

självbevarelse-drift vakar och styr över den politiska jämlikheten.

Avsnittet om demokratipolitiken, som påkallas av den fjärde frågan, bjuder på en intressant analys som kastar ett delvis nytt ljus över den allra senaste demokratihistorien. Den diskursiva kraften ska inte underskattas. Här fattas väl bara den lindbeckska demokratikommissionen och den birgerssonska demokratiutvecklingskommittén. Att inkludera dem skulle inte ha ruckat Nilssons teser, bara nyanserat dem och stagat upp dem. Det intressanta med den så kallade lindbeckskommissionen är ju att den relativiserar demokratin i förhållande till ekonomiska tillväxtnål, medan det intressanta med den birgerssonska expertutredningen är att den inte bara blir en förelöpare till de följande årens demokratidebatt utan att den också blir en ideologisk djärvare tolkning av samhällsutvecklingen än den kommunala förnyelsekommittén som lämnade sitt betänkande samma år, 1996. Redan där kan man förmodligen se fröet till den demokratiteoretiska kontrovers som blommar ut i början av 2000-talet.

Till sist bör mina kritiska invändningar vägas mot avhandlingens förtjänster. För det första drivs den av en ambition att försöka lägga ett pussel av bitarna i det lokala demokratiarbetet. Det är lovvärt och delvis framgångsrikt. För det andra ger sig författaren i kast med en central dimension i kontroversen kring demokratidealen som alla åberopar men få djuplodar. För det tredje är han övertygande i sin argumentation för att man måste relatera deltagandeformerna tydligare till den offentliga maktsfären. Hans egen tredelning är också både klargörande och pedagogisk. Slutligen gör Tom Nilsson en förtjänstfull jämförelse mellan demokratiretorik och demokratipolitik i den svenska demokratidoktrinen samtida utvecklingshistoria.

Anette Gröjer: Den utvärderade(n)de staten: Utvärderingens institutionalisering på den högre utbildningens område. Stockholm Studies in Politics 102, 2004.

EVERT VEDUNG¹

A process is said to be **path-dependent** if accidental events might have persistent effect on its course. ...

Consider for example the issue of *driving on the left or on the right*. Technically the two choices are equivalent. The initial choice might therefore have been due to accidental reasons. Once the choice was made, though, it became permanent because of the huge cost involved in modifying it. A similar issue is that of the *Rail gauge*.

http://www.fact-index.com/p/pa/path_dependency.html (22 okt 2004)

Utvärdering är för närvarande en otroligt populär styrmetod. Praktiskt taget allt skall utvärderas. Undersök noggrant resultatet av det du gjort och vägarna dit, så lär du dig att bättre orientera framåt, det är det enkla budskapet. Det är inte bara goda avsikter, ökade penninganslag och eggande visioner som räknas. Offentliga sektorn måste också leverera, uppvisa verkliga resultat. Den måste ge värde för pengarna.²

I sin rättframma form bygger utvärdering på en så generande enkel idé att man frestas tro att verksamheten alltid har funnits. Men faktum är att utvärdering kan-

1 Evert Vedung, professor emeritus i statsvetenskap särskilt bostadspolitik vid Uppsala universitet, är gästprofessor vid Mälardalens högskola i Eskilstuna.
E-post: Evert.Vedung@ibf.uu.se

ske bara har existerat sen 1950-talet i offentliga sektorn. Längre var också systematiska studier av utvärderingsfenomenet fåtaliga. Men nu är den tiden för längesen förbi.

Utvärderingens stadiga uppstigande på den internationella himlapallen speglas i den nyligen utkomna *International Atlas of Evaluation*. Ett diffusionsperspektiv anläggs där av Furubo och Sandahl (2002a). Utvärderingstänkande och utvärderingspraktik har spritt sig språngvis från en handfull pionjärländer 1960-70 till ett dusintal andra omkring 1990.

Ett tidigare verk med viss länderkomparativ inriktning var *Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations* som kom för femton år sen (Rist 1990). Även här fanns en essä med ett diffusionsperspektiv (Derlien).

Det främsta arbetet om utvärderingsfenomenet i ett nordiskt land är Hanne Foss Hansens stort upplagda *Evaluering i staten* (2003), som belyser utvecklingen i Danmark ur olika aspekter. I ett specialnummer av Nordisk Administrativ Tidsskrift (nr 2 2000) finns uppsatser om evaluering i fyra nordiska länder. Undertecknad har också själv gjort några tappra försök att skildra utvärderingens intåg och utveckling (Vedung 1992, 2004).

2 Artikeln återger i sina huvuddrag min bedömning efter min fakultetsopposition och Anette Gröjers försvar vid disputationen den 15 oktober 2004 i Frescati i Stockholm. Anette Gröjer har senare läst manuskriptet och lämnat omdömesgilla förklaringar. Under ett seminarium vid Mälardalens utvärderingsakademi, MdH, Eskilstuna, den 7 november 2005 lämnade Julia Jäder och Veikko Pelto-Piri uppslag som jag vidarefört. Ett hjärtligt tack till alla tre men främst, naturligtvis, till Anette Gröjer.

Välkommet bidrag

Anette Gröjers doktorsavhandling är ett välkommet bidrag till denna empiriska uppmätning och kartläggning av utvärderingsbøljan. Författarinnans ambition (s. 18) med *Den utvärdera(n)de staten: Utvärderingens institutionalisering på den högre utbildningens område* är att beskriva, analysera och förklara inte bara utvärderingstänkandet utan även utvärderingsaktiviteter och utvärderingsfunktionens organisering. Det har hon också lyckats med. Hon har åskådliggjort hur utvärdering, på diskussionsnivån, genom praktiskt utförda utvärderingar och genom organisatorisk inordning fått fäste i och utvecklats inom den nationella statliga myndigheten på den högre utbildningens område i Sverige under perioden 1964-1995.

Mitt syfte

Min recension är en kritisk granskning inom ramen för vad det är rimligt att begära av en nutida svensk doktorsavhandling. Jag ämnar därvid påtala skriftens styrkor men också dess otydligheter och svagheter. På vissa punkter tänker jag emellertid släppa normen "rimliga krav på en doktorsavhandling" och i stället utgå ifrån hur en författare rent allmänt skulle kunna nå längre i en analys av utvärderingsfenomenet. Därvid kommer jag att framlägga förslag till andra avgränsningar, breddningar och fördjupningar. Detta skall inte uppfattas som en negativ kritik av Anette Gröjers avhandlingsförfattari utan som propåer om hur exempelvis nästeför författare skulle kunna gå tillväga.

Gröjers huvudsyfte

Huvudsyftet med Anette Gröjers studie är att på den högre utbildningens område i Sverige skärskåda utvärderingens framväxt och utveckling med fokus på sysselsättningen vid de centrala myndigheterna

på den högre utbildningens område. Därvid analyseras utvärderingsverksamheten som en institution. Utvärderingens formativa moment på den högre utbildningens fält skall fastställas och utvärderandets fortsatta stigberoende fixeras. Fem nyckeltermer i Anette Gröjers traktat är med andra ord utvärdering, högre utbildning, institutionalisering, formativt moment och stigberoende.

Innebörden av "utvärdering"

Vad ord som "utvärdering" och "utvärderingsverksamhet" kan betyda dryftar Gröjer utförligt. Hon stannar för följande bestämning (s. 9): "Jag har valt att betrakta utvärderingsverksamhet som ett instrument för att noggrant bedöma genomförandet och resultatet av offentlig politik. Utvärderingens uppgift är att ge underlag för nya rationella beslut." Omedelbart därefter tillägger hon: "Denna både idealiserade och teoretiska utgångspunkt kan för vissa te sig verklighetsfrämmande men den passar väl in i den rationalistiska styrningsprocess som länge varit en del av förvaltningspolitiken."

För en generös bedömare är hennes definition någorlunda vettig. Precisionsgraden räcker för hennes syfte. Förvisso är utvärdering ett instrument för att noggrant och därmed systematiskt bedöma resultatet av vad den offentliga sektorn gör inklusive vägarna dit via genomförandet. Idealtypiskt kan utvärdering också sägas ha till uppgift att ge underlag för nya beslut.

Centrala myndigheterna inom den högre utbildningens område

Avhandlingens sökarljus riktas mot det utvärderingsarbete som bedrivits vid de centrala myndigheterna på den högre ut-

bildningens fält. Författarinnan har begränsat sig till tiden 1964–1995. Utvärderingsaktiviteterna vid följande ämbetsverk ingår därmed i studien:

1964–1976 Universitetskanslersämbetet (UKÄ)

1976–1992 Universitets- och högskoleämbetet (UHÄ) som bildas när UKÄ försvinner

1992 Sekretariatet för utvärdering av universitet som skapas då UHÄ läggs ner

1993–1995 Kanslersämbetet (KÄ) som uppstår genom att Sekretariatet ges detta namn fram till dess KÄ avskaffas och Högskoleverket uppstår 1995.

Avgränsningarna är hårdhänta. Så bör de också vara i en nutida doktorsavhandling i statskunskap. För att passera mållinjen inom den utmäta tiden krävs skyggglappar. I begränsningen visar sig mästaren. När jag ändå tar upp avgränsningsproblematiken, så är det mer för att föra ett principiellt resonemang än för att klanka på avhandlingen.

Först tidsspannet. Genom att slutåret sätts till 1995 lämnas nuvarande Högskoleverkets vittfamnande utvärderingsgärning utanför, trots att den är speciell, starkt systematisk, institutionaliserad och kritiserad.

Sen den rumsliga insnävningen. Varför definiera den "högre utbildningens område" som enbart det som vid varje given tidpunkt sysselsätter just den centrala nationella myndigheten? Därmed utesluter hon all utvärdering som de svenska universiteten och högskolorna själva genomfört. Om personalen vid Förvaltningshögskolan i Göteborg själv tar initiativ till att utvärdera sitt samlade utbud av kurser så kommer en dylik granskning inte med i studien. Varje lärarlags utvärdering av sin kursbukett beaktas inte, ej heller studen-

ternas kursvärderingar. Det hade varit spännande att få veta när och hur systemet med kursvärderingar kom till, slog igenom och motiverades. Men denna brukarorientering inom ämnen och inom högskolor faller utanför forskningsfältet. Uppåt utesluter hon sig också från statens utvärderingar av högskolesektorn genom SOU-systemet i regeringskansliet.

Genom den valda begränsningen till centrala myndigheter på högskoleområdet utesluts också all den utvärdering av forskning som utförts av Vetenskapsrådet och dess föregångare, av Formas och dess förelöpare Byggeforskningsrådet eller t.ex. av Naturvetenskapliga forskningsrådet.

En alternativ avgränsning skulle ha varit att täcka allt i sektorn från toppen ända ner till botten och strunta i begränsningen till en enda hierarkisk nivå, central myndighet. Detta skulle ha inneburit en principiell inriktning på alla offentliga organisationer i högskolesektorn som helhet från riksdag och regering över myndigheter ner till universitet och högskolor. Koncentration på några högskolor skulle då kanske ha blivit nödvändig. Kanske hade denna sektorssynpunkt varit något att tänka på även för Gröjer eftersom avhandlingens teoretiska utgångspunkt är institutioner och institutionalisering. I den lilla volymen *Institutions and Organizations* påpekar nämligen W Richard Scott att "the organizational field or sector [has been identified] as a new level of analysis particularly suited to the study of institutional processes" (s. 31).

Institutionalisering och historisk institutionalism

I sin teoretiska begreppsapparat knyter alltså Gröjer an till nyinstitutionalismen i statsvetenskapen. Hon skiljer mellan tre institutionalismer: den historiska, den so-

ciologiska och rational choice-institutionalismen. Vi får aldrig veta något om skillnaderna mellan dessa. Men klart är att Gröjer förankrat sin dissertation i den historiska institutionalismen.

Institution, vad är det?

Institutionsansatsen förutsätter att institutionsbegreppet måste preciseras, åtminstone nödortfögt. För Gröjer får en positionsbestämning av Hall (1986) bilda utgångspunkt. Hall menar att en institution består av: "the formal rules, compliance procedures, and standard operating practices that structure the relationship between individuals in various units of the polity and economy" (s. 19).

En institution är en mängd regler och typer av handlande som sätter sin prägel på det offentliga handlandet i statsförvaltningen (s. 17). Hall menar vidare att institutioner har en mer formell status än enbart kulturellt betingade normer (s. 16). Med andra ord: oskrivna kulturella betingade normer räknas inte som institutioner.

Utifrån ovanstående definition kan också utvärderingsverksamheten betraktas som en institution, hävdar Gröjer. Hon säger sig vilja analysera de formella regler, tillvägagångssätt och tillämpad praxis som utgjorde institutionens kännetecken (s. 17).

Som eftertänksam läsare blir man inte riktigt klok på hennes ordbruk i detta avseende. Uppenbart är att hon utesluter organisationer. Detta är ett väsentligt klarläggande ty uttrycket "institution" kan på svenska också betyda organisation (jfr Scott 1995). Rikets fängvårdsanstalter är organisationer liksom Statsvetenskapliga institutionerna vid Stockholms och Uppsala universitet. Men samtidigt kan termen "institution" på svenska också ha betydelsen av regel eller regelverk, och det är till denna mening som Gröjer knyter an.

I själva verket medger Gröjer att hon haft väldiga bekymmer med institutionsansatsen. Sina mödor härvidlag summerar hon så här (s. 189):

”Själva begreppet institution har varit besvärligt att få grepp om. De definitioner som jag valde att använda såg vid det första påseendet konkreta ut men tämligen omgående kom jag att uppfatta dem som abstrakta.

Allteftersom analysen av institutionen [utvärdering] fortskred klarnade begreppet något. Jag kom att studera institutionen [utvärdering] utifrån följande:

1. de idéer och principer som utvärdering teoretiskt sett bygger på
2. det praktiska innehållet i de utvärderingar som myndigheterna genomförde
3. de normer och regler som gäller för myndigheternas och tjänstemännens agerande
4. (4) de idéer och förväntningar som den politiska nivån haft på verksamheten.”

Tydligt är inte institutionen (1), (2), (3) och (4) utan den studeras *utifrån* detta. Men vad består då institutionen av? Det handlar inte om organisering och inte om kulturellt betingade normer. Men vad handlar det då om?

Det hade varit klagörande om Gröjer stringent samlat ihop sina insikter i begreppet och företeelsen institution i några väl genomtänkta, starkt komprimerade slutsatser.

Formativt moment (critical juncture, kritiskt vägval)

Ett fjärde grundbegrepp är formativt moment och dess synonymer (23 ff., själv föredrar jag uttryck som ”formativt ögonblick” eller ”politikfönster”). Ett *formativt moment* är beteckningen på den tids-

period då institutionaliseringsprocessen sätts igång och en institution börjar etableras. Momentet kan vara utdraget i tiden, menar författarinnan. I ett formativt moment fattas beslut som leder till stigberoende.

Stigberoende (path-dependence, spårbundenhet)

Stiganalys är att söka efter och förstå de kedjor av händelser, som lett fram till ett resultat. För denna händelsekedja använder Gröjer sitt femte nyckelbegrepp, stigberoende (path-dependence, 25 ff). Det historiska institutionella perspektivet fokuserar därmed två processer, säger hon. Den första är konstituerande och består av det formativa momentet, den andra är reproducerande och avser självförstärkningen. När institutionaliseringsprocessen kommit igång och ett handlingsalternativ är utvalt kommer de självförstärkande krafterna att befästa den fortsatta utvecklingen. Författarinnan hänvisar till Paul Pierson som påstår att varje ytterligare steg på den inslagna vägen ökar sannolikheten för ännu fler steg i samma riktning eftersom det i stigande grad blir svårare och svårare att återgå till utgångspunkten.

Utvärdering, institution, formativt moment, stigberoende, se där grundelementen i Gröjers teori bygge. Att det är en begreppsapparat, ett sökschema, är helt klart. Gröjer deklarerar också uttryckligen att hon skall bruka den historiska institutionalismen som analysverktyg för att studera verkliga processer och händelser (s. 16, 189).

En institutions början

För den empiriska undersökningen (46ff) är två frågor vägledande: 1) När etablerades utvärdering som institution på den

högre utbildningens område i den nationella myndigheten för högre utbildning i Sverige? 2) Varför uppstod denna institution?

Svaret på det första spørsmålet är att utvärderingsinstitutionen föddes i slutet av 1960-talet och början av 1970-talet. Viktiga händelser inträffade 1965 och 1972. 1965 (57ff) tillsattes universitetspedagogiska utredningen (UPU). Den initierade en mängd försöksverksamheter (s. 58). Den vurmade oerhört för metoder för utvärdering av detta experimenterande. År 1972 (s. 60) inrättades PU-enheten vid UKÄ. En grupp för utvärdering av forskningsprojekt initierades också av myndigheten.

Här vill jag själv lägga till att detta är tidigt i ett internationellt perspektiv. Sverige måste räknas till utvärderingens pionjäre-länder vad gäller området högre utbildning.

Drivkraft nr 1: en ny rationalism i förvaltningspolitiken

Att utvärdering uppstod berodde på händelseutvecklingen inom dels förvaltningspolitiken, dels den högre utbildningen (63ff).

Förändringar på det förvaltningspolitiska fältet innebar att en ny optik introducerades i den offentliga verksamheten. Särskilt starkt slog denna optik igenom i den ekonomiska styrningen genom programbudgeteringen. Rationalitet, produktivitet, effektivitet och helhetssyn på verksamheter stod i fokus. Utvärdering och utvärderingsliknande metoder passade väl in i denna styrfilosofi och infördes brett i den offentliga sektorn. Speciellt stora krav ställdes på att UKÄ och andra centrala förvaltningsmyndigheter skulle anamma och tillämpa det nya synsättet.

Här kunde hon ha lagt till att Riksräkenskapsverkets ombildning 1967 till Riksrevisionsverket var ett betydelsefullt led i denna introduktion av ett nytt synsätt. Den nya myndigheten fick bli en enhet för förvaltningsrevision, vilken i väsentliga drag sysslade med utvärdering.

Det tidigare arbetssättet i svensk offentlig politik i stort präglades av inkrementalism, skriver Gröjer. Förändringar i den förda politiken gjordes i små steg. Det senast fattade beslutet byggdes på lite grand men skilde sig inte så mycket från det tidigare. "Det handlade om direkt problemlösning utan samordning eller mål." Beslut fattades på ett ostrukturerat sätt, ett slags "muddling through" (eller trasslasig-igenom som Björn Wittrock kallat det).

Här tycker jag författarinnan har bara till hälften rätt. Må vara att det var små steg, en sorts tuffjätspolitik. Men inte var det, problemlösning i avsaknad av samordning eller mål. Det fanns en kompassriktning. Det hela gick ut på att öka det offentliga inflytande över samhällsutvecklingen på bekostnad av den privata egendomsrätten och det enskilda ägandet. Den offentliga sektorn skulle ta över mer och mer av det som tidigare sköttes av hushåll, näringsliv och civila samhället. Samhällsplanering skulle ersätta en tidigare mer spontan utveckling. Välfärdsstaten växte fram på alla områden.

I slutet av 1960-talet ersattes inkrementalismen av något som Gröjer kallar rätionalismen. Enligt den skulle politiken baseras på ett systematiskt tänkande. Programbudgetering, flerårsplanering, framtidsstudier, kostnadsintäktanalys, utvärdering och annat som brukar hänföras till policyanalys borde komma till användning.

Genom att tillämpa dylika vetenskapsliknande metoder skulle den offentliga

politiken bli mer välgrundad och effektiv. Kännetecknen var

- ekonomisk effektivitet
- programindelning, dvs. all verksamhet skulle delas in i program.
- mål-medeltänkande; politikerna skulle sätta målen för offentlig verksamhet medan utföra avgjorde vilka medel som skulle användas för att målen bäst skulle uppnås.
- uppmärksamhet mot mål och resultat, mindre mot processer och procedurer.

Rationalismen blev ett paradigm som var helt dominerande och det är ingen överdrift att påstå att rationalismen blev en institution, säger Gröjer. Även detta påstående reser frågor. Institutionaliserad, var då? Som tankesystem, tankefigur? I handling? I utfall?

Drivkraft nr 2: utvecklingen på den högre utbildningens område

Den andra drivkraften var reformer på den högre utbildningens område. Gröjer hänvisar till UKAS, PUKAS och studentrevolten 1968 som startade på Statsvetenskapliga institutionen vid Stockholms universitet och sen flyttade till Kårhuset vid Holländargatan.

Fanns det verkligen inte fler faktorer? Och varifrån emanerade egentligen grundidén?

Utvecklingen i grundskolan

Man kan undra över om den tidigare utvecklingen inom enhetsskolepolitiken hade någon inverkan. I början av 1950-talet utvärderades försöksverksamheter med enhetsskoleformen i betydande utsträckning, även med experimentell metod. Denna utveckling av utvärdering i lägre utbildning inverkar sannolikt på utvärdering i högre utbildning, menade

nestorn i svensk utbildningsutvärdering, Urban Dahllöf, vid disputationen:

Varför just utbildning?

Varför institutionaliseras utvärdering så tidigt just i utbildningssektorn? Frågan är spännande men faller utanför Gröjers avgränsningar. Överhuvudtaget skulle det vara intressant att få läsa jämförelser mellan utbildningssektorn och andra sektorer som t.ex. ulandsbistånd och energi. Här finns en uppgift för vidare forskning.

Stabiliseringsprocessen på UKÄ och stigberoendet

Efter det formativa momentet tar Gröjer itu med det följande stigberoendet som hon kallar det (s. 66ff). Hon nämner Mahoney och dennes typologi över självförstärkande krafter. Dennes grundläggande resonemang om stigberoende berör hon emellertid inte. Det skall jag däremot göra, utifrån en analys av Mahoney som genomförs av min kollega Bo Bengtsson (2004).

Enligt Bo Bengtsson definierar James Mahoney path dependence på följande sätt

”historical sequences in which *contingent* events set into motion institutional patterns or event chains that have *deterministic* properties” (BB:s kursiveringar).

”This implies,” förtydligar Bengtsson, ”that the identification of path dependence involves both tracing a given outcome back to a particular set of historical events, and showing how these events are themselves contingent and cannot be explained entirely on the basis of prior historical conditions.”

Att undersöka stigberoende innebär enligt Bengtsson att skriva historien bak-

länges. Så gör emellertid inte Gröjer. Hon skriver den framlänges. Hon börjar med de händelser som lägger fast stigen och undersöker sedan hur den stabiliseras och utvecklas. Fast egentligen demonstrerar hon inte hur valet av ursprunglig stig får konsekvenser för fortsättningen. Den granskning av effekt, som bör ingå i en prövning av om stigberoende föreligger, saknas.

Bo Bengtsson föreslår en svagare men i mitt tycke rimligare och mer fruktbar definition av stigberoende än Mahoney: [stigberoende föreligger när] "one event that is more or less contingent considerably changes the probability of other subsequent events or institutional outcomes." Han förklarar: "This weak concept of path dependence would transform Mahoney's strong distinction between contingency and determinacy into a matter of degrees on a scale going from contingency to determinacy, where prior events are considerably more contingent than subsequent. This weak definition would, of course, make path dependence more difficult to distinguish from other types of correlations, but certainly also more common in social life."

"Actors more or less deliberately design institutions at point A, which [for the actors] at a later point B define the rules of the political game." Detta innebär, skriver Bengtsson, att:

"Events at point A may at point B have an effect on

(I) *decision-making* (where actors choose other alternatives due to what happened at point A)

(II) *agenda-setting* (where other alternatives come up on the political agenda due to what happened at point A)

(III) *perceptions* (where other alternatives are conceivable to actors due to what happened at point A)."

Att Gröjer gjort en stiganalys förstår nog envar. Men är det en analys av stigberoende? Hon har inte visat, i vart fall på något tydligt sätt, att den ursprungliga stigen blev bestämmande för eller åtminstone ökade sannolikheten för den kommande utvecklingen. Hon har heller inte visat att det fanns alternativ till utvärdering omkring 1970, dvs. att det ursprungliga vägvalet då var "contingent".

Vilka mekanismer är verksamma vid stigberoende? Gröjer nämner följande fyra:

Utilitarian explanation: cost-benefit assessment of the institution by actors

Functional explanation: the function of the institution for an overall system

Legitimation explanation: the actors believe in the morality or appropriateness of the institution

Power explanation: the mechanism is supported by an elite group of actors

Hon tar fasta på maktförklaringen och pekar på att utvärdering stöddes av en maktbaserad elit. Till denna elit hörde Hans Löwbeer, Nils Eric Svensson, Lennart Lewin, Urban Dahllöf med ramfaktorteorin, Sven Moberg. Hon nämner också Ingvar Carlsson, Håkan Berg, Lennart Sandgren och Bertil Zakrisson.

Analysen av utförda utvärderingar

Jag skall nu gå över till att diskutera hur man skulle kunna karakterisera institutionen utvärdering genom att ange typiska, speciella eller viktiga drag i de utvärderingar som den centrala myndigheten har utfört över tid. Gröjer har gjort detta men jag skall föreslå förändringar. Oavsett om Gröjers analyser kan pressas in i den historiska institutionalismens begreppsschema eller ej så har författarinnan här gett ett beaktansvärt bidrag till forskningen om

utvärderingens former och innehåll på den svenska utbildningspolitikens område och i svensk förvaltning generellt.

Urval av myndigheternas skriftligt redovisade utvärderingar

Under perioden 1964-1995 har myndigheterna på fältet högre utbildning gett ut sammanlagt ca 650 stycken relevanta publikationer. Gröjer har satt sig i sinnet att ur denna mängd skriftligt redovisade studier vaska fram den delmängd som utgörs av utvärderingar. Hon begagnar sig därvid av tvenne selektionsprinciper.

Princip 1. För att vi skall kunna tala om en utvärdering så måste en studie uppfylla vissa kravspecifikationer som Gröjer själv utvecklat och ställt upp. Dessa villkor är (jag har nedan bytt ut ordet "utvärdering" mot "studien"):

- att studien ska ske ex post, dvs. att det som granskas satts igång och fått verka en tid eller avslutats,
- att studien ska innehålla något slags värdering eller bedömning av verksamheten,
- att studien ska avse genomförandet och/eller resultatet av verksamheten,
- att studien ska innehålla en jämförelse/mätning av resultatet gentemot något slags kriterium (uppställt mål e.d.),
- att studien ska ha ett systematiskt tillvägagångssätt.

Jag väljer därför att som minimum ställa krav på de tre första då dessa återkommer i många definitioner, konkluderar Gröjer (s. 35).

Princip 2. Det som myndigheten själv kallar utvärdering är en utvärdering. Inicialt är detta ett lätsamt kriterium att tillämpa. Utforskaren slipper hamra ut en egen begreppsapparat med kravspecifikationer och behöver dessutom inte penetrera var-

je rapport för att fastställa om den motsvarar de uppställda kraven. I slutänden kommer hon emellertid inte undan en egen analys menar jag.

Problemet med två urvalsprinciper är att de inte slår likformigt. Genom princip 2 urskiljs studier som kanske inte skulle passera princip 1, t.ex. utvärderingar ex ante. Därför vore det värdefullt att i efterhand ta reda på vilka undersökningar som ingår i den andra självutvalda gruppen. Om man som vetenskapsutövare låter forskningsobjektens språkbruk vara styrande för vad som väljs ut som utvärdering så måste man i stället ex post analysera vad den utvalda gruppen innehåller.

Analyschema för porträttering av utfört utvärderingsarbete

För att på ett systematiskt sätt påstå något upphetsande om faktiskt utfört utvärderingsarbete krävs ett analyschema med vars hjälp man kan uppsöka och gruppera på ett sätt som fånglar sinnet. Det schema som Gröjer använder är utan modifieringar övertaget från en studie av den unge Premfors (1987). Det ser ut så här (s. 36):

1. Syfte: Varför utvärdera?
2. Objekt: Vad utvärderas?
3. Organisering: Var och av vem utvärderas?
4. Kriterier: Vilka [värderings]kriterier används?
5. Metoder: Vilka metoder används?
6. Spridning: Hur och till vem sprids resultaten?
7. Användning: Vilka effekter får utvärderingen?

Tiden är inne att föreslå förbättringar av denna begreppsapparat. När Premfors publicerade sitt frågebatteri 1987 innebar det ett betydande framsteg. Märkligt vore det väl om det inte skulle kunna förbättras

när nu 17 år förflutit. Under de nästan två decennier som passerat har ju många nya former av utvärdering framträtt utifrån vilka apparaten kan göras utförligare men också smidigare och rikare på nyanser.

Utvärderingars övergripande syfte

Gröjer har valt att skilja mellan ett kontrollerande, inlärande (främjande), legitimerande och fördröjande syfte (108). Jag skall koncentrera mig på det kontrollerande och det legitimerande ändamålet.

Det *kontrollerande syftet* verkar rimligt definierat. Men det vore intressant att få veta vem en högskolemyndighet i så fall önskar kontrollera genom utvärdering. Vill myndighetens ledning kontrollera att underordnade enheter inom myndigheten sköter sig? Eller vill myndigheten kontrollera att underlydande organisationer, dvs. högskolor och universitet, sköter sig? Under inget av syftena får vi veta vem som skall kontrolleras, lära in, ges legitimitet och fördröjas.

Det *legitimerande syftet* handlar om att inför andra kunna visa på värdet av en verksamhet, säger Gröjer. Denna begreppsexplikation är emellertid för snäv. Legitimerande betyder att utvärderingen används för att stödja och förstärka ståndpunkter och åsikter som redan finns. Poängen kan vara att påvisa det värdefulla i en aktivitet. Men poängen kan också vara att uppvisa värdelöshet.

Överhuvudtaget verkar detta begrepps nät ha varit för finmaskigt för att skilja de stora fiskarna från de små. Alla 30 utvärderingar har hamnat i kategorin inlärande. Samtidigt har nätet varit så grovmaskigt att alla fiskar, stora som små, slunkit igenom. Så har ingen utvärdering hamnat i "fördröjande", för att ta ett exempel (s. 79).

Utvärderingsobjekt

Kategorin utvärderingsobjekt rör det som utvärderas (s. 39, 75–76, 115–121, 154, 182–184). En viktig indelning för Gröjer är om utvärderingarna orienterar sig mot den utvärderade interventionens genomförande eller resultat eller båda.

Dessutom urskiljer hon de specifika *objektklasserna* system, reformer, ämnen och utbildningslinjer. System är sådant som har med högskolesystemets totala funktion att göra, reformer sådant som kan knytas till de reformer som genomfördes under perioden. Men vad är skillnaden? Den är det inte så lätt att bli klar över (s. 75, även s. 116). Ett system kan väl bli föremål för en reform? Vad är det i så fall fråga om, ett system, en reform eller båda?

Såvitt jag begriper kan samma utvärdering hamna under System, Reform och Ämne/Utbildning. Denna trippelklassificering förefaller problematisk.

Vad gäller UKÄ:s utvärderingar visas följande fördelning på objektklasser (s. 76): System 11, Reform 14, Ämne/Utbildningslinje 16 och Övrigt 4. Vad säger egentligen denna spridning? Vad lär vi oss av att systemutvärderingarna är 11 och reformutvärderingarna 14? För att bli intressanta måste väl siffrorna jämföras med någon sorts förväntan om vad en högre myndighet bör eller skall syssla med. De skulle också kunna jämföras med hur fördelningen var inom den tidigare myndigheten eller åren innan i samma myndighet.

Just i fråga om utvärderingsobjekt har jag många kategorier som jag vill lägga till.

Tillägg 1. En spännande uppdelning är den på tidsbegränsad försöksverksamhet (projektarbete, ad hoc-aktivitet) å ena sidan och löpande verksamhet å den andra. Om det program som utvärderas är infört

på försök eller permanent kan göra en stor skillnad politikfilosofiskt. Det förekommer också i Gröjers material åtskilliga hänvisningar till försök och projekt (s. 77–78, 80). Här är författarinnan någonting på spåren, som hon skulle ha kunnat reagera häftigare på. Talet om försöksverksamheter och projekt avslöjar nämligen existensen av ett visst tillvägagångssätt för att göra den offentliga sektorn mer rationell. Offentlig politik borde bedrivas i två separata steg. Först borde man ordna försöksverksamhet i mindre skala i form av projekt vid sidan av linjen. Denna försöksverksamhet borde utvärderas. I ett andra steg borde utvärderingens resultat föras över till linjen och institutionaliseras i full skala i form av modifikationer av programmet eller ett nytt program (s. 58, 61).

Tillägg 2. Ett annat tillägg som borde bli fruktbart avser utvärdering av vanlig, normal utbildningspolitik (sakpolitik), av särskilda tidsbegränsade utbildningspolitiska satsningar och av förvaltningspolitik (118, 119). På ett ställe är det uppenbarligen förvaltningspolitik som utvärderas (118) men Gröjer har ingen kategori för detta. Hon resonerar i och för sig om distinktionen mellan förvaltningspolitik och sakpolitik. Dock tolkar jag henne så att hon inte ämnar ta hänsyn till detta i schemat utan helt begränsa sig till sakpolitik dvs. utbildningspolitik (s. 39). Hon påvisar också utvärderingar av särskilda satsningar utan att hon har någon särskild kategori för sådana.

Den positiva konklusionen är att Gröjer ser och analyserar saker trots att de faller utanför hennes schema. Den anbefallande konklusionen är att en mer nyanserad begreppsapparat bör användas, ty det ger en rikare bild av vad myndigheters utvärderingsverksamhet handlar om.

Tillägg 3. Om utvärderingar avser objekt innanför eller utanför den nationella myndigheten (s. 120).

Tillägg 4. Om utvärdering avser forskarutbildning eller grundutbildning (154).

Tillägg 5. Om utvärderingarna innehåller handlingsrekommendationer eller inte. Saken berörs i avhandlingen men kunde ha gjorts tydligare (s. 183). Om utvärderare skall avge rekommendationer eller inte är en stridsfråga i branschen.

Tillägg 6. Om utvärderingarna avser information om vissa utbildningar (utbildningsreformer) eller om de avser själva utbildningarna (utbildningsreformerna). Information om program A är inte detsamma som program A. Begreppsapparaten borde ta hänsyn till detta (s. 120).

Tillägg 7. Om utvärderingarna fokuserar utbildningarna och utbildningsreformerna eller om de fokuserar systemen för utvärdering av utbildningarna och utbildningsreformerna. Man kan nämligen också utvärdera utvärderingssystem för högre utbildning (s. 150, kallas metautvärdering). Lärosätenas kvalitetsarbete och kvalitetsprogram har också faktiskt utvärderats (155). Man utvärderar också utveckling av metoder för kvalitetsarbete, dvs inte bara faktiskt kvalitetsarbete. Det är stor skillnad mellan att utvärdera utvärderingssystem för högre utbildning och att utvärdera högre utbildning.

De tillägg jag nyss föreslagit är kanske något för framtida forskare att fundera över.

Utvärderingssubjekt (organisering)

Utvärderingssubjektet, det tredje stora inslaget i analyskemat, handlar om *vem*

som utvärderar (s. 40, 83-86, 125-128, 157-158, 185-186). Här redovisar Gröjer motsatsparet *löpande – ad hoc*. Med detta menar hon att utvärderingsfunktionen antingen ombesörjs av ett permanent organ eller inrättas tillfälligt (s. 40). En annan kategori är om utvärderingen utförs av egen personal eller av experter/fackmän utanför myndigheten. Hon arbetar också med begreppet uppdrags- och konsultbasis. Motsatsen måste vara ”inom myndigheten”. Det senare förefaller mig onödigt grovt. Uppdrag är en sak för sig. Vem som får uppdragen är en helt annan sak. Att ett uppdrag ges till ett universitet är något helt annat än att det ges till en konsult.

En viktig sak som Gröjers schema inte lyckas fånga in är det nuvarande systemet med självvärdering, kollegebedömning, platsbesök, rapportskrivning samt uppföljningar av rapporten. Detta system (121) innebär att utvärdering sker sekventiellt dvs. en grupp får utvärdera först och sen en annan. För att fånga detta måste det finnas en tidsaspekt i begreppsschemat. En dylik saknas nu.

Utvärderingskriterier

Bedömning är själva essensen i utvärdering. På den här punkten skriver Gröjer resignerat: ”Jag har inte på förhand hittat någon lämplig kriteriestruktur att använda varför själva analysen fått bidra till en klassificering av de kriterier som kunnat upptäckas” (42). Jag tolkar henne så att hon inte har något analyschema för värderingskriterier som går utöver det empiriska materialet vid varje tidpunkt (s. 40-42, 86-87, 158-160, 185-186).

I själva verket finns det rätt många taxonomier över värderingskriterier i utvärdering. En sådan visas här nedan (från Vedung 2006).

I Sakinriktade värderingskriterier

- 1 Mål
- 2 Värdekriterier för sidoeffekter
- 3 Klientkriterier
- 4 Intressenters farhågor och positiva förväntningar
- 5 Professionella kriterier: kollegers värdenormer, den professionelles egna värdenormer
- 6 Bidrag till lösning av underliggande problem (eller behov)

II Ekonomiska kriterier

- 1 Ekonomi (enbart kostnader)
- 2 Produktivitet
- 3 Kostnadseffektivitet

III Proceduriella kriterier

- 1 Legalitet
- 2 Rättssäkerhet
- 3 Offentlighet
- 4 Representativitet
- 5 Participativ demokrati
- 6 Deliberativ demokrati

Man behöver som vetenskapsidkare naturligtvis inte leta efter alla dessa värderingskriterier i utvärderingarna. Men en bred översikt är det fint att utgå ifrån när ett urval skall göras.

Sen skriver Gröjer: Det finns också mer specifika praktiska exempel på kriterier. I en handbok med amerikansk standard för utvärdering av utbildningsprogram finns en explicit kriterieuppsättning (120). Dessa kriterier är uppdelade i fyra huvudgrupper (utility, feasibility, propriety, accuracy) med ett flertal kriterier under varje huvudgrupp.

Men här blandar Gröjer ihop nivåerna. Detta är värderingskriterier för att bedöma kvaliteteten på utvärderingar av interventioner, inte för att bedöma kvaliteteten på interventioner. I det ena fallet frågas: vad är det som kännetecknar en god utvärdering? I det andra fallet frågas: vad är

det som kännetecknar en god intervention?

Man kan tänka sig utvärdering där endast ett kriterium tillämpas, t.ex. uppfyllelse av ett mål. Men det går också att föreställa sig att flera kriterier inom samma kriterietyper används, t.ex. fem mål. Likaledes kan vi tänka oss att flera kriterietyper, t.ex. brukarkriterier och mål, kan tillämpas i samma utvärdering. Värdekriterierna kan tillämpas samtidigt men också separat på samma objekt. Men de kan också tillämpas sekventiellt så att utvärderingen startar med brukarmål och brukarbedömningar för att sedan övergå till den berörda professionens mål och kanske till sist avslutas med politikernas mål.

Utvärderingsmetoder

Under denna rubrik (42–43, 87–89) grupperas utvärderingarna längs två dimensioner.

Gröjer tar sikte på om utvärderingsmetoden är standardiserad och/eller genomförs på löpande bas eller om metoden genomförs ad hoc och är mer kontextuell (43).

Utvärderingarna ordnas också på huvudsakligen kvantitativa – huvudsakligen kvalitativa. Vad gäller den senare distinktionen blir författarinnan abstrus. Hon säger: ”Att använda *kvalitativa* metoder innebär att data i form av texter (kvalitativa data) och talfördelningar (kvantitativa data) samlas in.” Ja, hon säger faktiskt att kvalitativa metoder innebär att samla in kvantitativa data. Är detta ett tyckfel? Allt nog, hon fortsätter: ”Det kan handla om texter som intervjuutskrifter och fältrapporter men också om enkätformulär som ger upphov till texter. Sådana enkla texter kan omformuleras till talfördelningar som ger överblick över större grupper. I de flesta undersökningar är det vanligt att

både kvantitativa och kvalitativa data används i olika blandning. Kvantitativa data är information som uttrycks i siffror eller andra mängdtermer. Metoder som används för att samla in och analysera kvantitativa data kallas *kvantitativa* metoder” (s. 121; se vidare kapitel 7 om Högskolleverket).

Under metoder vill jag föreslå ytterligare tre kategoriseringar: 1) typer av uppläggningar för att belysa kausalitetsfrågan, 2) metoder som innebär närhet respektive distans samt 3) typer av datainsamlingsmetoder.

En uppdelning enligt 1) skulle vara om reformen utvärderas i sitt naturliga sammanhang utan kontrollgrupp (fallstudie-metod) versus om reformen utprovas som en försöksverksamhet i liten skala med experimentgrupp och kontrollgrupp.

Egenskapen närhet står i fallet 2) för datainsamling genom växelverkande kommunikation med de undersökta. Man kan tänka sig att utvärderaren går in i dialog med de utvärderade, eller intervjuar dem. Distans står för observation när de utvärderade inte känner till i vilket större undersökningssammanhang de ingår.

Typer av datainsamlingsmetoder under 3) slutligen är dokumentmetoder inklusive statistik, interrogativa metoder (intervjuer, enkäter) och observationsmetoder (okulärbesiktning).

Utvärderingars spridning och användning

Gröjer överväger slutligen möjligheten att också anlägga en spridnings- och användningsaspekt på de utvärderingar som hon studerar (43 f). I stort sett avvisar hon den möjligheten. ”Det är ... möjligt att undersöka om utvärderingsverksamheten satt några spår i det politiska beslutsfattandet. Sådana spår kan sökas genom att leta efter

referenser till utvärderingsverksamheten i underlagen till de beslut som den politiska nivån fattar." Men detta avstår hon ifrån. "Däremot kommer de *hänvisningar* som statsmakterna gör i *det offentliga trycket* i samband med beslut om myndigheternas utvärderingsverksamhet i stort att uppmärksammas."

Om detta skall bara sägas att hon, i sin begreppsapparat, inte beaktar utvärderingars användning inom den högre myndigheten själv. Där har ju faktiskt indragning av examinationsrätt, påverkan på penningfördelning samt rankning av ämnen och lärosäten faktiskt diskuterats och i viss mån tillämpats (s. 133).

Till slut vill jag framföra ytterligare några styrkor med Gröjers avhandling.

Författarinnans empiriska handlag

Författarinnan visar gott empiriskt handlag. Hon tillämpar en väldigt stark systematik i den del av avhandlingen som avser gjorda utvärderingar under UKÄ, UHÄ och sekretariatet/KÄ: Samma schema tillämpas på utvärderingar från alla tre epokena.

Hon har också energiskt sökt upp och samlat ihop data. Universitetskanslersämbetets arkiv som finns vid Riksarkivet i Arninge har genomsökts i jakten på texter som kan förknippas med institutionen. Även annat skriftligt material som offentligt tryck och UKÄ:s anslagsframställningar har undersökts. (s. 66).

En forskare skall också ha tur. Man söker ett och hittar något annat. Serendipiteter kallas det. Av en händelse hittade Gröjer material om den konferens om utvärdering som UKÄ arrangerade i juni 1973 och som blev viktig för fortsättningen. (Utvärdering av en högskolereform, s. 215, s. 69).

En presentation av en utvärdering upptäckte hon av en annan slump i nyhetskriften UKÄ-aktuellt (70/71:8). Presentationen avser "Det första årets erfarenheter av den nya studieordningen (1969 års) vid filosofisk fakultet". (s. 234). Upptäckten medförde att författarinnan granskade myndighetens samtliga nyhetsskrifter i jakten på utvärderingar.

Gröjers avhandling reser många spännande problem. När kom utvärdering in i svensk statsförvaltning? Hur institutionaliserades utvärdering? Hur bedrevs utvärdering? Hon har också lämnat svar på dessa och liknande frågor, även om man naturligtvis skulle kunna borra djupare och systematisera tydligare. Hennes studie är mycket välkommen. Jag hoppas att hon får många efterföljare i utvärderingsbranschen.

Referenser

- Bengtsson, Bo. 2004. *Swedish Housing Corporatism – A Case of Path Dependence?* Gävle: Uppsala University, Institute for Housing and Urban Studies: unpublished conference paper.
- Dahler-Larsen, Peter. 2002. *Evaluering kortlagt: en undersøgelse af amters og kommuners evalueringspraksis på det sociale område*. Århus: Systime.
- Foss Hansen, Hanne. 2003. *Evaluering i staten: Kontrol, læring eller forandring*. Fredriksberg: Samfundslitteratur.
- Furubo, Jan-Eric – Sandahl, Rolf. 2002. A Diffusion Perspective on Global Developments in Evaluation, s 1–23 i Furubo, Jan-Eric – Rist, Ray C. – Sandahl, Rolf (red.), *International Atlas of Evaluation*. New Brunswick, NJ. och London: Transaction Publishers.
- Hansen, Anja – Foss Hansen, Hanne. 2000. Evaluering i Danmark, *Nordisk Administrativ Tidsskrift* 2, juni 2000, s 157-177.
- Karlsson, Ove – Kuivainen, Hannu. 2000. *Att utvärdera kvalitet i högre utbildning: Internationella jämförelser av system för kvalitetsgranskning*. Stockholm: Högskoleverket, Arbetsrapporter 2000: 15 AR.

- Nordisk Administrativ Tidsskrift*. 2000. 81 (2), s 107-177 (Hanne Foss Hansen, gästredaktör).
- Rist, Ray C. (red.). 1990. *Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations*. London: Transaction Publishers.
- Sandahl, Rolf. 1992. Evaluation at the Swedish National Audit Bureau, s 115-121 i Mayne, John – Bemelmans-Videc, Marie-Louise – Hudson, Joe – Conner, Ross (red.), *Advancing Public Policy Evaluation: Learning from International Experiences*. Amsterdam: Elsevier Science Publishers.
- Scott, W Richard. 1995. *Institutions and Organizations*. London: Sage.
- Trow, Martin – Nybom, Thorsten. 1991. *University and Society: Essays on the Social Role of Research and Higher Education*. London: Jessica Kingsley Publishers.
- Vedung, Evert. 1981. *Energi politiska utvärderingar 1973-1981 (Energy Policy Evaluations 1973-1981)*. Stockholm: Liber.
- Vedung, Evert. 1992. Five Observations on Evaluation in Sweden, s 71-84 i Mayne, John – Bemelmans-Videc, Marie-Louise – Hudson, Joe – Conner, Ross (red.), *Advancing Public Policy Evaluation: Learning from International Experiences*. Amsterdam: Elsevier Science Publishers.
- Vedung, Evert. 1993. *Utvärderingar av svensk energipolitik: En kommenterad bibliografi*. Stockholm: Närings- och teknikutvecklingsverket, NUTEK R 1993:34.
- Vedung, Evert. 2004. *Utvärderingsböljans former och drivkrafter*. Helsingfors: Stakes, FinSoc Working Papers 1/2004.
- Vedung, Evert. 2006. Evaluation, artikel för Peters, Guy – Pierre, Jon (red.), *Encyclopedia of Policy Analysis* utkommer 2006.
- Vedung, Evert – Furubo, Jan-Eric – Sandahl, Rolf. 2000. Utvärdering i det svenska politiska systemet: Åtta iakttagelser, *Nordisk Administrativ Tidsskrift* 81 (2), s 113-130.

Maria Strömvik: To Act as a Union: Explaining the development of the EU's Collective Foreign Policy. Lund: Lund Political Studies, 2005.

KJELL GOLDMANN¹

Maria Strömviks insiktsfulla och aktuella avhandling behandlar Europeiska unionens "kollektiva utrikespolitik".² Här har vi, skriver Strömvik, en grupp "suveräna, demokratiska och ibland motvilliga" stater som inte desto mindre allt mer kommit att agera "kollektivt som en union" (3).³ Att förklara detta ovanliga företag är avhandlingens huvudsyfte. En ytterligare ambition är att "kalibrera" och "fininställa" några välkända teoretiska "ramverk" (4): dels den s.k. realismen med dess beto- nande av frågor om makt och inflytande, dels vad som brukar kallas institutionalismen, den tidvis dominerande formen för realismkritik. Det är enligt Strömvik en poäng med avhandlingen att den för ihop studiet av EU:s gemensamma utrikes- och säkerhetspolitik (GUSP) med den gene- rella teorin om internationell politik.⁴

Studieobjektet, EU:s "kollektiva utrikespolitik", preciseras till det som EU:s

1 Kjell Goldmann är professor emeritus i statsvetenskap vid Stockholms universitet. Epost: Kjell.Goldmann@chello.se

2 Denna anmälan bygger på fakultetsoppositionen vid disputationen i Lund 24 september 2005.

3 Siffror inom parentes är sidhänvisningar till avhandlingen.

4 Jag kommer att använda uttrycket GUSP med avseende på hela undersökningsperioden fastän verksamheten tidvis haft annan benämning, liksom jag kommer att använda benämningen EU även med avseende på perioden före TEU.