVERSÆTTELSESKOMPETENCE | Oversætterens kompetence er en kritisk succesfaktor i implementeringen af en evalueringside. God oversættelse kræver viden, mod, tålmodighed og styrke. |
| BARRIERER FOR OVERSÆTTELSE: | Oversætteren kan opleve, at der er barrierer for en evalueringside's oversættelse, der skal overkommes. Barrierer kan være geografiske, kulturelle, politiske eller institutionelle. |
| EFFEKT | Effekterne af evalueringsideers rejse kan variere i omfang – fra begrænset til massiv – samt i karakter – fra succes til fejlslagen eller symbolsk. |

Tabel 2. Virus- og oversættelsesteori

Ifølge teorien kan evalueringsideer og -modeller anskues som virus, der rejser over lande, sektor og organisatoriske grænser. Virus bæres af aktører som for eksempel forskere, konsulenter og offentlige myndigheder, der ser et behov for eller direkte har til opgave at bidrage til, at evaluering indoptages i nye kontekster. Når ideer rejser kan de forandres for eksempel mutere eller gå i dvale for senere at blive reaktiveret. Lande, sektorer og organisationer kan som smittebærere og modtagende aktører være modtagelige eller immune overfor nye evalueringsideer. Aktørers immunitet kan efter en tid erstattes af modtagelighed. Inkubationstiden kan så at sige variere.

Teorien sonder mellem dekontekstualisering, det at aktører henter en ide ud fra en kontekst, og kontekstualisering, det at aktører indskriver en ide i en ny kontekst. Evalueringsideer og modeller er halvfabrikata, der videreudvikles, tilpasses og konkretiseres, når de oversættes og tages i anvendelse. Ideers oversættelighed kan variere.

De aktører, der rejser med ideer, oversætterne, kan vælge at arbejde i et reproducerende, et modificerende eller et radikalt modus. I et reproducerende modus kopieres en evalueringside ind i en ny kontekst. I et modificerende modus ændres en evalueringside enten ved at den forenkles (elementer trækkes fra) eller ved at den suppleres med nye elementer. I et radikalt modus forvandles evalueringsideen markant, idet den implementeres i en ny kontekst. Til tider kan forandringen være så radikal, at der snarere bliver tale om, at der lanceres en ny evalueringside.

Oversætternes kompetencer er kritiske for deres evne til at vælge et relevant modus og derved overvinde eventuelle barrierer af geografisk, kulturel, politisk eller institutionel karakter. Indoptagne evalueringspraksisser kan få varierende betydning fra at have begrænsede til massive effekter og de kan variere fra at være succesfulde, fejlslagne eller symbolske handlinger.

En evalueringsbølge i Everts forstand opstår, når evalueringsideer med givne karakteristika rejser massivt i form af, at flere aktører samtidigt fatter interesse for ideen ud fra logikken ”sådan kan vi med fordel også evaluere”.

ET ILLUSTRATIVT EKSEMPEL

I det følgende præsenteres et illustrativt eksempel på en evalueringsside, der rejser. Eksemplet viser, hvordan en evalueringsside over tid kan udvikles med inspiration fra flere af Everts evalueringssider samt hvordan Kjell Arnes Røvik's begrebsverden kan give indsigt i indhentningen, udviklingen og implementeringen af ideen. Konkret vedrører eksemplet introduktionen af test af elevfærdigheder i den danske Folkeskole. Sådanne test er i løbet af en årrække introduceret i mange landes skolesystemer. Eksemplet kan karakteriseres som en international sektoriel evalueringsside. Det vedrører således ikke en af Everts overgribende evalueringssider, men derimod en sektoriel side, der influeres af flere af Everts overgribende sider. Tabel 3 giver en oversigt over de vigtigste begivenheder i introduktionen og udviklingen af test i den danske folkeskole.

| ÅRSTAL | BEGIVENHED | KONKLUSION |
|------------|---|---|
| 2000 /2003 | PISA chock | PISA undersøgelser viser, at danske skoleelevers færdigheder er middelmådige. |
| 2004 | OECD rapport om grundskolen i Danmark | Evalueringsskulturen i skolen karakteriseres som svag. |
| 2006 | Folketinget vedtager lov om styrket evaluering og nationale test | Test indføres med første testrunde 2010. |
| 2006 | Det rådgivende ingeniørfirma Cowi A/S vinder opgaven med at udvikle og drive de nationale test. | |
| 2010 | Undervisningsministeriet beslutter selv at overtage driften for en periode. | Beslutningen kommer efter forsinkelser og systemnedbrud. |
| 2013 | Konsulenthuset Rambøll evaluerer testordningen. | De nationale test har en positiv effekt på elevernes faglige niveau. Der er forbedringspunkter, blandt andet udfordringer med at bruge testresultaterne fremadrettet. |

| ÅRSTAL | BEGIVENHED | KONKLUSION |
|--------|---|---|
| 2013 | Undervisningsministeren afgiver redegørelse til Folketinget. | Indstilling om ny evaluering efter yderligere 3-5 år. |
| 2018 | Nedsættelse af rådgivningsgruppe vedrørende evaluering af de nationale test. | Anbefalinger til evalueringstemaer. |
| 2019 | Styrelsen for IT og læring (STIL) under Undervisningsministeriet gennemfører teknisk analyse af testenes statistiske egenskaber og den adaptive algoritmes funktion. | STIL finder blandt andet, at der ikke er programmeringsfejl i testene og at testene er kriterievalide, idet der er sammenhæng mellem resultaterne og resultaterne af andre test og prøver. |
| 2020 | VIVE publicerer "Evaluering af de nationale test." Afrapportering af to undersøgelser: 1) Kritisk review af STIL's analyse; 2) Undersøgelse af brugen og betydningen af test. | Testene er et vigtigt styringsredskab på kommunalt og nationalt niveau, men opleves ikke af lærerne som et relevant pædagogisk redskab til undervisningsrettelæggelse. |
| 2020 | Rådgivningsgruppen for evaluering af de nationale test. Anbefalinger 2020. | Anbefalinger afgivet på basis af STIL's og VIVE's rapporter. Det adaptive princip foreslås erstattet af et lineært. Testene foreslås gennemført i starten af skoleåret for at øge mulighederne for fremadrettet brug. |
| 2021 | Politisk aftale om det fremtidige evaluering- og bedømmelsessystem i folkeskolen. | Folkeskolens nationale færdighedstest skal fra skoleåret 2026/27 erstatte de tidligere nationale test. I perioden 2022/23 til og med 2025/26 gennemføres overgangstests. |
| 2023 | Ekspertgruppen for Folkeskolens Nationale Færdighedstest. | Gruppen skal bidrage til at sikre et kvalificeret og solidt fagligt grundlag for den kommende udvikling og brug af test. Gruppen foreslår blandt andet, at testene knyttes til de øvrige nationale styringsredskaber samt at lærerne vejledes og understøttes i pædagogisk anvendelse af testresultaterne. |

Tabel 3. Indførelse af test i den danske folkeskole.

De nationale test blev besluttet indført i 2006, efter at to PISA undersøgelser (det såkaldte PISA chock) havde vist, at danske skoleelevers færdigheder var middelmådige. Testbatteriet bestod af 10 obligatoriske test i dansk, matematik, engelsk og fysik/kemi. Testene var it-baserede, selvrettende og adaptive, idet opgaverne varierede i sværhedsgrad alt efter den enkelte elevs niveau. Testene skulle gennemføres i slutningen af skoleåret.

Testsystemets udviklingsfase blev en skandale. Udviklingen blev forsinket, der var it nedbrud og den virksomhed, der var udset til at drive systemet blev sat fra opgaven. Da systemet kom i drift, blev kritikken især fra lærerkredse massiv. Det forlød, at de adaptive testresultater var vanskelige at aflæse for lærerne, ligesom testregimet medførte en risiko for, at lærerne omlagde deres undervisning fra at bygge på pædagogiske principper til at blive orienteret mod en "teach to the test" filosofi. Herudover kritiserede eksperter testenes validitet, hvilket gav anledning til tvivl om testenes pædagogiske anvendelsespotentialer.

Som det fremgår af tabel 3 blev der i kølvandet på kritikken søsat evalueringer i flere runder. I 2021 – efter 15 års diskussion – blev der så indgået en politisk aftale om det fremtidige evalueringssystem. De adaptive test bliver fremover erstattet af lineære test, hvor alle elever får samme opgavesæt, og testenes afholdelse flyttes til starten af skoleåret. Sigtet er at gøre testresultaterne mere forståelige og brugbare for lærerne som grundlag for planlægning af undervisningen.

Hvis vi tolker casen i Evert's bølgeperspektiv, hvad ser vi så?

Indførelsen af test i den danske folkeskole kan i et sektorielt perspektiv ses i konteksten af de internationale skolepolitiske nyliberale strømninger (Krejsler, 2021) og i et bredere perspektiv i konteksten af Evert's NPM-bølge med fokus på resultatstyring. Mens testsystemets 1. generations løsning baseret på det adaptive princip med test i skoleårets slutning blev udviklet med afstand til folkeskolens interesser blev 2. generationsløsningen baseret på det lineære princip med test i starten af skoleåret udviklet med inddragelse af folkeskolens interesser via den nedsatte

rådgivningsgruppe og den efterfølgende ekspertgruppe. Efter den massive kritik fra sektorens interessenter valgte myndighederne og politikerne således at tilrettelægge en proces, der kan tolkes som baseret på Evert's dialogorienterede evalueringsbølge.

Hvis vi tolker casen i Kjell Arne Røvik's perspektiv, hvad ser vi så? Ideen om at indføre et nationalt testsystem kan klart betragtes som en virus, der har rejst på tværs af landegrænser. Inspirationen kom til Danmark udefra. Testsystemer kan imidlertid designes på mange måder. Ideen er kompleks. Den er ikke let oversættbar. Som VIVE skriver i deres evaluering var den danske 1. generations-løsning på en og samme tid en typisk faglig test med den usikkerhed sådanne test har og en atypisk test sammenlignet med andre lande på grund af dels det dobbelte formål rettet mod både pædagogisk brug og brug som styringsredskab og dels det valgte adaptive princip (VIVE, 2020, s. 9). Ideen blev i den danske variant forsøgt indvævet i flere praksisser, hvilket efterfølgende viste sig at være yderst vanskeligt. Oversættelsen, der må karakteriseres som radikal – det adaptive princip anvendes i meget få test i udlandet – viste sig til dels at være fejlslagen. VIVE's evaluering viste, at testsystemet blev anvendt som styringsredskab især på nationalt og kommunalt niveau, i mindre grad på skoleniveau, mens lærerne ikke oplevede testresultaterne som et relevant pædagogisk redskab.

Hvad kan forklare den i første omgang delvis fejlslagne proces?

For det første er det risikofyldt at vælge et radikalt oversættelsesmodus. Der er ganske enkelt færre erfaringer at trække på. For det andet kan man stille spørgsmål ved, om oversætteren i denne case havde tilstrækkelig viden om skolens verden til at sikre en succesfyldt løsning. Endelig kan der peges på, at der synes at have været politiske og institutionelle barrierer for oversættelse. Den internationale skolepolitiske strømning er som nævnt ovenfor præget af tænkningen i Evert's NPM-bølge og dermed på en instrumentaliserings af skolen, som bryder med klassiske politiske og institutionelle danske skoleværdier. Den samlede proces over de

15 år kan betragtes som en kontekstualiseringsproces opdelt i flere faser. Mens ministeriet og Cowi var de primære aktører i første fase af kontekstualiseringsprocessen er aktørbilledet bredere i fase 2. Her er både praktikere fra skoleverdenen og eksperter inviteret ind som rådgivere i processen. Her ser vi, at aspekter af Evert's dialogorienterede bølge kan danne rammen for en mere konsensorienteret proces og en modificerende oversættelsesmodus (Dahler-Larsen, 2023). Mens testsystemets 1. generationsløsning blev indskrevet i en praksis baseret på det adaptive princip, er 2. generationsløsningen planlagt indskrevet i en praksis baseret på det lineære princip. Sigtet er at øge potentialet for pædagogisk anvendelse af testresultaterne. Spørgsmålet er blot, om dette vil ske, når testsystemets fortsat vil blive indskrevet i flere forskellige praksisser jævnfør det dobbelte formål.

KONKLUSION

Evert's teori om og analyse af evalueringsbølger er et interessant og meget inspirerende forskningsmæssigt bidrag. Det giver os frem for alt indsigt i sammenhængen mellem på den ene side udviklingen i forvaltningsmæssige styringsdoktriner og på den anden side udviklingen i evalueringstænkning. Mit bidrag her har reflekteret over, om og hvordan Evert's bølgetænkning kan videreudvikles, så de aktører og de dynamikker, der aktiveres, når bølgerne ruller over lande- og sektorgrænser kan tydeliggøres i analysen.

Konkret rejste jeg i indledningen følgende spørgsmål: Hvordan kan den empiriske analyse af evalueringsideernes udvikling og rejse over lande-, sektor- og organisatoriske grænser videreudvikles? Hvordan kan vi empirisk forklare evalueringsbølgernes dynamik og skiftende evalueringsideers implementering? Diskussionen af virus- og oversættelsesteorien og analysen af et illustrativt eksempel – indførelsen af test i den danske folkeskole – har vist, at en kombination af Evert's bølgetænkning og Kjell Arne Røvik's virus- og oversættelsesteori er et nyttigt greb.

Referencer

- Dahler-Larsen, P. (2023). Can we use deliberation to change evaluation systems? How an advisory group contributes to policy change. *Evaluation*, 29(2), 144-160.
- Ekspergruppen for Folkeskolens Nationale Færdighedstest. *Anbefalinger*, September 2023.
- Hansen, H. F. (2005). Choosing Evaluation. Models. A discussion on Evaluation Design. *Evaluation*, 11(4), 447-462.
- Hansen, H. F., Rieper, O. (2009). The Evidence Movement. The Development and Consequences of Methodologies in Review Practices. *Evaluation*, 15(2), 141-163.
- Krejsler, J. B. (2021). *Skolen og den transnationale vending*. København: Nyt fra Samfundsvidenskaberne.
- OECD (2004). *OECD-rapport om grundskolen i Danmark – 2004*. København: Undervisningsministeriet.
- Rambøll (2013). *Evaluering af de nationale test i folkeskolen*. København: Rambøll.
- Røvik, K. A. (2007). *Trender og translasjoner. Ideer som former det 21. århundrets organisasjon*. Oslo: Universitetsforlaget.
- Røvik, K. A. (2011). From Fashion to Virus: An Alternative Theory of Organizations' Handling of Management Ideas. *Organization Studies*, 32(5), 631-653.
- Røvik, K. A. (2016). Knowledge Transfer as Translation: Review and Elements of an Instrumental Theory. *International Journal of Management Review*, 18(3), 290-310.
- Røvik, K. A. (2023). *A translation theory of knowledge transfer. Learning across organizational borders*. Oxford: Oxford University Press.
- Rådgivningsgruppen for evaluering af de nationale test. *Anbefalinger*, Januar 2020.
- Vedung, E. (2012). Four Waves of Evaluation Diffusion. *Evaluation*, 16(3), 263-277.
- Vedung, E. (2016). Fire Evalueringsbølger. I T. Bredgaard (Red.). *Evaluering af offentlig politik og administration* (s. 37-68). København: Hans Reitzels Forlag.
- VIVE (2020). *Evaluering af de nationale test. Tværgående evalueringsrapport*. København: VIVE.

KAPITEL 11

De andra diffusionsvågorna: Utvärdering i förändring

Wolfgang Meyer & Reinhard Stockmann

INTRODUKTION

Utvärderingens historia är kort, men anmärkningsvärt framgångsrik. Rötterna finns inte enbart i USA, utan även i några andra länder som Kanada, Australien, Sydafrika och Sverige. I sin berömda systematisering nämner Alkin & Christie (2023, s. 13) tre olika grunder: samhällsvetenskaplig forskning, epistemologi och ansvarskrävande. Motsvarande uppställning kan också hittas i en uppsats av Evert Vedung (2010). Utvecklingen har kännetecknats av en snabb, världsomspännande spridning av utvärdering som ett instrument för offentliga administrativa åtgärder och konsekvensbedömning av lagstiftning. Utmärkande för denna spridning är till exempel uppkomsten av utvärderingsorganisationer, utvärderingspublikationer, utbildnings- och vidareutbildningsmöjligheter för utvärderare samt grundläggande regelverk med krav på och råd om utvärderingar i allt fler länder sedan början av 2000-talet (Meyer, 2015).

Detta är ett försök att skildra utvärderingens spridning i dess historiska förlopp, för att förstå drivkrafterna och deras betydelse för uppkomsten av utvärdering. Dynamiken beskrivs i termer av ut-

vecklingsvågor. Uppsatsen bygger på Evert Vedungs pionjärarbete, vilket kompletteras med resultaten från vår egen studie om den globala institutionaliseringen av utvärdering (Stockmann et al. 2020, 2022, 2023, 2024). Huvudfrågan är i vilken utsträckning Evert Vedungs slutsatser fortfarande är giltiga 14 år senare och i vilken utsträckning det finns nya, kompletterande eller motsatta utvecklingstrender. Vårt empiriska underlag är större än Evert Vedungs. Det bygger på empiriska resultat från mer än fyrtio länder i alla världsdelar och kan därmed för första gången ge en global översikt över utvärderingens ställning i de politiska, sociala och professionella systemen i de studerade länderna.

Först presenteras kortfattat Evert Vedungs centrala slutsatser om spridningsprocessen för utvärdering (avsnitt 2). Därefter följer en översikt av vårt eget arbete om den globala institutionaliseringen av utvärdering (avsnitt 3). Slutligen lyfts fyra nya trender fram och kontrasteras mot Evert Vedungs resultat (avsnitt 4).

DE FYRA VÅGORNA AV UTVÄRDERINGSSPRIDNING

Liksom Alkin och Christie (2023) och Guba och Lincoln (1989) beskriver Evert Vedung (2010) utvecklingen av utvärdering i Nordamerika och Nordeuropa som en sekvens av olika stadier, och använder bilden av vågor som överlappar och blandas. Från 1960-talet och framåt urskiljer han fyra vågor, som bestämde utvecklingen av utvärdering som ett instrument för offentlig politik med varierande intensitet och konsekvenser.

Den första av dessa vågor bygger på Max Webers tänkande och den radikala rationalismens ideal, som utvecklades mot vetenskapsbaserad planering och förvaltning i anslutning till "scientific management" av Taylor (2020 [1911]) och Gantt (2017 [1910]). Användningen av vetenskaplig kunskap fokuserar på planeringsfasen och därmed på den "främre delen" av policycykeln (Vedung, 2010, s. 266). I detta sammanhang representerar utvärdering resultatet av värdefri, vetenskaplig, neutral forskning som leder till objektiva resultat och fakta. På denna grund fattas sedan rationella beslut, som följs av alla deltagare i form av rationellt beteende.

I början av 1970-talet förändrades det offentliga tänkandet eftersom särskilt unga människor blev alltmer missnöjda med bristen på reformer och bristen på möjligheter att delta. I Tyskland ledde protesterna från "68-generationen" till att förbundskansler Willy Brandt lovade att "våga mer demokrati" (Schildt & Schmidt, 2019). Rätten till deltagande krävdes inte bara i planeringsfasen utan i hela styrningsprocessen (Vetter & Remer-Bollow, 2017). Samtidigt började olika discipliner utveckla "stakeholder analysis", i syfte att identifiera viktiga intressegrupper och ta hänsyn till deras perspektiv i beslutsfattandet (Brugha & Varvasovszky, 2000). Slutligen utmanades den deduktiva epistemologin med sin kvantitativa metodik stödd av naturvetenskapens framgångar, av konstruktivismen och den kvalitativa samhällsforskningen, vilket ledde till våldsamma metodologiska dispyter (Adorno et al, 1976; Kelle, 2009). Inom utvärderingen anknyter man till debatten om "deliberativ demokrati" och utvecklar förfaranden för deltagande av intressentgrupper och särskilda målgrupper i form av "empowerment evaluation".

Slutet av 1970-talet medförde en nyliberal vändning i många västländer, särskilt i USA, Storbritannien, Tyskland och Sverige, mot en ytterligare utvidgning av välfärdsstaten och föravreglering, privatisering och decentralisering. Särskilt reformen av den offentliga förvaltningen i riktning mot "New Public Management" (Lane, 2000), som hade sitt ursprung i Storbritannien, ledde till långtgående förändringar: Medborgarna ses nu inte längre som "hjälsökande" utan som "kunder". Administrativa åtgärder skall vara inriktade på resultat och styras i denna riktning decentraliserat av de direkt ansvariga cheferna ("resultatbaserad förvaltning"), vilket blev standard särskilt inom utvecklingssamarbetet (Binnendijk, 2000). Styrning blir viktigare än planering. Vedung (2010, s. 270f.) beskriver tre huvudelement i New Public Management: tron på ledarskap, indirekt istället för direkt kontroll ("navigering istället för rodd") och slutligen kund- istället för intressentorientering. Detta leder till en "ökad betoning av ansvar-utkrävande med sikte på resursanvändning" (Vedung 2010, s. 273, egen översättning).

Evert Vedung lokaliserar den fjärde vågen till slutet av 1990-talet och kallar den en “evidence movement”, som mäter statliga åtgärder efter deras framgång och producerar systematiska översikter av “evidence-based learning” med hänvisning till direkta effekter (“outcome level”). På sätt och vis representerar vågen en återgång till vetenskaplig forskning, men med nya och utökade uppgifter. Den kommer främst från det medicinska området och är enligt Vedung svagare än de tre vågor som beskrivits ovan.

Med dagens perspektiv slutar Vedungs beskrivning omedelbart före en stor global ökning av utvärdering. Som ett exempel: enligt IOCE:s (International Organization for Cooperation in Evaluation) statistik fanns det 1997 endast tio länder med utvärderingsnätverk och föreningar; när artikeln publicerades 2010 fanns det redan cirka åttio, och idag har detta antal stigit till långt över hundra (Meyer, 2023, s. 455). En beskrivning av utvärderingens utveckling kan därför inte längre begränsas till Nordamerika, Nord- och Västeuropa, särskilt eftersom dynamiken nu är högre utanför denna region.

SPRIDNING I VÄRLDEN – DET NYA ÅRTUSENDET

Forskning om utvärdering är fortfarande i sin linda. Därför vet man inte mycket om institutionaliseringen av utvärderingsstrukturer på nationell nivå runt om i världen. Furubo, Rist och Sandahl (2002) presenterade en första inventering för 20 år sedan. Den följdes av en replikering 2015 (Jacob, Speer, Furubo) – med samma indikatorer i ett slags poängsystem. Detta möjliggjorde en longitudinell jämförelse som bland annat visade två saker. Å ena sidan identifierades en “lyfteffekt”, dvs. att de flesta av de studerade länderna hade infört mer omfattande utvärderingsstrukturer och därför uppnådde en högre utvärderingspoäng än 2002. Å andra sidan visade det sig att de tidigare “föregångsländerna” (USA, Kanada, Australien och Sverige) i stort sett låg kvar på samma höga nivå, medan några nykomlingar (inklusive Spanien och Nya Zeeland) strävade efter att nå toppen.

I övrigt har arbeten om institutionaliseringen av utvärdering varit ganska sällsynta under de senaste 20 åren och endast foku-

serat på utvalda aspekter (Widmer et al. 2009, Rosenstein 2013, Stockmann & Meyer, 2016). Dessa studier kritiserades framför allt för att de saknade ett teoretiskt och metodologiskt koncept, att de fokuserade på Europa och Nordamerika och därmed uteslöt nästan alla länder i det globala syd, samt att de beskrev enskilda politiska eller tematiska områden.

Denna analys gav upphov till idén om CEval-Evaluation Globe, ett projekt med målet att skapa en systematisk, världsomfattande och landsjämförande översikt över institutionaliseringen av utvärdering samt att analysera dynamikerna och drivkrafterna bakom denna utveckling. Nationalstaterna valdes som analysenheter eftersom de är av central betydelse för social kontroll genom det politiska systemet (Hay, Lister & March, 2006), till exempel genom nationell lagstiftning och administrativa bestämmelser. Nationalstaten har också en ledande roll eftersom många offentliga uppgifter hanteras av dess administrativa organ.

För forskningsprojektet utvecklades ett teoretiskt koncept som anknyter till moderniseringsteorin (Zapf, 1996, Berger, 1996) och social differentiering (Luhmann, 1988, Schimank, 1996), enligt vilket olika funktionella delsystem utvecklas för att kunna hantera sociala uppgifter mer effektivt och ändamålsenligt. Ytterligare överväganden (Meyer, Stockmann & Taube, 2020) har lett till ett fokus på tre delsystem: det politiska systemet, det sociala systemet och systemet för professionalisering. På denna grundval undersöks den nationella institutionaliseringen av utvärdering (jfr figur 1).

Landexperter gjorde fallstudier och följde ett standardiserat, bindande analyschema. I vissa länder där utvärdering har påtaglig relevans gick det inte att hitta författare. I slutändan omfattar Europa-volymen 16 länder, Amerika-volymen 11, Asien-volymen 11 och för Afrika-volymen pågår 14 landanalyser. Dessutom belyser utvalda transnationella organisationer och nätverk sina bidrag till utvecklingen av nationella utvärderingsstrukturer i de enskilda regionerna.

| INSTITUTIONALISERING AV UTVÄRDERING I OLIKA DELSYSTEM | | |
|--|--|---|
| Politiskt system: Institutionella strukturer och processer | Socialt system: Spridning och acceptans av utvärdering i samhället | Yrkessystem: Utvärdering som en disciplin |
| Nationella lagar, förordningar och policyer | Det civila samhällets användning av utvärderingar | Akademisk utbildning och fortbildningspraxis |
| Parlamentariska och natio- nella revisionsstrukturer | Offentlig diskurs | Tidskrifter och kommu- nikationsplattformar |
| Organisationsstruktur | Det civila samhällets deltagande | Yrkesorganisationer |
| Utvärderingspraxis | Efterfrågan på utvärderingar | Förekomst av och efterlevnad av standarder |

Figur 1. Dimensioner för institutionalisering av utvärdering

Källa: Meyer, Stockmann, Szentmarjay, 2022, s. 19

DE ANDRA DIFFUSIONSVÅGORNA: FYRA NYA TRENDER

Baserat på resultaten från CEval Evaluation Globe kan fyra nya trender härledas som antingen bygger på Evert Vedungs identifierade vågor, kompletterar dem eller, åtminstone delvis, också går emot dem. Den första och uppenbart motsatta trenden är förmodligen den mest kända. I motsats till de rationaliseringstrender som Vedung med rätta betona, dominerar nu en motrörelse som producerar sina egna fakta och följer sina egna ideologiska idéer i det globala nord. Denna utveckling kan bäst beskrivas med hjälp av etiketten "fake news" som introducerades som ett slagord av USA:s tidigare president Donald Trump (Neo, 2021, Tandoc, 2019, Tambini, 2017). Donald Trump är förvisso en framträdande representant för denna utveckling, men långt ifrån den enda – det är en global utveckling under de senaste tjugo åren (Stockemer, 2019, Torre, 2020).

Populistiska partier vinner också mark i Europa, och de vänder sig kraftigt mot "etablissemangen" och dess rationalitet, underbygger dem med mer eller mindre logiska konspirationsteorier, samt försöker hävda sin överlägsenhet aggressivt, förtalande och ibland även med våld (Heinisch et al, 2020). Några av dessa partier sitter nu vid makten i ett antal länder – med början i de central-

och östeuropeiska länderna Polen och Ungern (Csehi & Zgut, 2020) har denna våg nu även nått västeuropeiska demokratier – och på senare år även Sverige, som länge har betraktats som ett bålverk mot högerpopulismen (Rydgren & Meiden, 2016).

Trots denna skrämmande utveckling i de europeiska demokratiernas politik har detta ännu inte haft någon märkbar inverkan på användningen och utformningen av utvärderingar. Även om populistiska krafter på sina ställen försöker att helt ta över centrala demokratiska institutioner som byråkratin eller rättssystemet, har dessa hittills visat sig vara mycket motståndskraftiga (se till exempel Helm & Nasu, 2021). I motsats till den politiska trenden, fortsätter utvärderingar att utvecklas nästan som i en administrativ parallellvärld. Drivkraften bakom utvecklingen av utvärderingen är just den institution som har upphöjts till populisternas fiende.

Särskilt i samband med utvidgningarna av Europeiska unionen i Syd- och Centralosteuropa (1986 och 2004) spelade strukturfondsåtgärder och *EU:s sammanhållningspolitik* en viktig roll för integrationen av de länder som anslöt sig (Blomeyer & Eussner, 2020, Stern 2009, Bachtler & Wren, 2006). Eftersom alla EU-åtgärder måste utvärderas enligt gemenskapens regler, har detta lett till utvecklingen av en utvärderingskultur, särskilt i Spanien, Tjeckien och Polen. Samtidigt utgör dock denna externt ålagda och externt drivna utveckling också ett problem eftersom engagemanget för utvärdering saknas i de inhemska institutionerna.

På samma sätt har den tredje nya vågen, som utgår från *multi-nationella organisationer* (särskilt Världsbanken, UNDP, UNICEF och OECD), främjat utvecklingen av utvärderingar över hela världen. Det har skapats nätverk som bland annat har samordnat verksamhet mellan enskilda utvärderingsenheter. De har kopplats samman med framväxande sammanslutningar för främjande av utvärdering i det civila samhället. Dessutom finns det ett stort antal direkta finansieringsåtgärder som särskilt syftar till kapacitetsuppbyggnad ("Evaluation Capacity Development" – ECD) och till installerandet av nationella övervaknings- och utvärderings-system ("National Monitoring- and Evaluation-systems" – NMES) (Meyer, 2023). Denna utveckling inleddes i slutet av förra seklet

och bidrog till en global spridning, särskilt under det första årtiondet. I motsats till EU, är kraven från de multinationella organisationerna på enskilda länder mindre bindande. De har därför också en ganska indirekt påverkan på framväxten av en oberoende utvärderingskultur.

Som en indirekt effekt av den utvärderingsvåg som skapats av multinationella organisationer utlöstes en ny utvecklingsvåg, som blir mer framträdande. Denna fjärde våg kan beskrivas som en *emancipatorisk våg*, eftersom många länder i det globala syd driver på för en oberoende utveckling av utvärderingar baserade på deras egen kultur. I Afrika strävar den kontinentala utvärderingsorganisationen AfrEA efter att främja en "made in Africa"-strategi för utvärdering och kräver i sina utvärderingsprinciper som publicerades 2021 att "Evaluation in Africa has to be rooted in Africa. The African evaluation community has to contribute to evaluation theory and practice while drawing from experiences and initiatives in other parts of the world." (AfrEA 2021, s. 4). En liknande diskussion om integrering av inhemsk kultur i utvärderingsstandarderna äger nu också rum i det latinamerikanska utvärderingssamhället och ReLac.

Krav på en "kulturellt responsiv utvärdering", som särskilt bör bygga på ursprungsbefolkningars kulturella kunskap, går i en liknande riktning. Denna rörelse är för närvarande särskilt stark i Nya Zeeland (Adams, 2018) och Kanada (Locklear et al, 2023), men är absolut inte begränsad till dessa länder. Det är dock fortfarande de västliga organisationerna som upphandlar utvärderingar och finansierar deras genomförande. Utvärderingar som beställs av nationella institutioner är fortfarande sällsynta undantag i det globala syd, även om de globala utvecklingsmålen SDG märkbart har ökat engagemanget för övervakning och utvärdering i ett antal ganska små länder och har fört fram intressanta egna sätt (D'Errico et al, 2020).

De få positiva exemplen från det globala syd som har utvecklat egen utvärderingspraxis är mindre inriktade på de typiska lokala, filosofiska, etiska, ideologiska eller kulturella särdragen i landet än på principerna för New Public Management och den administra-

tiva kulturen i landet. Det universella målet är modernisering och rationalisering av administrationen, vilket uppenbarligen är betydelsefullt oavsett andra faktorer i respektive land och representerar ett steg bort från ren inputkontroll (Stockmann & Meyer 2021, s. 517).

I många länder är det ACA ("General Accounting Office") som driver på utvecklingen av övervakning och utvärdering, ofta med budgetkontroll och effektivitetsvinster som huvudfokus (särskilt i Latinamerika och i asiatiska länder med mer planorienterad politik; Stockmann & Meyer, 2022, s. 518; Stockmann et al, 2021, s. 563ff.). Önskan om finansiell kontroll är de viktigaste motiven på nationell politisk nivå, som sedan kan stödjas av finansieringsåtgärder från internationella organisationer och stärkas i riktning mot organisatorisk utveckling. Utan en medborgarorienterad självbild i den offentliga förvaltningen och ett demokratiskt politiskt system kan dock utvärderingens inlärningsfunktion inte utvecklas och få motsvarande positiva effekter.

SAMMANFATTNING

Som beskrevs inledningsvis identifierade Evert Vedung fyra utvecklingsvågor för utvärdering, vilka har ställts mot fyra nyare vågor i denna rapport (figur 2). Dessa vågor har delvis fortsatt och förstärkt den utveckling som Vedung beskrev, men delvis har de neutraliserats eller utplånats av starka motrörelser.

| VÅG/TREND | UNGEFÄRLIG START | SOCIAL PÅVERKAN | PÅVERKAN PÅ UTVÄRDERINGSUTVECKLINGEN |
|-----------------------------|------------------|---|---|
| "Fake News" | 1990 | Avrationalisering och populistisk nationalism | Än så länge fortfarande liten |
| EU:s sammanhållningspolitik | 1986 | Inkludering av nya stater i EU | Spridning av utvärdering i Europa |
| Utvecklingspolitik | 1990 | Främjande av utvärderingskompetens | Stöd till nationell utveckling av utvärderingar |
| Emancipation | 2010 | Framväxten av nya utvärderingskulturer | Än så länge fortfarande låg |

Figur 2. De fyra nya vågorna och deras betydelse

Det mest motsägelsefulla är utan tvekan uppkomsten av högerpopulistiska partier och "fake news" i förhållande till den "evidensrörelse" som Evert Vedung noterade. Högerpopulistiska partier har funnits länge, och början på den nya rörelsen kan dateras till uppkomsten av Vlaams Blok i Belgien och Front National i Frankrike i början av 1980-talet. Betydande framgångar för dessa två partier, liksom för det post-högerinriktade FPÖ i Österrike och SVP i Schweiz, Dansk Folkeparti i Danmark, FIDESZ-partiet i Ungern och några andra partier, kan bara observeras från mitten av 1990-talet och sedan särskilt under det första decenniet av det nya millenniet. Antalet anhängare av högerpopulism har ökat stadigt sedan dess, liksom deras politiska betydelse. Men det har knappast fått något inflytande på utvärderingsutvecklingen.

Den andra av de nya vågor vi identifierat har däremot haft stor inverkan på utvecklingen. EU:s sammanhållningspolitik och kraven på utvärdering av de omfattande finansieringsprogram som följde av den, var en viktig impuls för utvecklingen av en utvärderingskultur i sydeuropeiska medlemsländer och Irland på 1990-talet. Liknande impulser följde på 2004 års EU-utvidgning i Central- och Östeuropa. För anslutningskandidaterna från de tidigare östblocksländerna var detta inte bara en viktig impuls för utvecklingen av deras egen infrastruktur, utan också startsignalen för utvecklingen av utvärderingen.

Evert Vedung nämner inte heller utvärdering inom ramen för utvecklingspolitiken. Där blev utvecklingspolitiska åtgärder från och med 1990-talet alltmer inriktade på effekter. Särskilt sedan millennieskiftet är genomslaget tydligt och utvecklingsländer erbjuds stöd för uppbyggnad av utvärderingskapacitet och institutionsuppbyggnad på nationell nivå. I vissa länder, till exempel Sydafrika, byggs nationella övervaknings- och utvärderingssystem (NMES) upp. Åtgärderna bidrar till utvecklingen av utvärdering i det globala syd och är ett skäl till den världsomspännande spridning som beskrivs i början.

Visst finns det kopplingar till den utveckling som Evert Vedung beskriver, särskilt när det gäller den dialogorienterade vägen. Här har FN:s Agenda 2030 gett nya impulser och med den vägledande

principen "Leaving No One Behind" satt inkludering på agendan, inte bara i samband med utvärderingar, utan som en övergripande princip för handling. Delaktighet har således blivit en central del av utvärderingsmetoderna, även om de förfaranden och metoder som används för att genomföra den fortfarande är ganska blygsamma.

En annan "bieffekt" av främjandet av utvärdering i det globala syd och intressenternas deltagande i denna leder för närvarande till en ny frigörande utvärderingsvåg som försöker ta upp inhemska kulturella element och skilja sig från västerländskt dominerade förfaranden. Denna intressanta utveckling är dock begränsad till vissa nischer och har knappast nått den globala utvärderingsdiskussionen. En orsak kan vara att utvärderingssamfundet märkbart tar avstånd från vetenskaplig forskning och den akademiska världen. Globalt sett fortsätter antalet lärosäten som erbjuder utbildning att öka, men forskningen om utvärdering är fortfarande betydligt underutvecklad och utvärderingsvärlden ifrågasätter alltmer behovet av universitetsanknytning. De starka akademiska band som Evert Vedung noterade har påtagligt försvagats.

Nyliberalismen som politisk drivkraft har också försvunnit under tiden, men New Public Management och idealen om rationell kontroll av administrativa åtgärder har under tiden utvecklat ett anmärkningsvärt eget momentum och kräver inte längre en politisk-ideologisk underbyggnad. I ett globalt perspektiv är ansträngningarna att modernisera förvaltningen, både genom finansiell kontroll och genom planerings- och styrorgan, de viktigaste krafterna i utvecklingen av utvärderingen. Denna våg tycks dominera alla andra.

Referenser

- Adams, D. (2018). An experience-based perspective on the relationship between indigenous and Western epistemic systems in research. *Evaluation Matters—He Take Tō Te Aromatawai* 4: 2018 (<https://doi.org/10.18296/em.0028>).
- Adorno, T.W. et al. (1976). *Der Positivismusstreit in der Deutschen Soziologie*. Frankfurt: Luchterhand.
- African Evaluation Association (2021). *The African Evaluation Principles. 2021 Version* (<https://afrea.org/AEP/new/The-African-Evaluation-Principles.pdf>).
- Alkin, M.C. & Christie, C.A. (2023). *Evaluation Roots. Theory Influencing Practice*. New York: Guilford Press (3.).
- Bachtler, J., & Wren, C. (2006). Evaluation of European Union Cohesion policy: Research questions and policy challenges. *Regional studies*, 40(02): 143-153.
- Berger, J. (1996). Was behauptet die Modernisierungstheorie wirklich und was wird ihr bloß unterstellt? *Leviathan*, 24(1): 45-62.
- Binnendijk, A. (2000). *Results Based Management in the Development Co-Operation Agencies: A Review of Experience. Background Report*. Paris: OECD (<https://www.oecd.org/development/evaluation/dcdndep/31950852.pdf>; retrieved 16-08-23).
- Blomeyer, R. & Eussner, A. (2020). European Union. In Stockmann, R., Meyer, W. & Taube, L. (Eds.), *The Institutionalization of Evaluation in Europe* (pp.463-479). Cham: Palgrave Macmillan.
- Brugha, R. & Varvasovszky, Z. (2000). Stakeholder Analysis: a review. *Health Policy and Planning* 15(3): 239-246.
- Csehi, R. & Zgut, E. (2021). 'We won't let Brussels dictate us': Eurosceptic populism in Hungary and Poland. *European Politics and Society* 22(1): 53-68 (<https://doi.org/10.1080/23745118.2020.1717064>).
- D'Errico, S, Geoghegan, T and Piergallini, I (2020). *Evaluation to connect national priorities with the SDGs. A guide for evaluation commissioners and managers*. IIED, London. <http://pubs.iied.org/17739IIED>
- Furubo, J. E., Rist, R. C. & Sandahl, R. (2002). *International atlas of evaluation*. Piscataway: Transaction Publishers.
- Gantt, H.L. (2017 [1910]). *Work, Wages, and Profits: Their Influence on the Cost of Living*. London: Forgotten Books.
- Guba, E.G. & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury et al: Sage.
- Hay, C., Lister, M. & Marsh, D. (2006). *The state: theories and issues*. Basingstoke: Palgrave.
- Heinisch, R., Massetti, E. & Mazzoleni, O. (eds, 2020). *The People and the Nation. Populism and Ethno-Territorial Politics in Europe*. London: Routledge.
- Helm, R.K. & Nasu, H. (2021). Regulatory Response to 'Fake News' and Freedom of Expression: Normative and Empirical Evaluation. *Human Rights Law Review*, 2021, 21, 302–328 (doi: 10.1093/hrlr/ngaa060).
- Jacob, S., Speer, S, & Furubo, J. E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation*, 21(1): 6-31.
- Kelle, U. (2008). Qualitative vs quantitative Forschung - Die Debatte. In *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte* (pp. 25-56). Wiesbaden: VS Verlag (2.).
- Lane, J.E. (2000). *New Public Management. An Introduction*, London: Routledge.

- Locklear, S., Hesketh, M., Begay, N., Brixey, J., Echo-Hawk, A. & James, R. (2023). Reclaiming Our Narratives: An Indigenous Evaluation Framework for Urban American Indian/ Alaska Native Communities. In: *Canadian Journal for Program Evaluation* 38(1): 8-26 (doi: 10.3138/cjpe.75518).
- Luhmann, N. (1988). *Soziale Systeme: Grundriss einer allgemeinen Theorie*. Berlin: Suhrkamp.
- Meyer, W. (2023). The Influence of Transnational Organizations: A comparison of effects in America, Asia and Europe. In Stockmann, R., Meyer, W. & Zierke, N. (Eds.), *The institutionalization of Evaluation in Asia-Pacific* (pp. 429-463). Cham: Palgrave Macmillan.
- Meyer, W. (2015). Professionalisierung von Evaluation. Ein globaler Blick. In *Zeitschrift für Evaluation* 14 (2): 215-246.
- Meyer, W., Stockmann, R. & Szentmarjay, L. (2022). The Institutionalisation of Evaluation: Theoretical Background, Analytical Concept and Methods. In dies. (Eds.), *The Institutionalisation of Evaluation in the Americas* (pp. 3-37). Basingstoke: Palgrave.
- Meyer, W., Stockmann R. & Taube L. (2020). The Institutionalisation of Evaluation Theoretical Background, Analytical Concept and Methods. In dies. (Eds.), *The Institutionalisation of Evaluation in Europe* (pp. 4-34). Basingstoke: Palgrave.
- Neo, R. (2021). The International Discourses and Governance of Fake News. *Global Policy* 12(2): 214-228.
- Rosenstein, B. (2013). *Mapping the status of national evaluation policies. Parliamentarians Forum on Development Evaluation in South Asia and EvalPartners*. <http://www.pfde.net/index.php/publications-resources/global-mapping-report-2015>
- Rydgren, J. & Meiden, S. van der (2016). *Sweden, now a country like all the others? The radical right and the end of Swedish exceptionalism*. Stockholm: Department of Sociology (Working Paper Series No. 25).
- Schildt, A. & Schmidt, W. (Eds.) (2019). »Wir wollen mehr Demokratie wagen«: *Antriebskräfte, Realität und Mythos eines Versprechens*. Berlin: Dietz.
- Schimank, U. (1996). *Theorien gesellschaftlicher Differenzierung*. Wiesbaden: Springer-Verlag.
- Stern, E. (2009). Evaluation policy in the European Union and its institutions. In Trochim, W.M.K., Mark, M. M. & Cooksby L. J. (Eds.), *Evaluation policy and evaluation practice. New Directions for Evaluation*, 123, 67–85.
- Stockemer, D. (Ed.) (2019). *Populism around the World. A comparative perspective*. Cham: Springer.
- Stockmann, R., Meyer, W. & Stockmann, T. (Eds.) (2024). *The institutionalization of Evaluation in Africa*. Cham: Palgrave Macmillan (in preparation).
- Stockmann, R., Meyer, W. & Zierke, N. (Eds.) (2023). *The institutionalization of Evaluation in Asia-Pacific*. Cham: Palgrave Macmillan.
- Stockmann, R., Meyer, W. & Szentmarjay, L. (Eds.) (2022). *The Institutionalization of Evaluation in the Americas: Evaluation Globe Vol. II*. Cham: Palgrave Macmillan.
- Stockmann, R., Meyer, W. & Taube L. (Eds.) (2020). *The Institutionalization of Evaluation in Europe*. Cham: Palgrave Macmillan.
- Stockmann, R. & Meyer, W. (2022). The Institutionalisation of Evaluation in Europe and the Americas: A Comparison. In *The Institutionalization of Evaluation in the Americas* (pp. 509-525), Cham: Palgrave Macmillan.
- Stockmann R. & Meyer W. (Eds.) (2016). *The Future of Evaluation. Global Trends, New Challenges, Shared Perspectives*. Basingstoke: Palgrave Macmillan.

- Tambini, D. (2017). *Fake News: Public Policy Responses. Media Policy Brief 20*. London: Media Policy Project, London School of Economics and Political Science.
- Tandoc, E.C. (2019). The facts of fake news: A research review. *Sociology Compass*. 2019;13:e12724 (<https://doi.org/10.1111/soc4.12724>).
- Taylor, F.W. (2020 [1911]). *The Principles of Scientific Management, River Denys*. Nova Scotia: Ocean of Minds Media House.
- Torre, C. de la (Ed.) (2020). *Routledge Handbook of Global Populism*. London: Routledge.
- Vedung, E. (2010). Four Ways of Evaluation Diffusion. *Evaluation* 16 (3): 263-277.
- Vetter, A. & Remer-Bollow, U. (2017). *Bürger und Beteiligung in der Demokratie*. Wiesbaden: Springer VS.
- Widmer, T., Beywl, W. & Fabian, C. (Eds.) (2009). *Evaluation: ein systematisches Handbuch*. Wiesbaden: VS Verlag.
- Zapf, W. (1996). Die Modernisierungstheorie und unterschiedliche Pfade der gesellschaftlichen Entwicklung. *Leviathan*, 24(1): 63-77.

KAPITEL 12

Att använda eller inte använda – det är frågan

Per Mickwitz

INLEDNING

Användning är ett av utvärderingens och utvärderingsforskningens fundament. Detta har Evert Vedung varit väl medveten om länge, även om han har klassificerats som en utvärderingsforskare inom kategorin metoder (Carden & Alkin, 2012). Att Evert sett användning som centralt bekräftas av hans definition av utvärdering som avslutas med orden *"vilken tänkes spela en roll i praktiska besluts-situationer"* (Vedung, 2009).

Även om Evert varit väl medveten om hur viktigt det är att utvärderingar används är det intressant att detta tema inte alls berörs i vår gemensamma publikation – Hildén et al. (2002). Eftersom jag under årens lopp funderat mycket på användningen av både utvärderingar och forskning, tycker jag att det är ett bra tema att behandla i denna skrift.

I detta kapitel kommer jag att presentera fyra fallstudier av användning. Jag har själv varit involverad i dessa, men min roll har varierat. I tre fall har jag haft en stark koppling till utvärderings-teamet, medan jag i ett fall blev en av de viktigaste användarna.

Inte helt överraskande kommer jag att starkt betona den specifika kontexten som möjliggör eller hindrar användningen av utvärderingar. Jag kommer att reflektera över den politiska dimensionen och se på hur gyllene tillfällen ("window of opportunity") utnyttjas eller missas.

STORA FÖLJDER AV ATT INTE ANVÄNDA EN UTVÄRDERING

Under de första åren på 2000-talet utredde man i Finland politiska åtgärder för hanteringen av hushållsavloppsvatten i områden utanför avloppsneten. Som en del av detta arbete tillsattes ett projekt för att bedöma de ekonomiska effekterna av ett förslag till en förordning som tjänstemän hade tagit fram. Konsekvensbedömningen, som på basen av Evert Vedungs definition inte skulle kallas för en utvärdering eftersom den gjordes ex ante, byggde på en analys av olika scenariers kostnadseffektivitet. Jag var ansvarig för projektet som utfördes av Marjukka Hiltunen (2003).

Konsekvensbedömningen kom fram till att förordningen skulle reducera vattenbelastningen, men att kostnaderna skulle vara höga, speciellt i förhållande till andra möjliga åtgärder för att minska utsläppen. De administrativa kostnaderna för att implementera förordningen skulle dessutom vara betydande. Slutligen gjorde Hiltunen också en analys av hur kostnaderna skulle fördelas, som visade att en stor del av kostnaderna skulle uppstå i glesbygden och att en betydande del av kostnaderna skulle belasta de fattigaste hushållen och ofta pensionärer. (Hiltunen, 2003)

Den gjorda konsekvensbedömningens kvalitet har senare bedömts som mycket hög och den har använts som ett exempel i bland annat rapporten som pilotprojektet för stöd till bedömningen av lagstiftningens verkningar publicerat (VNK, 2005). Processen kan också bedömas som god, speciellt hur den integrerade forskare från olika discipliner och från flera forskningsinstitut med tjänstemän i syftet att ta fram ett relevant och mångsidigt underlag. Det var bara ett problem. Konsekvensbedömningen användes inte.

Trots de kritiska resultat som presenterades i konsekvensbedömningen (Hiltunen, 2003) godkände regeringen förordningen "om

behandling av hushållsavloppsvatten i områden utanför vattenverkens avloppsnät” den 11 juni 2003 (542/2003).

I Finland har ingen annan enskild förordning fått så mycket uppmärksamhet i media och haft en så stor politisk betydelse som förordningen om ”behandling av hushållsavloppsvatten i områden utanför vattenverkens avloppsnät”. Bara under 2010 publicerade Finlands största tidning Helsingin Sanomat 79 artiklar om förordningen, varav tolv ledare. Förordningen har flera gånger förändrats och tidsfristerna för implementeringen har skjutits upp. Jan-Erik Enestam, som var miljöminister när beslutet om förordningen togs, har i en intervju sagt att han fick felaktiga uppgifter av en tjänsteman (Hufvudstadsbladet 26.8.2010). Enligt många bedömare var förordningen en viktig orsak till att partiet Sannfinländarna hade så stora framgångar i riksdagsvalen 2011 (till exempel ledare i Helsingin Sanomat 6.7.2019).

Varför användes då inte konsekvensbedömningen när man fattade beslut om förordningen? Jag kommer att kort sammanfatta min personliga uppfattning, som inte bygger på någon forskning, men på det jag såg, var med om och på diskussioner med flera, men inte alla inblandade. Min uppfattning är att beredningen av förordningen skedde i en mycket snäv krets, som bestod av en liten grupp tjänstemän. Dessa tjänstemän var kunniga i vattenkvalitetsfrågor och frågor kopplade till tekniken för vattenrening. Men en bredare syn på styrning, legitimitet och ett bredare miljöpolitiskt tänkande saknades. Konsekvensbedömningen kom också ur dessa tjänstemäns perspektiv för sent, de hade redan bestämt vad de tyckte borde göras.

Vem var då ansvarig för att utvärderingen inte användes? Först kan man konstatera att den unga forskaren Marjukka Hiltunen, som gjorde analysen, absolut inte var ansvarig. Hon gjorde ett utmärkt arbete. Man kan konstatera att ett stort ansvar vilar på tjänstemännen som beredde frågan och politikerna som fattade beslutet. Men att bara hålla dem ansvariga är lite för enkelt. Jag anser att jag själv som ansvarig för projektet och mera erfaren än Hiltunen borde ha gått ut i offentligheten med resultaten innan beslutet om förordningen fattades. När jag har presenterat denna

reflektion har flera andra personer också meddelat att de borde ha gjort mera för att föra fram resultaten i offentligheten. I denna krets av självkritiska personer ingår både min dåvarande chef och en av cheferna till de tjänstemän som beredde ärendet på det finska miljöministeriet.

När förordningen ”om behandling av hushållsavloppsvatten i områden utanför vattenverkens avloppsnät” blivit en politisk skandal fick detta en intressant sidoeffekt. Under flera år ökade intresset för utvärdering och konsekvensanalys samt deras användning markant på miljöministeriet i Finland. Jag var själv flera gånger inbjuden att föreläsa om dessa frågor. Tyvärr misstänker jag att intresset har svalnat igen, det institutionella minnet är ofta kort.

SMÅ DETALJER KAN VARA AVGÖRANDE FÖR ANVÄNDNING

Nu går jag över till att diskutera två exempel på utvärderingar på europeisk nivå. Dessa är kopplade till den Europeiska miljöbyrån (EEA) med huvudkontor i Köpenhamn. Jag har själv varit involverad i EEAs arbete i olika roller, men speciellt som medlem och ordförande för dess vetenskapliga kommitté (2014-2022). De två utvärderingar vars användningar jag kort kommer att diskutera är dels, den utvärdering av den Europeiska miljöbyrån som publicerades 2018 (SWD 2018) dels, den rapport ”The European environment — state and outlook 2020: Knowledge for transition to a sustainable Europe” som EEA publicerade 2019.

Utvärderingen av Europeiska miljöbyrån var den tredje utvärderingen av organisationen. Den initierades av EU-kommissionen och den omfattade perioden 2012 till 2016. Utvärderingen gjordes i enlighet med agendan för bättre lagstiftning, som var ett centralt politikområde under kommissionen som leddes av Jean-Claude Juncker. Kriterierna i utvärderingen var, i enlighet med agendan för bättre lagstiftning: ändamålsenlighet, effektivitet, relevans, samstämmighet och EU-mervärde. Utvärderingen startade 2016 och den slutliga rapporten publicerades i november 2018.

Jag hade själv många roller i utvärderingsprocessen. Jag blev intervjuad av konsulterna som gjorde bakgrundsstudien. Jag var

medlem i styrgruppen för utvärderingen och jag deltog i workshopen där de preliminära slutsatserna diskuterades. Jag var också med i de arbetsgrupper som skrev både den vetenskapliga kommitténs respons och responsen från den Europeiska miljöbyråns styrelse. Mina iakttagelser bygger på observationer jag gjort under dessa processer och sådant man kan läsa i de offentliga handlingarna.

Utvärderingen initierades och satte i gång under den kommission som leddes av Jean-Claude Juncker. Junker-kommissionen hade väldigt låga miljöpolitiska ambitioner, ordet ”miljö” nämns över huvudtaget inte i kommissionens program. Samtidigt var detta en period när Europeiska miljöbyråns resurser skars ner. Under 2016 och 2017 var de implicita och ibland explicita förväntningarna på den kommande användningen av utvärderingen så som de ansvariga lägre tjänstemännen vid EU-kommissionen uttryckte det, att fokusera och prioritera arbetet vid kommande nedskärningar.

Utvärderingen blev färdig i slutet av 2018 och hade då redan diskuterats på hög nivå i kommissionen. Våren 2019 var det val till Europaparlamentet och på sommaren utsågs Ursula von der Leyen till ordförande i Europeiska kommissionen, hon tillträdde den första december. Den nya kommissionens miljöpolitik är mycket ambitiös, vilket syns i att ”EU:s gröna giv” listas först bland kommissionens prioriteringar. Utvärderingen av Europeiska miljöbyrån användes dels som underlag för byråns nya ambitiösa strategi ”EEA-Eionets strategi 2021–2030” (EEA, 2021), dels för att prioritera nya tjänster och satsningar efter att miljöbyråns budget börjat öka från och med hösten 2019.

Enligt förordningen har Europeiska miljöbyrån som uppgift ”att vart femte år offentliggöra en rapport om miljötillståndet och om utvecklingstendenserna och framtidsutsikterna för miljön”. Dessa rapporter (som kallas SOER, från State and Outlook of the Environment Report) är dels en bedömning av tillståndet i miljön, dels en utvärdering av den förda politiken och slutligen en bedömning av vilken politik som skulle krävas för att nå de framtida mål som uppställts. I mars 2015 presenterades rapporten

”Europas miljö – tillstånd och utblick 2015” i Bryssel. Ett halvt år senare inleddes planeringen av nästa rapport som planerades för lansering under våren 2020 – förordningen stipulerar ju att rapporten ska publiceras vart femte år.

Även om det mesta i planeringsprocessen inför 2020 års rapport om Europas miljö tillstånd och utblick handlade om innehållet diskuterades också tidpunkten för lanseringen redan 2017. Inom den Europeiska miljöbyrån gjorde man en koppling mellan processen att ta fram och lansera SOER 2020 rapporten och den politiska processen inom EU. Man insåg att det kommer att vara val till Europaparlamentet i maj 2019. Detta innebär att en ny ordförande och en ny kommission kommer att utses under sommaren eller hösten 2019. Den nya kommissionen kommer därefter att ta fram och presentera sitt program som kommer att styra det politiska arbetet inom EU under åren 2020-2024. Om SOER 2020 rapporten skulle kunna ha en chans att användas vid formandet av den politiska agendan skulle det således vara alldeles för sent att lansera den i mars 2020. Man fattade alltså beslut om att tidigarelägga publiceringen till december 2019. Detta beslut måste tas i god tid eftersom rapportens innehåll måste vara färdigt ungefär ett halvt år före publiceringen för att först godkännas, och för att rapporten sedan ska hinna språkgranskas, layoutas och tryckas.

Efter valet till Europaparlamentet i maj följde intensiva förhandlingar som resulterade i att Ursula Von der Leyen den 3 juli 2019 valdes till kandidat för att bli ordförande för den nya EU-kommissionen. Hon presenterade sitt program för Europaparlamentet den 16 juli. Programmets titel är ”En ambitiösare union: Min agenda för Europa” och det första kapitlet börjar med meningen ”Jag vill att Europa ska bli ambitiösare genom att bli den första klimatneutrala världsdelen.” Hon valdes till ny kommissionsordförande av en majoritet i Europaparlamentet den 16 juli 2019. Den nya kommissionen tillträdde formellt den 1 december och presenterade färdplanen för den gröna given den 11 december 2019. Dessa detaljer var inte kända när beslutet om att tidigarelägga publicerandet av SOER 2020 togs, men en lansering i Bryssel den 4 december 2019 visade sig passa perfekt med den politiska tidtabellen.

Också innehållsmässigt fanns det en bra koppling mellan fokuset i SOER 2020 rapporten och Von der Leyen kommissionens politiska prioriteringar. Om man kort vill sammanfatta innehållet i den 499 sidor långa SOER 2020 rapporten är slutsatsen att traditionell miljöpolitik har nått visa resultat och så väl vatten- som luftutsläppen har minskat. Samtidigt kommer inte nuvarande politik att leda till att målen för 2050 uppnås. "Det är fortfarande möjligt att uppnå EU:s vision för hållbarhet till 2050, men det krävs att åtgärdernas karaktär och ambition förändras." (EEA, 2019b) Därmed är det inte så konstigt att SOER 2020 är en av fyra källor som hänvisas till i det första stycket i färdplanen för den gröna given (COM(2019) 640).

Att tidigarelägga publiceringen var inte den enda åtgärden som gjordes för att öka möjligheterna för att SOER 2020 skulle kunna användas. En viktig del av processen är att både EEAs medlemsländer och EU-kommissionen får en möjlighet att faktagranska rapporten innan den godkänns. En möjlighet som också EEAs vetenskapliga kommitté har. Denna process höjer kvaliteten, genom att många små fel upptäcks och korrigeras, men det ökar också förutsättningarna för användning. Redan före publiceringen presenterades rapportens huvudresultat för viktiga tjänstemän vid EU-kommissionen och olika implikationer diskuterades. Ett exempel på detta är att EEA, dess vetenskapliga kommitté och ledningen för avdelningen för forskningsfinansiering vid EU-kommissionen ordnade en workshop om vilka implikationer rapporten har för kommande kunskapsbehov och forskning.

Både utvärderingen av EEA och SOER 2020 utvärderingen av den europeiska miljöns och miljöpolitikens tillstånd visar hur viktigt det är när en utvärdering genomförs och publiceras i förhållande till politiska processer. Utvärderingen av EEA kom för sent för de som ville minska den europeiska miljöpolitikens ambitioner. Men den kom i rätt tid för dem som ville ta fram en ambitiös strategi för myndigheten. SOER 2020 rapporten publicerades i precis rätt tid i förhållande till Von der Leyen kommissionens politiska program. Just nu funderar EEA på när SOER 2025 ska publiceras, man kommer knappast att lyckas riktigt lika bra som vid lanseringen 2019.

ATT SOM BESLUTSFATTARE FÅ ANVÄNDA EN UTVÄRDERING

Den sista fallstudien av användning av en utvärdering gäller utvärderingen av forskningskvaliteten vid Lunds universitet som genomfördes 2019-20, den så kallade RQ20 (från titeln Lund University's Research Quality Evaluation 2020). I processen med RQ20 har jag haft många roller. Jag har lett en enhet som utvärderades, jag har ansvarat för en självutvärdering och jag har haft ett stort ansvar för hur utvärderingen ska användas. Jag kommer snart att återkomma till detta, men först kort om hur utvärderingen gjordes och vad den kom fram till.

Forskningskvalitetsutvärderingen RQ20 genomfördes i tre faser. I den första fasen gjorde 161 forskningsmiljöer (ibland institutioner, men ibland specialskapade enheter för RQ20) självutvärderingar som rapporterades enligt ett enhetligt, men ganska fritt format. I den andra fasen gjorde externa paneler en bedömning dels på basen av självutvärderingsrapporterna, dels på basen av intervjuer. Bedömningarna gjordes av 32 områdespaneler och fem tematiska paneler. Panelerna bestod främst av internationellt erkända forskare inom de olika områdena, men också vissa svenskar deltog. I den tredje och sista fasen sammanfattades materialet och slutsatser och rekommendationer togs fram.

Utvärderingen RQ20 konstaterade att Lunds universitet är ett stort och brett universitet som bedriver forskning som till en mycket stor del håller hög internationell kvalitet. Samtidigt innehåller RQ20 en rad rekommendationer till åtgärder som Lunds universitet bör vidta för att öka forskningens kvalitet. Dessa finns dels i utvärderingarna av de olika forskningsmiljöerna men de finns också sammanfattade i tio budord och 95 teser.

I november 2020 fick jag mycket överraskande frågan om jag var intresserad att bli vice-rector med ansvar för forskning vid Lunds universitet. Jag utnämndes och började arbetet som vicerektor i januari 2021. En av mina främsta uppgifter var att ansvara för processen för hur insikterna från RQ20 skulle tas till vara och användas, speciellt på den universitetsgemensamma nivån. RQ20-rapporten och slutsatserna av utvärderingen presenterades den 26 mars 2021.

Efter en intensiv diskussion och skriftliga inspel från fakulteterna, de centrala nämnderna och råden, samt en diskussion i universitetsstyrelsen fattade rektor den 24 juni 2021 beslut om fem prioriterade områden för det gemensamma arbetet med rekommendationerna i RQ20.

De fem prioriterade områdena på universitetsgemensam nivå på basen av RQ20 var:

- Rekrytering och karriärvägar
- Forskningsinfrastruktur
- Tvärvetenskaplig forskning
- Profilerings och starka forskningsmiljöer
- Uppföljning av hur insikterna i RQ20 tas tillvara och vilka åtgärder som genomförs.

Under 2022 och 2023 har universitetet gemensamt men också fakulteterna och institutionerna arbetat med dessa prioriterade frågor och med de mera specifika rekommendationerna. Arbetet har bland annat resulterat i en process för profilmråden som ledde till fem profilmråden för 2022-30. Dessa erhåll 4 miljoner kronor var för 2023-24. Implementeringen av rekommendationerna i RQ20 resulterade också i en ny handlingsplan för forskningsinfrastruktur med en årlig budget på 40 miljoner kronor.

Under 2022 genomfördes två strategiska processer vid Lunds universitet för vilka RQ20 utgjorde ett viktigt underlag. Dessa processer var framtagandet av den strategiska plattformen 2023-2024 och forskningsstrategin 2023-2026.

Forskningskvalitetsutvärderingen RQ20 initierades och genomfördes när Torbjörn von Schanz var rektor för Lunds universitet, men resultaten presenterades när Erik Renström just hade tillträtt som rektor. Detta möjliggjorde det för oss i den nya ledningen att inte bara ta hand om löpande frågor, utan att initiera en diskussion och genomföra åtgärder om de frågor som en gedigen utvärdering lyft som de allra viktigaste för att höja forskningens kvalitet. Efter som största delen av forskarna på Lunds universitet på något sätt

varit engagerade i utvärderingen, fans det intresse för att se vad man kommit fram till och hur detta skulle användas. Nu några år senare har detta intresse svalnat och det är snarare framgångar i utlysningar, nationella utvärderingar och indikatorer som mobiliserar diskussioner om behovet av utvecklingsarbete.

SLUTSATSER

På basen av de fyra fallstudierna av utvärderingar som antingen har använts eller inte har använts är jag beredd att dra vissa slutsatser.

Slutsats 1: Att potentiella användare är engagerade i utvärderingsprocessen främjar användningen av utvärderingen, men garanterar ingen användning. Detta är tydligt i både de två exemplen kopplade till den Europeiska miljöbyrån EEA och till RQ20 vid Lunds universitet. Samtidigt deltog de centrala tjänstemännen vid finska miljöministeriet i styrgruppen för projektet som gjorde konsekvensbedömningen av förordningen för hanteringen av hushållsavloppsvatten i områden utanför avloppsnäten. Trots att tjänstemännen var involverade valde de och vatteningenjörerna som var experter att inte använda resultaten.

Slutsats 2: Tidpunkten för när en utvärdering görs och när den publiceras är avgörande för om den kommer att användas. Detta är allra tydligast i exemplet med SOER 2020 som färdigställdes och publicerades just i tid för att Von der Leyen kommissionen skulle kunna använda den i sin färdplan för den gröna given. Det samma gäller RQ20 som kom precis i rätt tid för rektor Erik Renström och hans universitetsledning som tillträdde 2021. Samtidigt som exemplet med den finska konsekvensbedömningen av förordningen för hanteringen av hushållsavloppsvatten också kom före det formella beslutet togs, så gjordes den efter att de ansvariga tjänstemännen hade bestämt sig om hur förordningen skulle utformas.

Slutsats 3: Användningen främjas om potentiella användare har tillgång till resultaten innan de blir offentliga. Detta gör det möjligt för dem att sätta sig in i resultaten och planera för användning innan en offentlig debatt startar. I samtliga fyra fall som diskuterat har potentiella användare haft tillgång till resultaten

innan de har publicerats. Även om användning gynnas av att potentiella användare har tillgång till resultaten innan de publiceras är detta inte oproblematiskt ur ett rättvise- och demokrati-perspektiv.

Slutsats 4: Användningen av utvärderingen bör ske när utvärderingen är färsk. Detta är ganska paradoxalt och gäller även om innehållet av en ny utvärdering skulle vara detsamma under lång tid. Många resultat i SOER 2015 var fortfarande giltiga 2019, men deras politiska relevans var obefintlig. Rekrytering, forskningsinfrastruktur och tvärvetenskap hör fortfarande till de viktigaste frågorna för att höja forskningens kvalitet på Lunds universitet, men många tycker att det var länge sedan RQ20 genomfördes. De återkommande IPCC rapporterna är det kändaste exemplet på att utvärderingar är färskvara. Det som är nytt i den senaste IPCC utvärderingen är ganska lite jämfört med vad som redan var känt sex eller till och med tolv år tidigare, ändå får en ny rapport enorm publicitet och lyckas – åtminstone tillfälligt – mobilisera politiken.

Slutsats 5: Rollen för lägre tjänstemän och deras politiska tolkningar kan vara avgörande för användningen av utvärderingar. När Europeiska miljöbyrån skulle utvärderas var de tolkningar som tjänstemännen vid EU-kommissionen gjorde av Junker-kommissionens prioriteringar centrala för att diskussionerna främst gällde hur nedskärningar skulle riktas. Det var de ansvariga tjänstemännen vid finska miljöministeriet som valde att inte beakta konsekvensbedömningen när de beredde den slutliga förordningen för hanteringen av hushållsavloppsvatten. Min egen tolkning av Lunds universitets styrelses och rektor Erik Renströms ambition att arbeta för att förbättra förutsättningarna för forskning av världsklass vid Lunds universitet var avgörande för mitt arbete med att främja användningen av RQ20.

Även om det är angeläget att utvärderingar används och det är viktigt att man arbetar för att främja förutsättningarna för användning håller jag helt med Evert Vedung (2009) när han skriver: "Användning är centralt i utvärdering, men får inte över-skugga krav på kvalitet och respekt för fakta." I de fall som här

har studerats har de åtgärder som vidtagits för att förbättra möjligheterna att utvärderingarna ska användas inte påverkat kvaliteten. Den som utsattes för störst påtryckningar att tumma på fakta var nog Marjukka Hiltunen, men hon stod på sig. Historien gav henne sedan i stort sett rätt. När det gäller europeisk miljöpolitik och Lunds universitets forskning tummade inte utvärderarna på kvaliteten för att främja användning. Tiden får utvisa om utvärderingarna bidrog till bättre miljö eller forskning, men utan användning skulle de åtminstone inte ha gjort det.

Referenser

Benner, M., Bredenberg, M., & Ståhlberg, F. (Red.) (2021). *RQ20 - Lund University's research quality evaluation 2020*. Lund University.

https://lucris.lub.lu.se/ws/portalfiles/portal/101791672/210830_RQ20_Lund_University_s_Research_Quality_Evaluation_2020.pdf

Carden F. & Alkin M. (2012). Evaluation Roots: An International Perspective, *Journal of MultiDisciplinary Evaluation*, 8(17) 102-18.

COM (2019) *Den europeiska gröna given*. Meddelande från Kommissionen. Bryssel den 11.12.2019.

EEA (2015). *Europas miljö – tillstånd och utblick 2015: en sammanfattning*. Köpenhamn: Europeiska Miljöbyrån.

EEA (2019a). *The European environment — state and outlook 2020: Knowledge for transition to a sustainable Europe*. Copenhagen: European Environment Agency. <https://www.eea.europa.eu/publications/soer-2020>

EEA (2019b). *Europas Miljö – Tillstånd och utblick 2020: Sammanfattning*. Köpenhamn: Europeiska Miljöbyrån.

EEA (2021). *EEA-Eionets strategi 2021–2030*. Köpenhamn: Europeiska miljöbyrån. <https://www.eea.europa.eu/sv/publications/eea-eionets-strategi-202120132030>

Hildén M., J. Lepola, P. Mickwitz, A. Mulders, M. Palosaari, J. Similä, S. Sjöblom, & E. Vedung. (2002). *Evaluation of environmental policy instruments – a case study of the Finnish pulp and paper and chemical industries*. Helsinki: Monographs of the Boreal Environment Research, 21

Hiltunen M. (2003). *Taloustevesien käsittely viemäriverkostojen ulkopuolisilla alueilla – asetusedotuksen taloudellisten vaikutusten arviointi* [Hantering av hushållsavloppsvatten på områden utanför avloppsnäten – utvärdering av författningsförslagens ekonomiska effekter], Helsinki: Suomen ympäristökeskuksen moniste 275.

SWD (2018). *Evaluation of the European Environment Agency (EEA) and its European Environment Information and Observation Network (EIONET)*.

COMMISSION STAFF WORKING DOCUMENT

Vedung E. (2009). *Utvärdering i Politik och Förvaltning* (3:e upplagan). Studentlitteratur, Lund.

VNK 2005. *Tehokkaampaa, suunnitelmallisempaa ja hallitumpaa lainvalmistelua: valtioneuvoston lainvalmistelun suunnittelun ja johtamisen kehittämissuunnitelman mietintö*. [Effektivare, planerad och kontrollerad lagstiftning: rapport från regeringens utvecklingsgrupp för planering och ledning av lagstiftning] Valtioneuvoston Kanslian julkaisusarja, 13/2005.

KAPITEL 13

Taking side-effects seriously

Nicoletta Stame

INTRODUCTION

When, in the 1990s, most evaluation textbooks were proposing a goal oriented approach that ignored side-effects, and social science was dominated by the rhetoric of perverse effects, Evert Vedung (1997, p. 54) proposed an “effects tree” that defied both stances. Feeling at unison with the tradition of “unanticipated (or unintended) consequences” (Merton, 1936) and of that of bounded rationality, he took into consideration the effects of the program in the target area (main effects) as well as in the area outside the target (side-effects), and in both cases distinguished between beneficial and detrimental ones.

Although Vedung was working in a policy-making at the state level framework, this is a legacy that we can benefit from still today, when global challenges and complex policies tackle with unintended consequences of programs all the time, which is perceived in a pessimistic mood of total disaster.

Vedung’s take on the matter is different from other attempts at utilizing the two dichotomies of intended/unintended or anticipated/unanticipated consequences. The two were used indifferently by Merton himself (Merton & Barber, 2004 p. 234).

Other authors, on the contrary, have distinguished them: the reason to do so was to overcome the negative meaning that had been attributed to un-intended consequences as detrimental, perverse or un-desired¹: either by admitting of something that could be unintended-detrimental but anticipated (“lesser evil”, de Zwart, 2015²) or unanticipated but intended - beneficial (“possibilism”, Stame 2022).³

Yet, in the traditional version of the conceptual framework revisited by Merton both unanticipated and unintended consequences were different from their respective opposites not because detrimental vs. beneficiary, but because “the result of *human action and not of human design*”⁴. And the understanding that such consequences were positive even if not planned as such was prevalent for unanticipated or unintended alike. In any case, Merton (1936) states that “unanticipated consequences of *purposive social action* (...) unorganized or formally organized” even if “undesired” are not always “undesirable”.

In tune with Merton, Vedung uses the two dichotomies as overlapping (top of the effects tree), and solves the problem of the merit of the side-effects by distinguishing between beneficial and detrimental effects that can be both anticipated and unanticipated, intended or un-intended. In fact, this is implied in the same “rationale for doing policy evaluation”, because “state actions to some extent are unpredictable and regularly result in side effects not originally foreseen” (1997, p. 57). In this way he overcomes the limits of the goal-attainment evaluation approach:

1) Hirschman (1998, p. 93) considers the slippage of words from “unintended consequences” to “perverse effect” as “a betrayal of the idea of unintended consequences because it cancels the open-endedness (the open-endedness to a variety of solutions) and substitutes it by total predictability and fear”. Vedung, in his turn, says that perverse effect does not belong to side-effects, and “ought to be treated as a separate category” (1997, p. 50).

2) The “lesser evil” was also considered by Merton himself (1936, p. 896).

3) In Stame (2022) I have discussed the difference between de Zwart’s and mine respective tables that cross the two dichotomies. In both cases, it was a way of mapping side-effects, not main effects (see below).

4) Vedung (1997, p. 53). The distinction between human design and human action is echoed in Hirschman’s idea of change as a result of “voluntaristic change” or of “unintended side effects” (set in motion by actions with other purposes. See Stame 2022, p. 93).

Were the evaluators to confine themselves exclusively to re-searching the achievement of premeditated program goals, any serendipitous results or un-anticipated side-effects would not be included in the main evaluation process. The evaluation would provide a tunnel vision of events, and produce a biased, if not fundamentally wrong, picture of what the program has attained. In all likelihood a program generating some interesting spin-off effects must be better than a program producing several undesirable spillovers” (1997, pp. 45-6. Italics added) (...) Why it is so vital to pay attention to side-effects? Because side-products, whether detrimental or beneficial, are crucial factors in every inclusive judgment of the operation of an intervention (1997, p. 56).

Vedung then illustrates his approach to “side-effects evaluation”, that culminates in figure 4.4. (1997, p. 54): “Main effects, null effects, perverse effects and side-effects”. This is an outstanding contribution, that has not received the attention it deserves. Vedung then defines the evaluator’s tasks as those of *mapping* side-effects and *ascertain their value* (1997, p. 58): unanticipated side-effects are difficult to conceive, and their value is not as obvious as is that of the main effects (which is the goal of the program). Here is where he invites us to reflect. A challenge that I would like to take, in the company of two distinguished experts in unintended consequences, Robert Merton and Albert Hirschman. I will follow Vedung along three steps: discovering side-effects, mapping effects (main and side), ascertaining the value of side effects.

DISCOVERING SIDE-EFFECTS

Before “mapping” side-effects there is the need of “discovering” them, be they the result of human action or of human design. This is what Vedung himself refers to when he speaks of “serendipitous results of unanticipated side-effects”.

As Merton (1948) has masterfully elaborated⁵ “serendipity” is a “pattern in scientific inquiry” that consists of the observation of

5) Reported in Merton and Barber 2004, p. 195; p. 236.

“an *unanticipated, anomalous and strategic* datum that becomes the occasion for developing a new theory or for extending an existing theory”⁶. Each word is relevant here. *Unanticipated* means that it must have been detected by a keen observer, going beyond the data constructed according to the assumptions of the theory to be tested (the goals indicators). *Anomalous* means that it is inconsistent with the prevailing theory, provokes curiosity and asks to make sense of it. *Strategic* means that it “must permit implications which can bear upon theory”⁷. And this refers to the observer’s ability, “for it obviously requires a theoretically sensitized observer to detect the universal in the particular” (Merton and Barber, 2004, p. 237).⁸

In a similar vein, the work of who first talked about “the centrality of side-effects”, Albert Hirschman (1967), had originated from a way of scientific inquiry based on the observation⁹ of the unexpected (presumed obstacles that may turn inoffensive; assumed conditions that may be absent), that could turn positive and useful for the improvement of planned actions or for a general aspiration of society¹⁰. This attitude, that Hirschman calls “possibilism”, helps enlarging the alternatives open to operators and social innovators, and is opposed to “probabilism”, implied in theories built on what is causally expected to happen because it has already happened elsewhere or in other times, with no

6) Merton says that “unintended consequences” was his main sociological interest in the article of 1936, and only in 1940 he came across the concept of “serendipity”, which became central also in his methodological (the logic of inquiry) and substantive (the sociology of science) interests, together with his interests in neologisms (Merton and Barber, 2004, p. 235f.)

7) This could be linked to Vedung’s idea that one of the purposes of evaluation, beyond accountability and intervention’s improvement, is basic knowledge (1997, p. 110), again referring to Merton.

8) Merton had previously said: “scientists have to be prepared to make accidental discoveries” (Merton and Barber, 2004, p. 223)

9) This attitude is expressed even in the title of Hirschman’s work which reads *Development Projects Observed*.

10) This general attitude Hirschman had assimilated in his junior partnership with Eugenio Colorni, who had taught him, among other things, the importance of doubt, the observation of small ideas that can grow, and the need for “understanding” as more fruitful than “explaining” (Colorni, 2019).

guarantee that it will be repeated. Possibilism's working devices, as Hirschman (1971, p. 29ff.) clarifies, are "inverted sequences", "blessing in disguise" and, yes, "unintended consequences".

MAPPING SIDE EFFECTS

What are side-effects? Vedung defines them as "all consequence, intended or unintended, anticipated or unanticipated, not included among the main effects; in public intervention evaluation, impacts occurring outside the area of the intervention; also called spillovers, spin-off effects, by-effects" (1997, p. 309). Side-effects imply a spatial metaphor. What is side to what? Where do they occur? Another sphere of action? The context? In the example provided on p. 54 about a program for reducing energy use in Sweden, they refer to financial (reducing costs), environmental (water conservation), organizational (industrial relations) aspects: in fact, according to Vedung "main effects" and the "area of intervention" can be quite broad. Today we would include climate change, social inequalities, etc.

As anticipated above, Albert Hirschman had stated "the centrality of side effects" of development projects, in opposition to Cost-Benefit-Analysis and similar exercises that are limited on the rate of return of investment as the main metric of success, that were molded on industrialized countries, and could not catch the problems of the developing ones¹¹. Hirschman's map of side-effects is wide-ranging: "from the acquisition of new skills to great readiness, on the part of the consumers of the project's output, to produce for the market; from the stimulation of entrepreneurship to the learning of cooperation and discipline; from backward and forward linkages to greater propensity to engage in family planning (...) not to forget negative effects such as new or heightened social and ethnic tensions, fresh opportunities for spreading corruption, etc." (1967, p. 148).

11) It was a way of distinguishing himself from economists of the industrialized world who thought that in a steady economy side effects could be controlled, and that in any case it would have been possible to elaborate a metric of doing so inside a CBA. Hirschman on the contrary thought that in a developing world side effects were a mechanism for re-balancing the economy, hence to be favoured.

While not ignoring negative side-effects¹², Hirschman's attention is focused on the positive ones, and in particular on "forward and backward linkages": investments in other fields that were done as a consequence of the main intervention in order to balance the effects of the latter. In this way, some side-effects turned out to be "inputs essential to the realization of the project's principal effect and purpose" (1967, p. 149) (Vedung's effects in the main area), that is to say development effects properly. This led Hirschman to draw a distinction between pure side-effects and mixed side-effects, linkages being an example of the latter. They are mixed because "they are required for the proper functioning and growth of the project itself" (1967, p. 155), which is seen as contributing to a process of development, taking place in an unbalanced way, combining main effects and side effects in the long run. Starting from the observation of unanticipated positive side-effects, Hirschman then suggested that in making investment decisions anticipated side-effects (beneficial and detrimental) should be taken into account (1967, p. 148).

In our times, the concern for side effects has an opposite meaning: side-effects are seen as mainly as detrimental, regressive, even if it is sometime admitted that "they may even be positive". Hence the request of complex, "systemic" programs that anticipate all negative side-effects by enlarging the scope of main effects, regardless of the uncertainty that had been considered as the signature of our times. The accent is on overcoming the separation of the fields (silos), much less on the dynamics that can interact between main and side effects, and that can be detected only by a savvy side-effects evaluation approach. Why not, instead, look for cases where side-effects could be positively enlisted for addressing the global challenges? Once again, it is a matter of preparedness.

12) In fact, when mentioning negative side-effects, he is mainly concerned with dismissing what he considers an argument in favor of doing nothing.

MAPPING MAIN EFFECTS

The operation of mapping effects could not be complete if it did not include also main effects. In another table of the same chapter of Vedung (1997, p. 58, table 4.5), it is assumed that in the main area it is clear how effects should be positive (although with the question mark: Vedung's typical way of distancing himself from "goal-oriented evaluation" approaches). However, there is no parallel distinction, as in the side area, between anticipated positive, anticipated negative, unanticipated positive, unanticipated negative. This is questionable. In fact, beneficial results in the target area can be obtained also in a way different from anticipated, in the same way as anticipated could turn out to be detrimental. Something cherished (main effect) can be obtained through different ways than designed.

Hirschman's idea of the Hiding Hand (1967, p. 11) offers an extraordinary support to this stance. It holds that if known in advance unforeseen difficulties would have blocked the program, thus preventing the fact that actors might creatively have found other ways of solving the problem. A providential hand that hid-ed the difficulties would instead allow the discovery of new ways out. Hirschman's praise of creativity, something that social scientists seem not to keep in high consideration, is a counterpart of Merton's praise for curiosity.

This is linked to the importance of implementation, as has been stressed also in the positive thinking approaches (Stame 2014). Vedung thinks that not regarding implementation as a problem is a limit of the goal-attainment evaluation approach, that "concentrates on the consequences of the program, not of the implementation process", i.e. not on the "organization and the networks which are supposed to transform decisions into effects" (1997, pp. 47-8). And he continues: "consequences are produced by the application of the program, not the assumptions originally guiding the decision to adopt a program. Influencing the adoption of a program is entirely different from controlling its outcomes" (1997, p. 56).

UNINTENDED CONSEQUENCES OF HUMAN ACTION

Vedung had referred to the traditional thought of the Scottish moralists, that social phenomena are “the result of human action but not of human design” (1997, p. 53), but in fact what he talks about are the unintended results of something performed according to human design (the program).

However, if one looks at the situation where a program has been implemented, one will find that the effects listed in table 4.4. quoted may differ in their nature, and from beneficial to detrimental, because of the influence of other social actions (“human action”) external to the program (“human design”): other policies, people’s behaviour, cultural aspects, natural events (provoked by human treatment of nature), etc.

In this case, such “human actions” can be considered as contributory causes (Mayne, 2021) to the positive or negative effects in the main (and side) area where the program is implemented. All this could very well fit into the search for causality in the Theory of Change, that is the main concern for today’s evaluators, and that echoes the search for “basic knowledge”, that Vedung considers among the purposes of evaluation. This is a further contribution that Vedung’s reception of unintended consequences in evaluation research may bring to the current debate about evaluation approaches.

ASCERTAINING THE VALUE OF SIDE-EFFECTS

The problem of “ascertaining their value” is a consequence of having taken side-effects seriously. It involves a broader issue in evaluative research, that Vedung reminds us by endorsing the distinction drawn by Shadish, Cook, and Leviton (1991) between “descriptive” and “prescriptive” values (Vedung 1997, p. 248). Whereas foreseen consequences can be judged by reference to descriptive values (those declared in the policy), unforeseen consequences lack such reference. Should the evaluator try and work out values that could be assimilated to the descriptive ones, or should she judge them by recurring to prescriptive values? Moreover, may not even foreseen effects in the main area be judged by prescriptive values?

In another telling, Vedung (1997, p. 251, fig. 14.1) lists descriptive criteria of merit, which are mainly based on goal attainment in main area, and prescriptive criteria of merit. The latter are: (1) Contribution to problem-solving; (2) Client needs; (3) Equal distribution; (4) Public interest. It is not difficult to recognize here some of the main concerns of those who evaluate programs of sustainable development, and who most often denounce negative unanticipated consequences of programs or of other social actions.

Recent interest in evaluation ethics has brought to enlarging its perspective from “doing no harm” to “preventing bad” and “doing good” (van den Berg, Hawins, Stame, 2022). While the first precept is working within the traditional framework of results anticipated by the program design according to descriptive values, the other two precepts – aligned with prescriptive values – engage the evaluator in elaborating, together with other stakeholders, broader merit criteria, and to act for their enhancement accordingly. Underlying this elaboration is the now familiar idea that unanticipated consequences may be also favourable, not only detrimental, provided one makes explicit what those four points may mean in the specific situation under review.

THE SEARCH CONTINUES

It is true, “Unintended effects of intentional action fascinate”, as reads the incipit of Vedung (1997, xvii). I can testify that Evert keeps ruminating on them. On a recent occasion (the SES 2022 Conference in Stockholm), commenting my presentation Evert was wondering about how to be a reflective evaluator, intent at solving new problems with double loop learning. I remember having told him: “you should know, you have been the first to take unintended consequences seriously in evaluation”. He seemed incredulous, but became radiant when I quoted Merton’s idea of “being prepared”, as if he were ready to a new round of thinking, something from which we expect new illumination.

References

- Colorni, E. (2019). *Critical Thinking in Action*. New York: Bordighera Press.
- De Zwart, F. (2015). Unintended but not Unanticipated Consequences. *Theory and Society*, 44:283–297.
- Hirschman, A.O. (1967). *Development Projects Observed*. Washington DC: The Brookings Institution. Second ed. 2015.
- Hirschman, A.O. (1971). 'Political economics' and possibilism. In Hirschman A.O., *A Bias for Hope*. New Haven, Conn.: Yale University Press.
- Hirschman, A.O. (1998). *Crossing Boundaries*. Princeton NJ: Princeton University Press.
- Mayne, J. (2021). Contribution Analysis and the Long-term Perspective: Challenges and Opportunities. In Forss, K., Lindkvist, I., & McGillivray, M. (Eds.). *Long Term Perspectives in Evaluation*. New York: Routledge.
- Merton, R.K. (1936). The Unanticipated Consequences of Purposive Social Action. *American Sociological Review*, 1 (6): 894-904.
- Merton, R.K. (1948). The bearing of Empirical Research upon the Development of Sociological Theory. *American Sociological Review*, 13: 505-515.
- Merton, R.K., & Barber E. (2004). *The Travels and Adventures of Serendipity: A Study in Sociological Semantics and the Sociology of Science*. Princeton NJ: Princeton University Press.
- Shadish W. , Cook T., & Leviton L. (1991). *Foundations of Program Evaluation: Theory and Practice*. Newbury Park CA: Sage.
- Stame N. (2014). Positive Thinking Approaches to Evaluation and Program Perspectives. *Canadian Journal of Program Evaluation*, vol. 29, n. 2: 67-86.
- Stame N. (2022). *Possibilism and evaluation. Albert Hirschman and Judith Tendler*. New York: Peter Lang Publisher
- Van den Berg, R., Hawkins, P. , & Stame, N. (2022). *Ethics for Evaluation: beyond 'Doing No Harm' to 'Tackling Bad' and 'Doing Good*. New York: Routledge.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. New Brunswick NJ: Transaction Publishers.

KAPITEL 14

I huvudet på en utvärderare

Rolf Sandahl

INLEDNING

När Peer Gynt, i Ibsens drama (Ibsen 1976, s. 171f), sitter och skallar en lök för att finna kärnan (livets), upptäcker han efter ett tag, att löken inte har någon kärna; den består endast av en massa skal. Skulle denna metafor kunna appliceras på utvärdering? Skulle vi hitta en kärna om vi bestämde oss för att gå igenom åtminstone de mer namnkunniga böckerna, exempelvis ett gemensamt tanke-sätt, ett bästa metodval etcetera eller skulle vi endast finna en massa skal?

Av David M. Fettermans kapitelrubrik i *Evaluation Roots* (Alkin ed., 2013, s. 304) "Learning to Think Like an Evaluator" kan vi få intrycket av att det tycks finnas något gemensamt tankegods hos oss utvärderare. Marvin Alkin (2013, s. 290) menar emellertid att det som brukar räknas som den viktigaste variabeln i en bostadsförsäljning, det vill säga *läge, läge, läge* motsvaras inom utvärderingsområdet av *kontext, kontext, kontext*. Varje enskild utvärdering måste således anpassas utifrån rådande förutsättningar där tanke-sätt och metodval återfinns i en (oändlig) verktygslåda eller "skal" för att använda vår metafor igen.

Populära begrepp som utvärdering får lätt karaktären av "semantiska magneter", vilket Carol Weiss (1998, s. 3) påpekade redan för 25 år sedan, det vill säga många typer av aktiviteter eller studier kallas för utvärdering. Evert Vedung och jag har under årens lopp fört ständiga diskussioner om vad utvärdering är, till exempel om skillnaden mellan uppföljning och utvärdering, om utvärdering enbart är ex postbedömningar eller om även ex antebedömningar ska räknas hit. I Everts första upplaga av *Utvärdering i politik och offentlig förvaltning* (Vedung 1991) tror jag mig ha övertygat Evert om att utvärdering är en ex postbedömning. I dag verkar denna gräns inte längre existera, då ex anteutvärderingar är ett stående inslag inom till exempel EU:s analysverksamhet.

Går det att sätta någon gräns för vad utvärdering är, eller gäller den "risk", som Aron Wildavskys (1973, s. 127–153) såg när det gällde planering, nämligen "if planning is everything, maybe it's nothing" också för utvärdering? Innebörden är att när ett område blivit för omfattande, syns inte längre konturerna.

Syftet med nedanstående text är både att reflektera över området utvärdering och en inbjudan till Evert för fortsatta diskussioner, något som jag är övertygad om kommer att ske utifrån vetenskapen om Everts gränslösa nyfikenhet. I det följande är det tre områden som jag kommer jag att diskutera, nämligen begreppet utvärdering, utvärderingsteorier och utvärdering som profession.

BEGREPPET UTVÄRDERING

Bokstaven "e" i evaluation sägs komma från det latinska ordet *ex*, som betyder ut, och det latinska *valere* som betyder "av värde". Att evaluera skulle då betyda "att finna värdet av". Ska då ut i utvärdering kopplas till det latinska ordet *ex*, vilket kan verka logiskt, eller betyder det något annat?

I boken *Utvärdering i politik och förvaltning* (Vedung, 1998, s. 22 f) diskuterar Evert om *ut* i utvärdering verkligen kommer från det latinska *ex*. Med hänvisning till ett svar på en fråga i Dagens Nyheter skulle *ut* kunna tolkas som ett förstärkande prefix, som i utbildad, slita ut, sova ut etc., det vill säga *fullt ut*. En utvärdering skulle därmed kunna sägas vara en *noggrann* bedömning; en systematisk analys.

Peter Dahler-Larsen (2012, s. 5ff) talar om tre typer av definitioner; den begreppsmässiga-analytiska, den metodfokuserade och den syftesfokuserade definitionen. Den begreppsmässiga-analytiska definitionen är en mer neutral definition där vad som helst kan utvärderas (something). Den metodmässiga definitionen ser utvärdering som en aktivitet som använder forskningsmetoder i sin analys, medan de som definierar utvärdering utifrån ett syftesperspektiv betonar utvärdering som ett medel att förbättra samhället. Den konkreta situationen bestämmer vilka metoder som är bäst lämpade.

Den gemensamma grunden i många utvärderingsdefinitioner är att den är *systematisk* eller *noggrann*, den innehåller en *värdering* eller *bedömning* och den syftar till att *förbättra en verksamhet* eller åtminstone, enligt Evert Vedung, spela en roll i praktiska besluts-situationer.

Utvärdering har, liksom de flesta begrepp, diskuterats livligt under årens lopp. Redan i en av de tidiga utvärderingsböckerna, "Evaluative Research" (Suchman 1967, s. 27), menade Suchman att utvärdering var illa definierat och felaktigt använt. Kan vi, eller vill vi, enas om en definition av utvärdering?

Ett resonemang om hur snäv eller hur bred en definition bör vara utvecklar Michael Power i *The Audit Society* (1999, s. 4). Revision (audit), i det här fallet, utgör, enligt Power, ett vagt begrepp och det är svårt att skilja revision från exempelvis inspektion, utvärdering och bedömning, eftersom alla dessa element kan ingå i en revision. "It is wiser to speak of a cluster of definitions which overlap but are not identical". Det är därför intressantare, enligt Power, att fokusera på själva begreppet revision som "idé", det vill säga varför samhället ser revision som ett framgångsrikt koncept, i stället för att försöka definiera vad revision är, tekniskt sett.

Peter Dahler-Larsen (2012, s. 5 ff) för ett liknande resonemang gällande utvärdering: "Is it important to distinguish evaluation from commonsense judgements, from science, from politics from practice, from consultancy?" Utvärdering måste sättas i ett socio-historiskt sammanhang och inte definieras för smalt, då utvär-

dering kan ta olika former i olika kontexter. Han betraktar utvärdering som ett paraplybegrepp för olika aktiviteter:

Activities such as auditing, accreditation, quality assurance, benchmarking, etc., can to a great degree be considered within the same conceptual frame of reference as evaluation in general.

Det innebär inte, enligt Dahler-Larsen, att alla dessa aktiviteter är exakt lika, utan det finns olikheter i vilka sammanhang de pågår, vem som utför dessa och vilka konsekvenser de får. Många av dessa aktiviteter, som revision och certifiering, har vidgats med åren och blivit mer generella till sin natur utifrån hur koncepten såg ut från början.

Intrycket man får av att läsa tidiga utvärderingsböcker, är att utvärdering framställs som en ganska omfattande aktivitet med en mängd frågor att besvara. Utvärdering känns mindre "ambitiöst" i dag. Det beror möjligtvis på att det finns många likartade aktiviteter som pågår, att fler aktörer säger sig syssla med utvärdering, att löpande uppföljnings- eller utvärderingssystem, inte bara enskilda studier, blivit allt vanligare etc.

Ska vi betrakta detta som ett problem? Jag tror inte att definitionsfrågan upptar särskilt stort utrymme i utvärderingslitteraturen i dag. Det har alltid funnits olika syn på vad som ska menas med utvärdering. Därför blir det komplicerat, kanske inte ens möjligt, att urskilja utvärdering från andra likartade aktiviteter, som Power och Dahler-Larsen påpekar ovan.

Utvärderingssamhällets lösning på detta problem, är att benämna sina medlemsföreningar för *utvärderingsföreningar* och inte *utvärderareföreningar*. Föreningarna är öppna för alla som är intresserade av utvärderingsfrågor, utan att ställa några särskilda krav på deltagarnas kompetens eller bakgrund.

Vi får helt enkelt, sammanfattningsvis, acceptera och konstatera att utvärdering är svärdefinierat, vilket inte nödvändigtvis behöver vara något negativt. Det gör utvärderingsområdet till ett inkluderande, inte ett exkluderande, område. Det viktiga är, som Jacobsson et al. också påpekar i *Granskningsamhället* (2019), inte vad den

ena eller den andra granskningsaktiviteten kallas, utan vilken typ av kunskap som är intressant och användbar.

UTVÄRDERINGSTEORI

I den akademiska världen, där utvärdering diskuteras och kunskap förmedlas, finns alltid ett intresse för att utveckla teorier. Teorier handlar ju om att formulera begrepp som förklarar, systematiserar och begripliggör verkligheten och som sedan prövas mot verkliga förhållanden. Syftet är att vara till hjälp i det praktiska arbetet. Alla teorier är emellertid inte grundade på erfarenhet, som matematik, statistik eller logik, som utgör exempel på aprioriska vetenskaper.

Att indela teorier i deskriptiva eller preskriptiva (normativa) är vanligt förekommande. Deskriptiva teorier talar om hur det *är*, medan normativa teorier talar om hur det *bör* vara (en indelning som inte är helt oproblematisk men som får utvecklas i ett annat sammanhang).

Man kan också tänka sig att tala om teorier *i*, *för* och *om* någonting. Teorier *i* någonting kallas för *sakteorier* och beskriver och förklarar hur olika fenomen i verkligheten förhåller sig till varandra. Det kan till exempel gälla teorier inom fysiken, kemin eller medicinen uttryckta i form av naturlagar eller andra mindre lagbundna teorier. Sakteorier inom sociologin eller psykologin syftar till att förklara människans roll i samhället eller dess olika beteenden.

Teorier *för* någonting är hjälpteorier för att exempelvis finna säkrare sakteorier. Olika typer av metoder uttrycker inte något i sak men det finns metoder som olika användbara i olika sammanhang. Metoder har ofta en transdisciplinär karaktär, det vill säga de kan användas inom alla discipliner. Statistiken exempelvis har sina egna disciplinära teoriområden, som sannolikhetsteori och inferensteori, samtidigt som de kan användas inom alla discipliner.

Teorier *om* någonting innebär att man just uttalar sig specifikt om exempelvis utvärderingens roll i ett visst sammanhang. Har den ett främjande, kontrollerande eller ett rituellt syfte? Hur ser förhållandet ut mellan utvärdering och forskning? Vilken nytta eller mervärde har alla dessa utvärderingar på ett utpekade område?

De teorier vi diskuterar här är teorier *för* någonting (deskriptiva eller preskriptiva). Utvärdering har inga egna sakteorier. Programteori som ofta nämns i utvärderingssammanhang är en sakteori och ingen utvärderingsteori.

Inom vilka områden är det aktuellt att tala teorier för utvärdering? I "Foundations of Program Evaluation" (Shadish, Cook & Leviton 1991), en av få böcker som mer utförligt diskuterar utvärderingsteorier, talar man om utvärderingsteorier inom följande områden: *social programming, knowledge construction, use, valuing* och *evaluation practice*. Det ideala är naturligtvis att utveckla deskriptiva teorier som kan vägleda utvärderaren i den praktiska situationen, ett ideal som författarna menar inte går att uppfylla (Shadish, Cook & Leviton 1991, s. 30).

När Shadish et al. talar om teorier, är det svårt att urskilja vilka dessa egentligen är. Snarare är det en betoning på viktiga frågor som bör ställas, till exempel vilken kunskap som eftersträvas, vilken typ av användning man tänker sig, vems värderingar som ska framföras och enligt vilken måttstock, vad som är möjligt att åstadkomma inom given tidsram och budget etc.

Därför menar till exempel Marvin Alkin (2013, s. 4), att det som ofta benämns som utvärderingsteorier i utvärderingslitteraturen, inte når upp till den status som en (deskriptiv) teori har. I stället för teorier borde man, enligt Alkin, tala om ansatser eller modeller. Preskriptiva modeller råder det däremot ingen brist på inom utvärderingslitteraturen.

Varför detta gap mellan teori och praktik? Enligt Jean King (2003, s. 57ff) finns det flera förklaringar. Begreppet teori, precis som utvärdering, är inte alltid självklar. Utvärderingsteoretiker och praktiker kommer från olika bakgrunder, har olika kunskapsteoretiska antaganden, fokuserar på olika saker i sin utvärdering och har olika syften med utvärdering. Utvärderare måste vidare ta hänsyn till vad "klienterna" vill ha och försöker inte validera olika utvärderingsteoretiska modeller. Utvärdering handlar om att genomföra, inte att teoretisera eller forska om utvärdering som fenomen. Eftersom utvärdering inte är något "eget ämne" finns det inte

heller någon teoretisk uppbackning som det normalt sett finns inom en akademisk disciplin.

Ernest House (2003, s. 56) finner emellertid inget märkligt i detta gap mellan teoretiker och praktiker. Teoretiker söker efter nya koncept och idéer, som i vissa fall och efter en tid kanske kan omsättas i praktiken, i andra fall inte. Normalt sett ska det vara ett gap mellan teoretiker och praktiker enligt honom. En annan förklaring, enligt House, är den stora omsättningen på utvärderare. Det innebär att möjligheterna att ta sig an och förstå teoriområdet är begränsade och samtidigt att bidraget från utvärderarna till teoribildandet därmed blir litet. Likaväl som det brukar framhållas att inflytandet från teori till praktik är komplext, kan samma sak sägas om inflytandet från praktik till teori.

Michael Scriven (1991, s. 364) har emellertid länge rubricerat utvärdering som en *transdisciplin* eller alfadisciplin (Scriven 2013, s. 175), en hjälpdisciplin till övriga discipliner. Utvärdering är, enligt Scriven, därför jämförbar med transdiscipliner som exempelvis biostatistik och formell logik.

Skillnaden mellan till exempel statistik och utvärdering, som jag ser det, är att statistiken har egna teorier, som sannolikhetsteori och inferensteori, något som utvärdering saknar. För utvärdering finns helt enkelt inte någon bästa plan (Cronbach 1982, s. 321). Vi kan naturligtvis ha idéer om bästa eller lämpligaste metod i den enskilda studien, men inte som en generell rekommendation.

Den som vill veta hur stort urval en studie bör ha för att kunna dra slutsatser med en viss statistisk säkerhet eller hur man ska tänka i kausala termer, letar knappast i allmänna utvärderingsböcker. Hela evidensdebatten, ett tema som utvärderingssamhället fångat upp, handlar om en diskussion som länge först mellan olika discipliner gällande metod- och kunskapssyn, visserligen viktigt att diskutera, men där "disciplinen utvärdering" har lite att tillföra. Det är det motsatta förhållandet som gäller. När vi talar om glappet mellan teori och praktik, menar vi svårigheten att skapa generella utvärderingsteorier som kan tillämpas i den enskilda situationen. I den enskilda utvärderingssituationen däremot, behöver det inte

alls vara något glapp mellan teori och praktik, eftersom vi i många fall kan utgå ifrån respektive disciplins egna sakteorier.

Om det saknas generella utvärderingsteorier, vilket vi påstår ovan, hur påverkar det möjligheterna att göra utvärdering till en akademisk profession?

PROFESSION

Utvärdering är ett akademiskt ämne på många universitet och högskolor. På den amerikanska utvärderingsorganisationens hemsida (www.eval.org) listas de universitet inom USA som har en utbildning i utvärdering på master- och doktorsnivå. Ett drygt 80-tal utbildningsanordnare finns på denna lista, främst inom utbildning och utbildningspsykologi. Metodkunskap är vanligt förekommande som rubrik i dessa utbildningar.

I Sverige finns ett hundratal kurser med utvärdering i rubriken. Någon specifik examen, förutom kurspoäng, finns inte. Däremot kan man naturligtvis ha utvärdering som forskningsobjekt men inom ramen för något annat akademiskt ämne. För att område ska betecknas som profession, här menat akademisk profession, finns det några attribut som ofta nämns som konstituerar en profession (se exempelvis Forslund, 1993, s. 7ff.; Strömberg, 1994, s. 10 ff. & Brante et al, 2014, s. 18 ff.); nämligen *teoretisk kunskap; kunskapen är tillägnad genom akademiska studier; kompetensen säkras genom ett formellt examinationssystem; etisk kod* och *professionell kultur*.

I en profession ingår såväl kognitiva som sociala moment. De kognitiva momenten är de rent kunskapsmässiga, det man lär sig i utbildningarna och säkras genom en examen, medan de sociala är de moment som professionen medför rent arbetsmässigt, som graden av autonomi, etiska koder, organisering etcetera och vad detta i sin tur genererar.

Akademiska studier betraktas eller beskrivs ofta som att de genererar abstrakt och esoterisk kunskap, det vill säga *teoretisk kunskap* (Strömberg, 1994, s. 18). Komponenterna teori och beprövad erfarenhet utgör tillsammans professionen och ger den struktur i yrkesutövandet för att förstå hur saker och ting hänger ihop och därmed till problemlösningsförmågan. Den teoretiska

grunden är således tänkt till att leda till förmågan ”hur man tänker” till ”hur man gör”, det vill säga förhållandet mellan know-why (teori) och know-how (praktik) (Brante et al, 2014, s. 19).

Inom utvärderingsområdet finns otvivelaktigt en professionell kultur med en omfattande mängd böcker, artiklar och tidskrifter om utvärdering, nätverk och utvärderingsföreningar i många länder liksom ledande personer inom området. Kunskapen är bland annat tillägnad genom akademiska studier och det finns i många fall ett examinationssystem. Etiska koder är något som vissa utvärderingar anammar, men som rekommendationer. Den teoretiska delen i form av empiriska teorier saknas, preskriptiva modeller däremot, finns det ett överflöd av.

Rossi, Freeman och Lipsey (Rossi et al, 2003, s. 335 ff) har följande syn på förhållandet mellan utvärdering och profession:

In brief, evaluation is not a “profession”, at least in terms of the formal criteria that sociologists generally use to characterize such groups. Rather, it can be described as individuals sharing a common label who form a loose aggregate, who are not formally organized, and who may have little in common with one another in terms of the range of tasks they undertake or their competencies, work sites, and outlooks.

Rossi menar emellertid att det stora antalet representerade discipliner inom utvärderingsområdet leder till ett rikt intellektuellt utbyte mellan utvärderare. Att ”låna” metoder och idéer från varandra är positivt men bredden av metoder och idéer gör området svåröverskådligt för den enskilde utvärderaren. Den återkommande frågan när profession och utvärdering kommer på tal är: Vad är utvärderingskunskap? I stället för profession, där man tänker sig en gemensam grund, vilket många utvärderare inte alls förespråkar eller anser möjlig, talar professionsforskarna också om *professionalism*.

PROFESSIONALISM

Det väsentliga för området utvärdering är kanske inte att det blir en profession eller att utvärderare licensieras. När Brante (Brante et al, 2014, s. 254f.) i en enkätundersökning frågar representanter

för ett antal olika professioner vad professionalism innebär, är begreppen kunskap, kompetens och skicklighet de mest förekommande. Hänvisning till formell utbildning förekommer sällan. En förklaring skulle kunna vara att professionalism ses som en relation och ett fenomen på individnivå där den praktiseras, inte hur denna kunskap inhämtats.

Professionalism skulle därmed kunna uppfattas som en kontextuell förmåga, det vill säga anpassad till den enskilda situationen, och inte en allmän förmåga. Den viktigaste variabeln i utvärdering, enligt Marvin Alkin ovan, var just hänsynen till *kontext*, *kontext* och *kontext*. Kompetens kan ju innebära olika förmågor. Den teoretiska kunskapen, påståendekunskaper, får en mindre framträdande roll i professionalism än erfarenhetskunskaper, det vill säga förtrogenhetskunskaper som även inkluderar ”tyst” kunskap och sunt förnuft. Kunskap blir här liktydigt med kompetens, det vill säga ”... en individs potentiella handlingsförmåga i relation till en viss uppgift, situation eller kontext...”.

“Evaluation is an art” myntade Lee J. Cronbach (1982, s. 321) för många år sedan. Det kan tolkas som att det krävs kompetens som, till skillnad mot kunskap, innefattar fler moment än bara den kognitiva delen. Kompetens är hur den enskilde fungerar i det praktiska arbetet, medan kvalifikation är de formella krav som kan ställas i en befattningsbeskrivning (Brante et al, 2014, s. 247 ff).

I HUVUDET PÅ EN UTVÄRDERARE?

Vi kan sammanfattningsvis konstatera att det finns gemensamma moment i många definitioner av utvärdering, men att det i praktiken är svårt att upprätthålla någon rågång mellan olika aktiviteter. Om till exempel värderingsmomentet anses vara det moment som skiljer utvärdering från annan utredningsverksamhet, visade en undersökning för många år sedan att de tillfrågade utvärderarna satte just värdering som näst lägst (av 49) av viktiga utvärderingskompetenser (King et al, 2001, s. 245).

Betyder ovan sagda utvärdering requiem? Inte alls. Att bedöma olika offentliga insatser kommer vi alltid att behöva göra, oavsett om vi rubricerar bedömningen som utvärdering eller något annat.

Är Wildavsky syn på planering, det vill säga, "if *planning is everything, maybe it's nothing*" också tillämpbar på utvärdering? Att utvärdering blivit mer konturlöst och svår att urskilja från andra typer av verksamheter kan vi nog tveklöst konstatera, liksom att "utvärdering" inte har några egna generella teorier. Däremot förekommer "utvärdering", mer eller mindre strikt definierat, inom alla områden eller discipliner där egna sakteorier och metoder tillämpas och föredras i enskilda fall, beroende på kontexten. Det är bara så att *en* typ av kunskap finns i huvudet på *en* utvärderare, en *annan* typ av kunskap i huvudet på en *annan* utvärderare. Under sådana förutsättningar är teoribildning, på ett generellt plan, komplicerat. Utvärdering har kanske, som den lök Peer Gynt skalade, ingen kärna men en mängd skal.

Referenser

- Alkin, Marvin C. (2013). Comparing Evaluation Points of View. I Alkin, Marvin C. (Red.), *Evaluation Roots*, Second Edition (s. 3–10). Los Angeles: Sage.
- Alkin, Marvin C. (2013). Context-Sensitive Evaluation. I Alkin, Marvin C. (Red.), *Evaluation Roots*, Second Edition (s. 283–292). Los Angeles: Sage.
- American Evaluation Association (www.eval.org).
- Brante, Thomas, Johnsson, Eva, Olofsson, Gunnar & Svensson, Lennart G. (2014). *Professionerna i kunskapssamhället*. Stockholm: Liber.
- Cronbach, Lee J (1982). *Designing Evaluations of Educational and Social Programs*. San Francisco: Jossey-Bass Publishers.
- Dahler-Larsen, Peter (2012). *Evaluation Society*. Stanford: Stanford University Press.
- Fetterman, David M. (2013). Empowerment Evaluation. Learning to Think Like an Evaluator. I Alkin, Marvin C., *Evaluation Roots*, Second Edition (s. 304–322). Los Angeles: Sage.
- Forslund, Kenneth (1993). *Professionell kompetens. Fyra essäer om inlärning och utveckling för professionalitet*. Institutionen för pedagogik och psykologi, Linköpings universitet.
- House, Ernest R. (2003). Stakeholder Bias. I Christie, Christina A. (Red.), *The Practice-Theory Relationship in Evaluation* (s. 53-56). New Directions for Evaluation, nr ninety-seven. San Francisco: Jossey-Bass.
- Ibsen, Henrik (1976). *Peer Gynt*. Stockholm: Forum.
- Jacobsson, Bengt, Pierre, Jon & Sundström, Göran (Red.) (2019). *Granskningsamhället*. Lund: Studentlitteratur.
- King, Jean A., Stevahn, Laurie, Ghere, Gail, & Minnema, Jane (2001). Toward a Taxonomy of Essential Evaluator Competencies. *American Journal of Evaluation* 22(2), 229-247.
- King, Jean A. (2003). The Challenge of Studying Evaluation Theory. Christie, Christina A. (Red.), *The Practice-Theory Relationship in Evaluation*. New Directions for Evaluation, nr 97 (s. 57-67). San Francisco: Jossey-Bass.
- Power, Michael (1999). *The Audit Society*. Oxford: Oxford University Press.

- Rossi, Peter H., Freeman, Howard, E. & Lipsey, Mark W. (2003). *Evaluation – A systematic Approach*, Seventh Edition. Thousand Oaks: Sage.
- Scriven, Michael (1991). *Evaluation Thesaurus*, Fourth Edition. Newbury Park: Sage.
- Scriven, Michael (2013). "Conceptual Revolutions in Evaluation: Past, present, and Future".
- Alkin, Marvin C. (Ed.), *Evaluation Roots*, Second Edition, 167-179. Los Angeles: Sage.
- Shadish, William R., Cook, Thomas D. & Leviton, Laura C. (1991). *Foundations of Program Evaluation*. Newbury Park: Sage.
- Strömberg, Barbro (1994). *Läraryrket – en profession?* Rapport nr 1994:2. Vårdpedagogik – Omvårdnad. Institutionen för vårdlärarutbildning, Göteborgs universitet.
- Suchman, Edward A. (1967). *Evaluative Research*. New York: Russel Sage Foundation.
- Vedung, Evert (1991). *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur.
- Vedung, Evert (1998). *Utvärdering i politik och förvaltning*, 2 uppl. Lund: Studentlitteratur.
- Weiss, Carol H. (1998). *Evaluation*, Second Edition. New Jersey: Prentice Hall, Upper Saddle River.
- Wildavsky, Aron (1973). If planning is everything, maybe it's nothing. *Policy Science* 4(2), 127–153.

KAPITEL 15

Politiska intentioner och utvärdering i kohandelns tid: En pågående dialog med Evert Vedung

Bo Bengtsson

För snart 20 år sedan efterträdde jag Evert Vedung som professor i statskunskap, särskilt bostads- och urbanpolitik, vid Institutet för bostads- och urbanforskning (IBF). Jag har däremot inte gått i Everts fotspår forskningsmässigt. Min forskning har varit koncentrerad på att utveckla olika perspektiv på bostadspolitiken ("housing policy and housing politics"), medan Evert har en bredare repertoar med inriktning på offentlig politik, urbanpolitik, idéanalys och inte minst utvärdering.

Jag korsade dock Everts spår då jag på uppdrag av en offentlig utredning skulle identifiera och tolka de politiska intentionerna bakom den svenska hyrespolitiken som underlag för en utvärdering (Bengtsson, 1992).¹ I min rapport lanserade jag en modell för utvärdering baserad på statsmakternas institutionaliserade intentioner. I föreliggande kapitel försöker jag relatera den modellen till Everts kritik av måluppfyllelseutvärdering. Jag diskuterar också vilken relevans min modell kan ha i dagens

1) Senare lade jag också fram rapporten som lic-avhandling i statsvetenskap.

parlamentariska situation, där kohandel och partipolitisk positionering har fått allt större utrymme på bekostnad av argumentation i sakfrågor. Diskussionen relateras empiriskt till utvecklingen av den svenska hyrespolitiken.²

Även senare har jag ägnat en del intresse åt bostadspolitikens institutionaliserade intentioner, men jag har inte utvecklat den teoretiska logiken bakom min utvärderingsansats och inte heller diskuterat dess relevans i dagens politik. Jag har svårt att tänka mig ett bättre sammanhang för att ta mig an detta än en hyllningskrift till Evert Vedung.

POLITISKA INTENTIONER SOM UTGÅNGSPUNKT FÖR UTVÄRDERING

För att precisera hur en utvärdering av politiska intentioner kunde utformas tog jag hjälp av de svenska forskare som jag såg som de främsta auktoriteterna på utvärderingsteori. Förutom Evert, vars blivande klassiker *Utvärdering i politik och förvaltning* (Vedung, 1991) nyligen utkommit i sin första upplaga, konsulterade jag Lundafilosofen Tore Nilstun och hans ”problemram för utvärderingar” (Nilstun, 1984).

Till min förvåning fann jag inte så mycket direkt användbart i något av dessa verk. Nilstun förordar att utvärderingen ugår från *effekter* i stället för mål och då beaktar både utfall som anhängarna av en reform hoppas på och utfall som motståndarna varnar för (Nilstun, 1984, s. 14).

I sin första upplaga varnar Vedung för målbaserad utvärdering eftersom politiska mål ofta är vaga, mångtydiga och inbördes motstridiga. Dessutom framhåller han att en sådan ansats lätt bortser från oförutsedda bieffekter, liksom från aktörernas strategiska överväganden. Liksom Nilstun menar Vedung att en målutvärdering – om en sådan ändå ska göras – bör utgå från *olika* aktörers målbild och farhågor kring en reform (Vedung, 1991, s. 40f.).

2) Det finns inte här utrymme för att mer ingående beskriva den svenska hyrespolitiken med sitt unika kollektiva förhandlingssystem och de organiserade intressenas starka inflytande; jag hänvisar i stället till Bengtsson 2022.

Den till synes uppenbara lösningen att utvärdera en politik på basis av dess intentioner tycks alltså stöta på betydande hinder – och mitt uppdrag gällde alltså att undersöka just hyrespolitikens intentioner. Även om jag instämde – och alltjämt instämmer – i Vedungs kritik av enkel målutvärdering, ville jag ändå utgå från de formulerade politiska målen. Inspirerad av nyinstitutionalistisk teori valde jag då att betrakta de mål som kan härledas ur riksdagsbeslut som *institutionaliserade intentioner*, vilket enligt min mening gör dem särskilt relevanta som kriterier för utvärdering. Denna institutionalisering sker då via ”den representativa demokratins beslutskedja” som tänks löpa från medborgarnas preferenser och önskemål, via deras valda ombud i riksdagen och den parlamentariskt utsedda regeringen (Birgersson & Westerståhl, 1980).

Med en sådan utgångspunkt ses statsmakterna inte som en aktör med preferenser utan som en *institution*, det vill säga som ett formellt och reellt regelsystem av viss varaktighet. Institutioner bestämmer förutsättningarna för aktörers framtida handlande, de fungerar som spelregler som gör politiskt handlande förutsägbart. Om statsmakterna betraktas som institutioner för att fatta bindande demokratiska beslut innebär dessa beslut också en institutionalisering av en ståndpunkt om hur ett politiskt problem ska lösas. Denna ståndpunkt får därmed en självständig kraft oberoende av de aktörer som fattat beslutet.³

Institutionalisering av *politiska beslut* är naturligtvis ingen originell tanke. Grundtanken i den modell för utvärdering som lanseras i min rapport (Bengtsson, 1992, s. 20f.) är dock att även vissa *argument* för ett beslut kan institutionaliseras som *statsmakternas intentioner*, givet att de konfirmerats i enlighet med gällande parlamentariska regler. För att uppfylla detta krav måste argumenten ha anförts – eller åtminstone kunna härledas ur – den skriftliga *motiveringen* till riksdagsbeslutet. Det är således *publiceringen* av argumenten i beslutsunderlaget som gör att de institutionaliseras. Argument som politikerna inte finner legitima att lyfta fram explicit kan inte heller institutionaliseras som politiska intentioner.

3) Ett institutionaliserat parlamentariskt beslut behöver inte ens motsvara någon bestämd aktörs ståndpunkt i sakfrågan; det kan mycket väl vara resultatet av en kompromiss.

Institutionaliseringen av argumentationen är svagare än den sker av själva beslutet. Rimligen bör dock motiveringar som åberopas i det offentliga besluts materialet tillmätas större vikt än de argument (för eller emot) som enskilda riksdagsledamöter eller andra aktörer ger uttryck för i debatten inför (eller efter) beslutet.

Inte ens då jag skrev rapporten stod den representativa demokratin beslutskedja särskilt högt i kurs. Relationerna mellan politiker, medborgare och andra aktörer ansågs – och anses – ofta bättre kunna beskrivas som ett komplicerat spel om makt och resurser, och idén om den demokratiska beslutskedjan sågs mera som ett sätt att ge legitimitet åt den representativa demokratin. Kort sagt uppfattades modellen som förenklad som verklighetsbeskrivning och orealistisk som norm.

Som jag framhåller i rapporten (Bengtsson, 1992, s. 19) behöver emellertid modellens brister som verklighetsbeskrivning inte vara något avgörande problem för att ta den som utgångspunkt för utvärdering. Ska vi se de politiska intentionerna som institutionaliserad norm och kriterium för utvärdering måste vi utgå från att politikerna menar allvar i sak – åtminstone då de fattar och motiverar sina parlamentariska beslut. Detta är den normativa grundvalen för en utvärdering av politik mot institutionaliserade intentioner.

I rapporten pekar jag på några ytterligare premisser för modellen (Bengtsson, 1992, s. 21f.). Av dessa bör två särskilt framhållas. Låt oss benämna dem *antagandet om rimlig rationalitet* respektive *antagandet om giltighet tills vidare*.

Intentioner förknippas ofta med *rationalitet*. Även om statsmakterna inte betraktas som en aktör kan tanken om institutionaliserade intentioner ändå kopplas samman med rationalitetsbegreppet i den meningen att intentioner kan antas vara *rimligt konsistenta*. En avgörande förutsättning för analysen är att olika mål inte är inbördes motstridiga och att de anvisade medlen rimligen ska kunna tänkas leda till målen. Någon perfekt rationalitet behöver dock inte antas, utan de intentioner som institutionaliseras i den parlamentariska processen kan ha samma brister som andra mänskliga kalkyler. Antagandet om rimlig rationalitet innebär inte mer än att vi så långt möjligt väljer konsistenta tolkningar framför inkonsistenta.

I princip förväntas institutionaliserade intentioner *gälla tills vidare*, det vill säga så länge de inte explicit revideras i senare parlamentariska beslut. Detta följer av formkravet. Riksdagen har rätt att ändra sig, men om den vill ompröva de politiska intentionerna bör detta framgå i beslutsunderlaget till senare beslut – något som för övrigt skett vid olika tidpunkter med det svenska hyressättnings-systemet. I en rekonstruktion av rådande politiska intentioner ingår alltså att följa den historiska utvecklingen och pröva vilka argument som *inte förändrats* vid senare beslutstillfällen.

De politiska intentionernas eventuella roll i utvärderingen av politiken kan jämföras med förarbetenas roll i den juridiska tolkningen av lagstiftarens intentioner. Även om den juridiska innebörden är mera formell, och ibland bindande, gäller principerna om rimlig rationalitet och giltighet tills vidare i båda fallen.

VEDUNGS SENARE UPPLAGOR: MÅLUPPFYLLELSE-MODELLEN OCH DEN REPRESENTATIVA DEMOKRATIN

I rapporten instämmer jag alltså i Vedungs kritik mot tanken att rakt av ta politiska mål som utgångspunkt för utvärdering. Samtidigt framhåller jag idén om institutionaliserade intentioner som ett sätt att göra utgångspunkten mer entydig och demokratiskt förankrad.

I senare omarbetade och utvidgade upplagor från 1998 och 2009 tar Vedung upp den parlamentariska styrkedjan och dess betydelse för måluppfyllelsemodellen för utvärdering. Han påpekar att demokratiargumentet för måluppfyllelseutvärdering "hämtar sin näring ur föreställningen om den parlamentariska styrkedjan och representativ demokrati". Att beslut om mål fattats av ett folkvalt organ betyder att dessa mål har en speciell demokratisk legitimitet. Vedung betonar nu att det är måluppfyllelsemodellens stora förtjänst att den bejaktar att mål i offentlig sektor spelar en huvudroll i ett demokratiskt statsskick, något som är livsviktigt i ett medborgerligt perspektiv av ansvarsutkrävande (Vedung, 2009, s. 94f.).

Även om måluppfyllelsemodellen "hämtar sin näring" ur den representativa demokratin är Vedung fortfarande kritisk till modellen. Han radar upp och diskuterar inte mindre än sex olika motargu-

ment, av vilka han ser "oklarhetsargumentet" (att målen är diffusa) och "bieffektsargumentet" (betydelsen av oavsedda bieffekter av en reform) som särskilt allvarliga (Vedung, 2009, s. 97f.).⁴

I stället förespråkar Vedung vad han benämner *bieffektmodellen* som gör det möjligt att utvärdera även effekter som faller utanför på förhand angivna mål. Han betonar att även denna modell utgår från "den representativa demokratins teori och den parlamentariska styrkedjans primat". Jag tolkar honom som att de bieffekter som ska utvärderas i första hand ska vara oavsedda just i förhållande till de på förhand angivna målen – de institutionaliserade intentionerna i min modell.⁵ Vedungs bieffektmodell blir då en kombination av att utvärdera politiska intentioner och att testa ett antal hypoteser om andra tänkbara effekter. Samtidigt pekar han på problemen med att ställa upp hypoteser om bieffekter och kriterier att utvärdera dem mot (Vedung, 2009, s. 114f.).

Hade jag valt en bieffektmodell om de senare upplagorna av Everts bok hade existerat då jag skrev min rapport om hyrespolitiken? Troligen inte. En möjlig lösning hade varit att kombinera en utvärdering baserad på statsmakternas intentioner med en översiktlig bieffektsanalys – detsaknas sannerligen inte föreställningar om bieffekter i den hyrespolitiska debatten. Problemen med hypoteser och kriterier gör dock detta till en högst tids- och arbetskrävande ansats, samtidigt som den demokratiska styrkedjans primat blir otydligare.

På en central punkt uppfattar jag att min modell skiljer sig från Everts rekommendationer, även om vi tycks vara överens om betydelsen av den demokratiska legitimiteten hos statsmakternas beslut.⁶ Jag betonar just de *institutionaliserade* aspekterna av statsmakternas intentioner, det vill säga att de kan leva vidare oberoende

4) Som kommer att utvecklas i det följande ser jag för egen del oklarhetsargumentet som överordnat bieffektsargumentet. Om intentionerna är oklara blir det meningslöst att utvärdera bieffekter.

5) Han förespråkar dock samtidigt "en mer kreativ tillnärmelse" till värderandet av bieffekter (Vedung, 2009, s. 115).

6) Evert relaterar denna legitimitet enbart till de politiska målen, medan min modell inkluderar intentioner beträffande både mål och medel.

av om enskilda aktörer med tiden skulle ändra ståndpunkt. Detta betyder att man kan härleda rådande intentioner även från tidigare politiskt beslutsmaterial enligt principen om giltighet tills vidare.⁷

HYRESPOLITIKEN 1991 – MITT EMPIRISKA FALL

Den fråga jag brottades med då jag formulerade mina idéer om institutionaliserade intentioner var alltså den svenska hyrespolitiken. Att mina uppdragsgivare efterlyste just intentionerna bakom politiken var knappast resultatet av något medvetet val mellan olika utvärderingsmodeller utan snarare ett uttryck för att de såg det som den naturliga utgångspunkten i utvärdering av politik. Detta är i så fall en uppfattning som jag delar.

I min rapport identifieras och tolkas de politiska intentionerna bakom bruksvärdessystemet på hyresmarknaden som de såg ut 1991 före riksdagsvalet och regeringen Bildts tillträde samma år. De innehållsliga intentionerna i fråga om hyrorna avsåg besittningsskydd, hyresnivå och hyresstruktur. *Besittningsskyddet* ska vara såväl formellt som reellt, så att en hyresgäst inte ska behöva lämna sin bostad till exempel på grund av att hyresvärden försöker ta ut marknadshyra. Den *allmänna hyresnivån* på en ort ska inte pressas upp vid bostadsbrist. *Hyresstrukturen*, det vill säga hyresrelationen mellan olika lägenheter, ska återspegla tänkta marknadshyror på en bostadsmarknad i långsiktig jämvikt. Systemet ska *implementeras* genom hyresmarknadens funktionssätt, det kollektiva förhandlingssystemet för hyror, de allmännyttiga bostadsföretagens prisledande ställning och hyresnämndernas bruksvärdesprövning. Hyrespolitiken ska även bidra till avtalsfrihet, effektivitet och ett tillfredsställande fastighetsunderhåll samt till att det statliga bostadsstödet verkligen kommer hyresgästerna till godo. Den avtalsfrihet som eftersträvas måste för hyresgästernas del i huvudsak uppfattas som kollektiv (Bengtsson, 1992, s. iii).

I enlighet med principen om giltighet tills vidare rekonstruerade jag ovanstående bild av hyrespolitikens institutionaliserade intentioner genom en systematisk tolkning av motiven till de

7) Med mitt synsätt blir detta en helt central uppgift. Kan inga politiska intentioner – avsedda effekter – identifieras blir det också meningslöst att tala om bieffekter.

hyrespolitiska besluten från systemets införande 1968 fram till 1991. Jag gav också förslag om hur dessa intentioner skulle kunna användas som utvärderingskriterier (Bengtsson, 1992, s.119f.).

SENARE HYRESPOLITIK: SYSTEMSKIFTE, EU-ANPASSNING OCH KORPORATISM

Kan vi på samma sätt identifiera de politiska intentionerna bakom hyrespolitiken under den följande trettioårsperioden? Och hur har de i så fall förändrats sedan 1991? Här finns bara utrymme för en ytlig resumé, men förvånansvärt lite har faktiskt skett.⁸ Visserligen genomdrev systemskiftesregeringen Bildt 1991–94 betydande förändringar i bostadspolitiken, inklusive en snabbavveckling av Bostadsdepartementet (Persson, 2003). Emellertid genomfördes till slut, sannolikt av opinionsmässiga skäl, inga större förändringar i hyresförhandlings- och bruksvärdessystemet.

Den stora förändringen i hyressättningsystemet sedan dess var den *EU-anpassning* som skedde 2010, då den så kallade allbolagen ersatte de kommunala allmännyttiga bostadsföretagens självkostnadsprincip med regler om affärsmässighet, samtidigt som regelverket för hyressättning förändrades. Såväl de kollektiva förhandlingarna som bruksvärdessystemet behölls, men de allmännyttiga hyrorna ersattes som jämförelsenorm vid bruksvärdesprövningen med kollektivt förhandlade hyror i både privata och allmännyttiga fastigheter.

Innebar denna reform förändrade politiska intentioner för hyrespolitiken? För att besvara den frågan krävs en noggrannare granskning än vad som kan genomföras här. Vad som är tydligt är att intentionerna avseende politikens *medel* – främst allmännyttan och bruksvärdessystemet – förändrats i grunden av 2010 års beslut. Det framstår som mera oklart i vilken utsträckning även målen förändrats – propositionens argumentation handlar mest om att systemet måste uppfylla de krav som ställs av EU:s konkurrenslagstiftning om lika villkor för olika typer av fastighetsägare. Min försiktiga tolkning – i enlighet med principen om giltighet tills vidare – är att de tidigare målen om reellt besittningsskydd,

8) Se Bengtsson & Grander, 2023, kap. 8 för en grundligare genomgång.

allmän hyresnivå och hyresstruktur kvarstår. Möjligen bör man i enlighet med principen om rimlig rationalitet fråga sig om förändringarna av de hyrespolitiska medlen är konsistenta med dessa mål – men här ger förarbetena ingen vägledning.

Att den slutliga utformningen av EU-anpassningen var resultatet av en ”hyreskorporatistisk” kompromiss mellan hyresmarknadens parter (de privata och allmännyttiga fastighetsägarnas organisationer samt Hyresgästföreningen) är i sig i sig ingen invändning mot att tolka statsmakternas intentioner med stöd av den parlamentariska argumentationen. Men det bör noteras att analysen försvåras av att kompromisserna i den korporatistiska beredningen tenderar att göra de politiska målformuleringarna vaga.

DAGENS HYRESPOLITIK: POLITISKT SPEL SKYMMER INTENTIONERNA I SAKFRÅGAN

En korporatistiskt ordnad beredning och implementering försvårar således tolkningen men är ingen principiell invändning mot att basera en utvärdering på statsmakternas institutionaliserade intentioner. Detta lika lite som ett starkt reellt inflytande från andra aktörer utanför riksdagen, då det gäller hyrespolitiken framför allt EU och det internationella fastighetskapitalet. Det är fortfarande riksdagen som ytterst fattar de politiska besluten. Men hur hållbar är tanken om institutionaliserade intentioner i dagens svenska politik med januariöverenskommelser och Tidö-avtal, där en regerings linje i olika sakfrågor ofta framstår som resultat av mer eller mindre öppen kohandel och byte av preferenser? Blir inte sådana intentioner notoriskt vaga, mångtydiga och inbördes motstridiga?

Just hyrespolitiken är ett tydligt uttryck för detta problem. Januariöverenskommelsen 2019 mellan S, MP, C och L inkluderade en punkt om att införa ”fri hyressättning vid nybyggnation” (Januariöverenskommelsen 2019, punkt 44). När utredningsbetänkandet *Fri hyressättning vid nyproduktion* (SOU 2021:50) lades fram två år senare utlöste detta en regeringskris sedan en majoritet av riksdagen hade ställt sig bakom en misstroendeförklaring mot statsminister Stefan Löfven. Då den akuta regerings-

krisen var överstånden drogs betänkandet tillbaka och frågan om förändringar i regelverket för hyresmarknaden hade åter lagts på is.⁹

Det parlamentariska motståndet mot förslagen i SOU 2021:150, där Vänsterpartiet och Sverigedemokraterna var mest drivande, drogs alltså i gång innan utredningen hade remissbehandlats, och något regeringsförslag i frågan kom aldrig att läggas fram. Vi kan därför inte heller säga något om hur statsmakternas intentioner borde tolkas om förslaget hade lett till beslut. Så enligt principen om giltighet tills vidare får vi fortfarande gå tillbaka till de institutionaliserade intentionerna bakom tidigare beslut.

POLITISKA INTENTIONER EFTER KOHANDEL

Vad kan sägas mera principiellt om politiska intentioner efter kohandel? Kohandel skiljer sig från en vanlig kompromiss, där utfallet i en viss fråga hamnar mellan två ståndpunkter längs någon tänkt skala. Vid kohandel sker i stället ett byte av frågor på så sätt att den ena parten accepterar motpartens ståndpunkt i en fråga för att i gengäld få igenom sin egen ståndpunkt i en annan fråga. Den klassiska kohandeln i svensk politik, 1930-talets uppgörelse mellan Socialdemokraterna och Bondeförbundet gav de förra långsiktigt stöd för sin ekonomiska politik och de senare för sin jordbrukspolitik, varefter båda politikområdena kom att institutionaliseras för lång tid framåt (Lewin, 1984). Om vi tänker oss att vi skulle utvärdera riksdagsbesluten i dessa båda frågor mot de politiska intentionerna skulle vi förmodligen inte hitta något om själva kohandeln i förarbetena.

Liksom den klassiska kohandeln handlade januariöverenskommelsen om regeringsmakten. Men i stället för att byta två tydliga centrala frågor kastade partierna in ett brett smörgåsbord av frågor i köpsläendet. Så utvärderarens val står här mellan att bedöma det totala utfallet av januariöverenskommelsen – som dock inte kan betraktas som politiska intentioner i parlamentarisk mening – eller att behandla varje fråga för sig med utgångspunkt från det parlamentariska materialet i var och en av dem.

9) Se vidare Bengtsson & Grander, 2023, kap. 8, där också förslagen i SOU 2021:150 granskas kritiskt.

Då det gäller frågan om marknadshyror i nyproduktionen har de politiska aktörerna även avslöjat en kohandel på lägre nivå. I samband med att SOU 2021:50 offentliggjordes framhöll justitiedepartementets utredningschef Morgan Johansson (S) denna punkt i januariöverenskommelsen som en eftergift för att kunna behålla investeringsstödet till nyproduktion av hyresrätter.¹⁰ Om de båda bytesobjekten i denna bostadspolitiska kohandel hade gått vidare till lagstiftning borde, trots bytet av preferenser i beredningsskedet, argumentationen i respektive proposition också ha kunnat tolkas som politiska intentioner i min mening.¹¹

SLUTDISKUSSION – POLITISKA INTENTIONER SOM DEN UPPENBARA LÖSNINGEN

Den som läser Evert Vedungs standardverk inser att det finns ett nästan oändligt antal modeller för utvärdering, med olika utgångspunkter och perspektiv och olika för- och nackdelar. När det gäller politiska beslut framstår ändå de institutionaliserade intentionerna som den uppenbara lösningen. Att den aktuella politiska utvecklingen har försvårat tolkningen av intentionerna är en annan sak, men Vedungs bieffektsmodell löser inte just detta problem; även den utgår alltså från politiska intentioner.

Kan inga politiska intentioner identifieras, oavsett hur vaga, är detta ett demokratiskt problem, och ”den uppenbara lösningen” på utvärderingsfrågan är då inte tillgänglig. Möjligen kan man i en sådan situation – i Nilstuns anda – i stället utgå från effekter, men då går den demokratiska förankringen förlorad.

En kohandel i beredningsskedet med tydliga bytesvaror är dock inte något avgörande hinder för att utvärdera besluten på basis av de politiska intentionerna bakom de enskilda parlamentariska besluten – men det måste då göras för varje beslut separat. Rekonstruktionen av de politiska intentionerna försvåras dock

10) Denna bild bekräftas av en gemensam debattartikel skriven av januaripartiernas bostadspolitiska talespersoner (Johansson m.fl., 2020). Centerpartiet övergav mycket riktigt investeringsstödet sedan SOU 2021:150 hade kapsejsat i riksdagen.

11) Hyrespolitiken ingår inte Tidö-avtalet mellan de borgerliga regeringspartierna och SD, som blev resultatet av den senaste breda kohandeln för att skapa ett regeringsunderlag.

om beslutsfattarna vill skyla över att frågan varit föremål för byteshandel.

En bred form av kohandel (eller ”boskapsmarknad”) som januariöverenskommelsen och Tidö-avtalet kan däremot inte utvärderas mot institutionaliserade politiska intentioner utan på sin höjd i var och en av de medverkande partiernas perspektiv – kanske i speltermen av vinnare och förlorare. Utvecklingen mot sådana lösningar har möjligen gjort det svårare att hävda att en utvärdering som tar sin utgångspunkt i lagstiftarens rimligt rationella politiska intentioner ”hämtar sin näring ur föreställningen om den parlamentariska styrkedjan och representativ demokrati”. Trots detta, och trots Everts farhågor om vaghet och mångtydighet, är det svårt att se något rimligt alternativ om man vill ta den representativa demokratin på allvar när man utvärderar politiken.

Referenser

Bengtsson, B. (1992) *Hyra och bruksvärde – om hyrespolitikens intentioner*.

Gävle: Statens institut för byggnadsforskning.

Bengtsson, B. (2022) *Marknadsmakt och besittningsskydd. Om hyrespolitikens förutsättningar*.

Stockholm: Bostad 2030.

Bengtsson, B. & Grander, M. (2023) *Bostadsfrågan som politik och intressekamp*.

Malmö: Égalité.

Birgersson, B. O. & Westerståhl, J. (1980) *Den svenska folkstyrelsen*.

Stockholm: Allmänna förlaget.

Januariöverenskommelsen (2019) *Sakpolitisk överenskommelse mellan Socialdemokraterna, Centerpartiet, Liberalerna och Miljöpartiet de gröna*.

Johansson, M., Johansson, O., Hannah, R & Hult, E. (2020) Regeringen ska utreda fri hyressättning i nyproduktion. *Dagens Nyheter*, DN Debatt, 24 april 2020.

Lewin, L. (1984) *Ideologi och strategi. Svensk politik under 100 år*. Stockholm: Norstedts.

Nilstun, T. (1984) Problemram för utvärdering. I Nilstun, T. & Hermerén, G. (Red.), *Utvärdering och rättsliga reformer. Analys av orsaker och effekter*. Lund: Studentlitteratur.

Persson, T. (2003) *Normer eller nytta? Om de politiska drivkrafterna bakom Regeringskansliets departementsindelning*. Doktorsavhandling. Uppsala: Acta Universitatis Upsaliensis.

SOU 2021:50. *Fri hyressättning vid nyproduktion*.

Vedung, E. (1991) *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur.

Vedung, E. (2009) *Utvärdering i politik och förvaltning*. Tredje upplagan. Lund: Studentlitteratur.

KAPITEL 16

Typologies of policy instruments: 25 years of developments and challenges

Frans L Leeuw

BACKGROUND

Evert Vedung is a *gentle-man*, a *craftsman-social scientist-evaluator* and a (famous) *typology-man*. In this chapter I will discuss elements of typologies of policy instruments and will address a few challenges, all related to Vedung's work. The 25 years in the title of this chapter refers to 1998, when Evert wrote his chapter on typologies of policy instruments and when I now (June 2023) write this piece.

Why call Evert a "famous typology-man"? The first answer is his authorship of the often cited typology of policy instruments, published in 1998. A meta-analysis by Acciai and Capano (2020, p. 134) studied some 10 typologies and found that "Salamon is the most frequently quoted typology in our sample, followed by Vedung (1998)... Considered individually, the Vedung classification emerges as the most adopted classification [of policy instruments] for empirical analysis". The chapter was published in the book *Carrots, Sticks and Sermons: Policy instruments and their evaluation*, edited by him, Ray Rist and the late Marie-Louise Bemelmans-Videc (1998). The book is part of the Comparative Policy Analysis Series, now

published by Routledge (and earlier by Transaction Publishers). Ray C Rist, till 2023 Chair of Inteval, was the initiator and founding father of this series.

In his chapter, Vedung analyzed existing categorizations and their difficulties and then classified three basic instruments – regulation (sticks), economic means (carrots) and information (sermons) from the perspective of their degree of coerciveness. I also published a chapter in the book *The Carrot: Subsidies as a tool of government: theory and practice* (Leeuw, 1998). It was based on a government-wide evaluation done by the Netherlands National Audit Office and in particular its Department for Performance audits and Evaluation (directed by me, between 1987 and 1995). The study looked into characteristics of all 722 subsidies implemented by the Netherlands government in the second half of the 1980's. It discussed the goal-directedness of subsidies, the implementation costs, the ex-post evaluation of them and a few other items.

It was this study and the subsequent chapter in the book that made me meet Evert, in the Netherlands and in Sweden. Evert had very good personal contacts with Andries Hoogerwerf who was then professor of public administration and policy studies at Twente University and to whom I was also related. Evert and I discovered a joint interest in evaluations, social science research and tools of government; our collaborative interest and contact continues till today.

Vedung's fame not only regards policy studies and evaluation (in general), but also the field of environmental regulation. Here, Pacheco-Vega (2020) made the point that "Vedung's typology became widely used, because it simplified a rather complex discussion on instrument choice and target actor".

The second reason why Evert Vedung is a "famous typology-man" is that he also produced other typologies. One is the typology of *12 types of use of evaluations* (Vedung, 2021). Instead of limiting himself to the three standard types (instrumental, conceptual, and legitimizing), the article "is breaking loose from this three-times-one typology of use. One deviation is adding tactical, ritual, and constitutive product uses to the prevailing triad.

Even more fundamental is the second divergence, where all six uses are joined not only to product use but to process use as well. This means that all six types of use turn up in the form of both product use and process use, leading to a six-times-two pattern instead of the three-times-one-uses-of-evaluation typology. The result is the Twelve-part typology of evaluation uses” (Vedung, 2021, p. 101).

Another example is Vedung’s (2010) *categorization of four (types of) waves through which evaluation as a profession developed*. The *scientific wave* entailed that academics should test, through experimentation, appropriate means to reach externally set, admittedly subjective, goals. Public decision-makers were then supposed to roll out the most effective means. Faith in scientific evaluation eroded in the early 1970s. It has since been argued that evaluation should be participatory and non-experimental, with information being elicited from users, operators, managers and other stakeholders through discussions. In this way, the *dialogue-oriented wave* entered the scene. Then the *neo-liberal wave* from around 1980 pushed for market orientation. Deregulation, privatization, contracting-out, efficiency and customer influence became key phrases. Evaluation as accountability, value for money and customer satisfaction was recommended. Under the slogan ‘What matters is what works’ the *evidence-based wave* implies a renaissance for scientific experimentation (see also Foss Hansen in this volume).

The list of publications of Vedung also shows an interest in a few other typologies (dealing with types of theory-driven evaluations for example) but I leave it with this.

In this chapter I will focus on the typology of policy instruments. These are the leading questions for me:

- What do we know from recent meta-studies on typologies of policy instruments? (section 3).
- Which challenges are there when working with typologies? (section 4).
- How do we deal with these challenges? (section 5).

WHAT DO WE KNOW ABOUT TYPOLOGIES OF POLICY INSTRUMENTS?

To answer this question, I will make use of two recent meta-studies. The *first meta-study* is produced by Pal (2014)¹. Pal writes (p. 131) that "the history of attempts to classify governing policy instruments begins with Kirschen (1964). His system presented 62 different types of economic policy instruments". Vedung also paid attention to the Kirschen system in his 1998 chapter.

Pal discusses a number of other typologies. One is Doern and Phidd (1992, p. 97) which distinguishes five broad categories:

- self-regulation
- exhortation
- expenditure
- regulation (including taxation)
- public ownership.

Within these broad categories, Doern and Phidd identify as many as 26 finer "graduations of choice" such as grants and subsidies, guidelines, and speeches (p. 112). Hood (1983) and Hood & Margetts (2007) developed the "NATO scheme", standing for the different resources that governments have at their disposal to effect policy change:

- N stands for nodality or information resources;
- A stands for authority;
- T stands for treasure or money; and
- O stands for organization or personnel.

Margetts (2009) applied this typology to the digital world, which was innovative. Subsequently, Leeuw & Leeuw (2012) used

1) The part of this section on the results from Pal (2014) book is largely based on <https://www.atlas101.ca/pm/concepts/typologies-of-policy-instruments/> (visited June 8, 2023).

this categorization in their paper on “Cyber society and digital policies: Challenges to evaluation”.

Salamon (2002) defined a policy tool or instrument as “an identifiable method through which collective action is structured to address a public problem” (p. 19). He was particularly interested in the degree to which modern governments are third-party governments. His list includes:

- government enterprises
- economic and social regulation
- government insurance
- public information
- taxes
- charges and tradable permits
- procurement
- contracting
- purchase of service
- tort liability (giving people rights in law to seek compensation for wrongs).

Pal summarizes the findings from his comparative analysis (p. 133): “Given this rich variety of classifications and lists, how to proceed? Vedung (1998) defined policy instruments as the ‘set of techniques by which governmental authorities wield their power in attempting to ensure support and effect or prevent social change’ (p. 21). The important thing about this definition is that it reminds us that policy, programs, and ultimately the policy instruments that give them effect are about deliberately achieving some desired outcome, and that moreover, social change is ultimately a result of human behaviour. The question of instruments, therefore, is really about the resources and techniques that governments have at their disposal to achieve certain outcomes through affecting human behaviour”.

| | Vedung (1998) | Salomon (2000, 2002) | Schneider and Ingram (1990) | Doern and Phidd (1983) | Hood (1983) | Hood and Margetts (2007) | Howlett (2000) | Howlett (2011) |
|------------------------------|---|---|---|---|---|---|--|--|
| Classification criteria | Coercion | Coerciveness, directness, automaticity, visibility | Coercion, Individual motivations | Coercion | Government resources | Degree of activity of government/ government resources | Principle use/ government resources | Purpose of tools/ government resources |
| Meaning of policy instrument | Set of techniques by which governmental authorities wield their power attempting to ensure support and affect social change. They are ends in themselves because they represent the content of the political debate | Identifiable methods through which collective action is structured to address a public problem. They define the actors centrally involved in particular type of programmes and the formal role they will play | Techniques that address the problem that people are not taking the actions needed to ameliorate social, economic or political problems | Means for the management and manipulation of legitimate coercion | Means through which governments attempt to shape the lives of their constituents. They are about social control | Means through which governments attempt to shape the lives of their constituents. They are about social control | The techniques and devices governments use to implement policies and achieve their goals | The techniques and devices governments use to implement policies and achieve their goals |
| Type | Regulation, Economic, Information | Because of the multidimensionality of policy instruments, no single classification is possible, and these schemes will differ depending on which facet is used as the basis | Authority, Incentive, Capacity building, Symbolic, Learning | Self-regulation, Exhortation, Expenditures, Regulation (including taxation) public ownership. | NATO: Nodality, authority, treasure, organization | NATO Nodality, authority, treasure, organization | Positive/negative procedural NATO instruments | Procedural/substantive NATO instruments |
| Tool dimension | Action content: what the target population should do. Authoritative force: degree of power government employs to obtain compliance | Type of good delivery; delivery vehicle; the type of organization designated to provide the service; a system of rules defining the relationship among the actors involved | Different assumptions about how policy-relevant behaviour can be fostered by the target population by providing: legitimization of authority; tangible payoffs (positive or negative); resources to enable individuals to make decisions; alternative values and beliefs (persuasion) | Visibility, historical context. politicians prefer less coercive instruments | Detectors: instruments (used to gather information) Effectors (used to modify behaviour) | Detectors: instruments (used to gather information) Effectors (used to modify behaviors) | Procedural instruments indirectly influence policy outcomes through the manipulation of the policy process | What they do: substantive instruments modify the distribution of goods and services How they operate: procedural instruments indirectly influence policy outcomes through the manipulation of policy processes |

Table 1. Eight typologies more closely analyzed by Acciai and Capano (2020)

My second source is the *meta-analysis* published by Acciai and Capano (2020). They looked into 296 studies from the existing literature and applied qualitative content analysis to these studies. The subject matter of their study was the following overview of typologies (table 1).

The background of their study is that “although policy instruments have become increasingly central in public policy, it remains unclear whether their typologies have been used to explain real-world phenomena and, if so, how and to what extent they have been differentially applied according to the research goal and the policy sector under investigation”. Their study “provides a review of the most frequently used policy instrument typologies and, through a meta-analysis, it analyses how instrument

typologies have been differentially adopted to explain real-world phenomena. The results are:

- a high degree of heterogeneity in citation frequency;
- the polyhedral nature of the concept of policy instruments;
- a divide between typologies focused on governmental resources and those
- focused on drivers of expected behavior, and
- a practice of “labellism” (the authors do not define this concept but it boils down to persons that have not empirically investigated the original shape of the typology, but did “redesign and adapt [them] according to the individual preferences of the researchers and their research design”).

From that Acciai & Capano (2020, p. 133) draw the conclusion that there is “the urgency of a process of convergence towards a common framework”.

CHALLENGES OF WORKING WITH TYPOLOGIES

On the basis of my largely *informal* scanning of the literature, I see the following challenges that are (still) ‘hanging around’ the development and use typologies of policy instruments.

Missing informal policy instruments: Van Nispen (n.d.) made a point of “informal instruments” that are (not) addressed in these categorizations. He refers to “the change of membership or size of the decision making body, bargaining and negotiation processes, mobilization of political support as well as deceit and double crossing and even threats”. There is no place in the typologies for this type of ‘under water’ or ‘under the radar’ operating mechanisms. Although I agree with van Nispen on the importance of informal mechanisms, sometimes they are examples of formal instruments like ‘information’ or ‘regulation’. This could apply to his ‘change of membership’ instrument, though of course not when deceit and double crossing are at stake.

Assuming 'pure form instruments' while policy (instruments) mixes seem to be more 'talk of the day': Van Nispen also made the point that “policy instruments rarely appear in a pure form. They come bundled in programs combining various tools, each being a package containing a type of good or activity, a vehicle for the delivery of this good or activity and a delivery system”. This point relates to the idea of policy mixes. For evaluators these packages sometimes create a bit of a problem: what is the relative ‘weight’ of the respective components (= singular instruments) of the packages, when one wants to dissect packages in order to learn from them being evaluated. Pawson’s (2013, p. 128ff) “conceptual platform for behavioral change interventions” can help in unpacking these components.

Assuming the need for a common framework for typologies: According to Acciai and Capano (2020) there is a need for a common framework of typologies (or even a common typology). Their position raised the question why would working with a common framework be better, compared to having several frameworks available? Policy interventions change and they do so quite rapidly. Take as an example the development of neuro-interventions, neurolaw and – of a different nature – ChatGPT. Neuro-interventions are applied in the field of health, justice and education, but didn’t exist as ‘policy instruments’ 10 or 15 years ago. For neurolaw there is a similar story. ChatGPT has landed on the shores of policy making and implementation and is already capable of designing robots, creating music, writing essays and policy papers. Soon, ChatGPT will be able to design new policy instruments. Why would it be necessary to have, in a fluid world, only one common framework typology?

The risk of a Tower of Babel syndrome: Capalo and Engili (2020) make the point that “despite 40 years of scholarly debate and a flourishing of typologies, we are still at a loss when selecting one typology over another. (---) While labels of instrument categories sound similar from one taxonomy to another, they are nevertheless conceptualized in very different ways”. As Linder and Peters (1989, p. 40; see also Ingram and Schneider (1988)) state,

“there is as much variance within them as between them’. Capalo & Engli (2020) even refer to “jargonistic confusion” and illustrate this for two instruments: authority and capacity building². They see a risk of creating a Tower of Babel syndrome. Their answer to take care of this syndrome is simple: refrain from making new typologies.

Typologies are not theories: Although typologies are relevant and often needed to create order out of chaos and help decomplexify reality, from an epistemological perspective, typologies do not explain how the tools of government work and/or what the drivers (or mechanisms) are of subjects to comply with the instruments. Nor do they explain why some instruments are selected and not others in different types of situations. Therefore, theories are needed. Again, Vedung (1998) was an early adopter on this issue, as he in his 1998 chapter paid serious attention to theories and how they differ from typologies. This is how he put it: “A final test of a good classification is that it proves empirically and theoretically fruitful. Will the threefold classification pass this test? This must be left for future evaluations”. But he added that “in the present context, consider a few arguments in favor of the scheme”.

HOW TO DEAL WITH CHALLENGES?

A chapter in a Festschrift, even if that is written on behalf of a great thinker and typology-man, does not allow for a full-swing analysis of this issue. I will therefore focus on only two issues: (a) The assumption of working with pure instruments while instead mixes are ‘talk of the day’ and how to deal with that, and (b) the point that typologies are no theories and how to help integrate typologies with theories (on mechanisms underlying behavior and behavior change).

2) Capali & Engini (2020, p. 13) explain this as follows. “Governmental capacity to state rules and principles has been at the core of the traditional ‘command’ and has a legitimate home in all typologies—McDonnell and Elmore’s mandate, Vedung’s regulation, Hood’s authority, Schneider and Ingram’s authority and Salamon’s coerciveness. However, these categories do not include the same range of instruments. Sanctions are classified in Schneider and Ingram’s typology as incentives, while Vedung views them as regulations. A tariff is considered an economic instrument for Vedung but an authority instrument for Schneider and Ingram. One needs to return to the organising principle of both typologies to understand why Schneider and Ingram classify it as an incentive (the sanction induces a complying behaviour in order to avoid its application; the tariff stipulates a rule) and Vedung as a regulation (sanction as coercion)”.

Policy mixes rather than (singular) policy interventions are talk of the day: In fact, this challenge boils down to what is known as *complexity problems*. In the Good Old Days, governments worked with few instruments, that took years to be designed and decided upon and that had often a (rather restricted) coverage.

Over the last decades, things have changed. More policy instruments, more policy mixes, more need for ‘quick solutions’ and ‘fixes’. See as only one example the recent Covid 19 situation. The whole *gamut*, if not *shebang* of instruments ranging from *compulsory vaccination programs, lockdowns, digital covid certificates, covid social distancing apps, continuous governmental fear-arousal communication, streamlining if not deleting certain social media communications (on the basis of the assumption that it concerned ‘fake news’); subsidies for private organizations to continue when lockdowns were installed; building new hospital structures and others*, were implemented in a very short time period and based on hardly any (empirical, theoretical or practical) evidence.

For the Netherlands it was also made perfectly clear by the independent Netherlands Safety Board that the government’s interest in evaluating the impact (and costs and side effects) of this bag of interventions, was limited to zero³, let alone that there was any interest in the underlying theories on mechanisms that could possibly made the interventions ‘work’.

To evaluate how policy mixes work is a complicated task. Learning from complexity studies (Bamberger, Vaessen & Raimondo, 2017) and applying Pawson’s VICTORE model to unravel relationships between instruments, their history and competition and their working mechanisms, would probably help. See below for a shortcut of Pawson’s VICTORE model (based on Leeuw & Waltermann, 2022).

3) See the three reports on the Covid governmental response by the Dutch Safety Board (available on www.safetyboard.nl).

- **Address Volitions:** people make (behavioural) choices. Their agency is volatile and often unpredictable.
- **Do not forget Implementation:** what you see is not always what you get; implementation often implies a long-term chain of activities.
- **Contexts are important:** they are partly given and partly created while the intervention/arrangement is implemented; context addresses macro and micro issues.
- **Time is important to take on board:** interventions always come from somewhere and have a 'history' (in the minds of people, addressees). This adds to the complexity. Path dependency is also a feature of time.
- **Outcomes:** they can be desired, undesired (including adverse); they can be proximal or distal.
- **Rivalry:** during implementation of arrangements/interventions, dozens of other interventions and arrangements already exist and are relevant for the *Umfelt* of the one that is studied. There can be, and often is, rivalry between multiple existing interventions.
- **Emergence:** society, individuals and events change during the implementation process.

Table 2. A shortcut of Pawson's VICTORE-model

Typologies are no theories but integrating theories and typologies is possible, already done (in some fields) and relevant: Vedung (1998) discussed this issue, as I mentioned earlier. Schneider and Ingram (1990) made a substantial attempt to address this gap between theories and typologies and explore the behavioral assumptions underlying the tools. Olejniczak, Sliwowski and Leeuw (2020:4) presented this as follows. "Schneider and Ingram (1990) proposed five broad types of tools according to assumptions about the compliance *mechanisms* of policy subjects. These are:

- Authority tools that rely solely on legitimacy. Policy targets are expected to do what they are told. They are driven by their citizens' duty and inherent commitment to obey laws.
- Incentive tools rely on tangible payoffs that could be positive or negative. Policy addressees are portrayed as utility maximizers who will change their behavior driven by the net tangible payoffs.

- Capacity tools provide information, education, and resources to enable policy subjects to make decisions and carry out activities. Once deficits in information or resources are addressed, subjects will behave as expected.
- Symbolic and hortatory tools guide towards desired behavior by manipulating symbols and influencing values. People are motivated to take action based on their beliefs and values grounded in culture and social context. They will comply if the desired behavior is consistent with their beliefs.
- Learning tools assume that policy addressees do not know what needs to be done, or what is possible to do, and therefore policy tools are being used to promote learning, consensus building, and lay the foundations for improved policy.

In the paper Olejniczak, Sliwowski and Leeuw (2020) focus on the underlying behavioral assumptions of policy tools and presented a framework addressing the behavior of addressees, identifying problems that obstruct compliance, and choosing a tool with a distinctive change mechanism.

Although focused on health policies and health behavior, work done by Kok et al's (2016) developed a taxonomy of *Behavior Change Methods* and related them to their *Intervention Mapping approach*. They engaged in the description of the interventions as well as in articulating the relevant (scientific) behavioral (change) theories: "Recent attempts to establish a cumulative science of behaviour change have used taxonomies of behaviour change techniques to derive effectiveness of such techniques through meta-analysis of intervention evaluations... These have been widely applied and developed further, for example, geared towards physical activity and diet. Useful though they are, being taxonomies that are developed as a means to categories intervention content, 'evidence of "efficacy" or "effectiveness" is not part of the definition of BCTs'. This means that although these taxonomies contain effective behaviour change methods (techniques), they also contain ineffective methods, and may even contain counter-effective methods" (Kok et al, 2016).

Whereas taxonomies for coding interventions generally consist of descriptions of BCTs, taxonomies for developing interventions need to acknowledge a number of additional aspects of the dynamics of behaviour change. One of them is directly related to the role of theories explaining (and predicting) behavior: to realize assurance that when interventions/ tools are developed and implemented, the determinants (I would rather refer to mechanisms) addressed, “indeed predict the relevant behaviour”.⁴

CONCLUSION AND A PERSONAL WORD

Decades ago Evert Vedung started to work on policy instruments and typologies. Till today, this topic is still relevant and in need of further thinking. Issues raised by him, like the relationship between theories (that explain) and typologies (that order), 25 years ago, continue to be on the agenda. This chapter has mentioned some five challenges. At the same time, meta-studies allow for a better understanding of the (relative) importance of typologies, while evaluators and social scientists have also taken on board how to integrate theories with typologies. This chapter presented some examples.

Evert, I thank you for our discussions over the years and for your insights that you are always willing to share. But more than that, I have a vivid memory about your role as ‘non-certified medical advisor’, who ‘rescued’ me from a terrible flu that started the day before I had to give a keynote to the Danish evaluation society years ago. You gave me a pill, that you got from a Korean medical colleague, when you were visiting professor over there. With your reassuring friendly voice, you told me: “take this and in an hour you will feel like reborn (and will give an excellent speech). But, Frans, take it only once in your life.” I did. And survived.

Evert, all the very best for you and your loved ones.

4) See <https://osf.io/bpxwq> for in-depth information on the Taxonomy of Behavior Change Methods and Intervention Mapping Approach.

References

- Acciai C. & Capano G. (2021). Policy instruments at work: A meta-analysis of their applications. *Public Admin*, 99, 118–136.
- Capalo, G. & Engili, I. (2021). Using instrument typologies in comparative research: conceptual and methodological trade-offs. *Journal of Comparative Policy Analysis: Research and Practice*, 24, 2, 99-116.
- Doern, G. B., & Phidd, R. W. (1992). *Canadian public policy: Ideas, structure, process* (2nd ed.). Toronto, ON: Nelson.
- Doern, G. B., & Wilson, V. S. (Eds.). (1974). *Issues in Canadian public policy*. Toronto, ON: Methuen.
- Hood, C. (1983). *The tools of government*. London, UK: Macmillan.
- Hood, C., & Margetts, H. Z. (2007). *The tools of government in the digital age*. New York, NY: Palgrave Macmillan.
- Kirschen, E. S., et al. (1964). *Economic policy in our time* (3 vols.). Amsterdam, Netherlands: North-Holland.
- Kok, G. et al (2016). A taxonomy of behavior change methods: an intervention mapping approach. *Health Psychology Review*, 10, 3: 297-312.
- Leeuw, F (1998). The carrot: Subsidies as a tool of government—theory and practice. In Bemelmans-Videc, M. L., R. C. Rist & E. Vedung (Eds.), *Carrots, sticks and sermons* (pp. 77-102). Routledge. New Brunswick: Transaction Publishers.
- Leeuw, Frans L & Leeuw, Bastiaan (2012). Cyber Society and digital policies127: Challenges to evaluation. *Evaluation* 18(1): 111–127.
- Leeuw, F. L., & Waltermann, A. (2022). On identifying assumptions underlying legal arrangements: Some methodological considerations. *Law and Method*, (2), 1-24.
- Margetts H (2009). The internet and public policy. *Policy & Internet* 1(1): 1–22.
- Nispen, Frans van (n.d.). *Policy Instruments*. Paper, Robert Schuman Centre for Advanced Studies, European University Institute.
- Olejniczak, K. liwowski,P. & Frans Leeuw (2020). Comparing Behavioral Assumptions of Policy Tools: Framework for Policy Designers. *Journal of Comparative Policy Analysis: Research and Practice*, 22, 6, 498-520.
- Pcheco-Vega R. (2020). Environmental regulation, governance, and policy instruments, 20 years after the stick, carrot, and sermon typology. *Journal of Environmental Policy & Planning*.
- Pal, L (2014). *Beyond Policy Analysis – Public Issue Management in Turbulent Times*, Fifth Edition. Toronto: Nelson Education.
- Pawson, Ray (2013). *The Science of Evaluation*. London: Sage.
- Salamon, L. M. (Ed.) (2002). *The tools of government: A guide to the new governance*. New York, NY: Oxford University Press.
- Stewart, J. (2009). *Public policy values*. New York, NY: Palgrave Macmillan.
- Vedung, E. (1998). Policy instruments: Typologies and theories. In M.-L. Bemelmans-Videc, R. C. Rist & E. Vedung (Eds.), *Carrots, sticks and sermons: Policy instruments and their evaluation* (pp. 21–58). New Brunswick, NJ: Transaction.
- Vedung, E. (2010). Four waves of evaluation. *Evaluation* 16, 3, 263-277.
- Vedung, E. (2021). The Twelve-Part Typology of Evaluation Uses. *Zeitschrift für Evaluation* 20(1), 101-130.

KAPITEL 17

Mångtydiga idéer och den haltande implementeringens rationalitet – fallet deltagandebaserad samhällsplanering

Marianne Danielsson & Nils Hertting

Med distinktionen mellan ”innehållslig” och ”funktionell” idéanalys påpekade Evert Vedung redan på 1970-talet (1977; 2018) att idéer inte bara spelar en roll i en abstrakt och filosofisk tankevärld. Idéer fyller också en politisk *funktion* som orienterande mallar för politikens verkliga aktörer och vardagliga handlande. En idé och tankefigur som kommit att bli dominerande i ett sammanhang kan i så fall definiera och konstituera inte bara vad som är ett problem och vad som en lämplig lösning på problemet. Mer eller mindre direkt kan idén också peka ut vilka aktörer och grupper som problem placeras hos, vilka som antas ha resurser och positioner att genomföra lösningar och vad som krävs i form av styrning och styrmedel för att dessa ska bli genomförda (Danielsson 2010, s. 148ff.). I Vedungs distinktion ligger med andra ord tanken att vi i en idéns innehåll kan finna förklaringar till politiskt intressanta utfall.¹

1) Vedung har över tid växlat mellan formuleringen ”funktionell idéanalys” (1977) och ”förklarande idéanalys” (2018). Vi använder termerna här som synonymer för samma idéanalytiska begrepp.

Det är knappast förvånande att de idéer vars funktion får mest uppmärksamhet är de som får stort genomslag på centrala beslut och som leder till samhällsförändringar. Men det finns också politiska idéer som är beständiga över tid utan att någonsin få ett bredare stöd och som är hänvisade till en marginaliserad roll i maktens periferi. En tredje kategori av idéer är vitt omfattade och har en stabil popularitet över tid i politikens centrum men får ändå inte någon riktigt reell betydelse. Sådana idéer åberopas brett och ofta som vägledande mallar för gemensam politisk handling, men blir likväl aldrig omsatta i verkningfull praktik.

Idén om en deltagandebaserad samhällsplanering tycks vara ett exempel på den senare kategorin. Tillskyndarna tycks aldrig sina och i policy- och planeringsdokument återkommer ständigt uttryck för idén (se t.ex. Boverket, 2021; SKR, 2023). Samtidigt är den empiriska litteraturen om deltagarplanering närmast ett frosseri i ouppnådda ambitioner (Bäcklund & Mäntysalo 2010; Ganuza & Baiocchi, 2012; Kugelberg 2017). Försöken att implementera ett ökat deltagande är visserligen hela tiden återkommande men kännetecknas bäst av sina perifera, marginella eller instabila praktiker. Vi har således att göra med en idé som lyckas förbli både uppburen och återkommande, trots att den ter sig notoriskt svår att implementera och omsätta i resultat.

I detta kapitel ska vi hävda att problemet med att implementera deltagandebaserade samhällsplanering är en *funktion* av och kan *förklaras med* idéns själva *innehåll*, och närmare bestämt med dess *inneboende mångtydighet* och *spänningsfyllda relation* till en annan dominerande idé i västerländska demokratier. Vårt *vedungskt* inspirerade – och impregnerade – bud på förklaring till vad som skulle kunna kallas den haltande implementeringens rationalitet kombinerar med andra ord en tolkning av idéns innehåll med en tolkning av bärande idéer i dess kontext.²

2) Vårt argument kan formuleras i relation till den i samtida planeringsforskning populära "post-politiska" teorin om att deltagarplanering syftar till att avpolitisera en "nyliberal" planering och för den politiska eliten aldrig på allvar syftat att skapa nya möjligheter för deltagande (jfr Walker, McQuirre & Lee 2015; Dahlstedt, Kings & Tahvilzadeh 2018). En sådan förklaring är förmodligen rimlig i många fall. Men vårt argument är ett annat. Också när politiker främjar deltagande som mer än en symbolisk gest är idén svårimple-menterad eftersom den i sig är mångtydig och motsägelsefull, och därtill utmanar andra grundläggande idéer i den representativa demokratin.

FYRA TOLKNINGAR AV IDÉN OM DELTAGARBASERAD SAMHÄLLSPANERING

Populära idéer i offentlig politik är inte sällan både otydliga och mångtydiga (Pollitt & Hupe, 2011; Vedung, 2016). Idén om deltagarbaserad planering tycks vara ett gott exempel, och nedan presenterar vi en fyrdelad tolkning av idén om deltagandebaserad planering som var för sig ger rimliga och goda argument för ett ökat deltagande i samhällspaneringen. Syftet är att föreslå idealtyper för att beskriva den mångtydighet som finns i policyidén och därmed också ge en antydning om hur olika typer av aktörer i planeringsprocesserna – de ”radikala aktivisterna”, de ”reflekterande samtalarna”, de ”pragmatiska planerarna” och de ”förtroendevalda ombuden” – alla samtidigt kan skriva under på idén om deltagandebaserad planering, men med olika förväntningar på de roller och processer som följer av denna idé. Genom att renodla vad vi menar är fyra återkommande men distinkt olika *delidéer* om deltagande i samhällspaneringen försöker vi här visa hur idéerna som mallar för handling och interaktion drar åt olika håll. Även om idéerna var för sig har fog för sig och inte heller är logiskt oförenliga, det är vårt argument, så formuleras de alltför ofta med för lite uppmärksamhet på praktiska lösningar för hur de ska samexistera och komplettera varandra.³

Den *första idén* rekonstruerar vi med hjälp av demokrati-teoretiskt tankegods om deltagandebaserad samhällspanering. Vi kallar den *artikulativ* eller *intresseorienterad*. Deltagandet är här ett sätt att ge olika intressen – inte minst sådana som uppfattas som marginaliserade – en möjlighet att *direkt komma till tals*. Det grundläggande problemet i planeringen är utifrån denna idé bristen på representation av den mångfald av åsikter och intressen som finns i samhället. För att få demokratisk legitimitet och för att möjliggöra politisk respons på medborgarnas behov och preferenser (Dahl, 1971; Teorell, 2006) måste de traditionella kanalerna för representation och artikulation av intressen i samhällspaneringen

3) Albert O. Hirschman (1991, s. 152) – en av Everts favoriter – har kallat fenomenet med en överdriven tilltro till att idéer är förenliga eller ömsesidigt förstärkande bara för att de är goda för ”the synergy illusion”.

kompletteras. Genom att öka deltagandet bland marginaliserade intressen som annars riskerar att inte komma till tals, tänker man sig, görs det politiska systemet mer lyhört. Ambitionen kan formuleras som att kompensera för hegemoniska tendenser och strukturella orättvisor genom att tillhandahålla nya arenor där deltagare självständigt kan uttrycka sina intressen – i relation till andra deltagargrupper, det politiska systemet eller både och.

För den enskilde medborgaren blir ett ”artikulerat” deltagande i planeringsprocesser en slags motsvarighet till att gå med i en intresseorganisation eller ett politiskt parti. För politiker innebär det att de politiska uppdragen och mandaten i högre grad ges i och genom planeringens deltagarprocesser och av deltagarna i dessa, snarare än genom partiprogram och väljarstöd. På motsvarande vis innebär det för tjänstemän och planerare att de blir förmedlare av samhällsgruppers viljor i relation till politiker, snarare än genomförare av politikernas beslut (Danielsson, Hertting & Klijn 2017; Hertting & Hellquist 2022).

Också den *andra idén* om deltagandebaserad samhällsplanering finner sin tolkning med hjälp av demokratiteori. Vi kallar den *deliberativ* eller *reflexiv* planering. Här är tanken att deltagandet skapar en bättre förståelse mellan individer och grupper med olika perspektiv på och erfarenheter av att till exempel bo eller verka i en stadsdel. Om den artikulerativa idén handlar om att ge röst åt och prioritera mellan *olika* intressen som deltagarna har med sig in i processen så handlar den deliberativa idén om att genom processen skapa samförstånd och samhörighet, om än inte nödvändigtvis enighet. Man skulle kunna säga att den deliberativa idén ser polarisering i lokalsamhället som ett grundläggande problem för samhällsplaneringens möjligheter. Lösningen handlar om att, som en del i planeringen, skapa former där gemensamma ställningstaganden formuleras och rättfärdigas i en dialog som bygger på erkännande av, och förståelse för, andras perspektiv, erfarenheter och utgångspunkter. Deltagandet är därför utformat som ett samtal där man ”lyssnar på den andra sidan” (Mutz, 2006) och ”resonerar tillsammans” (Gutman & Thompson, 1996) i syfte att formulera nya gemensamma lösningar som är acceptabla för alla

(jfr Landwehr, 2010). En sådan process kräver att de medborgare som deltar i planeringen inte bara uttrycker sina preferenser, utan också anstränger sig för att förstå andras perspektiv och önskemål. Planerare och lokala politikers roll handlar om att både tillhandahålla processer där sådana värden tillgodoses, men också om att själva delta som lyssnande och resonerande aktörer (Danielsson, Hertting & Klijn, 2017). På så vis, tänker man sig, att mer legitima och därmed också ”rationella” ställningstaganden i samhällsplaneringen formuleras (Elster, 1998; Chambers, 2003).

Den *tredje idén* om deltagandebaserad samhällsplanering kallar vi för *funktionell*. Till skillnad från de två demokratiteoretiska idéerna ovan hittar vi denna i litteraturen om samhällsstyrning och *governance*. Förvisso delar den funktionellt deltagarbaserade planeringen den deliberativa idéns fokus på ”kollektiv rationalitet”, men den utmärker sig genom sin administrativa och pragmatiska snarare än demokratiska och principiella logik. Snarare än bristfällig kommunikation, avsaknad av ett integrerande samtal och gemensamt formulerade problemföreställningar handlar grundproblemet här om bristfällig *samordning* av lokalsamhällets resurser, verksamheter och aktörer (jfr Hertting & Vedung, 2012). Utifrån denna idé ses medborgarnas deltagande som ett verktyg för att mobilisera lokal expertis och utnyttja lokalsamhällets resurser mer optimalt i planeringen (Klijn & Koppenjan, 2016). Boende, politiker och planerare samspekar utifrån idéen i processer för en *apolitisk* problemlösning som antas gynna alla och *ex post* vinna legitimitet genom att leverera effektiva lösningar på medborgarnas *behov* (Scharpf 1997).

Den *fjärde och sista idén* om deltagandebaserad samhällsplanering kan karaktäriseras som *instrumentell*. Även om deltagarna måste få något i utbyte för sin medverkan är den primära bevekelsegrunden – eller *rationalen* – att beslutsfattare och planerare genom ett ökat deltagande ges bättre möjligheter att implementera intentioner och planer som redan har formulerats (Danielsson, Hertting & Klijn, 2017). Närmare bestämt kan man urskilja två varianter av den instrumentella idéen (Hertting & Hellquist, 2022). Deltagande anses underlätta en instrumentell implementering

top down dels genom att skapa *ökad kunskap* om lokala förhållanden av vikt för planeringen, dels genom att skapa *förankring* och *undanröja motstånd* som annars skulle försvåra implementeringen. Om deltagande utifrån den funktionella idén handlar om utveckla arrangemang för *samordning* på tvärs olika verksamheter med olika mål handlar det utifrån den instrumentella idén om ökad förmåga till *realisering av centralt formulerade mål*. Oavsett om det primära syftet är att inhämta information eller påverka attityder, så syftar deltagandet här till en effektivare implementering av beslut som redan antas ha skänkts demokratisk legitimitet genom den representativa demokratis reguljära institutioner.

Syftet ovan har varit att analytiskt beskriva att idén om deltagandebaserad planering närmare bestämt kan tillskrivas flera olika idéer. Empiriska studier i svensk kontext har också visat att den deltagandebaserade planeringens idé tenderar att bli mångtydig på ett sätt som låter sig beskrivas och tolkas med dessa begrepp (Danielsson & Hertting 2007; Gustafson & Hertting 2016; Hertting & Hellquist 2022). Vi återkommer nedan till vårt bud på hur den mångtydiga idén om deltagandebaserad planering i sin tur kan förklara kombinationen av en brett omhuldad idé i en haltande praktik. Först ska vi peka på hur den mångtydiga idén också riskerar att komma i konflikt med en annan väletablerad idé i samhällsplaneringens kontext.

IDÉN OM REPRESENTATIV DEMOKRATI SOM DEN DELTAGARBASERADE SAMHÄLLSPANERINGENS KONTEXT

Vårt argument här kan formuleras enkelt: Försöken att integrera mer direkta former av deltagande i samhällsplaneringen riskerar att komma i konflikt med den representativa demokratis normer och grundläggande idé. Utifrån en renodling av den representativa demokratimodellen får samhällsplaneringen sin demokratiska legitimitet genom en kontrollerad delegering av politiska mandat – från medborgares deltagande i val via valresultatens transformering till partimandat i folkförsamlingar och från mandat i folkförsamlingar till regeringsmakt.

Den representativa demokratins idé, dragen till sin spets, begränsar medborgardeltagande i syfte att uppnå politisk jämlikhet (Alonso, Keane & Merkel, 2011, Held, 2006). Politisk jämlikhet är möjlig i öppna och fria politiska val, där varje röst räknas lika. Det är den centrala tanken. Varje djupare inblandning i politiken innebär att ojämnt fördelade förmågor och resurser gör avtryck. Bara i en "tunn" demokrati, tänker man sig, kan effekterna av skillnader i medborgares materiella, sociala och intellektuella resurser minimeras (Parvin, 2021; jfr Danielsson, Hertting & Klijn, 2017).

Som idé om demokratins förutsättningar får den representativa demokratimodellen också implikationer för politikens *implementering* (Manin, 1997; Danielsson, Hertting & Klijn, 2017; Hertting, 2023). Vi preciserar vårt i grunden enkla argument genom att skilja mellan idén om röstningscentrerad representativ demokrati på det politiska systemets inputsida och idén om hierarkisk implementering på dess outputsida. Distinktionen görs för att förtydliga kopplingen: hierarkisk implementering har ett demokratiskt värde i ljuset av den representativ demokratins idé. Implementeringens hierarkiska karaktär är alltså kopplad till en idé om hur demokratisk legitimitet skapas och demokratiskt ansvar upprätthållas vars utgångspunkt är medborgarnas deltagande i öppna och allmänna val.

Utifrån ett sådant synsätt är det förtroendevaldas beslut – inte direktdeltagande aktörers prioriteringar – som ska *driva* implementeringen. Eftersom den representativa demokratin antas säkra medborgarnas politiska jämlikhet genom valmekanismen så kräver den som reell politisk ordning ett genomförande av politiska beslut *top-down*. För att medborgarna ska kunna hålla förtroendevalda ansvariga och för att politikerna ska kunna utkräva ansvar av implementerande aktörer krävs att det finns väl fungerande vertikala länkar som inte otydliggörs eller undermineras av deltagande "på tvärs" (Hertting, 2023; jfr Lowi, 1969).

Även om de mer precisa idéerna om ett ökat deltagande i planeringen tycks skilja sig åt i grad av förenlighet med den representativa demokratin – i annat sammanhang har vi hävdad att den artikulatoriska är den som framstår som mest främmande och den

instrumentella som mest kompatibel (Danielsson, Hertting & Klijn, 2017) – menar vi att varje försök att implementera idéerna om nya former för deltagande *riskerar* att utmana den representativa demokratins idé och de aktörer som ser sig som företrädare för och försvarare av denna. I kontexter som den svenska, där den representativa demokratins institutioner har ett förhållandevis starkt förtroende, kan vi också förvänta oss att den deltagandebaserade planeringen kommer möta särskilt stor misstänksamhet (jfr Hertting & Klijn, 2017).⁴

DEN HALTANDE IMPLEMENTERINGENS RATIONALITET

I det här kapitlet skissar vi på en förklaring till svårigheterna att implementera den deltagandebaserade samhällsplaneringens idé (och mer generellt fenomenet med populära politiska idéer som är svåra att omsätta i praktik). Vi gör det genom att sätta själva idéns innehåll i fokus och genom att uppmärksamma dess många olika idéstråk och dess spänningsfyllda förhållande till den grundläggande idén om representativ demokrati som utgör den kontext där deltagandebaserad planeringen ska implementeras. Närmare bestämt kan vi avslutningsvis peka på två vägar – två mekanismer – genom vilka planeringens mångtydiga innehåll förklarar en haltande implementeringspraktik. Den första formuleras utifrån ett *policyperspektiv* och den andra utifrån ett *spelperspektiv* (jfr Bengtsson, 1995). Policyperspektivet pekar på den haltande praktiken som en överordnad huvudmans *strategi*; spelperspektivet på den haltande praktikens som ett *utfall* av interaktioner mellan strategiska aktörer.

Policyperspektivet sätter fokus på inbyggda målkonflikterna i beslutsfattarnas intentioner. De fyra tolkningarna av deltagandepanering ovan är alla rimliga och legitima var för sig. Sammantagna borgar de också för ett brett stöd. Samtidigt innehåller de

4) I sammanhanget är det kanske talande att kommunstyrelseordföranden i Sverige är mer negativt inställda till deltagardemokrati jämfört med borgmästare i 16 andra länder (Bäck 2006), att fullmäktigeledamöter i Sverige är betydligt mer negativa till inslag av dialog och deltagande i kommunpolitiken än svenska medborgare generellt, samt att ledamöter som tillhör styrande partier är talande nog också mer negativa till ett ökat inslag av deltagande (Gilljam, Karlsson & Sundell 2010).

element som delvis är motstridiga. Också idén att korrigera den representativa demokratins vertikala ansvarandelogik med någon form av deltagande ter sig både rimlig och populär. Men inte heller detta förefaller spänningsfritt.

Utifrån föregångare som Eisenberg (1984), som talar om ”strategisk vaghet”, Vedung (2016, s. 70f.), som talar om en ”främjande otydlighet”, och utifrån Karl Weicks (1976) teori om ”loose couplings” vill vi här peka på vad som skulle kunna beskrivas som den haltande implementeringens rationalitet: Politiska beslut kan formuleras mångtydigt och implementeringen avsiktligt tillåtas härbärgera motstridiga värden i syfte att överhuvudtaget få någon förändring till stånd (jfr Hertting & Vedung, 2009). För att mobilisera stöd och undvika politiska och administrativa låsningar kommer beslutsfattare utifrån en sådan tolkning att fatta vaga beslut men också acceptera eller rent av bejaka en haltande – eller snarare hattig – implementering. Att söka stöd för en mångfacetterad och brett tilltalande idé, otydlig nog att förena grupper med olika önskemål och utan att på förhand precisera var eller exakt hur det direkta deltagandet ska komplettera den hierarkiska implementeringen, blir då en *strategi* för att alls kunna handla. Den haltande implementeringen skulle alltså bäst förstås och förklaras som en integrerad del av strategin för att förnya samhällsplaneringen genom att – över tid och från fall till fall – stegvis och växelvis göra olika situationsanpassade avvägningar i betoningen av deltagandes inriktning och karaktär (Thacher & Rein, 2004). Spåren av en sådan pragmatisk policyidé kommer knappast låta sig uttydas som ett konsekvent och programmatiskt skifte mot en mer deltagarbaserad planering.⁵

Utifrån ett *spelersperspektiv* är den haltande implementeringen inte en strategi utan ett *utfall*. Mekanismen sätter fokus på de komplicerade förhandlingar som aktualiseras och aktiveras i den mångtydiga idéstrukturens hägn. I denna *konstitutiva* idékontext mobiliseras en mängd aktörer – professionella och frivilliga – som med hänvisning till sina egna favorittolkningar av idén om deltagandebaserad planering rättfärdigar och legitimerar sitt deltagande och sina olika förväntningar på processerna. I en falang finner vi de

5) Vi har i annat sammanhang formulerat en snarlik tankegång i termer av den ”begränsade institutionaliserings logik” (Danielsson, Hertting & Klijn 2017).

skeptiker och kritiker som ser ett ökat deltagande som problematiskt i ljuset av den representativa demokratis idé. På sin höjd kan de tänka sig en instrumentell form av deltagande, men så fort idén ska omsättas i handling i samspel med aktörer som vill uppnå andra syften blir projektet obekvämt. I en annan falang finner vi ett knepigt spel mellan aktörer som alla vill se mer "horisontella" former av deltagande men som samtidigt sinsemellan skiljer sig åt med avseende på vilken typ av deltagande de vill se – artikulativt, deliberativt eller funktionellt. När kritiker och olika typer av anhängare ska omsätta den deltagandebaserade planeringens idé till praktik får vi ett krävande spel, i vars ljus den haltande implementeringen framstår som ett rimligt och förväntat utfall.

* * *

Inspirerad av Evert Vedungs blick för historiska ironier och motsägelsefullheter kan vi slutligen formulera kapitlets argument som ett paradoxalt dilemma: *Om idén om deltagandebaserad planering drivs ambitiöst nog för att kunna få reella effekter så kräver det prioriteringar och avstämningar, som i sin tur riskerar underminera den breda mobilisering som krävs för dess implementering. Att den deltagandebaserade samhällsplaneringen också när den tas på allvar kännetecknas av en haltande implementering är därför inte att förvåna. Det är en funktion av idéns innehåll.*

Referenser

- Alonso, S., Keane, J. & Merkel, W. (2011). *The Future of Representative Democracy*. Cambridge University Press.
- Bengtsson, B. (1995). *Bostaden – Välfärdsstatens marknadsvara*. Acta Universitatis Upsaliensis.
- Bäck, H. (2006). *Komparativ kommunal konstitutionspolitik: En kunskapsöversikt*. Sveriges kommuner och landsting.
- Bäcklund, P. & Mäntysalo, R. (2010). Agonism and institutional ambiguity: Ideas on democracy and the role of participation in the development of planning theory and practice—the case of Finland. *Planning Theory*, 9(4), 333–350.
- Chambers, S. (2003). Deliberative democratic theory. *Annual Review of Political Science*, 6(1), 307–326.
- Dahl, R. (1971). *Polyarchy: Participation and Opposition*. Yale University Press.
- Dahlstedt, Magnus, Lisa Kings & Nazem Tahvilzadeh (2018). Platskamp: inledande reflektioner. *Arkiv. Tidskrift för samhällsanalys*, (9), 7-25.
- Danielsson, M. (2010). *Fängslande idéer: svensk miljöpolitik och teorier om policyproduktion*. Acta Universitatis Upsaliensis.
- Danielsson, M., & Hertting, N. (2007). *Utvärdering av Stadsdelsförnyelsen: Slutrapport från Mälardalens utvärderingsakademi 2007*. Mälardalens högskola.
- Danielsson, M., Hertting, N., & Klijn, E. H. (2017). Tricky for good reasons: institutionalizing local participatory governance in representative democracy. I *Local Participatory Governance and Representative Democracy* (s. 18-48). Routledge.
- Eisenberg, E. (1984). *Ambiguity as Strategy in Organizational Communication*. *Communication Monographs*, 51(3), 227-242.
- Elster, J. (1998). Introduction. I J. Elster (Red.), *Deliberative Democracy* (s. 1–18). Cambridge: Cambridge University Press.
- Ganuza, E. & Baiocchi, G. (2012). The power of ambiguity: How participatory budgeting travels the globe. *Journal of Public Deliberation*, 8(2), 1–15.
- Gilljam, M., Karlsson, D., & Sundell, A. (2010). *Politik på hemmaplan. Tiotusen fullmäktigeledamöter tycker om politik och demokrati*. SKL Kommentus.
- Gutmann, A., & Thompson, D. F. (2009). *Democracy and disagreement*. Harvard University Press.
- Gustafson, P. & Hertting, N. (2017). Understanding participatory governance: an analysis of participants' motives for participation. *The American Review of Public Administration*, 47(5), 538-549.
- Held, D. (2006). *Models of Democracy*. Cambridge: Polity Press.
- Hertting, N., & Vedung, E. (2009). *Den utvärderingstäta politiken: styrning och utvärdering i svensk storstadspolitik*. Studentlitteratur.
- Hertting, N., & Vedung, E. (2012). Purposes and criteria in network governance evaluation: How far does standard evaluation vocabulary take us? *Evaluation*, 18(1), 27-46.
- Hertting, N., & Klijn, E. H. (2017). Institutionalization of local participatory governance in France, the Netherlands, and Sweden: Three arguments reconsidered. I Hertting & Kugelberg (Red), *Local Participatory Governance and Representative Democracy* (s. 168-194). Routledge.

- Hertting, N. & Hellquist, A. (2022). Om deltagarbaserad samhällsplanering och rätten till bostad. I Bengtsson, B., Holdo, M. & Holmqvist, E. (Red.) *Allas rätt till bostad. Marknadens begränsningar och samhällets ansvar*. Daidalos.
- Hertting, N. (2023). Implementering: perspektiv och mekanismer. I C. Dahlström (Red.), *Politik som organisation*. Studentlitteratur.
- Hirschman, A. O. (1991). *The Rhetoric of Reaction*. Harvard University Press.
- Klijn, E. H. & Koppenjan, J. (2016). *Governance Networks in the Public Sector*. Routledge.
- Kugelberg, C. (2017). Participatory Governance as an Embryonic Opportunity Structure: The Case of the Young Adult Center in Northern Botkyrka. I Hertting, N., & Kugelberg, C. *Local Participatory Governance and Representative Democracy* (s. 108-141). Routledge.
- Landwehr, C. (2010). Democratic and technocratic policy deliberation. *Critical Policy Studies*, 3(3-4), 434-439.
- Lowi, T. (1969). *The end of liberalism: The Second Republic of the United States, Second Edition – 40th Anniversary*. W. W. Norton & Company.
- Manin, B. (1997). *Principles of Representative Democracy*. Cambridge: University Press.
- Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press.
- Parvin, P. (2021). The participatory paradox: An egalitarian critique of participatory democracy. *Representation*, 57(2), 263-285.
- Pollitt, C. & Hupe, P. (2011). Talking about government: The role of magic concepts. *Public Management Review*, 13(5), 641–658.
- Scharpf, F.W. (1997). *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*. Boulder: Westview Press.
- Sveriges kommuner och regioner. (2023). *Delaktighetstrappan. Och liknande modeller*. Sveriges kommuner och regioner.
- Teorell, J. (2006). Political participation and three theories of democracy: A research inventory and agenda. *European Journal of Political Research*, 45(5), 787–810.
- Thacher, D. & Rein, M. (2004). Managing value conflict in public policy. *Governance*, 17(4), 457–486.
- Vedung, E. (1977). *Det rationella politiska samtalet: hur politiska budskap tolkas, ordnas och prövas*. Aldus.
- Vedung, E. (2016). *Implementering i politik och förvaltning*. Studentlitteratur.
- Vedung, E. (2018). Fyra typer av statsvetenskaplig idéanalys. *Statsvetenskaplig tidskrift*, 120(2).
- Walker, E. T., McQuarrie, M., & Lee, C. W. (2015). Rising participation and declining democracy. I *Democratizing inequalities* (s. 3-24). New York University Press.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21(1), 1–19.

KAPITEL 18

Rationalitet, energipolitik och policyprocess*Ylva Norén Bretzer*

INLEDNING

Som ivrig student i statskunskap i Uppsala under tidigt 1990-tal kom Evert att bli handledare för min C-uppsats som handlade om kommunal miljöpolitik i Danmark, i kölvattnet av den danska frikommunreformen. När jag senare fortsatte att läsa D-kursen i förvaltningspolitik var Evert ansvarig för kursen, vilken gav en gedigen introduktion till den internationella policy- och implementeringslitteraturen. Vi kursare bestod av ett tjugotal lika förväntansfulla elever, vars vägar har korsats på ett flertal olika sätt i det efterföljande yrkeslivet. Själv kom jag senare att hamna på statsvetenskapliga institutionen i Göteborg där jag senare doktorerade, men kontakterna med Evert har fortsatt. Någon gång besökte jag IBF i Gävle, någon annan gång sågs vi i Uppsala över en kopp kaffe och diskuterade utvärdering. Jag har fått värdefulla manuskomentarer av Evert, och något mail här och där. En bok-release, ett temanummer. I dessa professionella samtal nämnde Evert någon gång att "*...energipolitik... det är något att satsa på!*" Den uppmaningen inspirerade till att jag och några kolleger tillsammans sökte – och fick – ett FORMAS-projekt i ämnet (Norén

Bretzer, 2016). Men det ska jag inte resonera om här, utan i detta kapitel ska jag som en senkommen student redovisa på vilka sätt Evert är en pionjär i svensk energipolitisk forskning.

RATIONALITET SOM STUDIEOBJEKT OCH EFTERFÖLJANDE ENERGISTUDIER

På ett antikvariat hittade jag för flera år sedan boken *Det rationella politiska samtalet* (Vedung, 1977), vilken sedan dess har fått sin plats på min klassikerhylla. Den adderar till den *Upsaliensiska* skolans idéanalys och logik, där politikens argument kan utsättas för *inhållslig* respektive *funktionell* prövning utifrån preciserade analytiska principer. Sådana principer kan exempelvis vara huruvida argumenten är systematiska, mångtydiga eller vaga, relevanta, sanna, falska eller ohållbara (s. 24). Genom att bidra med detta vetenskapliga objektivitetsperspektiv argumenterar författaren att all politik, dess innehåll och genomförande kan prövas genom en systematisk och analytisk lins.

Det rationella politiska samtalet utgjorde en utgångspunkt för Vedungs nästkommande bok, *Kärnkraften och regeringen Fälldins fall*, som utkom två år därefter, 1979. Boken bygger på ett gediget material. Uppgiften var att nagelfara och rimlighetspröva de energipolitiska argumenten, vilket i arbetet förstås som ett exempel på rationalitetsprövning. Här visas de politiska partiernas ideologiska principer och möjliga handlingsalternativ, förhandlingar, strategier och möjliga kompromissutfall. Valet av analysmetodik – kärnkraftsfrågan – var vid denna tid principiellt intressant (*och är så fortsatt idag!*) då den representerade en sakpolitisk fråga som klöv högerblocket C, Fp och M mitt itu, liksom även vänsterblocket som vid den tidpunkten bestod av S och Vpk. Trots att C fick 25 procent av väljarnas stöd i valet 1973, och blev största borgerliga parti med 90 mandat i riksdagen, kom den nyvalde borgerliga statsministern Fälldin att stöta på patrull med sina koalitionspartners på grund av olika energipolitiska vägval. Vedung blottlägger genom motivanalys hur och varför partierna kom att göra de politiska vägval som gjordes, och varför regeringen Fälldin

senare havererade. Han visar också hur olika beslutskedjor leder fram till specifika utfall. Analysen har även drag av diskursanalys, baserat på ett gediget material.

Därpå följde arbetet *Energipolitiska utvärderingar 1973 – 81*, (1982) där Vedung kritiskt undersöker statens politik för energiberedskap och hur energihushållning utvärderas i Sverige. Rapporten genomförde en slags meta-utvärdering av 193 (!) befintliga utvärderingar för energihushållningsåtgärder i Sverige under den aktuella perioden efter energikrisen 1973 och oljekrisen 1979. Olika styrmedel hade stått i fokus för energiberedskapskrisen efter oljekriserna, främst inom industri, hushåll och service. Transportsektorn var undantagen. Vedung hänvisar till statsvetarprofessor Tarchys som vid denna tid hade konstaterat att ”att en systematisk och regelbunden utvärdering av den egna verksamheten är sällsynt.” (s. 61). Detta till trots konstaterade Vedung själv att ”Inget svenskt politikområde torde ha utvärderats så till den milda grad som energipolitiken.” (1982, s. 60ff, se även 1993, s. 34, s. 5). Intressant nog har utvärderingarna inom det energipolitiska området täckt såväl politik, policy som styrmedel. Varför var utvärderingarna så pass ymniga inom just energiområdet vid denna tid? Potentiella förklaringar som Vedung framhåller skulle kunna vara att policyfältet vid tidpunkten var ett relativt nytillkommet ämne. Andra förklaringar skulle kunna vara en förhållandevis bred definition av termen utvärdering, och en tredje skulle kunna spåras i 1970-talets hårdaste politiska kraftmätning i Sverige, nämligen kärnkraftsfrågan. En fjärde förklaring som inte nämns av Vedung (1982) men som tas upp av Everett (2003) är att utvärdering varit vanligt förekommande inom naturvetenskapliga discipliner och i princip har importerats till samhällsvetenskaperna därifrån. Det energipolitiska området är därför en intressant mix av samhällsvetenskap och naturvetenskaplig kompetens som möts i detta policysamtal.

Rapporten *Energipolitiska utvärderingar 1973 – 81* har ett stort läsvärde för den utvärderingsintresserade på 2000-talet, även om huvuddelen av den är en genomgång av utvärderingsteori. De

studerade rapporterna och evidensen återges endast översiktligt, men några av slutsatserna är till exempel att:

- Den myndighet som hade ansvar för statens insatser för energiberedskap, Överstyrelsen för ekonomiskt försvar (1962 – 1986), utvärderades inte.
- Att många myndigheter utvärderade sig själva (Bostadsstyrelsen, Bränslenämnden, Clearingnämnden, Elransoneringsnämnden, Energisparkommittén, Industriverket samt Transportnämnden)
- Universiteten intog en undanskymd plats inom utvärderingsområdet, främst var det privata konsulter och experter av olika slag som ägnade sig åt utvärdering
- Utvärdering planerades inte in när reformer, program och styrning lades upp utan kom ofta till i efterhand mer av en slump
- Målkonflikter uppmärksammades sparsamt
- Bieffekter synliggjordes inte (vare sig positiva eller negativa)
- Ofta mättes inte det utvärderarna utgav sig för att mäta (inferensproblemet)
- Utvärderingsdesignen var oftast otillfredsställande
- Program-, verksamhets- eller styrteorier (logiker) saknades

Denna "eländeslista" som Vedung avslutar sin utvärdering med skulle kunna tolkas på tre sätt; antingen behöver kunskaperna om utvärdering stärkas inom många discipliner och professioner för att kunna utföra utvärdering av god kvalitet – vi måste helt enkelt bli bättre utvärderare. En andra tolkning är att, ja, dessa problem finns inom utvärderingsfältet men vi kan inte bättre än så här. En tredje möjlighet är att vi inte ska hålla på med utvärderingar. Alls. Inte på myndigheter, inte inom akademi, inte inom konsultvärden. Ibland kritiseras utvärderingarna för att ha ett alltför rationellt grundantagande om offentliga (och privata) verksam-

heter, men sådan kritik kommer sällan med användbara egna svar. Om vi ger upp rationalitetsambitionen, då ger vi också upp ambitionen att ”lära av våra misstag” och förmågan att utvecklas som samhälle (Dahler-Larsen, 2020, s. 174). Min egen bedömning är att Vedungs erfarenheter 1982 i stort är representativ inom utvärderingsfältet idag, fyrtio år senare. Dels har termen *evaluation* en minskad användning inom den internationella forskningsdiskursen till förmån för forskningsteman som *governance* och *globalization* (Howlett & Lejano, s. 361, 365), begrepp som snarare är beskrivande än bedömande. Forskarna har delvis övergett utvärderandet och ägnat sig åt andra slags frågor under de senaste två decennierna, enligt Howlett & Lejanos. Det är dock en öppen fråga om och i så fall hur dagens policygenomförande och policyutvärdering behöver stärkas eller ej, eller i vilken grad den ersätts av *post truth* förgivettaganden (Patton, 2018).

Året efter, 1983, skrev Vedung rapporten *Lagen om kommunal energiplanering – tillkomst och politisk behandling*. Denna tar sin början i kommittén om kommunal energiplanering som tillsattes redan 1975 under regeringen Palme. I mars 1977 presenterade trepartiregeringen Fälldin en proposition med förslag till lag om kommunal energiplanering, en nätt skrivelse om en sida. Förslaget antogs av riksdagen 1 juni 1977, och från den första juli samma år började bestämmelsen ”kommun skall i sin planering främja hushållningen med energi samt verka för en säker och tillräcklig energitillförsel.” gälla (s. 9).

Vedungs kritiska öga identifierade att den nya lagregeln de facto tilldelar kommunerna fyra nya roller; a) som energihushållare i egna byggnader, b) pådrivare av allmänhetens energisparande gällande icke-kommunala fastigheter, c) kontrollinstans vid myndighetsutövning (beviljande av statliga lån och bidrag till energihushållning i bebyggelsen), samt d) att producera och leverera energi. Vedungs *rationalistiska* öga spårar i detta arbete politikens faktiska genomförande, kartlägger metodiskt varje förarbete, författningsskrivning och motion, tills genomförandet (implementeringen) utkristalliserar för läsaren i översiktliga utfallsmodeller. Staten styr här främst genom information; råd och upplysningar,

och att på olika sätt informera kommunerna om sin politik. Om kommunerna "lyssnar" eller hörsammar den statliga politiken får dock läsaren inte någon information om, men arbetet utgjorde troligen på sin tid en värdefull översikt av vilka statliga aktörer som förväntades agera, och hur.

Energifrågan har många ansikten och under våren 1993 pågick en intensiv debatt i tre Norrlandskommuner om hur befolkningarna ställde sig till förstudier om slutförvar av kärnbränsle. År 1998 redovisade Patrik Olofsson och Evert Vedung en gemensam analys, *Kärnkraftsavfall och lokalt beslutsfattande: Fallen Malå, Storuman och Överkalix* (1998) av hur kommunerna Malå, Storuman och Överkalix förhöll sig till den nationella kärnavfallspolitiken, en så kallad *bottom-up*-studie (Sabatier, 1986; Hjern 1982). Häri står inte staten och dess aktörer längre i fokus, utan tre kommuner och dess lokala beslutsfattare, lokalpolitik och lokala opinioner. Varför kom en förstudie att inledas i Malå och Storuman, medan förslag om detta drogs tillbaka i Överkalix, trots brett folkligt motstånd i alla tre kommuner? I studien prövas fem olika typer av förklaringar som bidrar till att förklara utfallet:

- De lokala motopinionernas styrka
- Skillnader i lokalt näringsliv och näringslivstraditioner
- Samspelet mellan den kommunala beslutsprocessen och den lokala motopinionen, som inleddes tidigt i Överkalix men senare under beredningsprocessen i Malå och Storuman

Denna uppsats breddar och kompletterar genomförande- och utvärderande perspektiven jämfört med tidigare arbeten. Häri finns en tydlig stat-lokal-dimension och analys av lokala opinioner samt kommunala beredningsprocesser visavi den centrala statsmakten.

År 2001 utkommer ett *magnum opus* inom energipolitiken, *Vattenkraften, staten och de politiska partierna*. Författarna är Evert Vedung och Magnus Brandel (Institutet för bostads- och urbanforskning, respektive Svenska torvproducenterna), och verket representerande en fruktbar kombination av teori och praktik inom ämnet. Bokens inledande fråga sökte svar på varför de under

1960-talet om vattenkraften samstämiga riksdagspartierna senare under tidigt 1970-tal kom att delas upp i Vpk och en samstämmig borgerlighet å ena sidan, socialdemokratin å den andra. Hur kunde det komma sig? Dessutom kom Folkpartiet från 1975 och framåt att framstå som den främste älvräddaren i svensk politik, hur kunde detta förstås och förklaras?

Flera centrala samhällsvetenskapliga teman tas upp till behandling; från partiteori, diskursanalys till institutionell teori och historia. Författarna visar hur olika juridiska regleringar vid olika tidpunkter ramar in (främjar eller hämmar) de nya teknikernas framväxt. Brytningstiden i svensk vattenkraftshistoria 1965-71 blottläggs, där skapande av institutionella ordningar för lång tid framåt reglerade vattenanvändningen i de svenska älvarna, men som kom att hantera älvarna något olika genom "språklig akrobatik" i utskottet. Författarna finner fem olika slags politiska regimer som verkade under vattenkraftens utbyggnad 1880-1993, där obruten utbyggnad rådde fram till 1961. Därefter tillämpades ett system av partsavtal mellan 1961-1972, för att övergå till riksplaneskydd 1972-1975. År 1972 försköts vattenkraftsfrågorna betydligt till älvräddarnas fördel (s. 408), medan från 1975 och framåt underställdes vattenfrågorna den mer övergripande energipolitiska regimen.

Våren 1973 föreslog regeringen Palme att svensk kärnkraft skulle byggas ut med 24 aggregat till år 1990, vilket förflyttade fokus från vattenkraftsutbyggnaden. 1978 rämnade dock regeringen Fälldin I på grund av oenighet i Kärnkraftsfrågan, och i mars 1979 inträffade Harrisburgkatastrofen. Folkstormen ledde till att regeringen Palme då föreslog en folkomröstning om kärnkraft som ägde rum i mars 1980, där kärnkraftsvännerna Moderaterna och näringslivet slöt upp bakom Linje 1. Socialdemokraterna och Folkpartiet stod bakom Linje 2 som menade att byggda och planerade reaktorer (6+6) skulle drivas vidare men sedan avvecklas till år 2010 (årtalet är ett senare tillägg, vilket senare blev regeringsbeslut). Linje 3 var helt emot kärnkraftstekniken och centrala aktörer här var Folkkampanjen mot kärnkraft tillsammans med kommunisterna och Kristdemokraterna (s. 411f). Trycket på vattenkraftsutbyggnaden

borde nu rimligen minska, när linje 2 hade segrat i folkomröstningen.

Socialdemokraterna inledde dock en vattenkraftsutbyggnads-offensiv i oktober 1982, och 1985 beslutades om en riksplan för de strömmande vattnen. När Tjernobylyolyckan sedan inträffade och orsakade radioaktivt nedfall över Sverige restes krav på kärnkrafts-avveckling. Regeringen Carlsson inledde partiöverskridande förhandlingar, vilket ledde till att startåret för kärnkraftsavvecklingen togs bort. När sedan regeringen Bildt tillträdde efter valet 1992 kom ett stort antal älvar att räddas enligt naturresurslagen, men med fortsatt kärnkraftsfokus.

Boken är ett centralt bidrag till svensk nutidshistoria, politisk historia och en analys av hur två energislag, vattenkraft och kärnkraft, underställs en lång rad politiska beslut, lagar och institutioner. Spelet om vattenkraften bestämdes delvis av spelet mellan partierna, men också av lokala opinionsspel och införandet av ny teknik. Boken avslutas ungefär där den småskaliga energitekniken börjar bli kommersiellt intressant och blir därmed även intressant för efterkommande energipolitiska förhandlingar. De långa tidsperspektiven i boken visar en unik överblick som sällsynt kommer till uttryck i dagens statsvetenskap. De stora perspektiven ger här en osedvanlig insikt om hur en allmännyttig resurshantering i mer än ett århundrade har utspelats, och drivkrafterna därbakom (jmf. Ostrom, 1992). Arbetet är en prestation av sällsynt art.

STUDIET AV SVENSK ENERGIPOLITIK

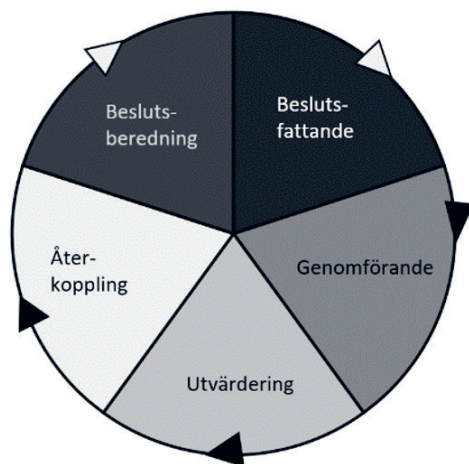
Genomgången av Vedungs studier på energiområdet visar att svensk energipolitik väl lämpar sig för statsvetenskapliga analyser och att de kan utformas på en mängd olika sätt. Ibland har det handlat om utvärderingar eller meta-utvärderingar. Andra gånger, som i det sistnämnda fallet, *Vattenkraften, staten och de politiska partierna*, genomgång av mer än ett sekels politiska ställningstaganden som korrigerar hanteringen av vattenresurserna på en lång rad olika sätt. *Det rationella politiska samtalet* lyser igenom de analyser som författaren ger sig i kast med under yrkeslivet och övriga arbeten hade sett annorlunda ut om inte denna grund hade

funnits. Tusentals av politiska beslut, förarbeten, motioner och reservationer har analyserats för att dessa böcker skulle kunna bli skrivna. Två centrala iakttagelser lyser fram i läsningen av denna energidiskurs.

För det första är det intressant att just energiområdet var så pass omgärdat av utvärderingar vid denna tid, 1970-talet. Det låg i tidens anda internationellt sett, i kölvattnet av de amerikanska uppföljningarna av 1960-talets *War on Poverty* som president Lyndon B. Jonson sjösatte, då forskare började följa "vart tar pengarna vägen" (Pressman & Wildavsky, 1973). Men att utvärderingarna blev så vanligt förekommande på just energiområdet tål att fundera på, hur kom det sig? Kanske var det slumpen. Konsekvensen blev dock att mycket av den utvärderingskunskap som importerades till Sverige kom via energiforskningen och dess utvärderingar, samt via den HFR-finansierade forskartjänst i utvärdering som Vedung fick, 1986-1992.¹ Energiforskningens och Vedungs bidrag till svensk utvärderingsforskning får därmed sägas vara avsevärd.

En andra iakttagelse är att Vedungs verksamhet har varit *omfattande*, där energifrågorna endast är en mindre del av hans omfattande gärning. Av de mer än 200 arbeten som Kungliga biblioteket listar i LIBRIS, har här endast en handfull valts ut som handlar om svensk energipolitik. Energiforskning får sägas utgöra en viktig – men en mindre del – av Vedungs totala produktion. Inte desto mindre har Vedung varit – och är – en banbrytande pionjär inom svensk energipolitikforskning. Före Evert var det inte många statsvetare som ägnade dessa frågor något större intresse, trots att energifrågan utgjorde och utgör grunden för hela vår civilisation (idag såväl som under 1900-talet)! Utan energi blir inte många hus värmda, och heller inte någon exportproduktion eller välfärd.

1) År 1986 till 1992 innehade Evert Vedung en särskild forskartjänst i utvärdering, finansierad av dåvarande HFR, Humanistiskt Samhällsvetenskapliga Forskningsrådet i Stockholm. Detta var Sveriges första (och hitintills enda) forskartjänst i utvärdering som tillsatts.



Figur 1. Snurrmodell över utvärderingens plats i beslutsprocessen

Kommentar: Figuren är hämtad från Vedung 1982, s. 7. Originalfiguren har här försetts med pilar och skuggade fält för att betona den politiska delen (svart) i förhållande till övriga delar av *policyprocessen*.

UPPTÄCKTEN AV POLICYPROCESSEN

I arbetet *Energipolitiska utvärderingar 1973 – 81* (1982) visar Vedung hur utvärdering bör förstås i det politiska systemet, nämligen med följande klassiska figur som finns väl representerad i den internationella litteraturen (Everett 2003; Davis et al. 1993; Premfors 1981; Wildavsky 1979, s. 129; Patton 1978; Lindblom 1959).

Figur 1 visar interaktionen mellan den politiska nivån i det offentliga systemet (beslutsberedning och beslutsfattande) och det som karaktäriseras mer av policy (genomförande, utvärdering och återkoppling).² Figuren visar att Vedung vid den aktuella tidpunkten är väl bekant med den internationella diskursen om offentlig förvaltning och kommer senare att bli en pionjär som förmedlar denna kunskap till svensk akademi, forskarutbildning och förvaltningsutbildningar i flera olika discipliner.

2) Möjligen kan även beslutsberedning hänföras till *policy* utifrån Kingdons tes om en "policy-soppa", varur politiken sorterar vilka problem som ska tas upp på den politiska dagordningen, och vilka problem som kvarstår hos andra samhällsaktörer såsom marknaden eller civilsamhället (Kingdon, 1984).

Figuren användes då, i en rapport 1982, och används fortsatt i den internationella litteraturen. Vedungs användning av figuren innebar en introduktion av modellen för svenska politik- och förvaltningsstudenter, av det som i dagligt tal kan sägas vara åtskillnaden mellan politik och policy.

Detta kan sägas vara ett tidigt bidrag till denna dikotomi eller sekvensförståelse av det rationella politiska beslutsfattandets faser, inklusive genomförande, utvärdering, återkoppling och förnyad beredning inför nästkommande politiska beslut. Vedung har senare utvecklat denna *idealtyp* till en *flerskiktmodell* där både utvärdering och återkoppling kan ske i flera olika delar av genomförandet (Vedung, 2009, s. 37; Hertting & Vedung, 2009, s. 18). Denna mer *cirkulära* syn på politiken som Vedung har introducerat är central för svensk statsvetenskap, som i decennier präglades av ett 'black-box-perspektiv' på hur politik fungerar (se diskussion Norén Bretzer, 2021, s. 25). Utan Vedungs insats i detta avseende hade vi kanske fortfarande trott att politiska beslut 'verkställs av sig själva'. I vart fall hade det tagit längre tid att upptäcka förvaltningens betydelse utan Vedungs insatser i detta avseende.

SLUTORD

Ett sista slutord, Evert, är att detta kapitel inte hade kunnat skrivas om jag inte hösten 1995 hade varit din student på masterkursen Förvaltningskunskap D i Uppsala. En tio veckors kurs kan få ett avtryck i en bok trettio år senare. Tack för allt du har bidragit med till mig och alla andra av dina studenter. Du har fått oss att växa, att kritiskt och rationellt vrida och vända på evidens, programteorier och utvärderingsmodeller. Och fått oss att fundera på det rationella politiska samtalets fundamenta. Tack också för att du öppnat dörren till den energipolitiska forskningen – en mycket betydelsefull insats, i det sociotekniska gränssnittet, i utvärderingsforskning och inom svensk samhällsvetenskap. Och att du importerade både policyprocessen och utvärdering till svensk akademi och förvaltning.

Referenser

- Dahler-Larsen, P. (2020). *Evaluering i udvikling – bevægelser i teori, metoder og praksis*. Hans Reitzels Forlag.
- Davis, G. J., Wanna, J., Warhurst, J. & Weller, P. (1993). *Public Policy in Australia*. Allen & Unwin.
- Everett, S. (2003). The Policy Cycle: Democratic process or Rational Paradigm Revisited?. *Australian Journal of Public Administration*, 62(2), 65-70.
- Hertting, N. & Vedung, E. (2009). *Den utvärderingstäta politiken*. Studentlitteratur.
- Hjern B. (1982). Implementation Research — The Link Gone Missing. *Journal of Public Policy*, 2(3), 301-308.
- Howlett, M. & Lejano, R. P. (2012). Tales From the Crypt: The Rise and Fall (and Rebirth?) of Policy Design. *Administration & Society*, 45(3):357.
- Kingdon, J. W. (1984). *Agendas, alternatives, and public policies*. Little, Brown and Co.
- Lindblom, C. E. (1959). The Science of Muddling Through. *Public Administration Review* 19(2), 79-88.
- Norén Bretzer, Y. (2021). *Sveriges politiska system*, 4:e uppl. Lund: Studentlitteratur.
- Norén Bretzer, Y. (Red.) (2016). *Urban välfärd – effektiv energi*. Boréa förlag.
- Olofsson, P. & Vedung, E. (1998). *Kärnkraftsavfall och lokalt beslutsfattande: Fallen Malå, Storuman och Överkalix*. Uppsala universitet: statsvetenskapliga institutionen.
- Ostrom, Elinor E. (1990). *Governing the Commons*. Cambridge University Press.
- Patton, M. Q. (2018). Evaluation Science. *American Journal of Evaluation*, 39(2), 183-200.
- Patton, M. Q. (1978). *Utilized-focused Evaluation*. Sage.
- Premfors, R. (1981). *Genomförande och utvärdering av offentlig politik*. Stockholm: Statsvetenskapliga institutionen, stencil.
- Pressman, J. L. & Wildavsky, A. (1973). *How Great Expectations in Washington are Dashed in Oakland*. University of California Press.
- Sabatier, P. A. (1986). Top-Down and Bottom-up Approaches to Implementation Research: A Critical Analysis and Suggested Synthesis. *Journal of Public Policy*, 6(1), 21-48.
- Vedung, E. (2009). *Utvärdering i politik och förvaltning*. Studentlitteratur.
- Vedung, E. (1983). *Lagen om kommunal energiplanering – tillkomst och politisk behandling*. Rapport R11:1983. Bygghälsningsrådet.
- Vedung, E. (1982). *Energipolitiska utvärderingar 1973 – 81*. Rapport 52 för Delegationen för energiforskning, Stockholm.
- Vedung, E. (1979). *Kärnkraften och regeringen Fälldins fall*. Rabén & Sjögren.
- Vedung, E. (1977). *Det rationella politiska samtalet: Hur politiska budskap tolkas, ordnas och prövas*. Aldus/Bonniers.
- Vedung, E. & Brandel, M. (2001). *Vattenkraften, staten och de politiska partierna*. Bokförlaget Nya Doxa.
- Wildavsky, A. (1979). *Speaking Truth to Power: The Art and Craft of Policy Analysis*. Little, Brown & Co.

KAPITEL 19

I framtidens skugga – om tiden i implementeringsprocesser

Stefan Sjöblom & Sebastian Godenhjelm

INLEDNING

Vi lever i en tid som präglas av exceptionellt stora utmaningar. Nya lömska problem ("wicked problems") förekommer på löpande band och dyker upp från höger och vänster (Head, 2019). I skrivande stund har världen till synes just avklarat en global covid-19-pandemi. Detta är såklart mycket glädjande, men glädjen är kortvarig då det samtidigt pågår krig i Ukraina och världen drabbas av den ena naturkatastrofen efter den andra. Kriserna och den moderna samhällsutvecklingen har satt oss i en situation där existerande samhällsstrukturer inte längre fungerar på önskvärt sätt. Vi lever med andra ord i osäkerhetens tid där samhällena kräver systemförändringar för att kunna överleva (Uusikylä & Jalonen, 2023). Som statsvetare undrar vi såklart vad man inom den offentliga förvaltningen kan göra för att åtgärda detta.

Någonting måste snabbt göras, men vad och hur? I dagens värld kan det vara lättare sagt än gjort. Existerande samhällsliga strukturer och institutioner kan inte ändras snabbt. Institutioner har bland annat en tendens att lida av institutionell tröghet, det vill säga motsätter sig förändring då nya komplexa utmaningar upp-

dagas (Munck af Rosenschöld, 2017). Blickarna riktar sig ofta till innovationer i hopp om nya lösningar. Detta väcker dock frågor om institutioner i verkligheten är mottagliga för exempelvis nya innovativa organiseringsformer såsom nätverk, hybrida eller temporära organisationer som antas producera lösningsmodellerna och ge möjligheter till snabba och flexibla reaktioner på uppkomna problem.

Offentlig politik har alltid präglats av beslutsfattarnas vilja att hantera såväl det förflutna som framtiden. Ambitionen märks vid olika tider i de flesta planerings-, styrnings-, och utvärderings-system. Möjligen är de tidsmässiga spänningarna i politiken och den offentliga förvaltningen idag mera påtagliga än förr. Inte minst gäller det framtiden, vilket bland annat märks av att anticiperande styrning blivit ett populärt koncept (OECD, 2020). I finländsk förvaltningspolitik är för övrigt den vedertagna benämningen i svensk översättning *föregripande styrning*. Även om 'anticipera' och 'föregripa' ofta används synonymt, förväntar sig förvaltningens reformatörer genom den här betoningen något mer än anteciperung i konventionell bemärkelse. Det handlar inte enbart om att förutse en intervention eller händelse och rätta sitt beteende därefter, utan om att kunna stoppa en framtida händelse eller utveckling som skulle komma att förutsätta en intervention. Betoningen följer möjligen av att föreställningen om smidiga (agila) och reaktionssnabba förvaltningsmyndigheter framhävts mycket starkt i det finländska reformarbetet. En sådan förvaltningsuppfattning följer delvis av den sociala accelerationens problem, eller dilemmat att lösa lömska problem och att hantera snabba sociala och teknologiska förändringar genom en långsam demokratisk politik.

Den centrala frågan är emellertid vilka konsekvenser de tidsmässiga spänningarna har för politiken. Har vi inte lärt oss något från tidigare framgångar och/eller misstag om hur vi kunde gå vidare? Finns det något sätt på vilket vi kunde förutse kommande samhällsliga utmaningar genom att blicka tillbaka och ta vara på det som vi fått genomgå tidigare utan att göra förhastade beslut? "Tidsaspekten i bemärkelsen det som hänt tidigare är oerhört viktig i implementeringsanalys", skriver Evert Vedung i sin bok *Implementering i politik och förvaltning*. Men tiden är en rymlig

kontext. Den inbegriper inte bara det förflutna och nutid utan även framtiden. Genom antecipering kan en intervention få effekter redan innan den antagits. Den kastar sin skugga in i framtiden, en annan implementeringsfaktor som han uppmärksammar i samma bok. Vad betyder egentligen de tidsmässiga spänningarna och deras uttrycksformer för vår förståelse av politiken och dess implementering? Det är huvudfrågan vi vill diskutera i det här bidraget.

TIDEN I DAGENS POLITIK

Genom historien har alla samhällen strävat att synkronisera sina aktiviteter. När standardiserad tid infördes 1913 blev tiden rationaliserad och separerad från lokala sammanhang (Adam, 1998, s. 107). Därmed blev det möjligt att använda tiden som ett abstrakt utbytesvärde. En grundläggande utgångspunkt exempelvis för hög-specialiserade produktionssystem som Taylorism och Fordism är att standardiserad tid kan användas vid beräkning av effektivitet och vinst. I vidare mening möjliggjorde standardiseringen av tiden inte bara global synkronisering, utan även säkrare produktionssystem, förbättrad kontroll och förutsägbarhet (jfr Adam, 1998, s. 107).

Idag utmanas de möjligheter som standardiseringen en gång medförde av att de grundläggande samhällsproblemen förändrats, vilket i sin tur fått svårhanterliga konsekvenser för de liberala demokratiernas storsystem. Klimatförändringar, energiförsörjning, befolkningsrörelser och många andra utmaningar förutsätter långsiktiga strategier men också omedelbara reaktioner och snabba beslut under stor osäkerhet. En del bedömare menar att det senare har övertaget i politiken, och att acceleration blivit ett hegemontiskt imperativ (Rosa, 2010). Andra har hävdats att senmoderna samhällen genomgår en massiv avsynkronisering eftersom politikens institutioner, marknader, medier och informationssystem fungerar inom mycket varierande tidsramar (Laux, 2011).

Men man behöver knappast hänge sig åt samhällsfilosofiska resonemang för att inse utmaningarna för samhällsstyrningen. Politikens tidsskalor blir helt enkelt mera differentierade. Ett av problemen är att förutsättningarna för intervention vid rätt tid

varierar beroende på samhällsproblem. Det svåra i att intervensera i rätt tid illustreras exempelvis i en utmärkt analys av europeisk reglering av farliga kemikalier (Munck af Rosenschöld et al., 2014). Den visar att de negativa effekterna av kemikalierna uppstår i en takt som lagstiftningsprocesserna inte kan hantera. Det beror framför allt på processernas sekvenser, särskilt kravet att regleringen ska föregås av vetenskaplig kunskap om effekterna. Därmed påverkas politikens tempo. Regleringen bromsas. Långsam kunskapsproduktion leder till långsam politik, vilket omöjliggör flexibel intervention på ett sådant sätt som kemikalieeffekternas dynamik skulle förutsätta. Studien visar med andra ord på betydande problem när det gäller att samtidigt hantera alla tidsvariabler; tidsramar, sekvenser, tempo och timing.

Ett annat problem är att den demokratiska procedurernas mekanismer förblivit i stort sett oförändrade över tid. Till skillnad från marknader, medier och informationssystem har demokratins procedurer bara i begränsad utsträckning kunnat dra nytta av exempelvis den teknologiska utvecklingen. Snabba interventioner är inte lätta att förena med krav på grundlagsenlighet, öppenhet och transparens, vilket hanteringen av covid-19-krisen visade (Niemikari & Raunio, 2022, s. 8). Om grundläggande samhällsfunktioner arbetar i varierande tempo ökar komplexiteten ur styrningssynpunkt. De formativa ögonblicken blir svårare att urskilja. Detsamma gäller möjligheterna att åstadkomma det som med engelskans något svåröversättliga begrepp kallas 'temporal fit', det vill säga en tidsmässig överensstämmelse mellan problem, intervention och slutmottagare.

NUET: OLIKA SYNER PÅ TID OCH DESS ORGANISERING

Enligt den protestantiska etiken uppfattas slöseri med tid som en av de dödligaste synderna av alla (Rosa, 2013, s. 299). Det finns aldrig tillräckligt med tid, "om jag än hade mera tid", "tid är pengar" och liknande utsagor är något som man ofta hör folk säga. Snabbhet ses som en dygd medan långsamhet ofta associeras med ineffektivitet och slöseri med resurser (Vakkuri, 2010). Alla är konstant upptagna, men gör vi rätta saker och vilken är egentligen vår syn på tid?

Forskningen kring tid och dess betydelse lyfter bland annat upp betydelsen av social acceleration, något som ofta sammankopplas med modernitet. Förändring syns bland annat som teknisk acceleration i form av allt snabbare produktion och kommunikation, som social förändring i form av allt kortare tidsperioder som representerar nuet, såväl som livets tempo, som ofta tar sig uttryck i att hinna med möjligast många saker inom en begränsad tidsperiod (Rosa, 2013, s. 301). Denna diskussion har uppenbara kopplingar till något som Nørmark och Fogh Jensen (2021) kallar "pseudoarbete" enligt vilket handlingar, hur irrationella de än må vara, ofta trumfar substans. Det viktigaste är att det uppstår handling (action), låt oss oroa oss över konsekvenserna senare. Få samhällsvetare skulle antagligen opponera sig mot möjligheten att i lugn och ro få bereda och tänka igenom åtgärder före beslut fattas. Idag representerar detta tyvärr närmast önsketänkande (se också Sverker Gustavsson om "demokratisk bearbetning" i denna volym).

Nya flexibla och effektiva organisationsformer som snabbt kan sättas upp och som kan agera utanför rigida strukturer kan utgöra synnerligen attraktiva alternativ, speciellt i situationer då snabba handlingar krävs. En sådan organisation är bland annat projektorganisationen vars huvudsakliga syfte är att åstadkomma handling i form av konkret och unik service, produkt eller resultat, inom en viss budget och under en förutbestämd tidsperiod (Project Management Institute, 2004). Forskningen kring den ökade användningen av projektorganisationer i den offentliga sektorn talar ofta om projektifiering, det vill säga den ökade användningen av projekt och organiseringen av diverse processer som om de vore projekt (Fred & Godenhjelm, 2023; Hodgson et al., 2019).

En central beståndsdel inom projekten är inte bara den begränsade tiden som finns till hands. Fragmenteringen, eller indelningen av tiden i olika sekvenser, utgör en central beståndsdel av projekt. Nedbrytandet av projektets mål i olika faser anses förtydliga dess syfte varvid projektets kapacitet och förutsättningar att uppnå dess mål och framgångsrik implementering ökar (Scott, 2023, s. 68).

Enligt Lundin och Söderholm (1995, s. 440) kan tiden ses som linjär varvid exempelvis organisationernas utveckling alltid har en

koppling till och bygger på tidigare kunskap. Enligt denna princip är nya organisationer alltid bättre än gamla eftersom antagandet är att nya organisationer bär med sig tidigare uppnådd kunskap. En annan syn på tid är den cykliska, enligt vilken handlingar upprepas om och om igen, men utan att någon markant förändring sker. En tredje syn på tid är spiralen, vilket indikerar att man aldrig återvänder till samma plats. Underförstått i detta synsätt är att det sker en gradvis, men konstant utveckling. Avsikten med projekt är med andra ord att genom frikoppling omforma den spirala uppfattningen av tid till en lineär form av tid. Vi har med andra ord skapat nya redskap och organisationsformer som förutsätter snabba reaktionstider och handlingar för att kunna fungera (Laux, 2011, s. 228).

Helt oproblematiskt är detta dock inte. Tidigare forskning på området har bland annat betonat spänningarna som uppstår mellan temporära och permanenta organisationer (Godenhjelm et al., 2019). De kontextuellt känsliga sammanlänkningsmekanismer som gör att långsiktiga resultat från tillfälliga insatser kan uppstå i ett politiskt-administrativt sammanhang, är ofta bristfälliga, vilket leder till att projekten inte bara förlorar sina flexibla och innovativa egenskaper, utan också kan försvaga de permanenta organisationernas förmåga att upprätthålla värderingar förankrade i den offentliga sektorn, såsom öppenhet och demokratisk ansvarighet, samt förmågan att upprätthålla politisk samstämmighet, samordning och kontinuitet. (Godenhjelm, 2016, s. 64; Munck af Rosenschöld & Wolf, 2017)

Det kan mycket väl vara så att tillfälliga organisationer och de lösningar som de producerar är överlägsna permanenta strukturer när det gäller att producera värdefulla resultat, men utan sammanlänkande mekanismer, som till exempel effektiv informationsutväxling mellan tillfälliga och permanenta strukturer, riskerar de inte bara att förlora sina flexibla och innovativa egenskaper, utan kan också försvaga de permanenta organisationernas förmåga att upprätthålla samordning och kontinuitet. (Godenhjelm et al., 2015, ss. 342–343). Tillfälliga organisationer kan med andra ord ur verksamhetens synpunkt leda till en rad oväntade såväl som ”perverse

effekter”. Denna typ av bieffekt definieras som en effekt som orsakats av insatsen och som har följd, som blir rakt motsatta dem som avsågs. De omfattas med andra ord av begreppsparet huvud-effekt-bieffekt. Det är dock värt att notera att ”perversa effekter” kan vara såväl positiva om negativa bieffekter (Vedung, 2009, s. 104). Ett illustrativt exempel på en negativ effekt är bildandet av Ängsö nationalpark 1909 vars syfte var att bevara de vackra ängarna och lövlundarna. Frånvaron av exempelvis betande kor hade dock en motsatt effekt vilket ledde till att ängarna och lövlundarna sakteligen övertogs av skog (Vedung, 2013). Tillfälliga insatser i en svårkontrollerad offentlig omgivning kan med andra ord få överraskande konsekvenser.

ATT BLICKA FRAMÅT: ANTECIPERING OCH IMPLEMENTERING

I ett framåtblickande perspektiv är frågan vad tidens rymliga kontext, som inbegriper det förflutna såväl som nutid och framtid, innebär för möjligheterna att bedriva och implementera en framtidsinriktad, föregripande politik. Framtidsinriktade ambitioner kommer inte minst till uttryck i anteciperande innovationsstyrning (Anticipatory Innovation Governance, AIG), särskilt sedan OECD (2020) lanserade konceptet som ett medel för att på ett konsekvent sätt handla med inriktning på framtiden ”så som den framstår idag”, för att citera rapporten. Orsakerna till dagens intresse för föregripande politik är många, från klimatförändring och covid-19 till en större geopolitisk osäkerhet. Med anteciperering avses här handlingar som ger kunskap om en osäker framtid (OECD, 2020, s. 7). Förenklat kan man säga att anteciperande innovationsstyrning i OECD:s tappning består av tre element (Erkkilä et al., 2023); framsyn utifrån svaga signaler (anteciperering), nya tekniker och styrinstrument (innovation) samt regeringarnas ’ideala framtider’ (styrning).

Det ligger i och för sig föga nytt i det här synsättet. Framsyn ingår som ett element i alla planerings- och styrningssystem. Ett visst mått av anteciperering ingår i all politik. Skillnaden är att anteciperering blir ett explicit och centralt element i den politik som

bedrivs med AIG som styrform. AI och stordata bidrar ytterligare till en föreställning om handlingsberedskap och möjligheter att kontrollera framtiden. I förlängningen förändras även vår syn på staten. Snarare än en samling institutioner blir staten en samling förväntningar eller sannolikheter. Därmed uppstår frågan var makten sitter utifrån en sådan styrningslogik, närmare bestämt: Hur skapas föreställningar om framtiden? Eftersom vi inte kan veta vad framtiden för med sig, blir det svårt att på ett övertygande sätt argumentera för andra framtider än den maktavarna representerar. Därmed ökar riskerna för konflikter och olika former av intressekonflikter. De gula västarnas proteströrelse i Frankrike är ett exempel på ouppfyllda förväntningar gällande statens skattepolitik, förväntningar som regeringen uppfattade som illegitima. De två parterna hade delvis oförenliga föreställningar om framtiden och om statens roll i sammanhanget (Erkkilä et al., 2023).

”I politiken spelar föreställningar om verkligheten större roll än verkligheten själv” skriver Evert Vedung och Magnus Brandel (2001) i sin stora studie om vattenkraften i Sverige. I den föregripande politiken intar föreställningar om framtiden samma roll. Frågan är som sagt hur dessa skapas. Förväntningar ställs till en bred mobilisering av aktörer och intressenter enligt samverkans-teorins antaganden om förutsättningar för en framgångsrik politik-utformning och implementering (Hertting, 2018). Det är givet att det politiska stödet för framtidsuppfattningarna och åtföljande interventioner får en avgörande betydelse.

I färsk analys argumenterar emellertid Niilo Kauppi för att de praktiska tillämpningar av AIG som utarbetats sannolikt är otillräckliga (Erkkilä et al., 2023). I Finland bygger de i hög grad på samverkan och kontinuerligt lärande mellan regering, myndigheter, intressegrupper och organisationer (OECD, 2022). I mindre grad uppmärksammas medborgarnas, politikernas och gräsrotsbyråkraternas roll i anteciperingen. Detta har speciellt stor betydelse för implementeringen av nya innovativa resultat som utgör en delikat och sårbar fas då entusiasmen som uppnåtts genom nya lösningar ska förankras bland gräsrotsbyråkrater och mottagare.

Implementeringen kan exempelvis ske av personer som inte haft någon roll i det gemensamma skapandet av lösningen. Uppfattningarna bland dem som har som uppgift att implementera resultaten kan avvika diametralt från dem som varit tätt involverade i utvecklingsarbetets tidigare skeden. Slutresultatet kan i sådana fall vara att initiativet förlorar sin drivkraft och sitt understöd (Ansell & Torfing, 2021, s. 125). Styrssystem som AIG borde med andra ord kunna beakta den variation i preferenser och intresseinriktningar som kännetecknar demokratiska system.

Dylika problem kan knappast lösas genom den teknologiska utvecklingens verktyg, även om de delvis banat väg för anticeperande innovationsstyrning. Nu som förr ligger mekanismerna för att hantera intresseinriktningar och alternativa framtider i det politiska samtalet och legitima procedurer. Politiken blir mera, inte mindre, viktig och de krav som bör ställas på demokratins procedurer är sannolikt särskilt stora inom ramen för en föregripande politik. Ur implementeringssynpunkt ändrar AIG och andra liknande koncept nämligen inte på det faktum att en förändring som beslutats i politisk strid blir svårare att genomföra än den som beslutats under politisk harmoni (Vedung, 2016, s. 121). De arenor där framtidsbilderna skapas är den föregripande politikens viktiga element.

POLITIK FÖR FRAMTIDEN?

Vi har här velat lyfta fram några av de utmaningar som tiden och föregripande ambitioner skapar i politiska processer. I den praktiska politiken kan en snabbhetskultur mycket väl förbytas i tröghet och inbromsning. Den demokratiska proceduren sätter gränser. De bör gälla också för en föregripande politik. Ur forskningssynpunkt är det skäl att framhäva tiden som en viktig källa till komplexitet.

Konsekvenserna av social acceleration skall dock inte underskattas. Det finns ett mervärde i tröghet. Politiska styrningsmekanismer och stabila arbetsförhållanden utgör exempelvis grunden för hållbar planering och utveckling, något som social acceleration kan tänkas underminera (Rosa, 2013, s. 303). Diskussionen för onekligen även tankarna till något som Daniel Kahneman (2011)

tar upp, nämligen betydelsen av att tänka snabbt och långsamt. Fokuserar vi med andra ord på det som Kahneman avser med "system 1-tänkande", det vill säga det som automatiskt, intuitivt och utan ansträngning dyker upp då problem uppstår. Eller agerar vi enligt det som Kahnemans kallar "system 2-tänkande", något som kräver fokus, deliberation och ansträngning för att lösa problem, men som tar mera tid? I dagens projektvärld tenderar tyvärr motsatt ordning råda, det vill säga tänka snabbt och agera långsamt (Flyvbjerg & Gardner, 2023).

Ur ett forskningsperspektiv lyfter den föregripande politikens idé upp viktiga teman med en stark förankring i implementeringsforskningen. Det gäller bland annat arenorna, processernas faser och förutsättningarna för intervention i stöd av en föregripande politik. Ingående studier av styrnings- och implementeringsprocesser blir möjligen ännu viktigare för att förstå de framtida villkoren för att genomföra politiska beslut.

Så vill vi avslutningsvis peka på Evert Vedungs viktiga insatser på området. Vi tänker inte enbart på böckerna om implementering och utvärdering utan inte minst på djupa och insiktsfulla analyser som den tidigare nämnda studien av svensk vattenkraft (Vedung & Brandel, 2001) eller processpåring av framgångsfaktorer för innovation (Vedung & Bjurulf, 2019). De står sig som förebilder för analyser av den offentliga politik som i flera bemärkelser sker i framtidens skugga.

Referenser

- Adams, B. (1998). *Timescapes of modernity: the environment and invisible hazards*. Routledge.
- Erkkilä, T., Chou, M.-H., & Kauppi, N. (2023). *Knowledge Alchemy: Models and Agency in Global Knowledge Governance (Transnational Administration and Global Policy)*. Bristol University Press.
- Flyvbjerg, B., & Gardner, D. (2023). *How Big Things Get Done: The surprising factors that determine the fate of every project, from home renovations to space exploration and everything in between*. Random House.
- Fred, M., & Godenhjelm, S. (Red.). (2023). *The Projectification of Organizations, Governance and Societies. Theoretical perspectives and Empirical Implications*. Palgrave Macmillan.
- Godenhjelm, S. (2016). *Project Organisations and Governance – Processes, actors, actions and participatory procedures*. Unigrafia.
- Godenhjelm, S., Lundin, R. A., & Sjöblom, S. (2015). Projectification in the Public Sector – The Case of the European Union. *International Journal of Managing Projects in Business*, 8(2), 324–348.
- Godenhjelm, S., Sjöblom, S., & Jensen, C. (2019). Project Governance in an Embedded State: Opportunities and Challenges. I D. E. Hodgson, M. Fred, S. Bailey, & P. Hall (Eds.), *The Projectification of the Public Sector* (s. 149-168). Routledge.
- Head, B. W. (2019). Forty years of wicked problems literature: forging closer links to policy studies. *Policy and Society*, 38(2), 180–197.
- Hertting, N. (2018). Implementering: perspektiv och mekanismer. I C. Dahlström (Red.), *Politik som organisation: förvaltningspolitikens grundproblem*. Studentlitteratur.
- Hodgson, D., Fred, M., Bailey, S., & Hall, P. (Red.). (2019). *The Projectification of the Public Sector*. Routledge.
- Kahneman, D. (2011). *Tänka, snabbt och långsamt*. Volante.
- Laux, H. (2011). The time of politics: Pathological effects of social differentiation. *Time and Society*, 20(2), 224–240.
- Lundin, R. A., & Söderholm, A. (1995). A theory of the temporary organization. *Scandinavian Journal of Management* (Vol. 11, Issue 4, 437–455).
- Munck af Rosenschöld, J. (2017). *Projectified environmental governance and challenges of institutional change towards stability*. Unigrafia.
- Munck af Rosenschöld, J., Honkela, N., & Hukkinen, J. I. (2014). Addressing the temporal fit of institutions: The regulation of endocrinedisrupting chemicals in Europe. *Ecology and Society*, 19(4), 0–9.
- Munck af Rosenschöld, J., & Wolf, S. (2017). Toward projectified environmental governance? *Environment and Planning A: Economy and Space*, 49(2), 273–292.
- Niemikari, R., & Raunio, T. (2022). Centralized leadership, ministerial dominance, and improvised instruments: The Governance of COVID in Finland. *Nordisk Administrativ Tidsskrift*, 99(2).
- Nørmark, D., & Fogh Jensen, A. (2021). *Pseudowork. How we ended up being busy doing nothing*. Gyldendal Notification.
- OECD. (2020). *Antecipatory Innovation Governance*. OECD Publications.
- OECD. (2022). *Anticipatory Innovation Governance Model in Finland. Towards a New Way of Governing*. OECD Publications.

- Project management Institute. (2004). *A Guide to the Project Management Body of Knowledge* (3 Uppl. Red.). Project Management Institute Inc.
- Rosa, H. (Red.). (2010). *High-speed society: Social acceleration, power, and modernity*. Penn State Press.
- Rosa, H. (2013). *Social Acceleration. A New Theory of Modernity*. Columbia University Press.
- Scott, D. (2023). Entering the World of Project Making: Mobilizing Assemblage Thinking to Unpack Projects as Political Constructions. I M. Fred & S. Godenhjelm (Red.), *The Projectification of Organizations and Societies. Theoretical Perspectives and Empirical Findings* (s. 57–73). Palgrave Macmillan.
- Uusikylä, P., & Jalonen, H. (Red.). (2023). *Epävarmuuden aika. Kuinka ymmärtää systeemistä muutosta*. Into Kustannus Oy.
- Vakkuri, J. (2010). Hitauden talousteoria. *Kunnallistieteellinen Aikakauskirja*, 38(3), 217–219.
- Vedung, E. (2009). *Utvärdering i politik och förvaltning* (3 Uppl. Red.). Studentlitteratur.
- Vedung, E. (2013) Side effects, perverse effects and other strange effects of public interventions. I L. Svensson, G. Brulin, S. Jansson & K. Sjöberg (Red.), *Capturing Effects of Projects and Programmes* (s. 35-62) Studentlitteratur.
- Vedung, E. (2016). *Implementering i politik och förvaltning*. Studentlitteratur.
- Vedung, E., & Barndel, M. (2001). *Vattenkraften, staten och de politiska partierna*. Nya Doxa.
- Vedung, E., & Bjurulf, S. (2019). Framgångsfaktorer för innovation. Samverkan genom projektifiering, kompetenscenter och kluster. I S. Godenhjelm, C. Jensen, & S. Sjöblom (Red.), *Innovation och projektifiering. Att styra och leda handlingskapaciteten i moderna offentliga organisationer* (s. 213–240). Nordic Academic Press.

KAPITEL 20

Lämna mig inte! Några reflexioner kring internationella skilsmässor*Li Bennich-Björkman*

1905 var ett dramatiskt år. Unionen mellan Sverige och Norge som existerat i nästan hundra år, sedan 1814, gick i graven. De båda länderna var hädanefter grannar, men inte under samma kung. De kom att bedriva – i synnerhet under Kalla kriget – helt olika utrikespolitik, vilken tidigare varit gemensam. Medan Norge blev medlem i Atlantpakten, senare Nato, höll Sverige fast vid sin långa tradition av alliansfrihet samtidigt som relationerna till den västliga försvarspakten var nära, ja närmast intima under alla efterkrigsdecennier (Holmström, 2011). För Sveriges del hade unionen gärna kunnat fortsätta. Det var Norge som ville klippa banden, bli ett helt självständigt land med egen kung och egen utrikespolitik. Kort sagt stå på egna, norska, ben. När upplösningen var ett faktum blev det en fredlig affär. Krig mellan de båda kontrahenterna utbröt inte, trots att bitterheten inom delar av den svenska eliten var avsevärd och det mullrade inom militären. Norge och därmed unionen sågs som mycket viktig för försvaret av Sverige, särskilt gentemot stormaktsgrannen Ryssland. Säkerligen förekom också känslor på den svenska sidan som emanerade ur uppfattningen att Sverige under de många år som gått gett ungtuppen Norge en

hel del; både säkerhet, viss tradition och glansen från den svenska adeln. Otack är, som alltid, världens lön, suckade därför förmodligen en del. Men det hjälpte inte. Skilsmässan mellan de båda var ett faktum.

Efter unionsupplösningen har Sverige och Norge genomlevt ett 1900-tal utan att tyngas av det särskilda trauma som ett krig mellan grannar, utgör. Tvärtom innebar det sekel som följde ett närmande och fördjupande av relationerna mellan de båda, som sedan länge är högst vänskapliga. Ett krig mellan Sverige och Norge är idag helt otänkbara, något som antyder att en stabil fred är för handen.

Men historiskt var Sverige och Norge fiender, och krig var gängse. Inom den omfattande vetenskapliga litteratur som växt fram inom statskunskap och historia hamnar det som Sverige och Norge var med om 1905 idag under det blodfattiga begreppet "secession", att bryta sig loss. Med ett mer vardagligt språkbruk hämtat från den privata sfären: skilsmässa. I litteraturen har forskarna sökt efter såväl orsaker till att unioner, stater eller federationer faller sönder, som i Sverige och Norges fall, som konsekvenserna av detta. Indien 1947, många av de afrikanska länder som varit kolonier under 1960-talet, Slovakien 1993, närapå Katalonien och Skottland i närtid; alla har de drivit fram skilsmässa eller stått på tröskeln att göra det. Exempelen kan mångfaldigas. Frågan om när skilsmässan äger rum under fredliga former eller när krig utbryter har varit central. Svaren på den öppnar för lärdomar för diplomati och internationella organisationer i strävan efter en fredligare värld.

Vad har då detta med Evert Vedung att göra? Mycket. Evert inledde nämligen sin akademiska bana genom att skriva en avhandling om just unionsupplösningen mellan Sverige och Norge: *Unionsdebatten 1905*. I den tradition som hade etablerats på Skytteanum under dåvarande professorn Carl-Arvid Hessler, att rikta sökljuset på hur offentliga aktörer talade och skrev om viktiga frågor, "debatten om..." passade Everts avhandling väl in. Han återvände till denna, ur många synvinklar, fascinerande skilsmässa också långt senare. I skriften *Varför ledde inte Norges secession 1905 till krig?* (2008), är Evert tillbaka där han en gång började och reflekterar kring ett antal faktorer som kan bidra till att förklara

varför ett krig den gången inte blev bister verklighet. En ambitiös sammanställning av korrespondensen som svenska (och fram till en viss punkt i tiden också norska) Utrikesdepartementet (UD) förde med beskickningarna på olika håll i världen under upplösningen kom ut av trycket 2017. Evert, tillsammans med historikern Gustav Jakob Pettersson och sin bror, Tage Vedung, har i denna spårat upp brev, notiser och depescher som skickades mellan Sverige och världen. De återfinns nu i den sjuhundra sidor långa *Secession och diplomati*. I fokus står de kritiska månaderna på hösten 1905 då relationerna mellan Sverige och Norge stod och vägde.

Som i alla skilsmässor, mellan individer såväl som mellan stater, är den som blir lämnad, "reststaten" kallad i litteraturen, och hur denne uppträder, en viktig komponent för att förklara varför krig eller fred blir sättet att lösa oenigheten. I fallet Sverige-Norge var de gemensamma organen, kungahuset och utrikesförvaltningen, som ju var svenska, i grunden för en fredlig lösning: "Något intresse från de unionella instansernas sida att på allvar sätta hårt mot hårt fanns inte. Tvärtom förordade de ett mildare agerande än de svenska riksorganen", konstaterar Evert (Vedung, 2008, s. 8).

Men inte heller detta var från början givet. Från att ha intagit positionen tidigt under vårvintern 1905 att unionen definitivt skulle bestå, ändrade kronprins Gustaf (sedermera Gustav V), ståndpunkt. Varur sprang det förändrade synsättet? Kronprinsen som delvis vistades i London kände den internationella opinionen på pulsen och skrev i telegram till kung Oscar II att inga villkor skulle ställas för Norge att få lämna ty: "Vi riskera annars förlora Europas sympatier som vi nu äga" (Vedung, 2008, s. 9). De närmast intresserade stormakterna, England, Tyskland och Ryssland, saknade alltså intresse att trycka på för en väpnad konflikt med Norge. Evert konstaterar att Norges utbrytning skedde vid en internationellt "maximalt gynnsam tidpunkt" (Vedung, 2008, s. 13). Några år därefter var redan positioneringarna som ledde fram till första världskrigets utbrott, med ett utbrett nätverk av allianser, i full gång.

Kronprins Gustaf var alltså redan innan han efterträdde sin far Oscar II en duva med avseende på Norge. Men i övriga kungahuset

rådde fortfarande uppfattningen att om Norge skulle få lämna det som från Sveriges sida sågs som ett ömsesidigt väl fungerande samarbete, fick det ske med villkor. Krig, däremot, avvisade Oscar II bestämt i Urtima riksdagen:

Nej, må det svenska folket låta sig ledas af klok självbehärskning och gifve Gud Sverige kraft och enighet att inom sina egna gränser återvinna hvad det genom en upplösning av unionen skulle förlora (Vedung, 2008, s. 10).

Oscar II talade om självbehärskning. Det är ett intressant ordval, som pekar på någon form av självbild där Sverige är en "förnuftig" reststat som inte låter känslorna styra. Från svensk sida visade man också besinning, trots besvikelsen och kanske till och med förödmjukelsen. Det fanns ett politiskt omdöme som ledde rätt; mot fred, inte krig.

När en annan union, Sovjetunion, gick i graven 1991 var det också till följd av en rad utbrytningar, alltså secessioner, från de ingående rådsrepublikerna. Sovjetunionen bestod av femton republiker, många av dem förvärvade med våld och tvång. I den domino-process som bröt ut under sensommaren och hösten 1991 var det en hel del av dessa som ville skilja sig, lämna unionen och Ryssland, och stå på egna ben. Inom de nya staterna som uppstod var det också på sina håll de som ville skilja sig från den politiska makten i den nya huvudstaden och bilda eget. I det skiljer sig unionsupplösningen mellan Sverige och Norge från den sovjetiska. Evert lyfter i sin analys fram att det inte fanns instängda enklaver med svenskar i Norge eller norrmän i Sverige, som kunde försvåra fredssträngningarna och som moderlandet kunde göra anspråk på att försvara – något som ofta varit den ryska taktiken gentemot rysktalande minoriteter i det tidigare Sovjetunionen. I det självständiga Moldavien blev det krig, eftersom Transnistrien ville bryta sig loss. Detsamma skedde i Georgien med områdena Sydossetien och Abkhazien. Och Nagorno-Karabach som tillhörde Azerbajjan, ville bryta sig ur den förbindelsen och återförenas med Armenien. Följden blev ett krig som ännu idag inte följts av någon stabil fred utan påverkar bägge länderna, och Nagorno-Karabach,

i högsta grad. Då och då blossar den upp, i skrivande stund har just detta skett och fredsförhandlingar är åter igång.

Resultatet blev i alla dessa fall att Ryssland fick och fortfarande har kontroll över utbrytarrepublikerna, som inte är erkända av omvärlden men väl fungerar som de-facto stater. Konflikterna med moderlandet kallas "frusna" och ligger som en tung filt över utveckling, ökad stabilitet och internationell samverkan.

Även i delar av tidigare Sovjetunionen som klarade sig undan väpnade konflikter kring utbrytning var det ibland skälvanande nära. I Estland fanns i det rysk-dominerade Ida-Virumaa med Narva som huvudstad på gränsen till Ryssland långtgående planer på att begära utträde. 16 och 17 juli 1993 anordnade man (och vilka "man" är, är något oklart men påminner starkt om ett ryskt *modus operandi*) därför folkomröstningar i de rysk-dominerade städerna Narva, Sillamäe och Kohtla-Järve, där en majoritet, 54 procent, röstade för skilsmässa. För dem som drev detta var den Transnistriiska modellen i Moldavien en källa till inspiration. Estniska regeringen, redan då kraftfull och beslutsför och utrustad med politiskt omdöme, förklarade emellertid resolut att omröstningarna var i strid med den estniska konstitutionen (Bennich-Björkman, 2022, kap. 11). Någon ytterligare skilsmässa, efter den mellan Sovjetunionen och Estland, blev det inte. Ida-Virumaas befolkning kan idag känna tacksamhet över detta. Estlands levnadsstandard utvecklades till att bli betydligt högre än Rysslands, och det står bra till med demokrati och ekonomi. Sedan 2004 är Estland medlem i Europeiska Unionen (EU) och dess invånare kan därmed röra sig fritt och arbeta i hela Europa. Till skillnad från sina ryska grannar på andra sidan floden Narva, i Ivangorod, slipper Idu-Virumaa-bornorna stå till svars inför omvärlden och inför Ukraina för ett ryskt blodigt och brutalt krig mot ett grannland.

I och med Rysslands angreppskrig på Ukraina har frågan om unionsupplösning, skilsmässa och krig återigen aktualiserats. "Ryssland, Ukraina och Belarus – tre slaviska länder med delvis samma ursprung, liknande språk och kulturer och med miljontals landsmän i respektive land – befinner sig nu i krig med varandra. Hur kunde det gå så illa? Vad fick Rysslands president Vladimir

Putin att bestämma sig för att angripa Ukraina? Och varför upplät Belarus ledare, Aleksandr Lukasjenko, landets territorium för ett anfall mot ett slaviskt broderland?” undrar Östeuropa-kännaren Jakob Hedenskog i en nyligen utkommen skrift. (Hedenskog, 2022, s. 4).

Ja, vad fick Putin att angripa Ukraina? Hur mycket har det med Nato och västvärlden att göra, och hur mycket med 1991 års skilsmässa, som Ryssland först gick med på men senare bittert ångrat. Annekteringen av Krim 2014 var ett första steg på väg mot krig. Snabbt följdes den av skilsmässotendenser i östra Ukraina, underblåsta eller till och med orkestrerade, från Moskva. Ett geografiskt begränsat krig med stora förluster i människoliv har sedan dess pågått, där Donetsk och Luhansk skulle avskiljas från övriga Ukraina och bli ryska, det var målet. Så enkelt var det emellertid inte, många som bor där har en ukrainsk identitet och ville inte – trots att man önskade fortsatt vänskapliga relationer till Ryssland – ha ryska pass. Ryssland strävade alltså att framtvunga en skilsmässa – i likhet med i Transnistrien och det misslyckade försöket i estniska Narva – utan att på åtta år lyckas med detta. Därefter var det hela Ukrainas tur. Är det nuvarande ryska kriget mot Ukraina därför att betrakta som ett krig om en skilsmässa, trots att Ukraina förklarade sig självständigt redan 1991? Jag vill hävda det. 2021, ungefär ett år innan 24 februari 2022, uttryckte Rysslands president Vladimir Putin på Kremls officiella hemsida att ukrainsk suveränitet endast kan bli möjlig i partnerskap med Ryssland (Hedenskog, 2022, s. 8). Vad är det, annat än budskapet ”jag släpper dig aldrig”.

Medan Ukraina under åren fram till 2014 betraktade sin självständighet som utgångspunkten för de fortsatta relationerna med Ryssland (och andra stater) och i såväl retorik som praktik underströk det, har Ryssland endast betraktat den som ett slags villkorad garnityr; ”egentligen är du för alltid min”. Hot, bestraffning och tillfälliga närmanden, ”the cold and warm treatment” har varit strategin för att visa Ukraina att man aldrig kommer att släppa taget. När det gäller den andra slaviska grannstaten Belarus har en ömsesidig ambition att förnya sina löften kommit till

uttryck formellt redan för över tjugo år sedan: ”Den 8 december 1999 skrev Lukasjenko och den ryske presidenten Boris Jeltsin under ett avtal om att upprätta en unionsstat mellan länderna. Avtalet hade en ambitiös agenda, där en gemensam konstitution skulle antas och ett gemensamt parlament upprättas tillsammans med gemensam försvars- och säkerhetspolitik, valuta, tull och skatter med mera” (Hedenskog, 2022, s. 21).

Indikatorerna på att Ukraina verkligen menade allvar med skilsmässan har emellertid blivit allt starkare sedan 2014, annekteringen av Krim och kriget i östra delarna av landet. Ett centralt steg i den riktningen – märkligt litet uppmärksammat i väst – var när den ukrainska ortodoxa kyrkans självständighet från Moskva-patriarkatet stadfästes av patriarkatet i Konstantinopel 2018, efter att kyrkan de facto varit uppdelad i ett Kyiv- och ett Moskva-patriarkat sedan självständigheten.

Skilsmässan mellan Ukraina och Ryssland ledde, till skillnad från den mellan Sverige och Norge, alltså i slutändan till krig. Den som blev lämnad, Ryssland, har inte, och har inte på många år haft, en ledning som visat politiskt omdöme, som besinnat sig eller slickat sina sår i skymundan. I stället är det död och förintelse av den otrogna som nu eftersträvas. Ukraina skulle bli så skadat att ingen annan vill ha landet heller. Men ack, vad Ryssland bedrog sig.

Unionsupplösningen 1905 är av intresse att analysera inte enbart utifrån det faktum att Sverige inte gick till angrepp mot Norge utan också för att det sedan dess upprätthållits korrekta, så småningom goda och till och med vänskapliga förbindelser mellan länderna. Sverige ångrade sig aldrig, som Ryssland, utan 1905 markerade verkligen slutet på en epok. ”Relationerna mellan de två länderna har förblivit hjärtliga ända sedan dess” citerar Evert Aldous Huxley (Vedung, 2022, s. 23). Ändå finns det taggar kvar. Mina erfarenheter av norska samarbeten, till skillnad från de i tidigare svenska Finland, är att normmännen ofta är korrekta men litet avmätta. De hyser inte samma nyfikenhet som finnarna på Sveriges göranden och låtanden, och känner sig helt enkelt inte lika nära. Helt förståeligt, och vad som utspelade sig under

andra världskrigets permittenttrafik har också bidragit. Men ingen betvivlar vänskapen, eller att den ska bestå.

Inte heller det var givet, kontrafaktiskt skulle det konservativa Sverige kunnat nära tankar på ett återupprättande av unionen – i likhet med dagens Ryssland när det gäller Ukraina – eller att i ett senare skede vilja straffa Norge för utträdet. I hur hög utsträckning kan det sätt på vilket förbindelserna stabiliserades efter 1905 förklara den långsiktigt gynnsamma relationen mellan de båda grannländerna? För att närmare diskutera det, knyter jag an till en Vedung-klassiker, *Det rationella politiska samtalet*, som handlar om hur man talar med varandra på ett fruktbart, kanske till och med intellektuellt hållbart, vis. Den boken kom ut redan 1977. Trettiofem år senare, 2012, gav Charles Kupchan, professor vid Georgetown, ut en bok, *How Enemies Become Friends. The Sources of Stable Peace*. I den lägger han fram en teori för hur fiendskap övergår till vänskapliga relationer på det internationella planet. Intressant i detta sammanhang är att den stabila freden i Kupchans värld tar sig uttryck bland annat i union, att två (eller flera) länder blir en enhet. Det är en slags slutpunkt, men i Sverige och Norges fall, vilket gör den relationen så intressant, var det istället en sönderbrytning av en stabil fred – om än med vissa tvingande förtecken – som ägde rum. Det ställde om möjligt de agerande politiska makthavarna inför ännu större utmaningar, eftersom Norges vilja till uppbrott satte ifråga den kulturella samhörighet – ”broderfolken” talades det ju om – som redan hade etablerats. Man skulle kunna hävda att vägen till unionen var en ”kupchansk” process, medan uppbrottet och vägen efter unionen, ja, vad var det?

Ömsesidig återhållsamhet reglerar sedan samarbetet och dämpar rivaliteten. Samhällsintegration följer, där personliga och institutionella band knyts mellan partnerstaterna. Den sista fasen innebär att nya berättelser om vänskap skapas och kompatibla, delade eller gemensamma identiteter formas. De mekanismer som är verksamma spänner över skilda traditioner från det rationalistiska till det sociologiska; strategiska förhandlingar inleder försoningsprocessen, samhällelig inte-

gration påskyndar uppkomsten av ett internationellt och förändringar i den politiska diskursen och skapandet av nya identiteter fullbordar upprättandet av en stabil fred. (Kupchan, 2010, s. 389)

Så sammanfattar Charles Kupchan sin teori. Och i det fjärde steget, den om en gemensam berättelse, kommer diskursen, sättet att skriva och tala om varandra och om relationen, in. Osökt går min tanke till Everts tidiga fokus på hur en diskursiv nivå har betydelse; något som sedan dess exploderat inom kulturvetenskaperna i form av diskursanalyser av skilda slag, inspirerade av bland andra den franske tänkaren Michel Foucault.

Charles Kupchan kan inte läsa svenska, men hade han kunnat göra det är det ingen tvekan om att han skulle vara högst intresserad av att ta del av Evert Vedungs analyser av den svensk-norska unionsupplösningen som ett av flera historiska exempel på internationella stabiliseringsprocesser. Visserligen var det inte längre två fiender som var inblandade. Sverige och Norge var, på det retoriska planet åtminstone, broderfolk, men fiendskap hade kunnat bli resultatet av en illa skött separation. Diplomati i olika former och på skilda nivåer förhindrade det. Som Kupchan konstaterar är det vare sig demokrati (statskick), ekonomiskt beroende och samarbete utan diplomati, som "är det avgörande elementet för att sätta igång en process där fiender kan bli vänner" (Kupchan, 2012, s. 3). Påfallande tydligt framgår det i både Vedung (2008) och *Seccession och diplomati* (2017). Det är hur man talar och kommunicerar med och om varandra, orden och den diskursiva nivån, som blir det allra viktigaste i att fläta samman Sverige och Norge under 1900-talet. Diplomati innehåller både ett pågående rationellt politiskt samtal i Evertska anda, inbäddat i berättelser om vilka som för dessa samtal, vad de har gemensamt, och vart de är på väg. Idag sker diplomatin mellan Sverige och Norge i ständiga kontakter på politisk och social nivå. Samtidigt, i Rysslands och Ukrainas fall, är det precis dessa som slitits sönder och vittrat bort under det krig som pågår. Om och när Ryssland och Ukraina kan gå från fiendskap till något annat igen är en helt öppen fråga. Terapeuter ger ibland rådet att man ska fundera på hur en eventuell

skilsmässa kan komma att te sig, innan man ingår äktenskap. Ex-
emplen också från den internationella politiken visar hur sant det
rådet är. Sverige svalde förnedringen av att bli lämnat, Ryssland
har inga sådana psykologiska resurser. Evert Vedung har hjälpt oss
att se detta, och också hur mycket som finns att lära av att nog-
grant granska historien.

Referenser

Bennich-Björkman, L. (2022). *Bakom och bortom järnridån. De sovjetiska åren och frigörelsen i Baltikum och Ukraina*. Stockholm: Appell förlag.

Hedenskog, J. (2022). *Ukraina, Ryssland, Belarus: Från brödraskap till fullt krig, Världspolitikens dagsfrågor*. Stockholm, Utrikespolitiska institutet.

Holmström, M. (2011). *Den dolda alliansen. Sveriges hemliga Nato-förbindelser*. Stockholm: Natur och Kultur.

Kupchan, C. (2012). *How Enemies Become Friends. The Sources of Stable Peace*. Princeton: Princeton University Press.

Vedung, E. (1972). *Unionsdebatten 1905*. Uppsala: Acta Universitatis Upsaliensis. Skrifter utgivna genom Statsvetenskapliga föreningen.

Vedung, E. (2008). *Varför ledde Norges secession 1905 inte till krig?* Uppsala universitet: Statsvetenskapliga institutionen.

Vedung, E., Peterson, G.J., & Vedung, T. (2017). *Secession och diplomati. Unionsupplösningen 1905 speglad i korrespondens mellan UD, beskickningar och konsulat*. Uppsala universitet, Statsvetenskapliga föreningens skriftserie, 199.

KAPITEL 21

Evert Vedung – en forskarbana med statskunskap, utvärdering och oväntade händelser¹

Joakim Tranquist

I ett kåseri för Institutet för bostads- och urbanforskning² skrev du en gång att "i min forskarbana har serendipiteter spelat en stor sannolikt avgörande roll". Jag är nyfiken på att höra mer om när och hur utvärdering kom in i ditt liv.

– Just det, det där om serendipiteter. Ja, jag har en historia att berätta om hur det där med utvärdering kom in i mitt liv. Det var så att jag disputerade i december 1971, en jämförelse mellan debatten i Sverige och Norge kring unionsupplösningen 1905. Sedan höll jag på med den jämförande metoden och lite annat, men jag fortsatte med idéanalys som ett sidospår. I början av 1975 startade Leif Lewin ett stort forskningsprogram på institutionen³

1) Intervjumaterialet som texten bygger på presenteras även i annan bearbetning i Tranquist, J (2015) "The Oral History of Evaluation: The Professional Development of Evert Vedung". I: American Journal of Evaluation. Vol. 36(4), sid. 570-583

2) IBFs nyhetsbrev. Nr 2, maj 2003. "Serendipiteter i mitt forskarliv".

3) Forskningsprogrammet "Politik som rationellt handlande" satte sin prägel på Statsvetenskapliga institutionen i Uppsala under tioårsperioden 1975–85. Detta berättar Leif Lewin i en artikel i Svenska Dagbladet: http://www.svd.se/kulturnoje/nyheter/artikel_2672269.svd.

som kallades "Politik som rationellt handlande". I det sammanhanget blev jag tvungen att välja ett sakområde – Lewin ville ha med alla docenter. Då valde jag energipolitik, för jag tyckte att där fanns en ny dimension i svenska partisystemet. Det var inte bara vänster/höger, utan det var det här med tillväxt/ekologi. Det där upptäckte jag för jag var lite "grön" i mina dagar, jag var intresserad av miljöfrågor. Det var kanske därför jag noterade den där dimensionen tidigt. Jag tror jag var först bland svenska statsvetare att se den.

Men sedan fick jag ett Fordstipendium⁴ och kunde åka till Amerika. Det var i augusti 1975. Jag fick också ett Kennedystipendium och därmed en associering till John F. Kennedy School of Government på Harvard. Det var inte en statskunskapsinstitution, utan en public policy school. Det var det jag ville ha. Så jag höll på med energifrågan där. Jag åkte på kärnkrafts-seminarier, men det var ingenting speciellt som hände. Jag fortsatte också med idéanalys, det som blev boken *Det rationella politiska samtalet*⁵ och senare *Political Reasoning*⁶. Det var först när jag kom hem, från och med då, som saker började inträffa.

Delegationen för energiforskning i Stockholm – den hade väl bildats 1975 för att utvärdera det stora energiforskningsprogram som då sjösattes – skickade en rundskrivelse till energiforskare om vi hade något att bidra med när det gällde energifrågan. Det var ju en slump det där för mig. Jag hade ju knappt börjat med energifrågor, men tydligen hade mitt namn kommit in via Riksbanksfonden i någon sorts översikt. Då började jag läsa, vad skulle jag kunna göra här? Och då kom jag ju naturligtvis in på policy, för jag hade ju varit på en Public Policy School. Så dök det här med utvärdering upp. Det visste jag inte vad det var riktigt, men jag började läsa om det.

4) Ett program hos Ford Foundation (www.fordfound.org) som gick ut på att ta svenska forskare över till USA.

5) Vedung, Evert (1977) *Det rationella politiska samtalet: hur politiska budskap tolkas, ordnas och prövas*. Stockholm: Bonniers.

6) Vedung, Evert (1982) *Political Reasoning*. Beverly Hills, CA: Sage Publications.

Carol Weiss bok kom jag på då, likaså Edward Suchmans och en antologi av Peter Rossi med ordet "evaluation" i namnet⁷. Och så upptäckte jag några svenskar faktiskt. Det var en Rolf Lind på Handelshögskolan som författade rapporter omkring 1977 med ordet utvärdering i titlarna⁸. Lars Söderström vid Statskontoret skrev en rapport med blå pärmar⁹. Jag började så smått begripa vad det handlade om. Och att det lät spännande. Och att jag skulle klara av det. Så jag föreslog till Delegationen för energiforskning att jag skulle kunna göra något i utvärdering.

Jag tog tåget till Stockholm för ett antal möten där på hösten 1976. Den här delegationen satt i andra skyskrapan i Hötorgscity på den tiden. Det var väldigt spännande att komma dit. Det var en politisk delegation som delade ut pengar och det satt företrädare för alla partier där. Jag togs emot av verkställande ledamoten Sigfrid Wennerberg. Han gav mig i uppdrag att författa en promemoria. Sagt och gjort. När jag kom ner igen sa Wennerberg: "Du måste nog förbereda det här om det här ska gå igenom. Det är bäst att du snackar med Olof Hörmander", minns jag han sa, "Han är ordförande i delegationen". Så jag besökte Hörmander, som satt i Johnson-koncernens stora hus vid Stureplan. Jag försökte övertala honom om att det jag ville göra var viktigt. Sedan sa Sigfrid Wennerberg: "Kontakta Lars-Göran Redbrandt, han är liberal politiker. Försök med honom också, så du får lite stöd i den här delegationen". Så jag bearbetade honom också, och så småningom fick jag ett anslag för att ta reda på hur den svenska energipolitiken implementeras och utvärderas. Det var så jag kom in på det här med utvärdering.

När jag tänker på det här efteråt, så inser jag att hade inte Delegationen för energiforskning funnits, det vill säga ett sektorsforskningsorgan och som till råga på allt var inriktat på utvärdering

7) Vedung syftar här på: Weiss, Carol H (1972) *Evaluation Research: Methods for Assessing Program Effectiveness*. New Jersey: Prentice-Hall, Inc. Suchman, Edward A (1967) *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*. New York: Sage. Rossi, Peter H & Williams, Walter (1972) *Evaluating Social Programs: Theory, Practice, and Politics*. New York: Seminar Press.

8) Rolf Linds Utvärderingsmetod: en översikt och Utvärderingsforskning: en litteraturoversikt från 1977 respektive 1979.

9) Lars Söderströms Utvärdering av rationaliseringsinsatser från 1975.

– om än av forskning – så hade jag inte kommit in på utvärderings-spåret. Det tror jag inte personligen. Och om jag inte nyss valt energifrågan som sakområde så hade jag inte blivit aktuell för DFE överhuvudtaget. Detta om någonting är väl serendipiteter.

Och detta var alltså efter att du varit i USA?

– Ja, det här var efter USA. Jag kom tillbaka på hösten 1976 och det var väl då den här rundskrivelsen kom. Det var också på hösten 1976 som jag hade mötena med Delegationen för energiforskning. Och då lunkade jag liksom på med det här. Jag förstod inte då att det skulle bli stort. Jag hade andra prylar jag skrev på – jag höll på med det här med Tingstens idéanalys, och det var ju en god sak i och för sig. Sverker Gustavsson, Leif Lewin och Reidar Larsson tyckte som jag. Men tongivande människor i statskunskap ansåg att idéanalys var ett egendomligt, rentav urmodigt, arv från den gamla svenska statskunskapens tid. Alla skulle ju vara positivister kring 1970 – man skulle inte hålla på med motiv, intentioner, idéer och föreställningar överhuvudtaget. Man skulle vara objektiv och inte värdera, ens mot mål som andra satt upp. Man skulle syssla med mätbara beteenden. Jag minns Terry Carlbom, som var min doktorandvän. Han blev liberal politiker sen, han blev bland annat ordförande i fullmäktige i Uppsala. Jag minns att han, när han höll på med sin doktorsavhandling, tyckte att man överhuvudtaget inte skulle bry sig om att spåra upp motiv och intentioner. Bara offentligt uttalade åsikter.

Jag blev ganska chockad över uttalanden av den här sorten. Jag delade inte den meningen, jag tyckte något annat. Det här gjorde i vart fall att det hos mig fanns en viss mental beredskap för värdering och utvärdering. Jag kunde lätt begripa att man kunde värdera mot mål till exempel, eller värdera mot intentioner, och att det måste vara något viktigt i politik. Så jag tror att det där idéanalytiska fanns också, i min bakgrund.

Det är inte bara serendipiteter som gör att man noterar och tar till sig nyheter. Det måste också finnas mottaglighet för det nya om det ska funka. Fast nog är det här med Delegationen för energi-

forskning en serendipitet. Jag hade ju från början ingen aning om det där spåret, att det skulle komma därifrån.

Hur påverkade tiden i USA dig, vad tog du i första hand med dig hem?

– Det var väl en sorts personlig djärvhet som USA gav mig. Men jag upptäckte inte det här med utvärdering i Amerika. Möjligen var jag på fel ställe, på Harvard Public Policy School. Där pågick många andra intressanta saker. De hade ett program, skämtsamt kallat Recycling of bureaucrats, på masternivå. Människor som jobbat i administrationen fem-sex år kom tillbaka och blev uppfräschade. Denna nära kontakt med administrationen fascinerade mig. Men om utvärdering fick jag inget speciellt med mig just då. Det skulle komma senare. Men då kom det något verkligt stort och det är en serendipitet igen.

Jag skrev ju då på den här studien – hur svensk energipolitik utvärderas – och i samband med det försökte jag följa upp politiken som beslutades 1975 kring energisparandet. I samband med det kom jag till Energisparkommittén – en kommitté i kanslihuset som liksom Delegationen för energiforskning också agerade som en sorts myndighet. I den kommittén bedrevs energisparkampanjer. Man gjorde också utvärderingar av olika slags energisparåtgärder. Där fanns en herre som hette Jan-Erik Furubo, som var en ung kille på den tiden. Jag kom i kontakt med honom, minns jag, och mötet med Furubo och med Energisparkommittén var viktigt. Jag fick ett anslag därifrån. Alltså från en annan kommitté, precis som Delegationen för energiforskning. Jag skulle utveckla utvärderingsmetoder för informativa styrmedel, minns jag. Det var Ulf Karlsson som gav mig anslaget.

Sen började Furubo och jag att träffas, han bodde på S:t Larsgatan i Uppsala. På vägen hem ibland på kvällarna så cyklade jag förbi, och så satt vi där och drack kaffe och thé och resonerade kring det här med utvärdering. Han hade inte skrivit någonting än, men vi satte igång en liten cirkel han och jag. Det var i slutet på sjuttio-talet, kanske på hösten 1978.

Sen höll jag på med det här hur svensk energipolitik utvärderas och jag hittade väl inte så många som var intresserade av det där ämnet och det där med utvärdering egentligen. Jag kan inte erinra mig om att statsvetarna tyckte det var något intressant. Men jag fick för mig att jag skulle åka till Amerika igen. Jag for till Vanderbilt ett par månader 1980, och sen en längre vistelse 1981–82 i Texas. Då var jag praktiskt taget klar med studien om hur svensk energipolitik utvärderas.¹⁰

Också den här gången åkte jag till en tvärvetenskaplig public policy school. Jag for till Lyndon Johnson School of Public Affairs i Austin, Texas. Tänkte jag skulle få se en annan del av Amerika. Jag ville också till en public policy school, men inte till samma. Och när jag satt där på mitt kontor – det här var en mycket viktig sak för mig som inträffade – så kom det plötsligt en lapp där det annonserades en konferens i Baltimore, i oktober 1982. Den där konferensen hette någonting på "evaluation". Den sponsrades av en organisation för "evaluation". Jag hajade till och tänkte: Herregud, finns det en sån organisation? Det här ska jag åka på. Jag har ju pengar från DFE. Jag blev jättenyfiken.¹¹

Jag såg att de hade försessioner och det var en som handlade om Stakeholder evaluation, så då tog jag den. Det var en hel dag och det var Egon G Guba och Yvonna S Lincoln som höll i den. Tidigare hade jag inte riktigt fattat det där med intressentmodellen, men efter den sessionen så fattade jag den och förde in den i mina anteckningar. Nu lärde jag känna de två och det fanns några andra där också. Det var ett tag där jag kände alla ordförandena i American Evaluation Association, för jag fortsatte att åka på de där mötena¹². Jag åkte egentligen varje gång på de där, och jag gillade de där försessionerna för då kunde man smyga iväg och lära sig någonting nytt. Så småningom fick jag med mig Jan-Eric Furubo och Rolf Sandahl, en annan upsaliensare. Även Sture

10) Vedung, Evert (1982) Energipolitiska utvärderingar 1973 – 1981. DFE Rapport nr 52. Stockholm: Liber.

11) Konferensen hette Evaluation 82: Joint Meeting of Evaluation Network and Evaluation Research Society.

12) Evaluation Network and Evaluation Research Society slogs senare ihop till American Evaluation Association (AEA).

Blomgren och Ola Nyquist vid Byggforskningsrådet kom med på sådana där konferenser. Vi åkte iväg tillsammans.

Så det där som inträffade på hösten 1982 fick en avgörande betydelse för mig och mitt utvärderingsintresse. Det var en viktig sak, men det var alltså en slump det också – en serendipitet igen. Jag visste inte att en sådan organisation fanns tidigare, utan det blev att jag hakade på det där och började lära mig väldigt mycket av amerikanerna. Jag började också förstå att det här med utvärdering kanske kan bli någonting stort. Det verkade bli viktigt det här. Men det var först då som jag började begripa den internationella storheten och kraften.

Om man tar de så kallade berättelserna¹³ om utvärderingshistorien i Sverige, så utgår de till stor del från ett strukturperspektiv, det vill säga att man förklarar uppkomst och tillväxt genom strukturer. När du berättar handlar det mycket om att olika aktörer träffas och så väcks idéer och man tar dem med sig. Hur har de här två spåren – aktörer och strukturer – hängt samman?

– Den typen av frågor är viktiga att ställa, för jag tror det är så att utvärdering också har drivits av typer av public policy som regeringar har hållit på med. Det har drivits av tänket i offentliga sektorn på något vis. På sextiotalet tänkte man ju att nu ska vi planera bort allt elände och det ska ske centralt. Vi ska ha en "Gosplan"¹⁴ i princip, och i denna Gosplan ska vi föra in experimentella resultat och på den vägen ska vi utrota allt tjafs, ungefär. Det var en vänstervåg som stod bakom utvärdering på det sättet, det skulle vara experimentella metoder. Det där finns absolut, fast de

13) I en artikel från år 2000 slår Vedung, Furubo & Sandahl fast att vi saknar bred kunskap om den svenska utvärderings historia. Tre berättelser är vad som står till buds: Berättelse 1 säger att Sverige importerade utvärdering från USA vid mitten av 1960-talet i samband med den stora expansionen på det socialpolitiska området ("War on Poverty" och "the Great Society"). Fel, säger Berättelse 2 som menar att det hela startade i början av 1950-talet med insatser från universitetsbaserade pedagoger på området utbildningspolitik. Även detta är fel säger Berättelse 3 som menar att åtminstone medvetenhet om vissa grundläggande metodproblem i utvärdering fanns långt tidigare (Vedung, Evert & Furubo, Jan-Eric & Sandahl, Rolf (2000) "Utvärdering i det svenska politiska systemet: åtta iakttagelser". I: Nordisk Administrativ Tidskrift 2/2000, s. 113-130).

14) Gosplan var den sovjetiska regeringens och därmed kommunistpartiets centrala ekonomiska planorgan åren 1921-1991 (Källa: Wikipedia).

perspektiven har inte riktigt varit levande för mig själv. Jag kom ju från universitetsvärlden.

En annan sak är New Public Management, det är ju också otvivelaktigt så att utvärdering har drivits av nyliberalism – en högervåg. Den kommer ju från höger, nyliberalismen. Det spelar ingen roll att det är Göran Persson och Tony Blair som har blivit utförarna av den i Europa. Den har varit riktad mot att man vill röja upp och få mera effektivitet. Den har drivits från socialdemokratisk höger, man ville effektivisera fast på ett annat sätt. För skulle man effektivisera kapitalismen genom storslagen planering. Sedan märkte man att det där inte håller riktigt. Nu skulle vi effektivisera kapitalismen med starkare ledarskap i offentliga sektorn och vi skulle kunna decentralisera till kommunerna. Kommunerna får ta hand om skolan, sa Göran Persson, vi ska inte ha statliga direktiv. Så det har hängt ihop med de här vågorna i public policy.

Om vi backar lite, så börjar du med att beskriva dig som att du "varit lite grön i dina dagar". Sedan blev du statsvetare med inriktning på energifrågor och du har skrivit om utvärdering. Vad ser du dig själv som i första hand?

– Ja, vad är man för någonting? Jag är en statsvetare som håller på med offentlig politik och mer speciellt utvärdering, det är jag nog. Statsvetarna har ett perspektiv på det här som de andra inte har och det är att vi kopplar ihop utvärdering med demokrati, med det demokratiska systemet och demokratiuppfattningar. Vi har en systemsyn, en helhetssyn.

Sen tycker jag nog att det finns mycket i statskunskap som är av mindre intresse om man är intresserad av offentlig politik. Det här med partier och väljare och massmedia och främmande länder och sånt. Som utvärderare är man också intresserad av det dynamiska perspektivet – förändringar och effekter. Statsvetarna, när de studerar offentlig förvaltning, tenderar att titta på hur organisationsstrukturen ser ut, hur tjänstemän rekryteras, vad de har för åsikter och så vidare. Men de har inte det tidsperspektiv och den dynamik som man har i utvärdering och implementering. Det skapas en policy och sen vill ju vi titta på hur den genomförs och vad får

den för resultat där ute. Så vi i utvärderingsbranschen har det här dynamiska perspektivet, tror jag, vid sidan av det demokratiska.

Men vad är man för figur egentligen? Jag är statsvetare, men med inriktning på utvärdering, implementering och styrning skulle jag nog vilja säga. Jag är ju inte ekonom och jag är ju heller inte pedagog. Det finns ju en mängd discipliner i utvärderingsgemenskapen och jag tycker nog att någon bör stå för det statsvetenskapliga. Vi är för få. Vi är inte många statsvetare i det här. Förr i världen... Nej, jag ska gå rakt på. Det här med politisk filosofi har blivit viktigt i statskunskap, kring rättvisa och jämlikhet och så vidare. Men statsvetarna har inte riktigt begripit att utvärdering är normativ statskunskap. Av någon anledning har de inte förstått det. Det här att värdera offentlig politik – normativ statskunskap – de har inte begripit det faktiskt. För det andra har statsvetare inte fattat det här med programteorimetoden, de har inte tagit den till sig. De gör det, de håller mycket på med det, men de förstår inte vilken källa det är att ösa ur, hur stor diskursen är på det där området.

När blev programteorin viktig för din egen del?

– Våldigt tidigt. Men här var en intressant sak, för jag blev avskräckt av statsvetarna faktiskt. Jag har ju alltid tänkt programteorin som ett redskap och inte som en avbildning av verkligheten. Det är ju en idealtypisk rekonstruktion av ett tänk, men det är ju inte en avbildning av hur det faktiskt ser ut. Utan tanken är att det där ska vara ett redskap med vars hjälp man sedan tar reda på hur det ser ut, och då skulle man kunna rita en annan figur av det.

Men till exempel Jan-Erik Lane, han sa till mig några gånger, att det där är för rationalistiskt. Och Benny Hjern för all del också, han sa: det där är top-down och det är helt fel, det ska vara bottom-up. För de fattade faktiskt inte det här med idealtyp, vad det var för något. Jag tror att det har trängt igenom sedan, men man begrep inte på 1980-talet att det här kunde vara ett redskap. Så jag blev snarare avskräckt av statsvetarna på den här punkten.

För min del kom jag in i det här med programteori och interventionsteori hemskt tidigt. Redan i den där studien av energipolitikens utvärdering på 1970-talet, fast jag kallade det kanske inte så på

den tiden. Jag märkte vilken kraft det var i programteoriidén. Om det skulle man ha kommit ut med en artikel väldigt tidigt... Åtminstone på svenska. Helst på engelska. Det är en annan sak som är lite roligt att tänka på nu på ålderns dagar, missade chanser på grund av att man inte förstätt värdet av det man då sysslade med. När jag ser tillbaka, i det här gänget som lever nu, så var nog jag var en av de första som kom på det här med programteori.

Om jag förstår dig rätt så har olika former av nätverk betydtt mycket för dig?

– Ja, det började med att jag fick en särskild forskartjänst i utvärderingsforskning. Det var 1986 till 1991. I Humanistisk-samhällsvetenskapliga forskningsrådet, HSFR. Den perioden blev utomordentligt viktig för mig för det var under den tiden som den här boken kom till – Utvärdering i politik och förvaltning¹⁵. Den borde ha kommit långt tidigare, jag sölade med den för länge. Men den där tjänsten var väldigt bra för mig. Dock var det var inte Uppsala universitet som tog initiativ, utan det var Göran Hermerén från Lunds universitet som satt i Forskningsrådet. Han hade nog tänkt att Tore Nilstun skulle få tjänsten. Kanske tänkte han lite på mig också. Ungefär när jag fick tjänsten höll vi på att planera den där gemensamma forskarkursen i Båstad som jag nämnde tidigare. Så han visste vem jag var. Jag hade förresten sökt upp honom tidigare, jag gillade honom. Jag hade också gått på en av hans kurser i vetenskapsteori i Uppsala.

Det där har med nätverkande att göra – det är otroligt centralt, det har jag förstätt. Jag förstod det ännu mera i efterhand, och det är viktigt att man använder nätverk själv också. För mig har det varit viktigt det här med NOPSA, det nordiska statsvetarförbundet. Jag har själv lett många arbetsgrupper i NOPSA kring utvärdering. Jag kommer inte ihåg hur många gånger, men det är många. Ibland i förening med en försession tillsammans med Sören Winter till exempel, eller Hanne Foss Hansen. Ja, det är många andra som

15) Vedung, Evert (1991/1998/2009) Utvärdering i politik och förvaltning. Lund: Studentlitteratur.

varit med på det där. På det sättet har man blivit spridd och känd i det nordiska samarbetet, så det har varit kolossalt viktigt med just den spridningskanalen.

Hur har nätverken påverkat ditt tänkande kring utvärdering?

– Ja, det har delvis kommit influenser från det. Lite grann under själva sessionerna, men också genom att man sen har läst vad de här personerna har skrivit. Till exempel Hanne Foss Hansen, hon kom med det där med forskningsevaluering som är någonting annat än utvärdering av offentlig politik. Så under de här sessionerna har det kommit upp en del, men så har man träffats personligen eller läst varandras verk i efterhand. Så det har varit viktigt på två sätt tycker jag.

Det europeiska nätverket har inte varit lika viktigt, det har inte haft samma betydelse för mig som det här med NOPSA. Sedan är det "Ray Rist-gruppen"¹⁶, den har hållit på i tjugofem år nu. Jag är inte med längre i den där för jag tröttnade på den, men jag var med ganska länge. Jag tror faktiskt det har haft betydelse för spridningen av min amerikanska bok¹⁷, den har blivit översatt till massa språk. Vet du vad jag tror det beror på? Naturligtvis inte bokens kvalitet, utan att de tror att jag är amerikan. Jag tror det faktiskt, det är lite lustigt.

"Utvärdering i politik och förvaltning" får man väl säga är utvärderingsverket framför andra i Sverige. Hur tänkte du när du skrev den?

– Ja, den har nått en viss position. Som grundbok i utvärderingsteori och utvärderingsmetod. Tanken var att skriva en bok för statsvetare, tänkte jag väl mera. Det här är ju en relativt okänd sak i statskunskap, tänkte jag. Sen blir det så att man tänker igenom saker själv – det är en självbildning också. Men jag tänkte nog mera

16) The International Institute of Administrative Sciences (IIAS) är en internationell akademisk organisation som specialiserat sig på administration av offentlig sektor. Arbetet sker till stor del genom arbetsgrupper, varav en – Policy and Program Evaluation – har letts av Ray C. Rist.

17) Vedung, Evert (1997/2003/2006) Public Policy and Program Evaluation. Piscataway, N.J.: Transaction Publishers.

för statsvetare faktiskt. Sedan har den ju tagits upp i andra discipliner också och kanske framförallt i kurser om offentlig förvaltning. Ambitionen är ju att ge en helhetsbild.

När jag ser tillbaka så tycker jag att jag höll på alldeles för länge med den. Den borde ha kommit 1984–85, för jag var i praktiken klar med den då. Den kom ut 1991, sen kom den andra upplagan 1998 och nu kom den tredje 2009. Den här tredje har också tagit alldeles för lång tid, den borde ha kommit 2004 någonting. Men den har ju fått konkurrens, på svenska också. Det finns massa andra böcker nu, det finns en hel del författare på det här området.

Nu har du varit verksam inom utvärderingsfältet under fyra årtionden. Vilka skillnader och likheter ser du om du jämför hur det varit att jobba med utvärdering under den här tiden?

– Du menar hur det var förr och hur det har blivit nu. Ja, det har ju gått i vågor, men det har blivit kvar sediment från alla vågor. Ingen våg har helt försvunnit, utan det har blivit kvar sediment.

För det första tycker jag man har blivit mer mycket mer sofistikerad när det gäller det här med användning. Det där med användning har liksom alltid stått i centrum i utvärdering. Inte i statskunskap, man har inte tänkt så. Utvärderingsskrået är otroligt mer sofistikerat när det gäller användning av kunskap, och ändå är det där otroligt viktigt i all forskningspolitik. Den svenska forskningspolitiken går också ut på att vår kunskap ska komma till användning. På den punkten har utvärdering bidragit till sofistikationen hos dem som har deltagit i den där diskursen, även statsvetarna.

För det andra tycker jag det är det där med mot vilka värdekriterier man kan bedöma en offentlig politik. Här har det skett en väldig utveckling, det är en ganska stor variation i det där med värdekriterier. Det är inte bara mål och inte bara intressentmål, inte bara brukarmål och inte bara produktivitet och effektivitet, utan det är liksom också hållbarhet, participation och demokrati. Man kan anföra en hel katalog. Det här vore ju någonting som normativ statskunskap skulle kunna ta till sig mera, tycker jag. För här är verkligen utvärderingsdisciplinen sofistikerad, tycker jag.

För det tredje tycker jag att utvärdering borde kunna inspirera förvaltningsforskningen mera i statskunskap. Man har nått fram till en stor sofistikerad gällande styrformer i utvärderingsbranschen. Det är ju inte bara det här med sektoriell politik och program inom den, utan vi har till exempel det här området integrerad politik. Du vet att jämställdhet ska inte vara en sektor, det ska integreras i alla sektorer. Hållbar utveckling ska inte vara en sektor, utan det ska integreras i alla. Integration av invandrare ska inte vara en sektor, utan det ska integreras i alla. Det är en styrform som har uppstått i offentlig politik, men som egentligen inte har kommit in i läroböckerna än, men som i högsta grad är levande bland utvärderare. En annan sak är samverkan. Horisontell tvärspektoriell samverkan.

Vad skulle du vilja säga att det innebär att vara utvärderare i första hand – vad är essensen av att vara utvärderare?

– Ja, bedömningen är ju en del av det. Men inte bara den. Essensen är ju också att förstå det här med offentlig politik, tycker jag. Att kunna typologisera den. Är det integrerad politik eller är det mål- och resultatstyrning? Vad är det underliggande problemet? Att typologisera olika inslag i offentlig politik tycker jag också tillhör den här branschen. Michael Scriven, han skulle säga att det bara är värdefrågan. Det tycker inte jag, utan det är också att se att det här hör ihop med demokrati och offentlig politik på något vis. Att kunna se den kopplingen.

Är det där som din bakgrund som statsvetare kommer in?

– Ja det tror, ja det tror jag. Scriven är en utomordentlig författare i utvärdering. Han är filosof, men han tänker som en kund på en marknad även när han skriver om utvärdering. Ofta hittar man inget politiskt system i hans skrifter. Det är nästan som om man inte behöver bry sig om det bakomliggande, det som finns bakom kundvalet. Men här kommer det statsvetenskapliga in. Jag tycker man ska koppla ihop det med styrkedjor och sånt där, att offentliga sektorn har att göra med demokrati, medborgare, representanter och förvaltning.

Ska vi försöka sy ihop detta lite? Om du ser tillbaka på din forskargärning, vad skulle du vilja säga är ditt största bidrag till utvärderingsdiskursen?

– Det tror jag är översikten, det är systematiken. I systematiken ligger min styrka, liksom överblicken över fältet. Både i Demokrati och brukarutvärdering¹⁸, översikten där med de sex argumenten för brukarinflytande, och i Utvärdering i politik och förvaltning, med översikten över modeller eller över hela problemfältet. Jag skulle vilja säga systematiken på något vis. Det är nog där jag ser det själv, min egen insats. Jag tycker inte jag har kommit på någon egen modell egentligen, utan det är snarare översikten.

Om du ska peka på någon person och någon händelse under karriären som har haft särskilt inflytande på dig vad skulle det vara?

– Ja, det var kanske Guba och Lincoln på den där försessionen 1982. Det var en viktig sak. Att Göran Hermerén ordnade en särskild forskartjänst i utvärderingsforskning. Att Sigfrid Wennerberg skickade runt den där skrivelsen till statsvetare i energisvängen om vad man skulle kunna bidra med och att jag sen då fick det där anslaget för den där studien om utvärderingsfunktionen i energipolitiken. Att jag hade börjat med energipolitik genom att Leif Lewin ville ha oss docenter med i Politik som rationellt handlande. Åtminstone de fyra är det väl.

Ja, när man tittar tillbaka finns det en massa serendipiteter, oväntade saker som är oplanerade, men där man ändå måste ha någon sorts beredskap för att möta dem. Vid den där NOPSA-diskussionen som pågick i Reykjavik var ett antal speciella danskar med, det tror jag har haft betydelse för min roll i Danmark. Bara den grejen till exempel att Henning Jørgensen var där. Där har du fem. Sen har vi det där med Furubo naturligtvis och Rolf Sandahl och tillkomsten av Svenska utvärderingsföreningen. Det är ett antal personer. Ove Karlsson Vestman till exempel på senare tid.

18) Dahlberg, Magnus & Vedung, Evert (2008) Demokrati och brukarutvärdering. Lund: Studentlitteratur.

Han som värvade mig till Mälardalens högskola efter min pensionering i Uppsala. Ja, det blir ett ganska stort galleri. För Danmark kan jag ju nämna Hanne Foss Hansen, Peter Dahler-Larsen, Morten Balle Hansen, Henning Jörgensen, Hanne Kathrine Krogstrup, Martin Steen Jensen, Poul Skov Dahl. För Finland Stefan Sjöblom, Per Mickwitz, Tuija Kotiranta. Det är ju alla studenter också, studenterna ska man inte förglömma. Det har varit en prioriterad sak för mig det här med studenter. Det var först på förvaltningslinjen, femte och sjunde terminen. Sen blev det pol.mag.programmet. Det var inte så förfärligt många studenter till att börja med. Men de blev i vart fall väldigt intresserade. Vi brukar säga så här att de kom in där och hade stort förakt för offentlig förvaltning som det tristaste som fanns, men det tog vi ur dem efter några dagar.

Om jag får avsluta med en fråga om du vill ge några tips och råd till framtida utvärderare baserat på dina erfarenheter och lärdomar vad är det?

– Jag tycker nog en viktig sak är det här med programteorimetoden, att man lär sig det ordentligt. Om man ska vara kortfattad, så personliga kontakter – det här med nätverk. Man ska inte tro att det bara är att åka på konferenser, man ska göra en rejäl presentation när man åker. För det är presentationen som är det viktiga, inte synpunkterna man får, för det får man nästan inga. Man ska göra en ordentlig presentation och ta det på allvar, den där spridningsfunktionen. Överhuvudtaget lyssna på alla, man vet aldrig vad det blir av dem... Var intresserad.

Linné hade ett valspråk *Omnia mirari etiam tritissima*, undra över allt även det mest all dagliga. Det var hans valspråk när han blev adlad och det finns på den svenska hundralappen. Ibland när jag är fakultetsopponent, så brukar jag roa mig med att förstora upp hundralappen så publiken ska se den annars osynliga texten.

KAPITEL 22

Om serendipiteter och Svenska utvärderingsföreningen

Sanja Magdalenic

Det var en serie serendipiter som ledde Evert Vedung till utvärdering. Det framgår i hans egna berättelser om sin forskarbana (till exempel Vedung, 2003). Serendipitet har kommit att bli ett ord med varierande betydelser. I Svenska Akademiens ordlista beskrivs den som "förmåga att göra värdefulla upptäckter av en slump". Sociologen Robert K. Merton (Merton & Barber, 2004) diskuterar emellertid serendipitet som ett psykologiskt, sociologiskt och teoretiskt fenomen relaterat till vetenskapliga upptäckter som kan handla om "the discovery by chance or sagacity of valid results that were not sought for", "looking for one thing and finding another" eller "observing an *unanticipated, anomalous, and strategic* datum which becomes the occasion for developing a new or extending an existing theory". I det här kapitlet vill jag rikta ljuset mot Svenska utvärdering föreningen (SVUF), där Evert, till följd av serendipiteter och tillsammans med andra satte sig själv – och andra – i arbete. Evert Vedungs gärning som utvärderingens linnean¹ är sammanflätad med SVUF:s tillkomst och utveckling.

1) Begreppet linnean används här i en vidgad betydelse att kartlägga, systematisera och ordna ett kunskapsområde, vilket kan relateras till den linneanska traditionen (Carl von Linné), i linje med hur begreppet används i boken Samhällets linneaner (Eriksson, B.E. & Qvarsell, R. 2000).

Fokus på SVUF är ingen tillfällighet även om jag också upptäckte utvärderingsfältet av en tillfällighet. Det är genom akademiska sammanslutningar som jag lärt känna Evert. SVUF, Sveriges sociologförbund, European Evaluation Society (EES) kommer till minnet som fora där han samtalande om utvärdering och visade konsten att nätverka. När detta skrivs sommaren 2023 är jag styrelseledamot i SVUF och SVUF fyller 20 år. Att det redan gått 20 år sedan föreningen grundades upptäckte jag av en slump när jag skulle betala medlemsavgiften för 2023.

Teoretiskt inspireras kapitlet av modern professionsforskning som intresserar sig för vad professioner och professionella gör i olika organisatoriska kontexter (till exempel Abbott, 1988; Evetts 2011). En typ av organisation som kan vara en del av professioners sociala struktur är professionella sammanslutningar. De bildas av olika anledningar och i olika syften samt bidrar till att skapa och reproducera gränsdragningar mellan och inom professioner, och är därför viktiga att studera (Magdalenic, 2004). Kapitlet grundar sig i empirin, främst arkivmaterial. Följande frågor belyses: Hur gick det till när SVUF grundades? Hur utformades gränsdragningar kring tänkt målgrupp i föreningen och kunskapsbasen, utvärdering? Finns andra gränsdragningar som kan identifieras?

BERÄTTELSE OM SVUF:S TILLBLIVELSE

Här följer en rekonstruktion av hur SVUF blev till som utgår från befintliga dokument om Svenska utvärderingsföreningen.

Det finns olika berättelser om utvärderingens utveckling som fenomen och om själva begreppet utvärdering (Vedung m.fl., 2000). Det visade sig att även olika berättelser om när SVUF grundades förekommer.

SVUF:s arkiv för perioden 2003-2007 förvaras på Riksarkivet. Arkivhandlingar för perioden 2003-2005 har Evert deponerat; de kommer från hans tid som föreningens förste sekreterare. I beskrivningen av innehållet i Svenska utvärderingsföreningens arkiv i Nationella Arkivdatabasen, under rubriken "Syfte och innehåll", finns *berättelse 1* om när SVUF grundades:

Föreningen grundades 2004 av Evert Vedung m.fl. Den har ca 250 medlemmar (2007). Ca 20 är institutionella medlemmar inklusive flera statliga verk. Föreningen medverkade i att bilda European Evaluation Society.

Mot bakgrund av dokument ur arkivet, som medlemmarna själva har skrivit, går det emellertid att argumentera att beskrivningen ovan inte är historiskt korrekt. I *Verksamhetsberättelse för 2003 – föreningens första år*, undertecknad av Jan-Eric Furubo och Evert Vedung, går exempelvis att läsa att diskussionen som ledde till ett upprop om att bilda en förening för utvärdering i Sverige fördes på European Evaluation Societies (EES) konferens som ägde rum i Sevilla i oktober 2002. Genom dokumentets titel meddelas *berättelse 2*, att föreningen bildades 2003, och i texten framgår att EES existerade innan uppropet om att bilda föreningen skickades ut. EES bildades drygt tio år tidigare, 1992, och konferensen i Sevilla var organisationens femte konferens.² Däremot deltog enskilda personer som kom att bli medlemmar i SVUF i bildandet av EES, exempelvis Evert Vedung.

Hur kommer det sig att föreningen bildades just då, 2003? Evert har i en intervju (se Tranquist i denna antologi) lyft ett internationellt tryck ”som vi kände på oss” som en utlösande kraft. Utöver EES hade en del andra länder, exempelvis våra grannländer Danmark och Finland, redan bildat sina nationella föreningar på utvärderingsområdet. Så här har Evert (ibid.) berättat:

Sen bildades en förening i Danmark och så bildades en förening i Finland. Och så satt vi några stycken svenskar där 2002 i Sevilla på europeiska utvärderingskonferensen: ’Varför inte bilda en svensk förening, vi som har varit med ända från början?’ Det här med utvärdering som är så viktigt i Sverige, vi måste nästan ha en förening för skams skull. Så satt vi där och drack öl i ett hörn. Jag tror Anders Hanberger var med och Ove Karlsson. Rolf Sandahl vankade omkring i lobbyen. Furubo satt där vid sidan om: ’Ja, vi tar ett initiativ’, sa han. ’Vi måste sätta igång. Kan vi inte bilda någon sorts grupp, så vi sätter igång med det där’. Och han satt ju på Riksrevi-

2) www.europeanevaluation.org

sionsverket i Stockholm, så då visste man att det finns uppbackning bakom.

I citatet ovan kan flera "ledtrådar" identifieras. Dels framställs att det har funnits en förhistoria, dels ett "vi"³ som har varit en del av den, "vi som har varit med ända från början". Vidare lyfts att tajmingen uppfattades vara rätt, "utvärdering som är så viktigt i Sverige", viljan tycktes finnas och organisatoriska förutsättningar för att "sätta igång" föreningens bildande upplevdes också existera. Sagt och gjort. En "stadgegrupp" bildades, i vilken Jan-Eric Furubo, Rolf Sandahl och Evert Vedung ingick, och utarbetade ett stadgeförslag, *Stadgar för Sveriges Utvärderingsförening (Swedish Evaluation Society)*. Uppropet om att bilda en svensk förening för utvärderare, som undertecknades av flera namn, skickades ut huvudsakligen via e-post under perioden november till december 2002. Bland annat användes Nätverket Utvärderarnas mejlista, som i januari 2003 uppgavs ha bestått av över 700 medlemmar.

I SVUF:s arkivmaterial finns ett mindre antal kommentarer till uppropet, i Everts urval. Bilden som träder fram genom urvalet är att responsen till initiativet var positiv. Några lyfte att det funnits tidigare diskussioner om att bilda en utvärderingsförening och att det fanns befintliga nätverk, uttryckligen nämns Utvärderarna och Utvärderingsringen. Vissa undrade över gränsdragningar kring den tänkta målgruppen, andra över gränsdragningar gentemot andra befintliga nätverk. Någon kom med tips om hur man bildar en förening, en annan erbjöd sig att hjälpa till med stadgarna. En kvinna⁴ som kommenterade uttryckte det så här:

Kul initiativ med en svensk utvärderarförening! Jag var ju med på några tidigare diskussioner om detta med Ingrid Munck⁵, men då fanns inte riktigt viljan eller kanske orken att dra igång detta. Nu hoppas jag att läget är annorlunda.

3) De som nämndes vid namn skulle senare ta styrelseposter i SVUF.

4) Citat från otryckta källor har avidentifierats.

5) Ingrid Munck ledde Nätverket Utvärderarna.

Jag dyker självklart upp! [...] Jag kunde tyvärr inte vara med i Sevilla (hur många rior tog det innan ni kläckte fram idén om en förening?).

Uppropet om att en ny förening var i pipeline skickades utöver professionella nätverk även till privata sådana. Ett exempel på det senare är ett e-postmeddelande med rubriken "Svensk Utvärderingsförening" adresserat till "family". Där skriver Evert så här: *"Hejsan Svejsan. Här får ni se vad jag tänker ställa till med. Kram Evert."* "Verkar mycket bra pappa! Mycket bra.", svarade en av Everts söner och kom med vad han kallade "spontana funderingar" om medlemskapsformer, vad medlemmarna ska få för sitt medlemskap, om föreningen ska ha en hemsida, vikten av att tänka på vilka värderingar som ska genomsyra organisationen och varför man bör vara med i organisationen. Några av dessa frågor diskuteras än idag.

Hur den nya föreningen skulle benämnas var inte givet. Av dokumenten i arkivet framgår att det till en början har funnits flera versioner av föreningens namn. I mejlen som avser uppropet förekommer varierande benämningar: Förening för utvärderare, Svenska föreningen för utvärderare, Svensk förening för utvärdering, Svensk utvärderarförening och Svensk utvärderingsförening. I ett stadgeutkast som stadgegruppen utarbetade och diskuterade på ett möte i december 2002 framgår att bokstäverna "ar" i den ursprungliga benämningen "Sveriges Utvärderarförening" ströks och ersattes av "ings", så att benämningen ändrades till "Sveriges Utvärderingsförening". I dokumentets högra marginal står också "Sverigeutvärderarna" skrivet med skrivstil. Det verkar som att både "utvärderare" och "utvärdering" övervägdes och användes på prov.

Konstituerande möte för en svensk utvärderingsförening ägde rum den 16 januari 2003 på Norra Latin i Stockholm. Totalt 72 personer närvarade inklusive stadgegruppen "på podiet", enligt en deltagarlista som finns i arkivet.

Vilka antaganden fanns det bakom föreningens uppkomst? Enligt protokollet som Evert har skrivit så lyfte han i inledande diskussion om bildande av en "oberoende svensk utvärderingsförening" fyra skäl:

- Medlemsunderlag till en förening ty utvärderandet har växt till utomordentligt kraftigt.
- Mångfalden av ansatser på utvärderingsområdet skapar behov av en fast mötesplats.
- Att utvärdering inte är ett ämne utan ett mångvetenskapligt fält vid universiteten på vilket kulturgeografer, ekonomer, pedagoger, sociologer, statsvetare, psykologer, statistiker, socialmedicinare m.fl. gör insatser talar för en samlingspunkt.
- Utvärdering lockar människor i olika världar: myndighetspersoner, forskare och konsulter.

När det gäller föreningens tänkta målgrupp beskrivs det finnas en stor skara som antogs behöva en "fast mötesplats". Utvärdering framställs som ett tvärvetenskapligt fält vid universiteten som antas "locka människor i olika världar: myndighetspersoner, forskare och konsulter". Efter diskussionen beslutade mötet att bilda en svensk utvärderingsförening.

Den nya föreningens namn var också en punkt på dagordningen. Fyra förslag, med enbart utvärdering i namnet, togs upp för diskussion: "Sveriges utvärderingsförening", "Svensk utvärderingsförening", "Svensk förening för utvärdering" och "Svenska utvärderingsföreningen". Mötet beslutade att föreningens namn skulle vara Svenska utvärderingsföreningen. I och med det föredrogs termen "utvärdering", vilket förefaller vara i linje med benämningar av motsvarande internationella sammanslutningar på utvärderingsfältet. Det verkar som att den engelska benämningen var lättare att komma överens om eftersom stadgegruppens förslag Swedish Evaluation Society beslutades.

Ytterligare en fråga som diskuterades var den nybildade föreningens relationer till andra existerande organisationer. Specifikt

nämndes Danska Evalueringsselskab, Finska utvärderingsförbundet, European Evaluation Society och Nätverket Utvärderarna.

På det konstituerande mötet valdes en interimsstyrelse, i vilken positionerna fördelades på följande sätt: Jan-Eric Furubo (ordförande), Ove Karlsson (vice ordförande), Rolf Sandahl (kassör), Evert Vedung (sekreterare) och övriga ledamöter (i alfabetisk ordning) Karin Alexandersson, Peje Bengtsson, Anders Hanberger, Per-Åke Karlsson och Sanna Palomaa.

Bland uppgifter som interimsstyrelsen arbetade med var att informera om den nya utvärderingsföreningen både nationellt och internationellt. I arkivet finns en text på engelska med titeln "The Swedish Association for Evaluation has been established", där Evert står som kontaktperson. Föreningens benämning väckte min nyfikenhet. Jag sökte en publicerad version av den här texten i tidskriften *Evaluation* och fann en notis om en survey som The International Organisation for Cooperation in Evaluation (IOCE) genomförde för att sammanställa vad som beskrivs vara en uppdaterad lista av alla aktiva webbplatser som tillhör nationella utvärderingsorganisationer. På listan, som är daterad december 2003, finns också Sverige och Swedish Evaluation Society. Men det är inte Svenska utvärderingsföreningen som lyfts. I *News from the Community* (2004:248) meddelas så här: *SES: Swedish Evaluation Society: The Evaluators – www.statskontoret.se/utvarderarna*.

Vid ett styrelsemöte längre fram togs även förkortningen av föreningens namn upp för diskussion. Evert skrev i protokollet att frågan väcktes av honom själv och att den aktualiserades i samband med styrelsens arbete med hemsidan. Tre förslag dryftades:

- SES, Swedish Evaluation Society
- SUF
- SVUF

SUF visade sig vara upptaget av Syndikalistiska Ungdomsförbundet och styrelsen beslutade att förkortningen skulle vara SVUF.

Vad var tanken med föreningen, vad skulle göras? Föreningens handlingsramar utformades och formaliserades i stadgarna som "Föreningens ändamål att främja" så här:

- En fortlöpande och mångsidig diskussion om utvärderingens och utvärderarens roll i samhället och om inflytandet över utvärderingsverksamheten
- Den praktiska och vetenskapliga kompetensen bland utvärderare
- Internationella kontakter

Gränsdragningar kring målgruppen konstruerades som öppna men villkorades: "medlemskap i föreningen är öppet för varje individ som vill främja föreningens ändamål". Även institutioner och organisationer kunde antas som medlemmar. Gränsdragningar kring utvärderingsområdet framställdes också som öppna. Mångfalden av olika perspektiv och förhållningssätt på utvärderingsområdet "bejakas" i föreningen, står det i stadgarna.

I det följande fokuseras på hur gränsdragningar kring utvärderingskompetensen utformades och förmedlades via en kurs inom ramen för SVUF:s verksamhet.

PRESENTATIONER AV UTVÄRDERING

I KURSBESKRIVNINGAR AV "UTVÄRDERINGENS ELEMENTA"

Ett återkommande inlag på SVUF:s konferenser har varit en kurs i utvärdering med titeln "Utvärderingens elementa". Evert har givit kursen från och med SVUF:s första konferens 2004 i form av en försession, i enlighet med beslut av SVUF:s styrelse. När detta skrivs har Evert tackat ja till att hålla kursen också på SVUF Jubileumskonferens i december 2023.

Kursbeskrivningar av "Utvärderingens elementa" ger inblickar i hur Evert tänkte att grunderna i utvärderingskompetensen är och hur det skulle presenteras för kursdeltagarna.

Det allra första kurstillfället ägde rum den 21 april 2004 och beskrevs i konferensprogrammet så här:

Kursen är avsedd för nybörjare och ägnas åt utvärderingens grunder. Följande teman behandlas. Olika utvärderingsbegrepp: *ex ante*, *ex post*, *ex nunc*. Utvärdering, uppföljning, utredning. Utvärderingsmodeller: målpuppfyllelseutvärdering, bieffektsutvärdering, brukarorienterad utvärdering, intresmentmodellen (participatorisk utvärdering) ekonomiska modeller samt självvärdering och kollegetbedömning. Uppläggningar såsom experiment, kvasiexperiment, tidserier, skuggkontroller och processinriktade fallstudier. Användning av utvärderingar: instrumentell, konceptuell, legitimerande, interaktiv och rituell användning.

Kursen strukturerades därmed kring fem grundkomponenter: utvärderingsbegreppet, förhållandet mellan utvärdering och närliggande aktiviteter, utvärderingsmodeller, utvärderingsupplägg och användning av utvärdering. Genom dessa komponenter konstruerades och förmedlades samtidigt en ram för utvärderingskompetensen – vad utvärdering är och vilka kunskaper personer som arbetar med utvärdering förväntas ha och använda vid utvärdering.

Evert delade ut åhörarkopior och rekommenderade försessionsdeltagarna att inför kursen bläddra i valda kapitel i hans bok *Utvärdering i politik och förvaltning*. Åhörarkopian "Utvärderingens elementa", kurstillfället 2005, omfattade sammanlagt 66 sidor.⁶ Innehållet bestod inledningsvis av litteraturtips om utvärderingslitteratur, mestadels författat av Evert men också i samarbete med eller av andra författare. Därefter presenterades en översikt över de fem grundkomponenterna och inslag som exempelvis interventionsteori/programteori. I beskrivningen av försessionen i konferensprogrammet för 2005 års SVUF-konferens lyftes specifikt även utvärderingars organisering, utvärdering i offentliga sektorns beslutsprocesser, programteorier, programteorimetoden, utvärdering som produkt, utvärdering som process, faktorer bakom användning och strategier för användning.

6) Anna Ahlms privata SVUF-pärm.

Tio år senare, 2015, presenteras innehållet på ett annat sätt i kursbeskrivningen i konferensprogrammet:

Försessionen vänder sig till dem som vill få en enkelt framställd, elementär orientering i vad utvärdering är och vilka frågor som ställs och belyses i utvärdering. Sessionen inleds med en definition av begreppet utvärdering bl.a. i förhållande till begreppet uppföljning. Ur innehållet:

- Hur utvärderaren bygger upp interventionsteorier (programteorier) samt på vilket sätt dessa kan användas i utvärderande.
- Utvärderingens grundläggande ändamål.
- Utvärderingsmodeller som exempelvis måluppfyllelseutvärdering, intressentmodellen, EU-modellen m.m.
- Effektlärans terminologi som exempelvis huvudeffekter, nolleffekter m.m.
- Effektfrågans grundläggande logik.

Effektfrågan lyfts särskilt även i kursbeskrivningen till 2017 års kurstillfälle liksom nyare inslag som "process tracing" och "realistisk utvärdering". Så här står det:

Effektlärans terminologi med avsedda - oavsedda och förutsedda - oförutsedda effekter samt huvudeffekter, nolleffekter, perversa effekter och bieffekter går igenom. Effektfrågans grundläggande logik klarläggs (evidensrörelsens slagord *what works? vad är det som funkar?*). Metodiska upplägg för att kasta ljus över effekter dryftas särskilt fallstudier med processpåring (process tracing) och realistisk utvärdering.

Kursbeskrivningar till och innehållet i "Utvärderingens elementa" tycks ha uppdaterats kontinuerligt. Så här låter det i kursbeskrivningen i programmet 2019 års SVUF-konferens där även styrningsformer berörs:

Försessionen ger en introducerande översikt över vad termen utvärdering står för samt en genomgång av utvärderandets

åtta fundamentala frågor (Åttapunktsprogrammet). För-sessionens styrka ligger i helhetsbilden och systematiken. Sessionen inleds med en bestämning av vad ordet "utvärdering" betyder (evaluation, da. evaluering). Begrepp som ex ante, ex post, formativ, summativ, enkel uppföljning, kvalificerad uppföljning och utredning och avstämning går igenom. Hur evidensvägen, löpande utvärdering, inflätad utvärdering och New Public Governance tränger delvis undan New Public Management och det nyliberala kommer att beröras. Användning av interventionsteorier (programteorier) anses vara en uppfinning gjord i utvärderingsbranschen. Den kan läggas till grund för utvärderingsmodeller som måluppfyllelseutvärdering, bieffektsutvärdering, brukarorienterad utvärdering, intressentmodellen, ekonomiska modeller samt självvärdering och kollegebedömning. Effektfrågan inklusive uppläggningar för att belysa den berörs. Sessionen behandlar slutligen olika typer av användning av utvärdering – instrumentell, konceptuell, legitimerande, interaktiv och rituell användning.

När jag började arbeta med utvärdering informerade mina kollegor mig att bland annat Everts bok *Utvärdering i politik och förvaltning* användes som en gemensam utgångspunkt vid genomförandet av uppdrag. Jag fick också Lena Lindgrens (2012) *Terminologihandbok för utvärdering* i välkomstpresent. Att ha gått Everts kurs "Utvärderingens elementa" användes som en slags kompetensmarkör. Kursen framställdes som en god introduktion för utvärderare i staten.

SVUF-KONFERENSER

En central del av SVUF:s styrelses verksamhet kan sägas bestå av att planera och genomföra en SVUF-konferens. Konferensens inriktning och övergripande tema brukar utformas i samspel mellan styrelseledamöternas kunskapsintressen och medlemmarnas förslag om konferenspresentationer. För en översiktlig bild av konferensverksamheten följer här en lista över konferensernas teman.

| ÅR | KONFERENSENS TEMA |
|------|--|
| 2004 | Utvärdering i Sverige – makt, kunskap och lärande |
| 2005 | Utvärderingssystem i Sverige – till vad och för vem? |
| 2007 | Utvärdering pågår! Pågår utvärdering? |
| 2009 | Utvärdering: Filosofi och hantverk |
| 2011 | Den goda utvärderingen; trovärdig, användbar, etisk och effektiv |
| 2013 | Utvärdering i politiken och politiken i utvärdering |
| 2015 | Utvärderingars värde – värden i utvärderingar |
| 2017 | Att utvärdera i en komplex värld |
| 2019 | Utvärderingens makt – maktens utvärdering |
| 2022 | Utvärdering för ett bättre samhälle |
| 2023 | Jubileumskonferens SVUF 20 år: Utvärderingens framtid och framtidens utvärdering |

Tabell 1. SVUF-konferenser 2004-2023

Som tabell 1 visar har SVUF:s styrelser mestadels valt utvärdering som övergripande tema för SVUF-konferenser. Konferensen 2017 är den enda där konferensens tema tycks beröra personer som ska göra jobbet.

Eftersom Evert har givit "Utvärderingens elementa" som för-session till SVUF-konferenser är det rimligt att anta att han även deltog på konferenserna. Ett antal deltagarlistor som finns i arkivet bekräftar det. Av konferensprogrammen framgår att Evert gjorde ett antal konferenspresentationer. Några teman var: Utvärderingsböljans former och drivkrafter (2004), Storstadssatsningens utvärderingssystem (med Nils Hertting) (2005), Implementering och utvärdering – skillnader och likheter (2017), Vittnesseminarium: 16 år med SVUF (2019), Projektifiering och utvärdering (2019) och Utvärderingens historia i Sverige (2022). Vidare framgår att Evert ställt upp som moderator, agerade kommentator eller initierade diskussioner inom ramen för "Open space". När det gäller det sistnämnda handlade diskussionerna om teman som "Löpande utvärdering" (2011), "Realistisk utvärdering" (2011) och "Utvärdering av innovationspolitik" (2015).

SVUF:s styrelse för 2018-2020 beslutade att utse Evert Vedung till SVUF:s förste hedersmedlem i samband med SVUF konferensen 2019. Det verkade som att styrelsen lyckades överraska honom med utmärkelsen. Styrelsen framförde också en sång till Evert på konferensmiddagen. När jag höll på att sammanställa texten till hyllningssången till melodin "It don't mean a thing if it ain't got that swing", som enligt en "källa" tillhörde Everts favoritlåtar, slog det mig att en hel del titlar till Everts böcker och artiklar passade utmärkt att sjunga till låten. Jag kunde gå omkring och nynna i huvudet *Ut-värd-er-ing i pol-it-ik och för-valt-ning, Ut-värd-er-ings-böljans form-er och driv-kraft-er, doo-ah, doo-ah, doo-ah, doo-ah, doo-ah, doo-ah, doo-ah!* Ut-värd-er-ing med swing!

AVSLUTANDE KOMMENTARER

Syftet med SVUF har varit att skapa ett forum för diskussioner om utvärderingens och, om än i mindre utsträckning, utvärderares roll i samhället. Samtidigt finns det individer och institutioner som arbetar med utvärdering som inte är med i dessa diskussioner. Utformningen och konsekvenserna av existerande gränsdragningar inom utvärderingsfältet behöver utforskas vidare.

EPILOG: ARKIVUPPROP GÄLLANDE SVUF:S ARKIV I SAMBAND MED TJUGOÅRSJUBILEET

Jag har ibland undrat om det kanske var så att Evert ville ha posten som SVUF:s förste sekreterare för att ta sig an arkivet. För tagit hand om arkivet har han gjort både under tiden som föreningens sekreterare och därefter. Under de senaste åren har Evert verkat som SVUF:s revisor med ansvar för verksamhetsrevision. Som SVUF:s sekreterare under samma period var min uppgift att skicka styrelseprotokoll till honom för revision. "Vill du ha styrelseprotokoll som jag ordnat och numrerat?", frågade han när han var klar med revisionsberättelsen, utövande systematiken även i protokollsammanhang. Det var bara att tacka och ta emot.

När Åsa Sohlman var revisor, lyfte hon inför tioårsjubileet viken av att SVUF:s styrelse ser till att materialet är "(någorlunda) komplett och hålla det uppdaterat till 2018 då det är dags för ett

nytt överlämnande till Riksarkivet”. Detta framgår av korrespondens från 2012. Jag skulle vilja ta tillfället i akt och göra ett nytt arkivupprop om SVUF:s arkiv i samband med tjugoförårsjubileet. Idag är det precis som vid tioårsjubileet viktigt att ta hand om SVUF:s arkiv och se till att det inte uppstår luckor i föreningens organisatoriska minne genom att välja vilka handlingar ska bevaras, hur ska de bevaras och när ska handlingarna lämnas över till Riksarkivet för att kunna vara sökbara inför framtiden. Detta är en styrelsefråga liksom föreningsmedlemsfråga. För att kunna skriva det här kapitlet var jag tvungen att komplettera med materialet som en kollega har sparat. SVUF:s arkiv behöver fyllas på och bevaras – inte lämnas åt slumpen.

Samtidigt handlar minnesbilder av utvärderingens och SVUF:s utveckling om mer än bara SVUF:s arkiv. Vilka bilder av utvärderingens utveckling finns det? Hur har personer som arbetat med utvärdering och varit medlemmar men inte styrelsemedlemmar i SVUF upplevt föreningens och utvärderingsfältets utveckling? Vilka bilder av samma utveckling har personer som inte varit med i SVUF? Det är angeläget att samla in fler berättelser om utvärderingens uppkomst och utveckling i Sverige.

Referenser

- Abbott, A. (1988). *The System of Professions. An Essay on the Division of Expert Labor*. The University of Chicago Press.
- Anna Ahlms privata SVUF-pärm, 2005-2022.
- Evetts, J. (2011). Sociological Analysis of Professionalism: Past, Present and Future. *Comparative Sociology*, 10, 1-37.
- Lindgren, L. (2012). *Terminologihandbok för utvärdering*. Studentlitteratur.
- Magdalenic, S. (2004). *Gendering the Sociology Profession: Sweden, Britain and the US*. Almqvist & Wiksell International.
- Merton, R.K. & Barber E. (2004). *The Travels and Adventures of Serendipity: a Study in Sociological Semantics and the Sociology of Science*. Princeton University Press.
- News from the Community (2004). Evaluation Website Survey – An IOCE Initiative. *Evaluation*, Vol 10(2), 247-250.
- Riksarkivet, Svenska utvärderingsföreningen, SE/RA/730790, 2003-2007
- Svenska utvärderingsföreningen, SVUF:s tre pärm, 2009-2014, 2016-2017, 2018-2019.
- Vedung, E. (2003). *Serendipiteter i mitt forskarliv*. IBFs Nyhetsbrev, 2003(2),s.6.
- Vedung, E. (2009). *Utvärdering i politik och förvaltning*. Studentlitteratur.
- Vedung, E., Furubo, J.E. & Sandahl R. (2000). Utvärdering i det svenska politiska systemet [1950-2000] – Åtta iakttagelser. *Nordisk Administrativ Tidskrift*, 81:2, 113-130.

KAPITEL 23

Intervention theory and outcomes of an international teaching collaboration: A personal reflection from Brazil

Luiz Pedone

INTRODUCTION

The Swedish government invited me in January 1993 to speak about the upcoming plebiscite where Brazilians would decide between a Parliamentary or Presidential system, as well as deciding on a Monarchic or Republican form of government (Pedone, 1994). At the time, I met Evert Vedung's graduate student, Mikael Roman. I had the chance to briefly talk with Evert at the 48th International Congress of Americanists, held at the Institute of Latin American Studies/Stockholm University, about Mikael Roman's research on environmental policies implementation included in the Treaty of Amazon Cooperation. In October 1998, Roman defended his doctoral thesis in Uppsala. As field research advisor I was invited, and again I met Evert.

With the creation of the ministry of Defense in 1999 the Brazilian interests in evaluation as a political science subarea increased. Both Lula I and II governments between 2003 and 2010 were marked by the rise on complex strategic defense programs. According to Eurico Figueiredo, the advanced industrial countries

give to the defense industry and to science and technology development “*an importance which corresponds to their strategic-political profile... and their international and national interests and objectives.*” (Preface to Pedone & Vedung, 2018, p. 16).

As a consequence, academic endeavors on defense and strategic studies grew exponentially in Brazil, in particular after the *Brazil Defense Network* was launched in 2005. Designed jointly by the ministries of Education and of Defense, it financed research and teaching graduate programs in civilian universities (for the first time) in collaboration with the military higher education institutions.

Eurico Figueiredo, then chair of the department of Political Science at University Federal Fluminense (UFF), was responsible for the pioneer and innovative Graduate Program in Strategic Studies in 2006. Both a master’s degree and a doctorate program were installed, later migrating to the newly created Institute of Strategic Studies/UFF in 2012.

After settling down in Rio in 2004, I was invited by Eurico Figueiredo, to pursue a second teaching career across the bay from Rio, at UFF-Niterói, immediately taking responsibilities for research methodology and public policy classes.

ORIGINS OF THE EXCHANGE PROGRAM (OR OUR “THEORY OF INTERVENTION”)

After writing on reduction of poverty and energy policies implementation, I began focus my analyses on defense policies in 2009. A Research Group on *Defense, Science & Technology and International Politics* (later renamed *Defense, Public Policies, and S&T* – short name *Defesa&Políticas*), was established under my coordination. This group had graduate and senior undergraduate students, colleague professors, researchers, and professionals from the military and civilian public service, and formed the future enrollment of Evaluation Courses between 2013 and 2017.

In February 2012, on his first trip to Brazil, Evert and I were having a dinner at the *La Fiorentina*, a famous meeting point for artists, journalists and intellectuals at Copacabana-Leme Beach.

Then we decided to advance an academic exchange cooperation between Sweden and Brazil, which included Evert's coming to UFF for the intensive courses in policy and program evaluation. On the counterpart, I planned research trips to Sweden (Uppsala and Stockholm Universities, later Försvarshögskolan – SEDU).

By the end 2012, the Research Group received a group of professionals from SAAB Gripen visiting Brasília, São Paulo and Rio for a Colloquium at UFF, which increased the student's interest in discussing policy evaluation and implementation analysis of military programs.

This context was perfect for evaluating the armed forces initiatives for greater defense capabilities through ambitious acquisitions programs with foreign defense powers. The negotiations taking place under the *FX 2 Program* included extensive science and technology transfer, especially because it was designed to substitute 30-years obsolete combat fighter operative on Brazilian Air Force and joint production. The two months experience in Sweden, when I interviewed defense policy-makers, industry executives (SAAB Gripen Program), and politicians like Björn von Sydow, also a former minister of Defense, was conveyed to students back home inciting them to have a hands-on experience. During the next 7 years, this intense exchange has resulted in an expressive academic production. The pandemics has brought everything to a grinding halt.

After research trip to Sweden in 2012-13, Evert and I planned the Intensive Evaluation Courses, which occurred every October between 2013 and 2017. After the period in Niterói, our dear friend and colleague at University of Brasília, Paulo Calmon, made arrangement for Evert's successful Evaluation presentations at governmental institutions, and seminars at the UnB Nucleus of Public Policy Research (NP3 – *Núcleo de Pesquisa em Políticas Públicas*), a successful center created by Calmon, myself and colleagues dedicated to research and offering continued education Policy Analysis courses to specialized professionals and public servants. That platform was fundamental to Evert's sojourns in Brasília, the political administrative center of Brazil.

HOW DID WE WORK (THE INTERVENTION)

Evert Vedung started visiting Brazil in our spring time (September-December) for our courses on *Policy Evaluation and Laboratory in Program Evaluation* at UFF. Between 2013 and 2017 about 15 students on average participated in a mix of UFF's graduate and senior undergraduate students, and professionals from the military and civilian public service.

The structure of the *Laboratory Defense, Public Policies and S&T* was used to launch the *Intensive Course in Public Policy Evaluation and Laboratory Practice of Governmental Programs Evaluation*. It was sequentially approved by the department, the graduate program, and finally registered at the Pro-Rectorate of Extension (Continuing Education). These steps were crucial to formalize Certificates to the concluding students.

Classes were in English, and the participants were required to make presentations in English for Evert's evaluation and suggestions for improvement and writing the final papers. We divided theoretical classes from directed studies ones. During the mornings, Evert presented the foundations of evaluation, from policy cycles and evaluation in the governance layers to the basic evaluation models (including side-effects) and the policy instruments-policy results – their obstacles and facilitators. The Directed Studies in the afternoon consisted of Laboratory of Evaluation Practice on a specific governmental program, a case study chosen by the students themselves or suggested by both Evert/myself to the students.

OUTCOMES AND RESULTS 1: STUDENT'S PAPERS AND HOW IT INFLUENCED THEIR CAREERS

Some persons are born with the capacity to change things around them, and unquestionably influence someone's life. Very few of them, as a matter of fact. Having this in mind, I can confidently state that Evert Oskar Vedung is one of them. During these years, and throughout the pandemics, I maintained contact with the participants of the courses. Many have written graduating thesis, master's thesis and doctorate dissertations utilizing evaluation

tools learned in the courses. In addition, I took this chance of this writing in celebration of Evert to conduct a very brief survey asking the participants two simple questions:

- How important was the Evaluation Course for the completion of their academic studies and for writing policy analysis/evaluation within the academic context.
- How important was the knowledge gained in the Evaluation Course for the professional work outside the academic context after the University.

Below you can find what some of them said about how learning evaluation theory and methods changed their lives and where they are now:

ROBERTO CEZAR ROSENDO, participating at a courses in 2015 (Docent at UFF-Campos dos Goytacazes, at the time Post-Doctoral Researcher, under my supervision, 2015-16) says: “I learned the foundations of policy evaluation and methods.” Rosendo is now teaching undergraduate and master’s in Development, Public Policies and Environment, advising master’s theses in housing and educational policies. “The knowledge acquired was very important for my challenging job as Director of the Institute of Sciences of Society and Regional Development (2017/2021).” He also attributes a book chapter he wrote with Silveira Lima (2018) to his participation in the course and published an article together with me in 2016. “I leave my immense appreciation to professors Evert and Pedone!”

YANCA BORGES, who participated in 2016, is another example. She is an undergraduate and Niteroi City Hall Director of strategic projects and cooperation with international organizations and universities. She has been accepted for a Master’s degree in Local Development at the University of Padova, Italy. She says: “Evaluation was the front door to understand methodologies of policy evaluation, reorienting my academic and professional life. I was part of the Work Group who wrote the Guide to Public Policy Evaluation, used in training of evaluators in Niterói’s Evaluation

System (...) Professor Evert and you were fine conductors of the course, and certainly influenced many students to follow a career in this area.”

Yet another example is KARINE DE PAULA, participant 2017, now undergraduate in Public Safety/UFF. She has a master's in Anthropology, UFF and a MBA in Municipal Public Management, UFF. She said: “The course was essential for all my trajectory, point of departure to knowledge in policy evaluation. (...) Today, I belong to the international network of young evaluators -- Eval Youth. Knowledge obtained in the course were fundamental for my current work with evaluating public safety programs in Niterói City Hall.”

Some more comments from previous students are listed in the table to the right.

| |
|---|
| <p>BARBARA OTTERO – 2013 and 2014 – Master’s student. Project manager, monitoring the implementation State Culture Programs. Executive assistant to the Director of Rio’s Municipal Theater. “All this was due to the course in Evaluation of Public Policy.” Book chapter with Violante (2018).</p> |
| <p>ALEXANDRE VIOLANTE – 2014 – Navy Captain. Master’s. PhD Student in Strategic Studies (2019-2023). Dissertation: evaluations and analysis of implementation of Brazilian Nuclear Propelled Submarine Program. Currently Post-Doctoral scholar, under my supervision. Book chapter with Ottero (2018). Published together two articles (2020; 2023), and book (2021).</p> |
| <p>GIORDANO BRUNO A RONCONI – 2015 – Master’s. Policy Analyst, Municipal Secretariats of Finance in higher strategic public management positions in Niterói and Rio de Janeiro. Last May 2023, took office as External Control Auditor, career at the Federal Court of Accounts for auditing federal economic development policies. “It was a very difficult selection examination, which demanded knowledge I learned in our Course Evaluation of Public Policies.” Book chapter (2018). Article published jointly (2017).</p> |
| <p>GUILHERME BERRIEL – 2014 – Master’s. Professional lawyer. Doctorate in Law candidate Universidade NOVA de Lisboa/Law School (2023 expected). “It was extremely important to my master’s thesis as it gave me the methodology for evaluating transfer of technology in three defense programs, receiving two awards. This would not have been possible without the course (...). After UFF, I lectured public policies at the Brazilian Bar Association (...) At NOVA de Lisboa I am managing a Project on Science and Research & Cooperation, based on my knowledge of the evaluation culture, concepts and methods (...) My compliments to professor Evert!” Joint article published, 2017.</p> |
| <p>LUCAS P. PINHEIRO DA SILVA – 2013 and 2014 – Undergraduate. Scientific Initiation Young Talent for Science. Group Research Member. Master’s, 2019. Currently PhD candidate in IR Program at University of São Paulo. “From the 2013 course resulted an article on AMX, a Poster of Scientific Initiation Evaluation Project, and a paper presented at Policy Work Group at ENABED 2014 (...) Prof Pedone coordinated a FAPERJ (Research Support Foundation of the State of Rio de Janeiro) named Evaluation of Public Policies of the Large Scale Defense Projects, which resulted in articles, monographs and dissertations after students took the courses with Vedung (...) On my Master’s thesis, I worked on Brazil’s decision-making process adhering to the Non-Proliferation Treaty (NPT) linked to Public Analysis as a whole. A great part of the form I analyze processes within the public sector is based on the knowledge acquired in the courses.”</p> |
| <p>VICTORIA V S GUIMARÃES – 2015 – Undergraduate. Research Group member. Graduating monograph and Master’s thesis received awards in 2017 and 2022. Currently a Doctorate candidate, International Relations Program, Universidade Estadual Paulista (UNESP). CAPES-Ministry of Education scholarship for Sandwich Period at University of Queensland, Australia (2023-24), where she is researching the AUKUS Pact. “I would like to congratulate and thank you and professor Evert Vedung for all the knowledge transmitted on the Course, which I consider a watershed in in my academic trajectory. The course stirred my interest for the defense and security area, what influenced my decisions, and it was fundamental to arrive where I am now (...) The Evaluation Course was my first contact with policy implementation and analysis/evaluation of governmental programs in defense, which led me to write on the theme. I worked with the FX2-Gripen program and the nuclear propulsion submarine, utilizing tools of policy analysis and evaluation offered in the course (...) I think on my object of study as a public policy (...) In my doctorate I belong to a Defense and Security Study Group and to a Research Network on Strategic Autonomy, Technology and Defense of the multi institutional project PROCAD-DEFESA, where I mostly apply policy analysis/evaluation on defense and security placing these contributions for the adjournments on the South America Defense Atlas.” Victoria published articles on policy analysis and evaluation coauthored with Lucas Pinheiro and myself, and a book chapter with Roberta Helena Moreira dos Santos (2018).</p> |

Figure 1. Comments from students.

OUTCOMES AND RESULTS 2: ACADEMIC MEETINGS WITH EVERT AND OTHER ACADEMIC, POLICY MAKERS AND PROFESSIONALS FROM SWEDEN AND BRAZIL

Another sphere of outcomes concerns academic meetings of different sorts, illustrated below. In 2014, I invited Gunnar Hult (from Swedish Defense University, SEDU) and Martin Lundmark (then FOI, now SEDU) to the National Meeting of the Brazilian Association of Defense Studies in Brasília. Moreover, several roundtables and workshops originated from the collaboration were organized at the EBERI – Brazilian Meeting of Strategic Studies and International Relations. Michael Mohr (vice-president of SAAB) and Ulf Hammarström (then director of FXM, Swedish Defense and Security Export Agency) have participated in 2015, as well as Brazilian researchers and Evert have participated on the 2016 and 2107 editions.

OUTCOMES AND RESULTS 3: BOOKS PUBLISHED

At the end of the 2014 Course, Evert and I decided that his classic book *Public Policy and Program Evaluation* (1997) book was crucial to Brazilian and other Portuguese speaking students, and needed an urgent translation.

On his sojourns as a visiting professor and researcher not only in Denmark and Finland, but also countries in Europe and Asia, and the United States, Vedung prepared a book that was not “a mere translation of my Swedish book of 1991, but an entirely revised product”, as he himself says on the foreword to the English edition (1997, pp. xiv). His book translation was published by Transactions Publishers, becoming an important, influential, definitive representative of scholarly work in public policy evaluation.

This time, Evert Vedung again not only rewrote, but also enlarged it almost entirely. Whole chapters have been reworked, specific topics added. Theory of intervention appeared upfront, and was characterized specifically as an evaluation tool, a foundation for the subdiscipline. The evaluation models were expanded and embodied in three new chapters.



Luiz Pedone, Evert Vedung and Eurico Figueiredo, 2013.

The initial translation was carried out by exceptional students linked to the Research Group Defense, Public Policies, S&T: Mariana Guimarães, Lucas Pinheiro, Urias Fernandes Jr, Leandro Pinotti and Pedro Henrique Peres. After many revisions, I tackled the most difficult task of revising paragraph by paragraph, comparing it with the original, and with Evert's revision and expansion. Particularly, adapting the language to the terms commonly used in political science, public policy, public administration in Brazil was intense, but very rewarding.

Evert used a partly new bibliography that did not exist when his book was published in English. This involved an extensive electronic survey to obtain the complete references of the books, papers, theses, research reports published after 1997 in Swedish, Danish, Finnish, English, and comply them to ABNT (Brazilian Technical Norms Association) rules. At the end of the book, I added a complementary list of Public Policy Periodicals and Professional Associations in Brazil, showing that, since the 1985 re-democratization, the field has grown significantly in Brazil. One of the most rewarding aspects of academic life is to broaden the reach of knowledge and see the dissemination of the fundamental concepts of this field of public policy. To achieve this goal, only an extremely generous, remarkable person like Evert Vedung could have offered me the co-authorship of this work now published in Portuguese. (Vedung & Pedone, 2021).

Over a beer on Clarion Amaranthen Hotel in Kungsholmen, in 2017, Evert and I came up with the idea about collecting the first-rate evaluation studies done by our students into a book. Our main

purpose was to disseminate evaluation knowledge on Defense and International Security programs in Brazil. Seven of the most significant works written were invited as contributions in what eventually became *Avaliação de Políticas Públicas: Programas Militares Complexos* (Pedone & Vedung, 2018).

Different procedures utilized in UN Peace and Stabilization Missions in Haiti and Ivory Coast are evaluated in one of the chapters. Another chapter evaluates defense planning processes in the period 2003-2015. More specifically, five chapters focused on the large-scale ongoing armed forces defense projects: 1. Submarine Program; 2. Surface Navy Program; 3. FX 2 (Gripen Program) as compared to the AMX; 4. Military Helicopter Program; 5. Armored Car Guarani. All chapters are referenced in figure 2.

CONCLUDING REMARKS

The indisputable amount of quality academic work produced by the participants show how important were the evaluation courses we offered during the period. By a simple examination of the results achieved not only during the courses, but also in their following studies and/or professional careers we arrive at the conclusion that much more than simply fundamentals learning of policy evaluation was achieved.

I could not be more thankful to Evert Vedung's compelling enthusiasm in coming down to Brazil and giving himself to students. He was responsible for creating a challenging and inspiring policy and program evaluation environment.

Evert was always accessible and caring with each one of the participants, listening carefully to their projects, sometimes in an English hard to understand.

His passion for public policy evaluation, and educating students provide the utmost quality in learning.

I hope our cooperation contributed to broaden knowledge and enabled more effective analyzes and evaluation activities throughout academic and professional sectors in Brazil.

Um forte abraço e muito obrigado, Evert!

Pedone & Vedung (Eds.) (2018). *Avaliação de Políticas Públicas: Programas Militares Complexos*. [Compilation of best papers presented by participants of the Evaluation Courses 2013-2017 at Inest-UFF]. Rio de Janeiro: Luzes.

1. Rosendo & Silveira Lima: Programa de Desenvolvimento de Submarinos (PROSUB) – Avaliação do processo de transferência de tecnologia e nacionalização da produção
2. Martins, Pinotti & Pinheiro da Silva: Evaluating the AMX and FX-2 Programs – Two Programs, Two Side-Effects: One Real, The Other Expected
3. Violante & Ottero: Avaliação e Implementação do Programa de Obtenção de Meios e Superfície (PROSUPER): obstáculos ao Projeto Estratégico de Defesa da Marinha do Brasil
4. Santos & Guimarães: Operações de Paz: Teoria da Intervenção e Resultados
5. Serrão & Pacheco: The Offset Policy evaluated through the HX-BR Compensatory Perspective
6. Pinotti: Avaliando a Implementação do Programa VBTP-MR Guarani
7. Ronconi: Considerações sobre uma avaliação do planejamento de defesa no Brasil no período 2003-2015

Figure 2. Book chapters as outcomes from the teaching.

References

- Berriel (2016). *Uma avaliação dos processos de transferência de tecnologia do PROSUB, Guarani e H-XBR*. Thesis (Master's). UFF. Advisor; Luiz Pedone. Published by Ed. Dialética, 2021. (Best Thesis Rene Dreifuss Prize, 2017. Third place in National Contest, Ministry of Defense, 2016.)
- Hernandez & Pedone (2020). A Reconstrução da Estação Antártica Comandante Ferraz: uma análise de políticas públicas de defesa. *Revista Brasileira de Estudos Estratégicos*, v. 12, p. 37-55.
- Guimarães (2017). *O Programa FX2: Uma avaliação dos processos de transferência de tecnologia*. Monograph (Undergraduate program in International Relations). UFF. Advisor: Luiz Pedone. (First prize Ministry Defense National Contest on best monograph, 2019.)
- Guimarães (2021). *O Programa do Submarino de Propulsão Nuclear Brasileiro e o Regime Internacional de Não Proliferação de Armas Nucleares: em busca de compatibilizar dois projetos divergentes*. Thesis (Master's). UFF. Advisor: Renato Petrocchi. UFF. (Second place on Ministry of Defense/Capes Contest Dissertations and Theses on National Defense (Tiradentes Prize, 2022).)
- Guimarães & Pinheiro da Silva (2022). Politicization, Foreign Policy and Nuclear Diplomacy: Brazil in the Global Nuclear Non-Proliferation Regime after the NPT. *Conjuntura Austral. Revista do Núcleo Brasileiro de Estratégia e Relações Internacionais da UFRGS*, v. 13, p. 48-63.
- Missaglia & Guimarães (2020). Recuo, avanços e continuidade do Programa CBERS (1999 - 2019). *Revista da UNIFA*, v. 33, p. 6-16, 2020.
- Nogueira (2016). *O Gripen Sul-Africano: um estudo dos resultados alcançados pelo processo de transferência de tecnologia na África do Sul*. Thesis (Master's). UFF. Advisor: Luiz Pedone.
- Ottero (2016). *Programa de Veículos Lançadores: uma avaliação da etapa de implementação*. Thesis (Master's). UFF. Advisor: Luiz Pedone.
- Pacheco & Pedone (2016). Incentivos governamentais e indústria de defesa. *Revista Brasileira de Estudos de Defesa. Porto Alegre*, v. 3, n° 2 (Jul/Dez), p 177-196
- Pedone (1994). Brazilian Political Development Confronting the Plebiscite. In: L. E. Aragón, W. Karlsson & Å. Magnusson, *Science, Development and Environment in Brazil*, p. 101-137. Stockholm: Institute of Latin American Studies.

- Pedone (2016). The Challenges for Sweden's Defence and Security Policies: old and new. *Revista de Estudios Europeos*, v. 68, p. 75-93.
- Pedone (2017). Science, Technology, and Innovation for Defense in Brazil. *Brasiliana: Journal for Brazilian Studies*, v. 5, p. 429-469, 2017.
- Pedone, Pinheiro da Silva & Guimarães (2018). Avaliação de Políticas Públicas para Defesa: uma análise dos principais programas governamentais para o setor aeroespacial brasileiro entre 2012-2018. *Revista Brasileira de Estudos Estratégicos*, v. 10, p. 13-40.
- Pedone & Vedung (Eds.) (2018). *Avaliação de Políticas Públicas: Programas Militares Complexos*. [Compilation of best papers presented by participants of the Evaluation Courses 2013-2017 at Inest-UFF]. Rio de Janeiro: Luzes.
- Pinheiro da Silva (2018). Política Externa como Política Pública: o processo de 'politização' da política externa brasileira na Nova República. *HOPLOS*, v. 1, p. 80-95, 2018.
- Pinheiro da Silva (2019). *Tomada de Decisão na Política Externa Brasileira: um estudo da adesão do Brasil ao Tratado de Não Proliferação de Armas Nucleares*. Thesis (Master's) - UFF. Advisor: Luiz Pedone. (Best Thesis Rene Dreifuss Prize, 2020.)
- Ronconi (2017). *O processo político-decisório do planejamento de defesa: impactos sobre a indústria de armamentos em dois momentos (1973-1985/2003-2015)*. Thesis (Master's) - UFF. Advisor: Luiz Pedone.
- Santos (2018). *Avaliação do Programa de Proteção de Civis da Missão das Nações Unidas para a Estabilização da República Democrática do Congo (MONUSCO): 2010-2015*. Thesis (Master's) - UFF. Advisor: Luiz Pedone.
- Serrão (2015). *Indústria de Defesa, Complexidade Econômica e Inovação: estudo do spin-off baseado em modelagem fuzzy*. Dissertation (Doctorate Political Science/Strategic Studies). UFF. Coadvisor: Luiz Pedone. (First prize Best Dissertation on VII Competition of Theses and Dissertations on National Defense/Ministry of Defense, 2016.)
- Vedung (1991). *Utvärdering i politik och förvaltning*. Studentlitteratur
- Vedung (1997). *Public Policy and Program Evaluation*. Transactions Publishers
- Vedung & Pedone (2021). *Avaliação de Políticas Públicas e Programas Governamentais: fundamentos e modelos*. Rio de Janeiro: Luzes.
- Violante (2017). *Política Externa, Política de Defesa e Cooperação Sul-Sul Como Grande Estratégia na África Ocidental: um estudo de caso em Cabo Verde e São Tomé e Príncipe*. Thesis (Master's). UFF. Advisor: Luiz Pedone. UFF.
- Violante (2023). *O Submarino Convencional com Propulsão Nuclear Brasileiro no Planejamento Espacial Marinho: Análise e avaliação da implementação (2012 a 2022)*. Dissertation (Doctorate Strategic Studies). UFF. Advisor: Eurico Figueiredo.
- Violante. & Pedone (2020). Avaliação da Implementação da Cooperação Sul-Sul do Brasil com São Tomé e Príncipe nos Governos Cardoso, Lula e Rousseff. *Brazilian Journal of International Relations*, v. 9, n. 3 (set-dez), p. 491-530.
- Violante. & Pedone (2023). Programas Estratégicos Complexos: Análise da Avaliação da Implementação do Submarino Convencional Riachuelo, 2008-2018. *Revista de Estudos Internacionais*, v. 14, n. 2, p 139-151.

Extension projects

Pinheiro da Silva & Pedone: On two occasions teaching of the mini-course Evaluation of Public Policies on Defense at the 2018 and 2019 Annual Brazilian Meeting on Strategic Studies and International Relations, Inest- UFF.

KAPITEL 24

Evert – et eksempel på den gode mentor i evaluering

Olaf Rieper

INDLEDNING

Dette kapitel har en personlig touch, idet jeg vil berette om, hvordan Evert blev min mentor i evaluering – og en ven. Det bygger på min egen skrøbelige hukommelse støttet af egne dagbogsnotater og andre skriftlige kilder. Det er således mit perspektiv, mine erfaringer og oplevelser, der gengives. Men jeg håber, at Evert (og måske andre), kan nikke genkendende til fortællingen. Afslutningsvis gør jeg mig nogle tanker om den gode mentor generelt set. Det er således mit håb, at andre i evalueringsfeltet vil blive bevidste om, hvor meget en ældre kollega kan betyde for de yngre, der engagerer sig i evalueringsteori og -praksis.

SVENSK OG DANSK EVALUERINGSFORSKNING

For at sætte en ramme for det følgende vil jeg kort omtale nogle aspekter af evaluering i de to lande i de sidste årtier af 1900-tallet med et udkik til USA.

I faglitteraturen er der enighed om, at Sverige var før Danmark med hensyn til at udvikle evaluering. Det gælder Bemelmans-Videc (1992), som viser, at Sverige kan kategoriseres som et "første bølge

land” med hensyn til evaluering, mens Danmark betegnes som et ”anden bølge land”. Det samme vises i International Atlas of Evaluation, hvor 21 lande rangordnes efter ni forskellige indikatorer for evalueringskultur: evaluering forekommer i flere politikområder, udbuddet af evaluatore i forskellige fagdiscipliner, national diskurs om evaluering, professionel evalueringsforening, grad af institutionalisering af evaluering i henholdsvis regeringen og parlamentet, mangfoldigheden af evaluatore inden for hvert politikområde, evaluering inden for rigsrevisionen, og endelig andel af resultat evalueringer i forhold til output og procesevalueringer. I forhold hertil rangordnes Sverige som nr. 4 og Danmark som nr. 8 (Furubo og Sandahl, 2002).

Evert var således en del af og i høj grad bærer af en evalueringskultur, som på mange områder var foran Danmark, hvis vi accepterer forestillingen om en fremadskridende evalueringskultur. Således havde Evert en viden, som jeg ikke umiddelbart kunne hente i det danske miljø.

EVERTS INTERNATIONALE UDSYN

Førertrøjen i evaluering på den internationale scene havde og har USA. Fra 1960-80 skete der en kraftig udvikling af evaluering i USA både på føderalt og statsligt niveau og i den anvendte samfundsvidenskab (Rist og Paliokas, 2002). Evert var en observatør og deltager i denne udvikling. Han var tilknyttet amerikanske universiteter i 1975-1976, blandt andet the Department of Political Science at the University of California, Berkeley og i 1982-1983 the Lyndon B. Johnson School of Public Affairs ved Texas Universitet, Austin. Evert har beskrevet (Vedung, 1983), hvordan den amerikanske statskundskab påvirkede Sveriges forskning, og han så klart de tendenser inden for policy analyse, herunder evaluering, der kom til at præge evalueringsfeltet i Sverige – og i Danmark. Han nævner meta-evaluering, problemer med anvendelse af evalueringsresultater og debatten om lodtrækningseksperimenter som det eneste saliggørende (jf. den senere evidensdebat). Han kommer med konkrete anbefalinger til, hvordan forskere kan komme direkte i kontakt med amerikanske forskere

og evaluatorer, blandt andet gennem specifikke konferencer og sommeruniversitetskurser.

HVORDAN DET HELE STARTEDE

Min første kontakt med Evert startede netop på et sådant kursus, Nordisk sommeruniversitets kursus i evalueringsforskning, som Evert var medarrangør af. For mig og mange af deltagerne blev det en legendarisk uge i sommeren 1985 i det charmerende Båstad i Sverige. Der var fremtrædende amerikanske foredragsholdere, bl.a. J.S. Wholey og M.Q. Patton. Og mange andre prægante oplægsholdere. Frem for alt blev kurset præget af en atmosfære af fællesskab og kollegialitet, når den er bedst, med berigende diskussioner. Og flere af os holdt kontakten i mange år efter. Jeg var samme år blevet ansat i Amternes og Kommunernes Forskningsinstitut (AKF), i dag VIVE, og kurset inspirerede mig til at få evalueringsforskning og -praksis styrket i AKF og mere bredt i amter (regioner) og kommuner.

EVERTS BETYDNING FOR MIG GENNEM ÅRENE

I maj 1989 mødtes vi til AEA-konferencen i San Fransisco, også kaldt jordskælvs-konferencen, fordi der var et voldsomt jordskælv lige før konferencen startede. Evert var kommet nogle dage før, og da vi spiste sammen under konferencen, fortalte han om den voldsomme oplevelse, det havde været for ham. Netop da det første skælv kom, befandt Evert sig i Universitetets bibliotek, hvor de tunge reoler kurede rundt.

Samme år opfordrede Evert mig til at deltage på en konference om policy and program evaluation, som fandt sted i Haag december 1990 (Mayne et al. (eds.), 1992). Konferencen fokuserede blandt andet på forskning i relation til ledelse og organisering af program evaluering på tværs af nationale, regionale og lokale myndigheder samt på anvendelse af evalueringsresultater. Jeg kom hjem med megen inspiration og mine kontakter på konferencen førte til, at jeg blev opfordret til at deltage i en international forskergruppe om evaluering ledet af Ray C. Rist. En gruppe, jeg stadig medvirker i 30 år efter. Igen Everts fortjeneste. Og også i gruppen (INTEVAL)

spillede Evert en vigtig rolle. Vi mødtes en gang årligt i gruppen, udvekslede ideer og erfaringer og skrev bøger om evalueringsteori og -metoder.

Everts bog, *Utvärdering i politik och förvaltning*, som blev udgivet første gang i 1991 og senere kom i adskillige oplag (og oversat til flere sprog), var til stor inspiration for mig. Her åbnede Evert med sit præcise sprog og begrebsanvendelse for mange af de emner, der senere skulle blive relevante inden for evalueringsfeltet, såsom evidensdebatten, anvendelse af evaluering, forholdet mellem politik og evaluering og (demokratiske) værdier.

Som Evert skriver (s. 17): Evaluering er ikke bare en del af den rationalistiske kuling, som blæste hen over politik og forvaltning i 1960'erne, men også en reaktion mod visse kraftige vindstød (den ekstreme rationalisme) i denne kuling. Med drøftelsen af værdier bliver evaluering løftet op over det tekniske niveau.

Everts inspiration kommer tydeligvis fra Nordamerika og Sverige. Få litteratur referencer er til danske publikationer, men her viser Evert, at han er velbekendt med danske bøger om evaluering.

Også på andre måder støttede og anerkendte Evert mig. Min afdøde kone, Laila, og jeg skrev en bog om metoder i samfundsforskningen, som er udkommet siden 1988. Evert og Laila kendte hinanden godt, vi sås på konferencer og spiste sammen. Jeg sendte ham bogen, og han skrev tilbage:

”Käre Laila og Olaf! Kärlek och samhällsvetenskaplig metod, hör de ihop? Uppenbarligen, att döma av er interessanta bok Forskning om og med mennesker. Detta är verkligen inget ungdomsarbete. Den bär erfarenhetens avklarende prägel.” Og han fortsætter med rosende ord.

Senere gav Evert også nyttige kommentarer til næste udgave af bogen. Sådan er Evert: anerkendende, følsom og faglig skarp.

EVERT I DET DANSKE EVALUERINGSMILJØ

I sagens natur havde og har Evert tætte forbindelser til kolleger i det danske evalueringsfelt både blandt både forskere og konsulenter. Det kan man se bl.a. af referencer til hans publikationer

i danske bøger om evaluering og af, at han var gæsteproffesor ved Københavns Universitet 2005-2010 og er æresmedlem i Dansk Evalueringsselskab (DES), hvor han har været foredragsholder flere gange. Han står også på programmet for DES' årskonference september 2023.

Evert har således givet inspiration ind til det danske forskningsmiljø, men også formidlet dansk evalueringsslitteratur ud til andre lande.

EN MENTORS BETYDNING FOR UNGE

I EVALUERINGSMILJØET

Med dette essay har jeg villet berette om den ganske store betydning, som Evert har haft for min faglige udvikling i evalueringsfeltet.

I det større perspektiv er det også en påmindelse om, hvor vigtige mentorer generelt er for unge – i og for sig også inden for andre fagområder. Vi bør som ældre evaluatore huske at varetage sådanne mentoropgaver med respekt for og indlevelse i de unge. Og give os tid til dem. En mentor har opgaver, der rækker ud over den faglige kritik og kræver forståelse for den unges særlige forudsætninger og drømme. Og på den anden side bør den unge være sig bevidst om, hvad en mentor kan bibringe. Mange gange er det tilfældigt, hvem den unge møder som mentor. Men det kan jo også gøres mere bevidst, selvom man næppe kan eller skal rationalisere alting. For mig var det et lykkeligt tilfælde!

Baseret på min erfaring med Evert vil jeg generelt opsummere nogle egenskaber ved den gode mentor. Mentoren er anerkendende over for den yngre kollega og lever sig ind i den yngres situation, både karrieremæssigt og øvrige livsomstændigheder. Mentoren inspirerer ved sit eget eksempel og ikke kun ved ord og hjælper den yngre med at danne netværk.

Bør mentoren være både kollega og ven? Det er vanskeligt at sige noget generelt om. Jeg mener, at et venskab kan berige virket som mentor, men det er ikke et must. For mig blev den venskabelige relation til Evert en berigelse.

Referencer

- Bemelmans-Videc, M.L. (1992). Institutionalizing evaluation: International perspectives. In Mayne, J., M.L. Bemelmans-Videc, J. Hudson and R. Conner (Eds.), *Advancing public policy evaluation. Learning from international experiences*. Elsevier Science Publishers: Amsterdam.
- Furubo, J-E. og Sandahl, R. (2002). Coordinated pluralisme – The Swedish case. In Furubo, J-E., Rist, R.C. and Sandahl, R. (Eds.), *International atlas of evaluation*. Transaction Publishers: New Brunswick and London.
- Vedung, E. (1983). American political science: A Swedish import. In: Lindmark, S. and Tallroth, T. (Eds.), *Swedes looking west: aspects on Swedish-American relations* (p. 122-143). Stockholm: Timbro.
- Vedung, E. 1991. *Utvärdering i politik och förvaltning*. Studentlitteratur: Lund.
- Rist, R.C. and Paliokas, K.L. (2002). The rise and fall (and rise again?) of the evaluation function in the U.S. government. In Furubo, J-E., Rist, R.C. and Sandahl, R. (Eds.), *International atlas of evaluation*. Transaction Publishers: New Brunswick and London.

KAPITEL 25

Serendipity in place*Terry Hartig*

Evert and I began working at the Institute for Housing and Urban Research about the same time, he as a new professor in 1995, me as a new lecturer in 1996. It had a simpler name then, the Institute for Housing Research, and we still refer to it as IBF, the acronym for that name in Swedish (Institutet för bostadsforskning). At the time we met, IBF had its offices in Gävle, in facilities built specifically for the former National Institute for Building Research (Statens institut för byggnadsforskning, SIB). That institute had been shut down a few years earlier by a conservative government, for reasons explained in a research report by our sociologist colleague Jim Kemeny (1996). Some engineers and others with a more technical background then got organized as a department within the Royal Institute of Technology, and Uppsala University stepped in to create IBF with some of the social scientists. That both of these new institutions remained in Gävle rather than moving to Stockholm or Uppsala owed in part to the fact that so many of the surviving employees lived in Gävle, and in part because of the continuing significance of regional labor market concerns. Those concerns and related political pressures had led to the de-centralisation of SIB from Stockholm to Gävle two

decades earlier; creation of the new institutions was not a reason for re-centralisation. The new institute was however not only staffed by former members of SIB; Evert and I were two new recruits.

So, there we were, working in Gävle and indeed with living quarters in the same building in Gävle, an apartment house built as part of the Million Program, on Hospitalsgatan, not far from where we had our offices on Brynäs. Evert rented a 2-room apartment, and he stayed there only part time, as he had long since established a home in Uppsala. I had a 1-room apartment, 37 square meters of efficient organization, and I would have spent more time there if I wasn't newly arrived from a postdoc at Berkeley and so spending a lot of time in the office working on papers and otherwise getting established as a junior researcher.

I might have tried to find a place to live in Uppsala or even Stockholm and then commute to Gävle. Some friends and colleagues had advised me to do, suggesting I would find Gävle rather boring in comparison to the possibilities in the larger urban alternatives, especially after my years in the San Francisco Bay area. But I disliked the idea of having to make a long commute every day, knowing I would want to spend much of my time in the office. That aside, I was ready to give Gävle a chance – I knew that many of my new colleagues lived there, and some of them had spoken warmly of the city and what it had to offer. From what I could gather from different sources, I understood it had a rich history, and with its industrial heritage it would offer interesting contrasts to the temples and sensibilities of state, religion and academia that I would encounter when in Uppsala for my teaching there.

Evert shared my sense of the importance of being present in the workplace. That and a shared work ethic meant that we saw a lot of each other at IBF's offices. After IBF's move to the old town hall in the center of Gävle in 1998, we also frequently walked to work together from our apartment house, usually following the river that flows through the city just before spilling into the inner harbor where sailing ships once came in with coffee and went out with wood and iron. That and other features of Gävle that we

noticed in passing, like the sculpture *Sjöfartsmonumentet*, three anchors arranged on the pedestrian bridge over the river, provided stuff for our recurrent conversations about *genius loci*. Yes, we were not only there to work; we also shared an interest in the spirit of this place where we had come to spend so much time, and curiosity about what one could discover there.

Our joint explorations were perhaps not so extensive, but they were memorable. Not least, I have clear memories of sitting with Evert in one and another of Gävle's pubs, talking not only about work and the city, but also about circumstances in our personal lives. In these conversations, Evert more often took on the role of patient, selfless, and sympathetic elder, listening carefully and offering words of kind encouragement and solace to a junior colleague far from his homeland and the important relationships there. He had after all had much foreign experience by that time, and he knew what it was like to live in places that sometimes challenged one's understanding.

Reflecting on those conversations now, I am reminded that *genius loci* has two quite different meanings, and that I can associate both of them with Evert. The one meaning I have already indicated: put simply, the essential character of the place, following modern Western usage. The other, older meaning, from the classical Roman, refers to a protective spirit that watches over a place. Now Evert himself may well protest – in my mind's eye I can see him laughing as he does so – but thinking of those years in Gävle, he was in my experience of that place something of the protective spirit.

Genius loci is not the only concept to which our conversations have returned time and again. "Spirit" is not far removed from "sprite," etymologically as well as phonetically, and "sprite" in turn is not far removed from "sprightly," which denotes liveliness, alertness, qualities of being animated, keenly alive, indeed spirited – to borrow just a few meanings from the Merriam-Webster dictionary. These words seem fitting for the kind of person who would, with a smile and a twinkle in his eyes, speak warmly of serendipity: "the faculty or phenomenon of finding valuable or agreeable things not

sought for,” again according to Merriam-Webster. I like this definition because it refers not only to seemingly chance phenomenon, but also to a faculty one might acquire, one that Evert possesses. It’s similar to what Ingmar Stenmark spoke of many years ago, along the lines of “I don’t know anything about luck, but that the harder I train the luckier I get.” Similar, but not the same; a faculty for finding serendipity is not the same as preparing for luck through training; I can’t see it so much as a matter of training as one of remaining open and sensitive to the possibilities that inhere to the set of circumstances with which one is presented. Training along a particular line may indeed reduce openness, narrowing one’s vision and so limiting possibilities.

That Evert has recognized serendipity in the past is well attested. One only need consult the oral history of his professional development that Joakim Tranquist (2015) wrote based on an interview (see also his contribution to this volume). In that exchange, Evert stated, “Looking back, I realize that there have been several serendipities in my career. Unexpected things that happen, that you need some sort of personal preparedness for in order for something to occur” (p. 581). Also in that interview, he noted the importance of a receptiveness for the novel as a condition for serendipity. This openness to and curiosity about possibilities is something we have enjoyed discussing, in Gävle and since, after IBF moved to Uppsala. We might have compared notes on our respective experiences in Odense and Berkeley, for example, and perhaps we still will someday, but so far we have taken these other paths. Serendipity seems to follow Evert around in some way, and through his example he happily shares with others the key to a faculty for it. That I could make his acquaintance those many years ago in Gävle was itself serendipitous.

References

Kemeny, J. (1996). *Abolition of the Swedish Institute for Building Research: Social change and the politics of compromise* (Research Report 1996:2). Gävle. Institute for Housing Research, Uppsala University.

Tranquist, J. (2015). The oral history of evaluation: The professional development of Evert Vedung. *American Journal of Evaluation*, 36, 570-583.

OM FÖRFATTARNA

MORTEN BALLE HANSEN är professor i offentlig förvaltning vid Syddanmarks universitet, Odense.

PER ØYVIND BASTØE är utvärderingsrådgivare och tidigare utvärderingsdirektör vid Norges biståndsmyndighet (Norad).

BO BENGTSSON är seniorprofessor i statskunskap vid Institutet för urban- och bostadsforskning, Uppsala universitet.

LI BENNICH-BJÖRKMAN är professor i vältalighet och statskunskap vid Uppsala universitet.

PETER DAHLER LARSEN är professor i statsvetenskap vid Köpenhamns universitet.

MARIANNE DANIELSSON är fil. dr i statskunskap och utredare på Brottsförebyggande rådet.

VERNER DENVALL är professor i socialt arbete vid Lunds universitet.

KIM FORSS är ekon. dr, utvärderingskonsult och författare.

HANNE FOSS HANSEN är professor emerita i statskunskap vid Köpenhamns universitet.

JAN-ERIC FURUBO är författare och tidigare chef inom Riksrevisionsverket och revisionsråd vid Riksrevisionen.

SEBASTIAN GODENHJELM är universitetslektor i statskunskap med förvaltning vid Svenska social- och kommunalhögskolan, Helsingfors universitet.

SVERKER GUSTAVSSON är professor emeritus vid statsvetenskapliga institutionen, Uppsala universitet.

TERRY HARTIG är professor i miljöpsykologi vid Institutet för bostads- och urbanforskning, Uppsala universitet.

NILS HERTTING är docent i statskunskap och prefekt för Institutet för bostads- och urbanforskning, Uppsala universitet.

FRANS L. LEEUW är professor i juridik, offentlig politik och samhällsvetenskap vid Maastricht Universitet.

LEIF LEWIN är professor emeritus i vältalighet och statskunskap vid Uppsala universitet.

MATS LINDBERG är professor emeritus i statsvetenskap vid Örebro universitet.

SANJA MAGDALENIĆ är fil. dr i sociologi och utredare på Socialstyrelsen.

WOLFGANG MEYER är professor i sociologi vid Saarland Universitetet, Saarbruecken och vice-direktor för Center for Evaluation (CEval).

PER MICKWIZ är professor i miljöpolicy och vicerektor vid Lunds Universitet.

YLVA NORÉN BRETSER är docent i offentlig förvaltning och lektor vid Förvaltningshögskolan, Göteborgs universitet, samt ordförande i Svenska utvärderingsföreningen.

LUIZ PEDONE är professor i offentlig politik och internationella relationer vid Institutet för strategiska studier, Fluminense Universitetet i Rio.

OLAF RIEPER är frilansande konsult, tidigare forskningsledare vid Danska institutet för regional- och lokalforskning.

ROLF SANDAHL är fil. dr i statskunskap och författare med mångårig erfarenhet av utvärderingsfrågor inom statsförvaltningen.

STEFAN SJÖBLOM är professor i kommunal förvaltning vid Helsingfors Universitet.

NICOLETTA STAMNE är professor emerita i sociologi vid Sapienza Universitetet i Rom

REINHARD STOCKMANN är professor i sociologi vid Saarland Universitetet, Saarbruecken, samt grundare och direktor för Center for Evaluation (CEval).

JOAKIM TRANQUIST är utvärderingskonsult och författare.

Skrifter utgivna i serien Uppsala Housing and Urban Studies,
ACTA UNIVERSITATIS UPSALIENSIS

1. Erola, J. m.fl., Norms, Social Structures and Policy Changes.
Essays in honor of Hannu Ruonavaara.
201 p., 2021.

2. Danielsson, M. & Hertting, N., Samtal om idéer, utvärdering
och samhällsförändring. En vänbok till Evert Vedung.
312 s., 2024.

