

Utvärderingsmodeller

evert vedung

Vilka frågor kan man egentligen besvara genom en utvärdering och hur skall man gå tillväga för att utvärdera? I artikeln presenteras en översiktlig bild av olika utvärderingsmodeller. Med en historisk överblick typologiseras utvärderingsmodellerna efter grundläggande värdekriterier.

En ambitiös utvärdering försöker besvara två frågor. Motsvarar resultatet det som man kan förvänta eller önska sig utifrån något värdekriterium? Är detta resultat på något sätt en produkt av interventionen?

När utvärdering började praktiseras i Sverige inom skolpolitik på 1950-talet och kom upp som en nyhet i USA omkring 1965 besvarades dessa båda frågor på ett mycket karakteristiskt sätt. Frågan om kriterieuppfyllelse uppfattades som en fråga om interventionen uppfyllde sina egna mål. Motsvarar de resultat som uppnås de i förväg uppställda målen med interventionen? Utvärdering innebar att besvara denna måluppfyllelsefråga. Problemet med interventionens effekter löstes på ett liknande karakteristiskt sätt. Lösningen bestod i att använda sig av bästa möjliga naturvetenskapliga metoder. Den uppläggning som favoriserades var det randomiserade eller matchade experimentet. På slumpmässiga grunder borde akademiska forskare skapa två likvärdiga grupper, en

experimentgrupp och en kontrollgrupp. Forskarna borde sedan utsätta experimentgruppen men inte kontrollgruppen för interventionen och skrupulöst nog mäta utvecklingen i båda grupperna före, under och efter interventionen. Om det uppstod några skillnader mellan grupperna borde de kunna tillskrivas interventionen eftersom allt annat utom interventionen genom randomiseringen var lika.

Båda dessa ganska stränga synsätt finns fortfarande kvar i utvärderingsdiskursen. Det som hänt är att de trängts tillbaka. Alla medger att offentliga insatser kan utvärderas också på många andra sätt än som uppnående av på förhand uppsatta mål. Man kan utvärdera mot ekonomiska mål som produktivitet och utfallseffektivitet, mot intressenters mål och mot brukarmål för att nu nämna några. I utvärdering av högre utbildning och forskning har det uppstått en specialitet som utvärderar mot de professionellas (forskarnas) egna kvalitetsmål i självvärderingar samt mot forskarkollegers kvalitetsuppfattningar. Det har också skett en förskjutning från att se utvärdering som forskning mot att se utvärdering som

Evert Vedung är professor i statsvetenskap och arbetar på Institutet för bostads- och urbanforskning och Statsvetenskapliga Inst. vid Uppsala universitet.

demokratisk i bemärkelsen participativ och deliberativ, som en angelägenhet för vanliga berörda.

Den översikt som här skall ges sker utifrån ett smalt perspektiv. Utvärderingsmodeller skall ordnas på basis av de grundläggande värdekriterier som används för att bedöma verksamheten.¹ En taxonomi över utvärderingsmodeller grupperade efter grundläggande värdekriterier presenteras i figur 1.

Figur 1.

Utvärderingsmodeller.

Substansmodeller

Måluppfyllelsemodellen

Bieffektsmodellen

Brukarmodeller

Intressentmodeller

Professionella modeller: kollegiebedömning

Professionella modeller: självvärdering

Ekonomiska modeller

Produktivitetsmodellen

Effektivitetsmodeller: kostnads-effektivitetsanalys

Effektivitetsmodeller: kostnads-intäktsanalys

Substansmodeller är helt och hållet inriktade på interventionens, verksamhetens, sakinriktade aspekter. Det är resultaten eller verksamheten i sak som står i cen-

trum. I vilken riktning och hur mycket har energipolitiken påverkat energiflödena i samhället? Vilka effekter på miljön har det miljöpolitiska programmet fått? Även ekonomiska modeller är inriktade på saken men inte enbart. De undersöker också vad det kostar att producera saken. Skillnaden mellan ekonomiska modeller och substansmodeller är att de förra även beaktar kostnader.

Substansmodellerna utgör en mångfacetterad grupp. Det urval som tas upp är måluppfyllelsemodellen, bieffektsmodellen, brukarorienterade modeller, intressentmodeller och professionella modeller. Viktigast bland ekonomiska tillnärmelser är produktivitets- och effektivitetsmodellerna.

Måluppfyllelsemodellen

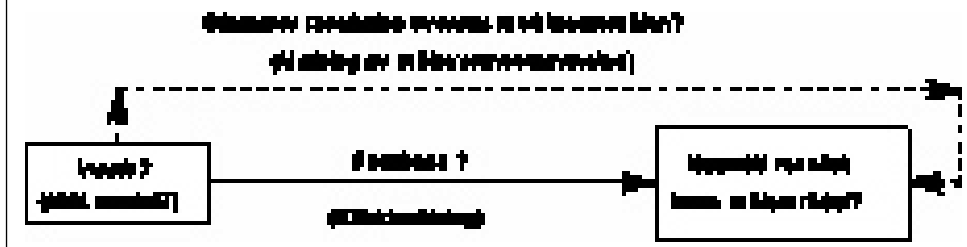
Vid måluppfyllelseutvärdering undersöks om de resultat som insatsen faktiskt producerat inom målområdet motsvarar insatsmålen. Måluppfyllelsemodellens grundfrågor är två. Stämmer resultaten inom målområdet överens med de beslutade insatsmålen? Beror detta i så fall på insatsen? Det första momentet kallas här mätning av målöverensstämmelse, det andra effektmätning.

Måluppfyllelsemodellen är föredömligt tydlig och enkel. Först måste en utvärderare klara ut hur målen för styrningen ser ut samt undersöka om målen har uppnåtts inom målområdet. Sedan kan hon ställa frågan om insatsen i så fall bidragit till måluppfyllelsen. Måluppfyllelseutvärderingens enkla anatomi illustreras i figur 2.

¹ Framställningen utgör en kraftigt förkortning och skärpning av kapitel 4 i min Utvärdering i politik och förvaltning (Lund: Studentlitteratur, 1998, 2:a upplagan, www.studentlitteratur.se) och kap 4 i Public Policy and Program Evaluation (New Brunswick, N.J.: Transaction Publishers, paperback 2000, www.eurospan.co.uk).

Figur 2.

Måluppfyllelsemodellen för utvärdering



Ett gripbart exempel må anföras. Det institutionaliserade målet i 1978 års svenska energisparplan för befintlig bebyggelse var att »nettoenergiförbrukningen« i 1978 års byggnadsbestånd år 1988 skulle vara cirka 35 TWh lägre per år än det var år 1978. Detta innebar en reduktion på 30 procent. För att uppnå detta förbluffande klart uttryckta mål anslog riksdagen betydande belopp för isolering av fastigheter samt till rådgivning för energisparande. En måluppfyllelseutvärdering av energisparplanen skulle innebära att vi undersökte om energianvändningen i 1978 års byggnadsbestånd verkligen var 35 TWh lägre per år vid 1988 års utgång samt om det uppmätta resultatet på något sätt hängde samman med den av riksdagen antagna planen.

Måluppfyllelseutvärdering avser nästan alltid kollektiva aktörers mål. Det betyder att utgångspunkten är mål som institutionaliserats genom att uttryckligen återopas i beslutssituationen. Modellen kan avse offentliga interventioner på alla tänkbara politiska och administrativa nivåer från den subkommunala över EU-program till den globala nivån.

Måluppfyllelsemodellen kan också tillämpas på interventioner av olika aggre-

geringsgrad på samma hierarkiska nivå. Evaluanden kan vara den samlade beslutade rikspolitiken i en politisk sektor, ett stort programbeslut bland flera i en sektor eller en liten konkret myndighetsåtgärd i en sektor.

Måluppfyllelsemodellen kan tillämpas på alla led i implementeringskedjan. Vilka är målen för inflödet i verksamheten (eller för förvaltningen, slutprestationerna, brukarnas mottagande, brukarnas åtgärder, samhällsutfallet, utvärderingen och återkopplingen)? Hur ser det faktiska läget ut beträffande inflödet (förvaltningen, slutprestationerna, brukarnas mottagande, brukarnas åtgärder, samhällsutfallet, utvärderingen och återkopplingen)?

I en inflödesutvärdering jämförs önskat inflöde i form av pengar, kompetens och materiel med faktiskt. Hur stora intäkter förväntade sig exempelvis myndigheten kamma in genom att sälja tjänster och hur mycket flöt faktiskt in? En förvaltningsutvärdering ställer önskad omvandling mot faktisk. Vad planerade exempelvis myndighetens ledning att göra för att sprida sina intentioner ner till den operativa fältpersonalen och vad åstadkoms i praktiken? En slutprestationsutvärdering ställer frågan

om de faktiska slutprestationerna i form av tjänster eller andra styrmedel gentemot adressaterna motsvarar vad som krävs. I en brukarutvärdering ställs värdekriterier för brukarnas mottagande av styrningen, t.ex. vilken grupp av individer man ville nå fram till, mot det faktiska mottagandet, t.ex. vilka individer som faktiskt nåddes av insatserna. Åtgärdsutvärderingar jämför önskade brukaråtgärder med faktiska. I en utvärdering av samhällsutfallet tar utvärderaren reda på det faktiska utfallet och värderar det i ljuset av något bedömningskriterium. Uppgiften för en metautvärdering respektive en återkopplingsutvärdering slutligen är att undersöka om de önskemål som förelegat eller föreligger med utvärderingen respektive återkopplingen har realiserats.

Av pedagogiska skäl skall jag i fortsättningen utgå från fallet att mål avser samhällsutfall.

Måluppfyllelsemodellens styrka

Måluppfyllelsemodellen har överösts och överöses med störtfloder av kritik. Nära nog alla som förespråkar någon annan modell gör det i polemik mot måluppfyllelse. Här skall två skäl för måluppfyllelsemodellen nämnas: demokrati- och objektivitetsargumentet.

Det första skälet, demokratiargumentet, hämtar sin näring ur föreställningen om den parlamentariska styrkedjan och därmed den representativa demokratins perspektiv. Den förda politikens egna officiella mål är inte vilka hugskott som helst.

De har i konstitutionell ordning fastställts i beslutande församlingar av folketsrepresentanter. Målen är institutionaliserade och offentligt fastlagda. Detta gör att statens, landstingets eller kommunens mål för ett politikområde har en annan status och legitimitet än exempelvis intressentgrupperns mål och förväntningar, riktade mot samma politikområde. Legislaturers mål får också en särställning gentemot andra intressenters genom att de kommit till under ansvar. I princip måste ju politiker ta hänsyn till inte bara vad de vill utan även vad som är resursmässigt möjligt att genomföra. Särintressen kan alltid driva krav och sätta upp mål utan att som ansvariga politiker behöva tänka på helheten och den finansiella situationen.

Detta gäller också andra organ i offentliga sektorn såsom regeringar och myndigheter, vilka handlar på delegation från beslutande församlingar. Om en myndighet beslutar inrätta ett program i syfte att uppnå vissa mål, så får dessa mål sin legitimitet av att regeringen delegerat denna beslutanderätt till myndigheten och att regeringen i sin tur fått denna rätt delegerad till sig av riksdagen.

Måluppfyllelsemodellen har alltså sina poänger ur den representativa demokratins synvinkel och därmed indirekt i ett medborgarperspektiv.

Det andra skälet, objektivitetsargumentet, baseras på föreställningen att måluppfyllelsemodellen erbjuder en objektiv lösning på måttstocksproblemet vid utvärdering. Eftersom interventionsmålen är uttryckligen angivna eller åtminstone tydligt underförstådda i texter och förarbeten, kan de i princip fastställas empiriskt.

Genom att ta dessa som värdekriterier - deskriptiv värdering - behöver utvärderaren inte ställa upp egna värdenormer vid bedömningen - preskriptiv värdering. Utvärdering med deskriptiv värderingsmetod blir en bedömning av andra ordningen, en bedömning med utgångspunkt i andras mål och värden. Då värdegrundsfrågan påstås kunna lösas objektivt, kan hela utvärderingen bedrivas objektivt. Eftersom målen är exogent givna, består utvärderingsuppgiften i att undersöka om anvisade medel bidragit till faktisk måluppfyllelse. Detta studium av medel blir ett rent empiriskt spörsmål, som kan bedrivas objektivt med vetenskapliga metoder.

Måluppfyllelseutvärdering har således starka sidor, främst ur den representativa demokratins synvinkel. Men det är också uppenbart att den lider av besvärliga brister, mot vilka läsarens uppmärksamhet nu skall riktas.

Måluppfyllelsemodellens svagheter

De viktigaste invändningar, som framförs mot måluppfyllelsemodellen som substansmodell är oklarhetsargumentet, bieffektsargumentet, argumentet om strategiska motiv och toppen-nerargumentet.

Det första skälet, oklarhetsargumentet, går ut på att insatsmål på grund av sin dimmighet är otjänliga som värderingsgrunder. Att politiska och administrativa mål är diffusa är en truism, som i den förvaltningspolitiska diskursen upprepas till leda. Här skall en suddighetsform, oprioriterade målkataloger, och de oklarheter detta

skapar uppmärksammas. Problemet med målkataloger gör sig gällande dels inom ett och samma program, dels mellan program.

I samband med varje större reform redovisas regelmässigt en hel uppsättning mål. Visserligen kan måhända ett av dem pekas ut som huvudmålet men samtidigt framhålls att det skall vägas mot alla de andra, kanske potentiellt motstridande målen utan att det exakt sägs hur vägningen skall gå till. Portalparagrafen i 1979 års svenska skogsvårdslag (SFS 1979:429) innehöll en sådan oförlöst målkonflikt mellan virkesproduktion å ena sidan och naturvårdens och andra allmänna intressen å den andra. Texten löd:

Skogsmarkmeddesskogskallgenomlämpligt utnyttjande av markens virkesproducerande förmåga skötas så att den varaktigt ger en hög och värdefull virkesavkastning. Vid skötseln skall hänsyn tas till naturvårdens och andra allmänna intressen.

Mer sades inte i skogsvårdslagen. I skogsstyrelsens förordning (SKSFS 1986:6) däremot utvecklades naturvårdens och andra allmänna intressen till åtta mål (jag har markerat dem med siffror):

Skogsbruket skall bedrivas med hänsyn till skogens betydelse för (1) växter och (2) djur, för (3) vattenbalans och (4) lokalklimat samt för (5) friluftsliv och (6) rekreation. Hänsyn skall tas till (7) värdefull kulturmiljö och (8) landskapsbild.

Exakt hur avvägningen mellan virkesproduktion och de åtta andra hänsynen skulle ske sades inte. Oklarheten var säkerligen

befogad. Det är svårt att i centrala beslut på kontor i Stockholm och Jönköping specificera hur prioriteringen skall gå till vid varje enskild skogsavverkning i Ekshärad i Värmland eller Vemdalen i Härjedalen. Denna klokhet skapar emellertid problem för måluppfyllelsemodellen. Något klart förväntat utfall kan nämligen inte utpekas, vilket innebär att programmålen inte ger någon säker vägledning i utvärderingsarbetet.

Oklarhetsargumentet är en bärig invändning mot måluppfyllelsemodellen. Om politiker eller administratörer inte klarat ut vad de vill uppnå, så hamnar utvärdering enligt måluppfyllelsemodellen i ett intellektuellt moras. En lösning är naturligtvis att ta målen ett för ett och undersöka hur de uppfyllts.

Samtidigt bör understrykas att offentliga mål inte nödvändigtvis behöver vara oklara. Ibland finner politiker för gott att sätta klara, t.o.m. kvantifierbara mål, vilket vårt ovan anförda exempel 1978 års energisparplan visar. I dessa fall träffas naturligtvis inte måluppfyllelsemodellen av oklarhetsargumentet.

Ännu starkare är den andra invändningen, bieffektsargumentet. Medveten offentlig styrning får i allmänhet konsekvenser, som inte förutsågs av reformbärarna i beslutsögonblicket. »Det är svårt att sia, särskilt om framtiden«, som Niels Bohr skämtsamt uttryckte det. Om utvärderingsom måluppfyllelsemodellen koncentreras på de reformbärande krafternas ursprungliga mål med reformen, kommer den per definition inte att spåra upp oförutsedda bieffekter. Detta kan ge en skev bild av vad reformen åstadkommit. En aktivitet som ger många spinoffeffekter måste

väl rimligen vara bättre än en verksamhet, som inte utlöser någon extra energi alls.

Förekomsten av oavsiktliga bieffekter i avsiktligt handlande utgör för övrigt det kanske starkaste skälet för att överhuvudtaget göra utvärderingar. Att lägga upp dem på ett sätt som omöjliggör spårandet av sådana bieffekter vore då helt förkastligt.

Det tredje skälet mot måluppfyllelsemodellen, argumentet om strategiska motiv, påstår att denna utvärderingsform är oförmögen att hantera aktörernas strategiska överväganden. Utvärderingar som utgår från officiella sakmål ser bara toppen på ett isberg; sakmålen avslöjar blott en obetydlig del av de överväganden, som kan ligga bakom interventionen. De syften som redovisas offentligt kan vara av symbolisk karaktär och inte avsedda att förverkligas, medan de reella, outtalade motiven pekar åt andra håll. Kanske var huvudmotiven bakom reformen att hålla samman partiet, vinna väljare eller bereda marken för en koalitionsregering. När dessa mål uppnåtts, är man ointresserad av att genomföra reformen i sak.

För statsvetare tillhör denna analys av drivkrafterna bakom offentliga ingripanden det intellektuella allmängodset. Som kritik av måluppfyllelsemodellen är argumentet om strategiska motiv emellertid i ett avseende inte träffande. Rent logiskt måste det gå att bruka måluppfyllelsemodellen och ändå ta hänsyn till reformbärarnas strategiska syften. De kan nämligen användas teoretiskt, som förklaringar till varför det avsedda sakresultatet blev som det blev. Det faktum att initialt deklarerade sakmål ej blivit uppfyllda kan bero på att dolda strategiska övertygelser varit de verkliga drivfjädrarna bakom reformen. Att ge

sådana förklaringar är en huvudpoäng i utvärderingsforskningen och med en viss utvidgning skulle de kunna rymmas inom ramen för måluppfyllelsemodellen.

Det fjärde och sista skälet, toppener-argumentet, säger att måluppfyllelsemodellen utgår ifrån en konventionell syn på förhållandet mellan huvudman och exekutiv enligt vilken exekutiven (förvaltningen) även i praktiken lydigt verkställer huvudmannens (politikens) beslut.

Toppener-argumentet är emellertid felaktigt. Måluppfyllelseutvärdering tar inte för givet att förvaltning och andra genomförare faktiskt också är beslutsfattarnas tjänare. Det finns ingenting i måluppfyllelsemodellen som säger att insatsen gett upphov till det avsedda utfallet. Tvärtom är det just detta kausalsamband, som skall undersökas i måluppfyllelsemodellens effektanalys. Måluppfyllelsemodellen består ju av en uppsättning frågor som ställs vid utvärdering, inte en samling svar på dessa frågor. Kontentan av studien kan bli att insatsen inte haft någon inverkan på utfallet. Invändningen att modellen tar för givet att program de facto är målsökande robotar drabbar således inte måluppfyllelsemodellen ty det är just detta som modellen siktar till att granska.

Måluppfyllelseutvärderingens styrka är att den tar officiella mål på allvar. Att detta är en god egenskap hänger ihop med den representativa demokratins teori. I ett styrperspektiv behöver både medborgare och politiker målutvärdering för att kontrollera om deras exekutiver sköter sig.

Men kraftiga invändningar kan också riktas mot måluppfyllelsemodellen. Av dessa är bieffekts- och oklarhetsargumen-

ten särskilt allvarliga. Det är nu hög tid att presentera en annan målorienterad modell, som tar hänsyn till den tunga bieffektskritiken - bieffektsmodellen.

Bieffektsmodellen

Margaret Thatcher lär ha sagt: »There is only one rule in politics: the unexpected will happen.« Problemet oförutsedda bieffekter skapar problem för måluppfyllelsemodellen. Hur kan icke föregripna bikonsekvenser upptäckas och värderas i en utvärdering, som tar sin utgångspunkt i och enbart är inriktad på i förväg angivna mål? För att lösa detta problem måste måluppfyllelsemodellen utvidgas till att omfatta sökande efter följer vid sidan av målområdet. Detta kallas här bieffektsmodellen.

Bieffektsmodellen liknar måluppfyllelsemodellen i så måtto att på förhand uppställda mål bibehålls som viktigaste värdekriterium. Det nya med bieffektsutvärdering är att letande inom målområdet efter resultat kompletteras med sökande utanför målområdet efter sidokonsekvenser. Att bieffektsmodellen ändå, liksom måluppfyllelsemodellen, är baserad på mål framgår av uttrycket »sidoeffekt«. Sidoeffekter kan endast finnas i relation till huvudeffekter, vilka pekas ut av avsikterna i insatsen. Men mål kompletteras som värdekriterium av kriterier för bieffekter. Den underliggande idén är att offentliga ingrepp kan åstadkomma annat än avsedda resultat. De kan ge upphov till glada överraskningar, men också skapa nya problem. Lösningar på problem blir problem, som måste lösas.

Ett exempel på en lösning som blivit ett

problem är värmepumparnas roll i svensk energiförsörjning. Med början på 1970-talet utbetalades statliga pengar till installation av värmepumpar för återvinning av spillenergi som en lösning på oljeproblemet och kärnkraftsproblemet. Värmepumparnas snabba utbredning i landet var en konsekvens av det statliga stödet. Efter en tid upptäcktes att köldmediet CFC i värmepumparna kunde läcka ut och med lång tidsfördröjning skada jordens skyddande stratosfäriska ozonlager. En liten del av skadorna på ozonskiktet kan ses som en oavsiktlig bieffekt av det svenska statliga stödet till värmepumparna. Från 1980-talets mitt blev därför värmepumparna ett miljöproblem, som måste lösas politiskt.

Att offentliga interventioner kan ge upphov till bieffekter är en självklarhet. I interventionsbärarnas perspektiv kan bieffekt definieras som en i vart fall partiell faktisk konsekvens av insatsen, som inte kan räknas till de eftersträvade huvudeffekterna. Huvudeffekter kan definieras som de faktiska effekter som i vart fall delvis produceras av insatsen och som reformbärarna medvetet önskar uppnå med insatsen i fråga. Huvudeffekt är således i just detta fall knuten till insatsens sakinriktade mål samt till vad insatsens införare trodde sig kapabla att uppnå. Vidare är den per definition både förutsedd och positivt värderad av interventionens tillskyndare.²

2 Vissa personer använder termer som interna och externa effekter, varvid de förra hänför sig direkt till insatsen, exempelvis eliminering av mygg som en intern effekt av ett myggutrottningsprogram medan tillkomst av ett nytt rekreationsområde blir en extern effekt. Se Vedung 1998:60.

Offentliga interventioner kan också få perversa effekter. Perversa är sådana följder av offentliga ingrepp, som blir rakt motsatta de som avsågs. Eftersom dessa självmotverkande konsekvenser uppträder inom målområdet för ingreppet, så kan de inte betecknas som bieffekter. De är heller inga huvudeffekter, eftersom de inte är medvetet önskade av interventionens tillskyndare. Perversa effekter av avsiktligt handlande bör behandlas som en separat kategori.

Perversa effekter är också något annat än nolleffekter. Nolleffekter betyder att insatsen inte får några konsekvenser alls inom målområdet. Vid perversa effekter uppträder konsekvenser av interventionen, men rakt motsatta de som avsågs.

Eftersom perversa effekter och nolleffekter uppträder inom målområdet, har målluppfyllelsemodellen inga problem med att spåra dem. Men detta betyder också att modellen inte kan upptäcka och fastställa sidoeffekter, eftersom dessa per definition faller utanför det förspecifierade målområde, som målluppfyllelsemodellen täcker.

Bieffekter kan vara förutsedda och beaktade i den kalkyl, som föregick programbeslutet. De kan vara både positiva och negativa.

Efter oljekrisen 1973 inrättades i Sverige ett nationellt stödprogram för energihushållning i bostadshus. Genom utvidgning förvandlades det år 1978 till den ovan omtalade energisparplanen för den befintliga bebyggelsen. Genom statliga bidrag och lån för tilläggsisolering av fastigheter ville regeringen Fäldin I spara bort kärnkraften. Det avsedda utfallet var naturligtvis att spara energi. År 1988 skulle

energianvändningen vara 35 TWh lägre per år i 1978 års bebyggelse. Samtidigt förutsågs att stödet skulle kunna få en ogynnsam inverkan på förmögenhetsfördelningen i befolkningen, ty välsituerade människor som villaägare och ägare av stora hyreshus kunde förväntas utnyttja medlen i högre grad än ekonomiskt svaga grupper. Det rörde sig här om en bieffekt, som reformens tillskyndare inte ville åstadkomma, men som de var beredda att ta för att befrämja det överordnade målet, energieffektivisering. »Man får ta det onda med det goda«, var en fras som förekom vid denna tid.

Ytterligare en förutsedd negativ bieffekt var att kulturhistoriskt värdefull bebyggelse skulle kunna förstöras genom att putsade fasader och träpaneler kläddes in med plåt i samband med tilläggsisolering.

Samtidigt förutsågs positiva sidokonsekvenser. Stödprogrammet skulle höja den ekonomiska temperaturen och öka sysselsättningen. Genom tätning runt fönster, i väggar och på vindar skulle draget minska, vilket skulle leda till ett behagligare inomhusklimat. Väggisolering och treglasfönster skulle hålla buller ute.

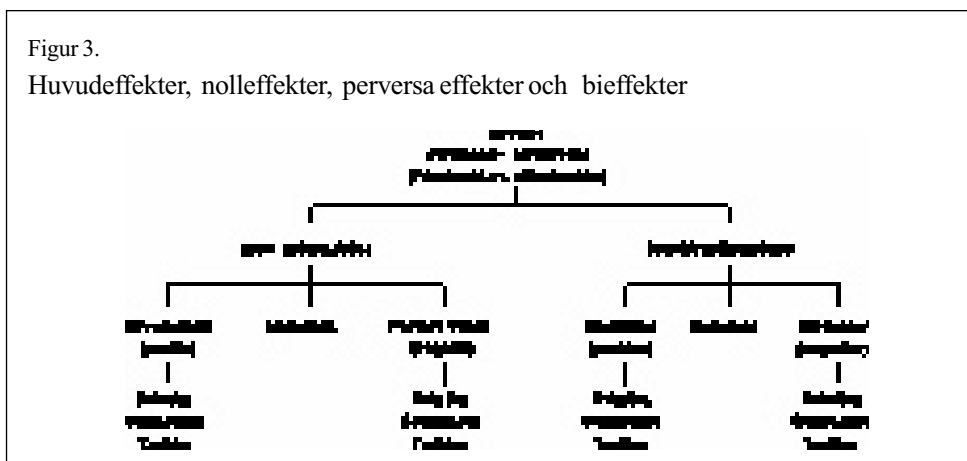
En studie som undersökte programmets effekter på energianvändning men även på förmögenhetsfördelning och sysselsättning och som dessutom uppmärksammade stadsmiljöaspekter och komfortfrågor skulle vara en bieffektsutvärdering.

Hittills har jag uppehållit mig enbart vid förutsedda effekter. Men många bieffekter är oförutsedda. Liksom sina förutsedda motsvarigheter kan de vara både positiva och negativa. Som exempel på en negativ oförutsedd sidokonsekvens kan vi anföra att tilläggsisoleringen innebar att husen blev mycket tätare vilket ledde till ökad radioaktivitet framför allt genom radon (denna oförutsedda konsekvens är särskilt intressant eftersom den första fälldinregeringen med sitt energisparprogram ville spara bort kärnkraften, som ansågs farlig genom den radioaktiva strålningen).

Resonemanget om huvudeffekter, perversa effekter, nolleffekter och bieffekter sammanfattas i klassifikationsträdet i figur 3, som visar vilka konsekvenser som kan studeras i utvärderingar.

Perversa effekter kan inträffa långt fram i styrkedjan, t.ex. i fjärde, femte eller längre

Figur 3.
Huvudeffekter, nolleffekter, perversa effekter och bieffekter



bort liggande led. Dessutom uppträder de många gånger efter mycket lång tid.

Rekommendationen att utvärderarna skall beakta perversa effekter är truistisk. Om insatsen får rakt motsatta effekter måste det vara något fel på den. Men varför är det så viktigt att beakta bieffekter? Naturligtvis därför att de är centrala element i varje helhetsbedömning av en offentlig verksamhet. Om det visar sig att sidoverkningar, som på förhand varit kända och positivt värderade, inte kommit till stånd, trots att styrningen varit i kraft under avsedd tid, så bör detta få följder för reformprogrammet. Om sysselsättningseffekten av ekonomiskt bidrag till isolering av hus är mycket mindre än man på förhand kalkylerat med, så finns det mindre anledning att behålla styrmedlet än förut, även om förväntningar om energibesparingar blivit infriade.

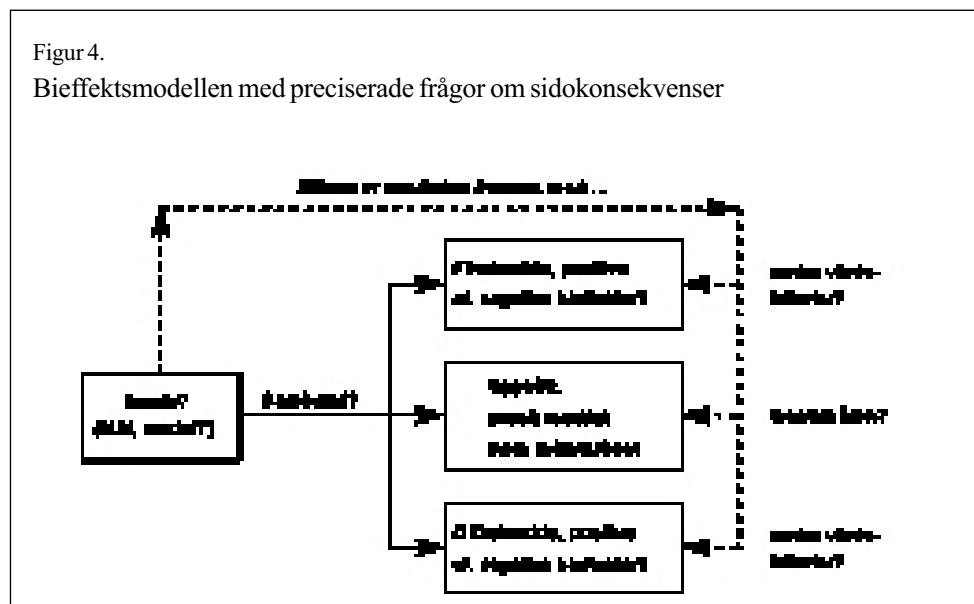
Därmed skulle vi ha avlägsnat oss långt från insatsens ursprungliga intentioner.

Betoningen ligger på samtliga effekter, avsedda och oavsedda, negativa och positiva. Men fortfarande gör utvärderaren en distinktion mellan förutsedda och oförutsedda effekter, vilket förutsätter att på förhand uppställda mål används som organiserande princip. Även den bieffektsbeaktande utvärderingen är delvis målbaserad.

Om verksamhetens samtliga effekter inom och utanför målområdet undersöktes och bedömdes, skulle utvärderingens struktur på resultatsidan bli som i figur 4.

Jag rekommenderar starkt bieffektsutvärdering framför måluppfyllelseutvärdering. Ett av de kraftfullaste argumenten för att överhuvudtaget göra utvärderingar är, som redan framhållits, insikten att resultaten av offentliga ingrepp är oförutsebara, eftersom oavsiktliga händelser alltid inträffar. Det måste vara en viktig uppgift vid utvärdering att klarlägga dessa bieffekter.

En utmaning för bieffektsutvärderaren är vilka värdekriterier som skall användas



vid bedömningen. Idealt skulle hon vilja balansera värdet av den avsedda huvudeffekten mot värdet av de lyckliga och skadliga bieffekterna. Negativa sidoeffekter skulle då kunna bli en motvilligt accepterad kostnad för att uppnå den övergripande, värdefulla effekten. Starka positiva bieffekter skulle också öka godtagbarheten av program med svag måluppfyllelse. För att kunna genomföra denna kalkyl krävs kriterier för huvudeffekten, varje typ av bieffekt samt för avvägningen dem emellan.

Den deskriptiva värderingsteori som anvisas av måluppfyllelsemodellen - att insatsens egna mål skall utgöra måttstocken - är uppenbart otillräcklig av följande skäl. Om vissa bieffekter inte är förutsedda, så är naturligtvis varken värdekriterier för bedömningen av dessa var och en för sig eller för sammanvägningen av dem till ett allomfattande värde på hela insatsen inte angivna. Därför är officiella mål otillräckliga för att värdera oförutsedda bieffekter.

En möjlig lösning på kriterieproblemet ligger i en mer kreativ tillnärmelse till värderandet. Utvärderaren kan kartlägga om huvudeffekten och förutsedda bieffekter verkligen inträffat och värdera dem i ljuset av angivna mål. Vid sidan härav kan hon också lokalisera oförutsedda bieffekter men överlåta till olika användare av utvärderingen att själva göra den sammanvägande värdebedömningen. Eftersom modellen litar till andras värdekriterier, vilar den fortfarande på en deskriptiv värderingsteori. Men beaktandet av oförutsedda bieffekter tvingar utvärderingens nyttjare att beakta andra värden än dem, som från begynnelsen var inkorporerade i interventionen.

Brukarmodeller

Brukarorienterade utvärderingsmodeller inriktar sig på mötet mellan verksamheten och dess klienter. Huvudfrågan är om den offentliga verksamheten i leveransögonblicket motsvarar den kvalitet, som mottagaren kräver eller behöver. Brukarmodeller använder antingen adressaternas förväntningar (önskemål) eller deras behov som normativt bedömningskriterium. Att utvärdera mot behov och att utvärdera mot förväntningar är emellertid principiellt skilda saker, eftersom de förra ofta anses vara något annat än utsagda önskemål. För enkelhetens skull utgår jag ifrån att brukarmodellen tar sin utgångspunkt i slutmottagarnas önskemål och förväntningar.

En parallell till konsumenters inköp på en marknad anses ibland kasta ljus över brukarutvärdering. När konsumenten köper i butiken, är det inte producentens syfte utan hennes egen bedömning av varans värde som avgör. Brukarorienterad utvärdering bygger på föreställningen att förvaltningen producerar för kunder på en marknad.

Trots att svenska språket rymmer uttryck som »fängelsekund«, kan parallellen med kunders inköp på marknadsplatsen inte drivas för långt. I motsats till kunden betalar brukaren oftast inte personligen för den service hon får, i vart fall inte i full utsträckning, utan det får de skattskyldiga göra. I tänkandet kring brukare ingår därför också en deltagardemokratisk aspekt, som innebär att brukarna skall kunna göra sina stämmor hörda gentemot serviceproducenten och därmed påverka eller ta ansvar för tjänstens innehåll.

Brukarmodeller innehåller ett diskursivt, resonerande, diskuterande och influerande moment, som bör få komma till uttryck vid brukarorienterad utvärdering. Brukarmodeller hänger ihop med ett brukardemokratiskt synsätt. Om måluppfyllelsemodellen bygger på den representativa demokratins teori innehåller brukarmodellerna inslag av participativ demokratisk teori.

Det finns många varianter av brukarorienterad utvärdering. Brukarmodeller kan använda adressaternas förväntningar (önskemål) eller deras behov som normativt bedömningskriterium. Att utvärdera mot behov och att utvärdera mot förväntningar är principiellt skilda saker, eftersom de förra ofta anses vara något annat än utsagda önskemål. Brukarutvärdering kan utföras av brukarna själva som självvärdering men också beställas och utföras uppifrån. Låt mig resonera utifrån det senare fallet.

Vid brukarevaluering är utvärderarens första uppgift att lokalisera klienterna. Ibland finns det flera klientgrupper; vid utvärdering av akademisk utbildning kan vi dels tänka oss studenterna, dels studenternas framtida arbetsgivare som klienter. Eftersom samtliga troligen inte kan kontaktas, måste ett urval göras. Sedan pekar den klientdrivna modellen i olika riktningar. En möjlighet är att undersöka om den faktiska klientgruppen avviker från den avsedda, en annan att granska brukarnas bedömningar av insatsen. Brukarmodellerna säger ingenting om vilka verksamhetskomponenter som bör studeras. Det mesta är tillåtet. Klienterna kan bli ombedda att fälla omdömen om utfallet, slut prestationerna, tjänstens tillgänglighet, tjänstens kvalitet, tjänstens administration eller tjänstens kostnader.

Ett annat karakteristiskt drag är att effektfrågan, som är så central i måluppfyllelseutvärdering och bieffektsutvärdering, ibland inte reses i brukarorienterad utvärdering. Utvärderarna nöjer sig med att be klienterna värdera någon aspekt på den service de får.

Ett annat notabelt inslag i brukarmodellen är värdepluralism. Olika brukare bedömer servicen utifrån skilda värdegrunder. De kan vara djupt oeniga vid bedömningarna. Brukarmodellen tillåter en konfliktsyn på offentliga program och deras mottagande.

Det har tagit oerhört lång tid innan brukarna upptäcktes i offentlig förvaltning. Idéerna om den politikerstyrda förvaltningen och den parlamentariska styrkedjan - förvaltningen som politikernas exekutiv - har dominerat tänkandet. Numera används emellertid brukarmodeller i en rad utvärderingssammanhang, särskilt när det gäller offentlig serviceproduktion inom barnomsorg, sjukvård, äldreomsorg och biblioteksväsen eller program där adressaternas aktiva deltagande är nödvändigt för hela verksamheten. Vid universitet får de studerande lämna synpunkter på kurslitteraturen och undervisningen. De bedömer lärarnas prestationer utifrån exempelvis systematisering av lärostoffet, inställning till debatt och diskussion, förmåga att stimulera studenternas intresse, motivation och självständiga tänkande, omsorg om studenterna och entusiasm.

Här följer några exempel på marknadsundersökningar som görs eller har gjorts inom den statliga sektorn: Integrationsverket undersöker hur invandrare ser på den offentliga introduktionsverksamheten; Riksskatteverket vill veta hur de skattskyldiga ser på skatteväsendet, bemötandet,

personalens kompetens och utförandet av olika tjänster; Arbetsmarknadsstyrelsen frågar hur kunderna upplever arbetsförmedlingarnas service; Byggnadsstyrelsen vill ha sina kunders (hyresgästers) syn på hur verket sköter lokalförsörjningen; socialdepartementet spørjer ett stort antal föräldrar i hela landet hur de ser på barnomsorgen; Naturvårdsverket undersöker vilken trovärdighet miljömyndigheterna har hos allmänheten; och Patentverket tar reda på hur uppfinnarna upplever kötiderna.

För att ytterligare konkretisera kan vi ta biblioteksundersökningar som exempel. Besökare kan tillfrågas om bibliotekets tillgänglighet, bokbestånd, utrymmen och serviceanda: a) är öppettiderna godtagbara eller har biblioteket stängt när folk är lediga från arbetet och skulle kunna låna? b) är böckerna uppställda på ett begripligt sätt? c) finns den rätta blandningen av gammalt och nytt, barnböcker och vuxenlitteratur? d) är reservationstiderna för långa? e) tar det för lång tid innan återlämnade böcker utplaceras i hyllorna? f) är lokalerna tillräckligt rymliga? g) är ljusförhållandena tillfredsställande? h) blir låntagarna trevligt bemötta av personalen?

Brukarmodeller bör ses som ett komplement till de tidigare presenterade modellerna, eftersom de reser andra problem. Kravet att förvaltningen skall vara brukarorienterad även om brukarna inte betalar för servicen är sunt i vissa avseenden och vissa fall. Det förefaller väldigt rimligt att ta hänsyn till boklånarna när det gäller öppettider och personligt bemötande, för att ta två exempel. Generellt bör brukaråsikter få spela en större roll vid bedömningen av offentlig service än vid utvärdering av

myndighetsutövning eller myndighetsinformation.

Men brukarkrav kan inte i alla lägen tillåtas väga tyngre än principen att förvaltningen skall vara följsam mot de folkvalda och därmed ytterst medborgarna. Detta beror på att brukarmodeller har begränsningar. Enligt teorin om rationella val vill brukarna ha någonting för ingenting. De vill ha bättre service därför att de själva kommer i åtnjutande av den samtidigt som de i stor utsträckning slipper betala eftersom kostnaderna sprids ut över en stor mängd skattebetalare. Brukarna har fiskala illusioner. De har begränsad information om vad tjänsterna kostar, vilket gör att de underskattar kostnaderna samtidigt som de kräver mer tjänster. Brukarna kan således inte förväntas anlägga ett totalt kostnads-intäktsperspektiv, vilket det politiska systemet tvingas göra. Detta gör att även ekonomiska modeller har en legitim roll att spela i offentlig politik, ett spørsmål till vilket vi skall återkomma.

Intressentmodeller

Intressentmodellernas värdekriterier är farhågor och undringar hos interventionens berörda. Detta är något helt annat än att ta sin utgångspunkt i insatsens mål. Det är också något annat än att fråga brukarna. Skillnaden är att intressentmodeller i princip beaktar samtliga berörda under det att måluppfyllelsemodellen och brukarmodellerna tar hänsyn endast till vardera en grupp, nämligen beslutsfattare respektive brukare.

En uppställning av tänkbara intressenter i ett offentligt program presenteras i figur 5.

Figur 5.

Intressenter i en offentlig insats och dess utvärdering

INSATSINTRESSENTER:

Medborgarna:	Folket som väljer representanter till beslutande församlingar på olika nivåer i det politiska systemet
Högsta beslutsfattare:	Makthavare direkt ansvariga för att insatsen införts, fortsättes, trappas av eller avslutas
Politiska opponenter:	Insatsens politiska motståndare
Myndighetsledning:	Personer ansvariga för ledning och samordning av insatsens förvaltning
Handläggare centralt:	Personer som på central nivå direkt arbetar med insatsens förvaltning
Reg. o lokala myndigheter:	Personer som regionalt och lokalt är delaktiga i insatsens förvaltning
Gräsrotsbyråkrater:	Operativ fältpersonal
Enskilda mellanhänder:	Enskilda som på uppdrag av offentliga instanser eller på annat sätt deltar i insatsens förvaltning
Slutmottagare:	Insatsens målgrupp (adressater, brukare, klienter)
Andra slutmottagare:	Aktörer, kollektiva eller individuella, som (oavsiktligt) berörs av insatsen utan att vara avsedda slutmottagare

De tio grupper som listas i figur 5 representerar de viktigaste parter, som berörs av en insats i offentliga sektorn. Listan är komplicerad med tanke på en nationell reform, beslutad av högsta legislaturen. Självfallet skulle en förteckning över intressenter i kommunala verksamheter se annorlunda ut. Intressentmodellen är för övrigt öppen och ger inget klart svar på frågan vilka de berörda är.

Intressentutvärdering, som på engelska kallas »stakeholder approach«, kan ske på olika sätt. Intressenterna kan själva utgöra utvärderingslaget och själva exekvera utvärderingen. Vi kallar detta den svenska SOU-modellen. Utvärderingen kan också utföras av särskilt utsedda utvärderare, som kan vara fristående forskare, konsulter eller anställda vid myndigheten. Poängen är att alla intressenter skall beaktas men att själva utvärderingen utförs av särskilda

utvärderare. Detta är den nordamerikanska intressentmodellen. I fortsättningen skall jag resonera utifrån den nordamerikanska modellen.

Alla tio intressentgrupper i figur 5 är tänkbara bidragsgivare i en utvärdering. Även om utvärderaren i praktiken måste göra ett urval förutsätter den fortsatta diskussionen att samtliga intressenters förväntningar beaktas i intressentmodellerna. Intressentansatsen tolkas som en holistiskt, inriktad på hela spektrum av berörda.

Utvärderaren börjar med att ta reda på vilka som har intressen i insatsens tillkomst, genomförande och konsekvenser. Hon spårar upp insatsens utarbetare, finansierare och genomförare, ringar in målgruppen och de personer och institutioner, som drabbas av kostnader och sidokonsekvenser. Hon pekar ut människor, som utan att veta om det har intresse i interventionen.

Sedan de berörda blivit identifierade, kan den nordamerikanska intressentmodellen fungera på olika sätt. I *Effective Evaluation* föreslår Guba och Lincoln (se Vedung 1998:77) att alla berördas »farhågor« (concerns) och »frågor« (issues) får bilda utgångspunkten. En »farhåga« kan vara något som en part befarar kan medföra nackdelar eller något som den gärna vill få närmare belagt. »Virtually any claim, doubt, fear, anticipated difficulty, and the like expressed by anyone with a legitimate basis for making such a representation could be entertained as a concern.« En »fråga« å andra sidan är »any statement, or focus that allows for the presentation of different points of view; any proposition about which reasonable persons may disagree; or any point of contention«.

För att komma fram till en lämplig uppläggning av undersökningen utesluter intressentmodellens förespråkare inget tillvägagångssätt. De hyser emellertid förkärlek för kvalitativa metoder. Nyckelorden är direkt observation, samtal, dialog och växelverkande sökprocess. Utvärderaren måste observera och samtala med klienter, programadministratörer och andra intressenter. Vad hon ser eller får till svar inverkar på vad hon gör därefter. Efter en tid börjar hon kanske upptäcka både uppgivna och verkliga syften med verksamheten och de farhågor och frågor, som olika intressenter har rörande både styrningen och utvärderingen. Efterhand blir utvärderaren mer involverad och kan själv börja argumentera för vilka farhågor och frågor, som skall tas med i undersökningen. Den slutliga inriktningen bestäms av utvärderaren. För varje farhåga och varje fråga specificerar

utvärderaren vilken typ av information som behövs och hur den skall samlas in.

Det är typiskt för intressentmodellen att utvärderaren tänkes leta ganska länge efter problemen eller uppläggningsen. Måluppfyllelse t.ex. pekar på ett helt annat sätt än intressentmodellen ut vad utvärderingen skall handla om. Den nordamerikanska intressentmodellen är påfallande öppen. Tanken är att utvärderaren måste vara lyhörd för de berördas farhågor och undringar och låta dessa styra nästa steg i sökprocessen. Hon skall genom växelverkande kommunikation med de berörda successivt komma fram till vilka farhågor och frågor, som skall tas på allvar och utredas. Utvärderingens uppläggning bestäms gradvis. Intressentutvärdering är lyhördhetsutvärdering.

För den slutliga datafångsten kan alla strategier komma i fråga. Men för det mesta föredras observationsmetoder och intervjumetoder framför dokumentmetoder. I många fall sätter intressentutvärderarna mest värde på direkt observation genom platsbesök och personliga iakttagelser. Djupintervjuer är en annan favoriserad metod för materialinsamling. Poängen är att programmet studeras i sitt naturliga sammanhang utan kontrollgrupp. Det skall vara fallstudier. Resultatrapportering, slutligen, kan variera från berörd till berörd. Nyckelordet tycks vara »porträttering«, som kan vara skriftlig men lika gärna muntlig eller visuell. Det sammanhängande helhetsintryck som förmedlas av ett porträtt är viktigt. Utvärderarens egna bedömningar spelar en mindre roll eftersom både val av värdegrunder och värderandet skall skötas av intressenterna. Normalt måste det bli

fråga om flera värdegrunder och många olika bedömningar. Intressentmodellen blir därmed i praktiken utpräglat pluralistisk. Till sist avgörs om resultaten skall rapporteras i en monografi, vilket inte alls är nödvändigt.

Intressentmodellen har många förtjänster. Demokratiska argument, kunskapsargument, användningsskäl och målhante- ringsargument talar till dess fördel.

De demokratiska argumenten för intres- sentmodellen utgår från participativa och deliberativa synsätt. Demokrati innebär inte bara att medborgarna i val röstar fram eliter, som skall fatta beslut på medbor- garnas vägnar. Medborgarna bör också få delta i offentligt beslutsfattande mellan valen. Dessutom är deliberation och debatt viktiga värden i en demokrati ty härigenom kan människor få klarhet i sina preferenser. I högre grad än måluppfyllelsemodellen och bieffektsmodellen tillgodoser intres- sentmodellerna dessa participativa och deliberativa värden

Enligt kunskapsargumentet vore det dår- aktigt av utvärderaren att avskärma sig från de insikter i reformen, som dess berörda besitter. Intressenterna hyser dessutom ofta farhågor om bieffekter, implemente- ringsbarriärer och samspel med andra pro- gram, som kan ge ytterst värdefulla uppslag för fortsatt undersökning. Det är därför lätt att instämma i rekommendationen att nästan varje utvärdering bör inledas med en aktörsbestämning och intervjurundor.

Detta kunskapsskäl är emellertid inget huvudargument för intressentmodellen. Ett sådant är däremot påståendet att den ökar resultatens användbarhet. I den amerikanska debatten kring utvärdering

diskuteras intressentmodellen helt och hållet i detta avnämarperspektiv. Resultat som tas fram i måluppfyllelseutvärderingar och bieffektsgranskningar kommer sällan till användning. Rapporterna staplas på var- andra i beslutsfattarnas bokhyllor, olästa och bortglömda. Även om viljan är god så nyttjar parlamentariker, generaldirektörer, byråchefer och lokala politiker inte det framtagna informationsunderlaget till att fatta mer genomtänkta beslut. Utvärdering lider, så kan kritiken formuleras, av ett elak- artat skrivbordslådesyndrom.

Detta till synes grovt irrationella hand- lande har förbryllat utvärderingsforskarna. Varför betar sig de styrande så oklokt? Huvudanledningen sägs vara att utredarna arbetar i suverän isolering, utan samfärdsel med avnämarna. Därför tenderar de att penetrera problem, som intressenterna inte känner något behov av att få utreda. Intressentmodellen ökar sannolikheten för att frågor tas upp som berörda grupper finner relevanta. Därmed kommer resulta- ten lättare till verklig nytta.

Intressentmodellen kan också ses som en strategi att hantera notoriskt oklara och osammanvägbara mål. Intressentmodellen erbjuder en praktisk lösning på bekymret med många mål samtidigt. Intressent- modellen hör ihop med en intressegrupps- demokratisk syn på offentlig verksamhet.

Målen för offentliga verksamheter är nästan undantagslöst diffusa. Framför allt kännetecknas de av att många olika mål skall balanseras mot varandra utan att någon på förhand kan avgöra hur detta skall gå till. Avvägningen kan därför inte göras på vetenskaplig väg utan måste ses som en politisk uppgift. Därför gäller det att låta

de många intressenternas många mål på många områden bilda utgångspunkter för utvärderingen. Att intressenterna då visar sig djupt oense måste anses normalt och naturligt i en fri stat. Intressentmodellerna är till sin natur politiska och pluralistiska.

Det finns också uppenbara nackdelar med intressentmodellerna som substansmodeller. De är kostsamma och omständliga eftersom varje intressent måste kontaktas och utfrågas.

Intressentmodellerna är oskarpa (Karls-son 1995). De har inget entydigt svar på frågan vilka som egentligen är berörda av ett offentligt program. De skulle behöva kompletteras med en politisk teori om urval av berörda. Intressentmodellerna likställer vidare alla bemälda, hur de nu än valts ut. Men i ett konstitutionellt politiskt system måste väl ändå de folkvalda politikerna - de överordnade beslutsfattarna - vara viktigare än experter på sakområdet eller oavsedda slutmottagare? Intressentmodellen saknar regler för hur dylika prioriteringar bör gå till. Den är ett uttryck för kålsuperi. Risker finns att de bäst organiserade och mest engagerade intressenterna konsulteras medan vaga och mer svårdefinierbara intressenter lämnas utanför. Även på denna punkt skulle intressentmodellerna behöva baseras på en mer problematiserad politisk filosofi. Intressentmodellerna och deras intressegruppsdemokratiska grundval kan komma i konflikt med måluppfyllelsemodellens synsätt, som hänger ihop med föreställningen om representativ demokrati.

Samtidigt skall konstateras att intressentmodellen ger utvärderaren möjlighet att gå förbi de stora korporativa, byråkratiska och partipolitiska etablissemangen

på nationell nivå och koncentrera uppmärksamheten på nätverken i de lokala sammanhangen. Detta är en synpunkt som särskilt framhållits av den amerikanska utvärderingsforskaren Robert Stake.

Intressentmodellen reser frågetecken kring vetenskaplig objektivitet i resultatredovisningen. Hur skall utvärderaren handla om hennes ståndpunkt går stick i stäv mot mäktiga intressenters? Fakta kan bli en fråga om förhandlingsstyrka snarare än rationell diskussion och empiriska belägg. Det finns också risk att utvärderingen sker mot utvärderarens egna mål på ett icke redovisat sätt.

Intressentmodellen är alltså suddig i konturerna och relativt kontroversiell. Trots detta talar demokratiska argument, användningsargument och mångmålsargument till dess fördel. Ett annat starkt skäl för den är kunskapsargumentet: att den kan användas som sökstrategi för att komma fram till en lämplig uppläggning i utvärderingens inledningsskede, när det gäller att i grova drag få grepp om reformens innebörd, genomförande och utfall.

Professionella modeller - kollegetbedömning

Särskilt i utvärdering av högre utbildning och forskning har det utbildats ett ganska unikt tillvägagångssätt - kollegetbedömning - som skiljer sig kraftigt från särskilt mål- och bieffektmodellerna men även de ekonomiska modellerna som vi senare skall ta upp. Kollegetbedömning innebär att medlemmar av en profession utses att utvärdera professionsmedlemmarnas

arbete och verksamhet utifrån yrkeskårens egna kvalitetskriterier. Advokater utvärderar advokater, kirurger kirurger, ingenjörer ingenjörer och professorer professorer. Utvärderingen utförs av kolleger, som ju per definition är professionella likar. Även professionell utvärdering kan utföras på flera sätt. Självvärdering är en form. Den professionelle utvärderar då sina egna insatser eller professionella på en institution utvärderar tillsammans institutionens insatser. En annan form är kollegial utvärdering genom utomstående likar. Idealt sett bör likarna i detta fall vara lite bättre inom sitt område än de kolleger som skall utvärderas. De utvärderades förtroende för utvärderarna är viktigt. Ofta kombineras den interna, självvärderande typen av professionell utvärdering med den externa.

Den mest bekanta kollegemodellen är s.k. peer review, som används vid utvärdering av vetenskap. Tanken är att renomméerade forskare på ett område får i uppgift att utvärdera hur ett forskningsprojekt, ett forskningsprogram, en universitetsinstitution eller en hel disciplin i ett land står sig kvalitetsmässigt och relevansmässigt. Vetenskapsmännen och deras insatser utvärderas av kolleger (för exempel, se Allardt m.fl. 1987 sociologi, Danielsen m.fl. 1988 historia, Öhman & Öhngren 1991 psykologi och Engwall 1992 ekonomi).

Termen »peer review« har på senare tid vanligen syftat på en procedur för att välja ut bidrag till vetenskapliga tidskrifter. Till en tidskriftsredaktion sänds anonymt artiklar som redaktionen vidarebefordrar till någon kollega för peer review i syfte att utvärdera om artiklarna håller sådan kvalitet att de förtjänar att publiceras. Peer review

används också för att ge råd till forskningsfonder om vilka forskningsprojekt som skall ges stöd. Ansökningarna skickas ut till en grupp kolleger för granskning. Kollegial bedömning innebär i detta fall att forskningen bedöms i förväg, ex ante. Det är forskningsplanen som värderas, inte de uppnådda resultaten. Peer review har vidare använts för att utreda och bedöma brott mot god forskningsetik och gott forskaruppträdande. Här avser jag med kollegial bedömning en procedur för att i efterhand utreda och bedöma genomförande, prestationer och utfall i offentlig verksamhet.

Kollegebedömning pekar inte i detalj ut de kriterier och standarder som bör vara vägledande vid granskningen eller sakområden som bör granskas. Frågan om värdekriterier och standarder tillåts variera från profession till profession.

Vad har egentligen kollegebedömning med utvärdering i offentliga sektorn att göra, kanske någon undrar. Svaret ligger förborgat i principen om den professionsstyrda förvaltningen. På vissa områden i den offentliga sektorn är målsättningar och kvalitetsnormer så invecklade och föränderliga att politikerna funnit för gott att överlämna deras närmare utformning till vetenskapligt utbildade professioner. Professionens egna medlemmar får genom kritisk debatt och diskussion utveckla bedömningsgrunder och kvalitetskriterier. Exempel på professionella som denna ställning är arkitekter, domare, professorer, lärare, läkare, veterinärer och ingenjörer. Detta gör det naturligt att också anförtro utvärderandet åt professionerna.

Utvärdering av sektorsforskning och

universitetsvetenskap kan ses som bedömning av statens forskningspolitik. Forskning ombesörjs inte av en webersk byråkrati, ej heller av korporationer eller politiker. Den sköts av professionella vetenskapsidkare. Detta får stundom till följd att även forskningsutvärdering sker i kollegiala former. Vid forskningsutvärdering arbetar experterna oftast på någons uppdrag, t.ex. ett forskningsråds. Forskningsrådet skriver direktiv, som i allmänna termer innehåller de problem och frågor man vill ha belysta. De utvärderingar som startades under 1987 av Byggforskningsrådets vetenskapliga nämnd uppmanades beakta

- a relevansen av problemval och analysuppläggning
- b analysmetodernas tillämplighet
- c hållbarhet hos genomförda resonemang och konklusioner
- d arbetets inplacering i förhållande till diskursen på området
- e möjligheterna att praktiskt tillämpa forskningsresultaten
- f värdet av den genom projektet eventuellt uppbyggda forskningsmiljön
- g överensstämmelse mellan ursprungliga intentioner och slutligt resultat.

Därefter kontaktas lämpliga sakkunniga, vilka bör ha högre kompetens inom området än kollegerna som skall utvärderas. Experterna måste vara oberoende av de utvärderade; de bör inte ha samarbetat med varandra inom det aktuella sakområdet. En viktig skillnad går mellan kollegebedömningar där de utvärderade får föreslå och på förhand godkänna urvalet av kolleger och sådana där detta inte sker.

Sedan expertgruppen - kollegiet - utsetts organiseras det fortsatta arbetet. Oftast ombes de utvärderade sända in relevant forskningsmaterial, så att de sakkunniga kan läsa in sig. Detta tyder på att någon form av självvärdering ligger till grund för den kommande kollegebedömningen. Därefter besöks varje forskare och forskargrupp för presentationer och samtal. Kollegegruppen diskuterar kanske vilka kvalitetsgrunder som skall appliceras på forskningsresultaten eller forskningsprocesserna. Efter en tid skrivs en preliminär rapport. Den cirkuleras till berörda forskare som kan beredas tillfälle att yttra sig. Därefter komponeras den värderande slutprodukten. Men även om utvärderingskollegiet interagerar med de forskare som utvärderas så är det kollegiet som är ansvarigt för slutrapporten och därmed bedömningen. Hela proceduren brukar vid Byggforskningsrådetsvetenskapliga nämnda ta ca 18 månader.

Kollegebedömningens styrka ligger i val och tillämpning av värdekriterier. För detta bör de professionella själva vara mest lämpade. Det är svårt för politiker att sätta kvalitetsmål på forskning. Dessutom torde måttstockarna liksom forskningen utvecklas över tid i takt med kunskapsmassans tillväxt. Utvärderingar bör emellertid inte begränsa sig till problemet om forskningen motsvarar krav på kvalitet i resultat eller i process utan även komma med förklaringar av varför exempelvis resultat inte motsvarar eller inte uppfyller krav på kvalitet. Otvivelaktigt torde kolleger t.ex. från ett annat land vara skickade att ställa upp lämpliga kvalitetskriterier och utföra bedömningar med ledning av dessa. Mer

tvetsamt är om de också är mer kompetenta än andra att leta upp förklaringar. I den mån framgång och misslyckande beror på verksamhetens organisering borde kollegerna kanske kompletteras med experter på forskningsorganisation.

Man skulle på lite tydligare sätt kunna kombinera kollegebedömning med självutvärdering. De utvärderade skulle då först ombes bedöma värdet av sin egen produktion. På grundval av dessa bedömningar men naturligtvis även från egna utgångspunkter skulle sedan de sakkunniga göra sin betygsättning. Innan de slutliga omdömena slås fast skulle de utvärderade beredas tillfälle att yttra sig över de sakkunnigas resultat.

Ekonomiska modeller - produktivitet

Gemensamt för alla substansmodeller är att de inte tar in kostnader i kalkylen. Samtidigt som de undersöker kriterieuppfyllelse, oavsett om det är insatsmål eller brukares och intressenters önskemål och farhågor, så beaktar de inte uppoftningar för att uppnå värdenormerna. Hänsynstagande till kostnader å andra sidan ingår alltid i de ekonomiska utvärderingsmodellerna. Den rena kostnadsmodellen som enbart beaktar kostnader skall inte beröras här. Sökarljuset skall riktas mot produktivitet och effektivitet, två ekonomiska modeller som beaktar såväl substans som kostnader.

Produktivitet definieras som prestationer dividerade med kostnader. Den ofta citerade svenska programbudgetutredningen formulerade denna tanke så här:

Med produktivitet avses alltså förhållandet mellan prestationer och uppoftningar inom en myndighet eller en del därav, dvs. produktiviteten utgör kvoten mellan ett mått på mängden utförda prestationer och ett mått på använda produktionsresurser (output dividerad med input).

Produktivitet kan exakt uttryckas med den matematiska formeln i figur 6.

Figur 6.
Begreppet produktivitet

$$\text{Produktivitet} = \frac{\text{prestationer (output)}}{\text{insatser (input)}}$$

Kvotformeln i figur 6 kan självklart operationaliseras på olika sätt. Låt mig ta exempel från en finsk biblioteksstudie (Vedung 1998:81). Vid analysen av produktivitetens utvecklingen i finska kommunbibliotek användes följande mått:

$$\frac{\text{antalet boklån}}{\text{boklånens kostn. i mark}} = \text{kostnadsprod.}$$

Ett alternativ till kostnadsproduktivitet kunde ha varit arbetsproduktivitet, vilket kan illustreras med uttrycket:

$$\frac{\text{antalet boklån}}{\text{antal arbetade timmar}} = \text{arbetsprod.}$$

Skillnaden är att kostnaderna i det förra fallet anges i pengar, i det senare fallet i antal arbetstimmar, dvs. i fysiska termer. Det bör understrykas, att kostnader kan

kalkyleras på båda sätten vid produktivitetanalys. Andra valmöjligheter som kan övervägas är antal böcker i biblioteket: kostnader, antal invånare i kommunen: kostnader och antal låntagare: kostnader. Tidsenheten kan vara budgetår, kalenderår, eller eventuellt också kalendermånad.

Semantiskt och retoriskt intressant med produktivitet är att den också betecknas som »inre effektivitet«. Detta återspeglar förhållandet att produktivitet är ett internt mått. Produktivitet fångar ett förhållande inom organisationen, men beaktar inte verksamhetens effekter på omgivningen eller omgivningens bedömning av effekterna.

För att säga något klokt om en myndighets produktivitet behövs referensfall. Ett biblioteks produktivitet måste jämföras med något för att ge ett intryck av hur god produktiviteten verkligen är. Detta kan tekniskt uttryckas så att utvärderarna måste ha standarder på produktivitetskriteriet för att ange vad hög och låg produktivitet verkligen innebär. Åtskilliga standarder brukas: jämförelse med tidigare prestationer, med liknande institutioner i landet, liknande institutioner i andra länder, brukarförväntningar och intressentmål.

Låt oss återvända till biblioteken för att illustrera problemen med att hitta giltiga indikatorer på slutprestationer. Är verkligen antalet utlånade böcker ett relevant och heltäckande prestationsmått? Relevant är det nog. Att låna ut böcker är rimligen bibliotekets viktigaste uppgift. Men heltäckande är måttet inte. En finsk utredning konkluderar att endast 30-50 procent av bibliotekens klienter är boklånare. De övriga besöker biblioteket för att läsa

dagstidningar och tidskrifter men detta avsätter inga spår i utlåningsstatistiken. De frekventerar referensavdelningen för att slå i ordböcker och encyklopedier eller musikavdelningen för att lyssna. Biblioteken svarar också för informationsservice. I den finska undersökningen diskuterades självfallet problemen innan forskaren beslöt sig för att stanna för boklån. Att det är en begränsning går dock inte att komma ifrån.

Den finska studien skulle naturligtvis ha kunnat använda alla dessa delmätt och försökt aggregera ihop dem till ett mått på den totala produktiviteten. Då skulle undersökarna emellertid ha tvingats ta ställning till besvärliga vikttningsproblem.

En annan svårighet vid produktivitetstutvärdering är att kvaliteter ofta förbises. Böcker är av varierande kvalitet. Hur kan detta mätas i en produktivitetstutvärdering?

Produktivitet som mått på offentliga sektorns varor och tjänster kan också kritiseras från principiella utgångspunkter. Viktigast är kritiken att det mäter ett internt förhållande inom organisationen och inte det vi helst vill veta, nämligen de resultat den offentliga interventionen gett upphov till hos slutmottagarna eller ute i samhället, vad detta resultat är värt samt om förtjänsterna uppväger uppoffringarna. I biblioteksfallet är det inte boklånen i och för sig som är intressanta; folk kan ju låna böcker, lägga dem på hög i hemmet och efter påstötningar lämna dem tillbaka olästa. Viktigare är läsandet av de lånade böckerna. Avgörande är emellertid läsarnas upplevelser av läsandet. Upplevelserna kan delas in i bildning och förströelse. Det som

den kostnadsmedvetne, finkulturellt inriktade utvärderaren kanske vill åt är

värdet av bildning och
förströelse genom lånade böcker

kostnader för att förvärva denna bildning

Därmed är hon inne i ett effektivitetsresonemang.

Förespråkarna för produktivetsmätningar kan inte komma ifrån att produktivitet mäter slutprestationer och att detta inte är någon idealisk mätpunkt för att bedöma den offentliga sektorns verksamheter. Den offentliga institutionen kan ju göra fel saker, dvs. slutprestationerna kanske inte leder till det önskade utfallet.

Ekonomiska modeller - effektivitet

De andra ekonomiska modeller som används vid utvärdering av offentlig verksamhet är effektivitetsmodeller. Effektivitet (»efficiency«) mäts i utfallsledet och kan registreras på två sätt: i kostnads-intäktsanalys (K-I-effektivitet, »cost-bene-

fit efficiency«) och i kostnads-effektivitetsanalys eller synonymt kostnads-nyttoanalys (»cost-utility«).

Om effektivitet mäts i en kostnads-intäktsanalys ställs insatsens intäkt i kronor i förhållande till kostnaden i kronor för intäktens förvärvande. K-I-effektivitet kan anges som kvoten av det monetariserade värdet av de resultat som verksamheten producerat och de monetariserade kostnaderna. Om effektivitet däremot likställs med vad som mäts i kostnads-effektivitetsanalyser, ställs insatsens fysiska avkastning i förhållande till kostnaden för avkastningens förvärvande. Vid kostnads-effektivitetsanalys beaktas med andra ord de monetariserade kostnaderna precis som vid kostnads-intäktsanalys medan däremot värdet av effekterna anges endast i fysiska termer. De två effektivitetsmåten kan uttryckas med formlerna i figur 7.

Det vi i formlerna menar med »insatsens effekter« är sådant, som just insatsen gett upphov till. Det vi är ute efter är de effekter, som just insatsen åstadkommit.

Som alla andra utvärderingsmodeller ger produktivets- och effektivitetsmät-

Figur 7.

Effektivitet mätt genom kostnads-intäktsanalys och kostnads-effektivitetsanalys

$$\text{Effektivitet (kostnads-intäktsanalys)} = \frac{\text{värdet av insatsens effekter (i kr, euro)}}{\text{insatsens kostnader (i kr, euro)}}$$

$$\text{Effektivitet (kostnads-effektivitets-
analys, kostnads-nyttoanalys)} = \frac{\text{insatsens effekter i fysiska termer}}{\text{insatsens kostnader (i kr, euro)}}$$

ningar endast delperspektiv. De bortser från andra krav, som normalt ställs på offentlig verksamhet, exempelvis rättssäkerhet, representativitet, deltagandevärden och offentlighetsvärden. Hur avvägningen mellan dessa värden och produktivitet/effektivitet skall ske, kan ingen utvärdering med vetenskapliga anspråk ge besked om. Avvägningen kan endast ske med hjälp av offentlig debatt, opinionsbildning, kompromisser, majoritetsbeslut och maktspråk, dvs. politik.

Slutord om val av utvärderingsmodell

Den genomgång som här har gjorts har skett utifrån ett smalt perspektiv på utvärdering. Vi har valt att organisera framställning efter värdekriterier som används vid utvärdering. Därmed har vi i stort sett förbigått en mängd andra problem, som t.ex. om utvärdering skall ske som akademisk forskning av utbildade vetenskapsidkare eller om den skall genomföras av berörda professionella, intressenter eller brukare. Vi har i stort sett bara snuddat vid om utvärdering skall ske som experiment eller om den skall ske i form av fallstudier. Vi har inte heller systematiskt behandlat om utvärdering skall vara demokratisk i bemärkelsen participativ och deliberativ och ske i dialog med de berörda.

Genomgången har visat att den totala uppslutning, som i 1960-talets nordamerikanska och europeiska utvärderingsvåg rådde bakom måluppfyllelsemodellen och därmed den representativa demokratins princip, har brutits upp och ersatts av en

situation där många modeller konkurrerar. Till en del beror detta på att läran om offentlig politik utvecklats mot större sofistikerad, t.ex. vad gäller bieffekter av offentlig planering, vilket lett till att måluppfyllelsemodellen övergått i bieffektmodellen. Vidare har under 1990-talet den offentliga sektorns kostnader kommit i förgrunden, vilket resulterat i en renässans för ekonomiska utvärderingsmodeller. Andra intressenter än folkets valda representanter har med ökad styrka börjat resa krav på deltagande i och inflytande över förvaltningsprocesser, vilket lett till att brukarmodeller och intressentmodeller kommit till användning. Den ökade användningen av professionella modeller är lite speciellt eftersom krav på detta kommer uppifrån och inte från professionerna själva. Allt i allt har utvärdering utvecklats från uniformitet till pluralism.

Det finns en stark tendens i nutida utvärderingslitteratur att helt avfärda måluppfyllelsemodellen. Från en utgångspunkt är denna totala nedvärdering oberättigad. Ur den representativa demokratins synvinkel är nämligen måluppfyllelsemodellen helt central, eftersom den grundas på idén om den parlamentariska styrkedjan. Mål satta på hög politisk nivå (eller på förvaltningsnivå men på delegation från den politiska nivån) är inte vilka mål som helst. De har framkommit genom konstitutionella procedurer, som gjort att de institutionaliserats som representativa församlingars mål. Medborgare och valda ombud har legitim anledning ta reda på huruvida dylika mål har uppnåtts på fältet. I annat fall kan det representativa demokratiska styrelseskicket inte fungera.

Det representativdemokratiska argumentet till försvar för måluppfyllelsemodellen kan emellertid inte göra något åt det faktum, att modellen råkar in i svårigheter särskilt med målkataloger. Dess största analysmässiga nackdel är dock att den ej förmår beakta sidoeffekter. Av denna anledning föredrar jag bieffektsutvärdering framför måluppfyllelseutvärdering. Bieffektsutvärdering kan ses som måluppfyllelseutvärdering utbyggd med sidoblickar på sidokonsekvenser. Bieffektsmodellen hämtar också sin styrka ur den representativa demokratis princip.

Ekonomiska modeller har från slutet av 1980-talet vunnit terräng, bl.a. därför att de offentliga finanserna försämrats. Även dessa modeller hänger samman med den parlamentariska styrkedjan och den representativa demokratisynen. Medborgarna och deras valda representanter behöver inte bara kunskap om sakinriktade resultat av offentliga åtgärder utan även om kostnader för att uppnå dessa resultat. Häri ligger också de ekonomiska modellernas styrka. En annan styrka är att de förmår reducera en verksamhets värde till enkla, lätt begripliga tal. Vi bör emellertid hålla i minnet att ekonomiska modeller i likhet med andra modeller ger endast ett partiellt perspektiv. Faran med ekonomiska modeller är att beslutsfattare fascinerats av deras matematiska precision och felaktigt tror att de ger allomfattande, slutgiltiga svar.

Om måluppfyllelsemodellen, bieffektsmodellen och ekonomiska modeller hämtar sin grundläggande legitimitet ur en representativ demokratisyn synes intressentmodeller och brukarmodeller bygga på en deliberativ och participativ demokratisyn.

Medborgarna skall inte delta i offentliga sektorns angelägenheter enbart via de allmänna valen. De bör delta även i sin kapacitet av slutmottagare eller annan intressent. Intressentmodeller och brukarmodeller skall kännetecknas av öppen debatt genom överbevisning och saklighet. De tänkes upprätta arenor för överläggningar och informerade samtal där folkvettet är direkt involverat. På det sättet kan de komplettera det representativa demokratiska systemet. Brukar- och intressentmodeller kan komma i konflikt med de modeller som drivs av föreställningar om den representativa demokratin. I denna framställning har de senare modellerna getts ett principiellt försteg. Med hänvisning till den parlamentariska styrkedjans primat för den offentliga sektorn kan brukar- och intressentmodeller endast komplettera men inte ersätta ansatser, som tar sin utgångspunkt i program mål och offentliga kostnader.

Styrkan i kollegetbedömning och självvärdering är att de kan fånga in och bedöma kvaliteter. Detta är oundgängligt inom områden som kan kännetecknas av ytterst specialiserad kunskap. Paradigmfallet är naturligtvis akademisk forskning. Även dessa professionella modeller måste avvägas mot representantmodellerna.

Faran med alla utvärderingsmodeller är att de används alltför okritiskt och att beslutsfattare felaktigt tror att en enda modell kan lämna slutliga svar på centrala frågor. Därför bör det hållas i minnet att varje modell endast ger perspektiv och partiella svar. Av denna anledning rekommenderas att flera modeller får komma till användning. Pluralism är den hållning, som för närvarande ter sig mest rimlig.

Litteratur

- Allardt, Erik & Sverre Lysgaard & Aage Bøttger Sørensen. (1987), *Sociologin i Sverige: Vetenskap miljö och organisation*, Uppsala: Swedish Science Press.
- Danielsen, Rolf. (1988) *Historia i belysning: Sex perspektiv på svensk historisks forskning*, Uppsala: Swedish Science Press.
- Engwall, Lars. Red. (1992) *Economics in Sweden: An Evaluation of Swedish Research in Economics*. London: Routledge.
- Karlsson, Ove. (1995) *Att utvärdera – mot vad? Om kriterieproblemet vid intressentutvärdering*, Stockholm: HLS Förlag.
- Vedung, Evert. (1997) *Public Policy and Program Evaluation*, New Brunswick, N-J.: Transaction Publishers.
- Vedung, Evert. (1998) *Utvärdering i politik och förvaltning*, Lund: Studentlitteratur.
- Öhman, Arne & Bo Öhngren. Red. (1991) *Two Faces of Swedish Psychology: 1. Frontiers in Perception and Cognition: An Evaluation of Swedish Research in Cognitive Psychology*, Uppsala: Swedish Science Press.

Summary Evaluation models

The present classification, based on value criteria used in evaluation, has demonstrated that the total agreement which once existed about the appropriateness of the goal-attainment model has been replaced by a situation where several models compete. Evaluation has evolved from uniformity to pluralism. Yet, every model provides partial perspectives only. Combinations of several models are recommended.

There is a strong tendency in contemporary evaluation literature to recommend stakeholder and client-oriented evaluation and debate particularly the goal-attainment model. In many areas this is entirely justified. On one account, I take exception to this tendency. From a representative

democracy point of view the goal-attainment model and particularly the side-effects model are important, since they are based on the conception of the parliamentary chain of influence. Policy goals, set by parliaments and governments, and program and project goals set by lower-level administrator on delegation from parliaments and governments are not just any goals. Established through a constitutionally determined procedure, they are institutionalized as the collective goals of the state. Principals of various kinds have legitimate reasons to ascertain whether goals set by their agents have in fact materialized in the field. Otherwise, they cannot function as principals in the representative system of government.

The major drawback with the goal-achievement model is its lack of focus on side-effects. For this reason, I prefer side-effects evaluation to goal-achievement evaluation.

Client-centered evaluation has a role to play, particularly as far as government services are concerned. The strength of peer review and other professional models lies in their capacity to capture and judge qualities. Their paradigm area of application is academic research and higher education. The stakeholder model provides the broadest view possible of government interventions, promises to take all involved into consideration, and may give legitimacy to programs.

Economic models will stay with us fore-

ver in public policymaking. It must be kept in mind, however, that like other designs they provide partial perspectives only. The danger with economic models is that decision-makers are fascinated by their mathematical precision and wrongly believe that they provide comprehensive, final answers. While the goal-attainment model, the side-effects model and economic models get their fundamental justification from the theory of representative democracy, stakeholder models and client-oriented models relate to deliberative and participative democracy. One of their rationales is to establish deliberative arenas to supplement representative democracy with some citizen participation and deliberation.