



Det här verket har digitaliserats vid Göteborgs universitetsbibliotek och är fritt att använda. Alla tryckta texter är OCR-tolkade till maskinläsbar text. Det betyder att du kan söka och kopiera texten från dokumentet. Vissa äldre dokument med dåligt tryck kan vara svåra att OCR-tolka korrekt vilket medför att den OCR-tolkade texten kan innehålla fel och därför bör man visuellt jämföra med verkets bilder för att avgöra vad som är riktigt.

This work has been digitized at Gothenburg University Library and is free to use. All printed texts have been OCR-processed and converted to machine readable text. This means that you can search and copy text from the document. Some early printed books are hard to OCR-process correctly and the text may contain errors, so one should always visually compare it with the images to determine what is correct.



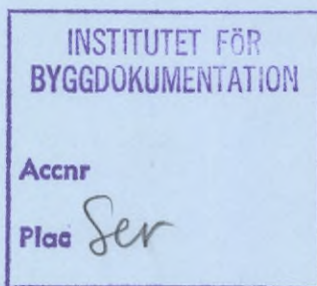
Rapport

R46:1987

**Reformförsök och
planeringsansträngningar**

Utvärdering av planering och politik

Anders Gullberg



Byggforskningsrådet

R46:1987

REFORMFÖRSÖK OCH PLANERINGSANSTRÄNGNINGAR

Utvärdering av planering och politik

Anders Gullberg

Denna rapport hänför sig till forskningsanslag 850849-5
från Statens råd för byggnadsforskning till Hb SALG,
Stockholm.

REFERAT

Reformförsök och planeringsansträngningar - vad förmår egentligen politiken?

Försök att med hjälp av vetenskapliga metoder bedöma resultat av offentlig verksamhet har nu pågått under många år. Syftet med s.k. utvärdering har ibland varit att förbättra verksamheten men inte mindre sällan att finna försvar för eller angreppspunkter mot densamma.

Erfarenheterna av denna slags erfarenhetsåterföring är nu så omfattande att de väl lämpar sig för utvärdering. Hur fungerar utvärdering? Vilka är problemen? Hur kan de avhjälpas?

Något av en utvärderingens vagga kan återfinnas i det amerikanska kriget mot fattigdomen. Erfarenheterna av dessa utvärderingar undersöks i rapporten. Utvärderingsmodets mottagande i Sverige redovisas också och några svenska utvärderingsförsök granskas. Slutsatsen blir att utvärderingsgenren knappast utvecklat en vana att använda goda och tillfredsställande metoder. En viss slagsida åt intressentanpassning, även raffinerad sådan, kan registreras.

Föra att mildra effekterna av tillräcklig metodmedvetenhet och anpasslighet förordas inträttandet av fristående utvärderingsorganisationer och en satsning på teori- och metodutveckling efter riktlinjer som skisseras i rapporten.

I Byggeforskningsrådets rapportserie redovisar forskaren sitt anslagsprojekt. Publiceringen innebär inte att rådet tagit ställning till åsikter, slutsatser och resultat.

R46:1987

ISBN 91-540-4734-X

Statens råd för byggnadsforskning, Stockholm

Svenskt Tryck Stockholm 1987

INNEHÅLLSFÖRTECKNING

Upptakt	5
Inledning	6
Nytt och gammalt inom utvärdering - USA som föregångsland	9
Akademiska discipliner i utvärdering och utvärderingens genrerikedom	19
Utvärderingsforskningen som politikens redskap och som sanningens apostel - erfarenheter från den amerikanska fronten	31
Utvärdering i Sverige - inhemska modell och amerikainport	43
Den amerikanska evalueringsforskningen presenteras i Sverige	49
Två halvpermanenta expertgruppers försök att bedöma effekter	62
Metodresonemang	65
Att utvärdera arbetsmarknadspolitik (SOU 1974:29)	68
Arbetsmarknadspolitik i förändring (SOU 1978:60)	76
Att främja regional utveckling (SOU 1978:46,47)	79
Hur långt har den offentliga effektmätningen nått?	82
Dubbelstrategi för en vetenskaplig utvärdering	85
Institutionella arrangemang	85
Metodologiska överväganden	88
Utvärdering i praktiken	92
Efterskrift och förslag	93
Appendix	95
Referenser	97

UPPTAKT

Att försöka lära av erfarenhet, egen och andras, är en tilltalande tanke. Att låta bli framstår som huvudlöst. Därför säljer sig idén om utvärdering av reformer, offentlig verksamhet och många andra aktiviteter nästan av sig själv. Ambitionen att systematiskt bedöma utförda handlingar, identifiera misstag och finna de lyckade greppen för att nästa gång göra bättre möter knappast gensagor. Dessa försök till erfarenhetsåterföring i utvärderings-skepnad har i åratal varit så frekventa att vi nu kan företa en utvärdering av vad dessa utvärderingar fört med sig i form av förbättrad kunskap och/eller politik. Det är till och med möjligt att med ledning av dessa erfarenheter formulera rekommendationer och förslag om hur utvärderingsforskningen skall reformeras för att bättre kunna fylla sina uppgifter.

I det följande kommer i första hand två frågor och deras inbördes sammanhang att dryftas: utvärderingens institutionella inplacering och utvärderingsstudiernas metodologiska problem. Den här presenterade texten har, med avbrott, utarbetats under perioden 1978 till 1986. En sista bearbetning inför publiceringen har möjliggjorts genom stöd från Statens råd för byggnadsforskning.

Stockholm i juli 1986

Anders Gullberg

1Se efterskriften för närmare redogörelse av omständigheterna i samband med rapportens tillkomst.

INLEDNING

Inom det samhällsvetenskapliga forskningsområdet finns en typ av frågeställningar som är lika centrala som de är sorgligt försummade. Sociala förändringar och orsakerna till samhällelig omvandling är förvisso inga okända teman inom socialvetenskaperna. Däremot röner de speciella - föregivet medvetna - omvandlings- och/eller bevarandekrafter som politiska program, reformer och planeringsansträngningar utgör knappast någon större uppmärksamhet, i varje fall inte vad gäller deras effekter¹. Detta är så mycket märkligare som den praktiska nytta som allt mera intensivt avkrävs forskning inom envart område på ett utmärkt sätt förefaller kunna tillgodoses just genom att reformer och politiska program studeras vad beträffar deras verkningar. Enligt min uppfattning är det planeringsforskningens centrala uppgift att utveckla forskningsstrategier för dylika undersökningar och att pröva dessa genom empiriska studier².

Det är t.o.m. överraskande att den relativa reformvänlighet som varit förhärskande och den reformpraxis som utvecklats i Sverige under efterkrigstiden inte givit upphov till en forskningsgenre som just tagit sig an att utvärdera effekterna av genomförda reformer. ERU kan sålunda konstatera att effektmätningar är en relativt sen företeelse i Sverige (SOU 1978:47 s. 171). Internationellt och då i synnerhet i USA har dock omfattande forsknings- och utredningsaktiviteter växt fram i anslutning till de stora

¹Texten här är ursprungligen författad 1979. Se efterskriften för en redogörelse.

²Det är genom ett studium och en kritisk granskning av planeringslitteraturen som jag kommit att intressera mig för vad som blir resultatet av reformförsök och planeringsansträngningar (se Gullberg 1981). Det visar sig nämligen att planeringsteori och planeringstänkande är suveränt "ointresserade" av om - och i så fall hur - planering verkligen gör någon skillnad. Istället upprepas gång på gång samma rationalistiska tankefigurer inom planeringsteorin. Dessa är i sig så självklara att erfarenheter knappast biter på dem. Det är som en utväg ur den abstrakta, formalistiska återvändsgränd planeringstänkandet befinner sig i som jag föreslår ett inträngande studium av planeringens resultat och inte bara av dess intentioner. Men det är klart att detta tema har intresse långt utöver planeringstänkandets väl och ve.

socialpolitiska reforminsatserna på 60-talet - Great Society, War on Poverty, etc. Det kan därför tyckas att jag talar mot bättre vetande när jag inledningsvis hävdar att studiet av effekter av reformer och planering är ett eftersatt område. Eftersatt är det emellertid. Det är nämligen först på den allra sista tiden som dessa forskningsaktiviteter lett fram till enstaka fruktbara, men ofullgångna formuleringar av effektproblematiken. Som helhet förefaller också de här aktuella undersökningsgenrerna uppvisa en tämligen medioker standard¹. Endast en liten proportion av samtliga empiriska studier och teoretiska reflexioner är av allmänt intresse.

För Sveriges vidkommande kan - inte utan rätt - hävdas att det statliga utredningsväsendet fyller uppgifter motsvarande dem som den amerikanska evalueringsforskningen gör anspråk på. I vissa avseenden tycks den svenska utredningsmodellen t.o.m. (beroende på var man ställer sig och betraktar eller vad man har i kikaren) stå framom den amerikanska och då särskilt när det gäller politisk förankring av utvärderingsverksamheten. Effektproblematiken har emellertid inte - annat än undantagsvis - varit föremål för intresse inom svenskt utredningsväsende, detta till skillnad från särskilt vissa akademiska reflexioner över utvärderingsverksamhetens möjligheter företagna på andra sidan Atlanten.

Syftet med denna framställning är att dra lärdomar av erfarenheterna inom den amerikanska utvärderingsforskningen och att undersöka hur denna forskning mottagits i Sverige. Vidare skall några försök till effektmätningar i svenska statliga utredningar granskas och de metodologiska strategier som är gängse i utvärderingsforskningens effektmätningförsök kritiseras. På den punkten kommer jag att föreslå en annan forskningsstrategi. Slutligen diskuterar jag de speciella svårigheter som uppstår genom att utvärderingsverksamheten vanligen blir direkt indragen i den politiska kampen om olika reformer. I varje fall gäller det

¹"/F/ew sound evaluations ... have been undertaken." Rossi & Williams (1972:xiii). "The methodological and 'philosophical' literature on evaluation is distressingly repetitious." Glass (1976:10). Se också Gordon & Morse (1975:339ff).

användningen och tolkningen av dess resultat. Jag kommer alltså att föreslå en strategi för hur man skall kunna undersöka effekterna av reformförsök och planeringsansträngningar.

NYTT OCH GAMMALT INOM UTVÄRDERING - USA SOM FÖREGÅNGSLAND

När ryktet om intelligensindustri och expansiva branscher når våra öron från det stora landet i väster är det nog explosionen inom datorvärlden vi i första hand kommer att tänka på. Men ryktet skulle också, under nästan halvtannat decennium¹ ha talat sanning om det gällt en helt annan bransch: Den som producerar tjänsten att i olika tempi besvara frågan "vad förmår politiken?" Det gäller - kanske är språkbruket mera bekant - utvärdering av offentliga åtgärder, politiska program eller reformförsök. En lång rad institutioner vid och fristående från universiteten i USA erbjuder sig att under olika etiketter undersöka hur ett visst program kommer att lyckas eller varför ett annat, redan genomfört, inte gav det utfall politiker och planerare hade hoppats. Branschen är full av växtkraft och man talar till och med om den som en social rörelse (Freeman 1977:22), och om "the evaluation industry" (Floden & Weiner 1978:12). De flesta instituten finner också uppdragsgivare. Annars hade det naturligtvis aldrig blivit någon tillväxtbransch i Wall Streets mening².

Nu är emellertid kunskapsproduktionen inom detta område av intresse även utanför den krets av politiker och tjänstemän som önskar rättfärdiga sin egen versamhet eller finna argument mot fiender och akademiker som siktar mot kommersiell karriär. Frågeställningen om politikens effekter (vilka är politikens effekter och hur skall vi kunna upptäcka dem?) är nämligen av grundläggande samhällsvetenskapligt intresse. Den torde också kunna räkna med viss uppmärksamhet från en intresserad allmänhet som kanske inte alltid låter sig nöja med politikernas och tjänstemäns försäkringar. Den berör spörsmålet om hur samhället förändras och tangerar den gamla drömmen från upplysningstiden att människan skulle kunna bli sin egen historias herre, att mänskligheten själv skulle kunna frambringa önskade förändringar av samhället. I sina

¹Formulerat 1979.

²Freeman (1977:19). Några data angående omsättning återges i Rein & White (1977:239ff).

allra bästa stunder skulle denna forskningsgenre kunna ge en liten fingervisning om hur långt borta - eller nära - denna (mar?)dröm är.

Men forskningsområdet rymmer stora metodologiska och teoretiska problem som uppfordrar till självprövning. Nya vägar har också sökts. Hur en sådan omprövning - utvecklandet av en ny forskningsstrategi för studiet av planeringens och politikens effekter - borde gå till kommer jag att beröra nedan.

En samhällsvetenskap som utvecklats till storindustri i politikens och förvaltningens tjänst blir själv - på gott och ont - indragen i politiken (C.H.Weiss 1973, Freeman 1977:22). Den institutionella placering som utvärderingsprojekt erhåller visar sig ha stor betydelse för vilka resultat som produceras (Gordon & Morse 1975:353). Detta komplex av beställarnas förväntningar, utvärderarnas förhandsinställningar¹, resultatens innehåll och användningen av utvärderingsundersökningarna skall jag försöka ordna upp en smula med hjälp av den amerikanska debatten.

Att ställa samhällsvetenskaperna i tjänst hos makthavare är ingen ny idé². Den har förfäktats både av samhällstänkarna själva, som t.ex. Comte³, som hade oturen att erbjuda sina tjänster utan att vinna några samtida prominenters öron, och av furstar och revolutionsbekämpande imperialiststater⁴. Samhällsvetenskaplig kunskap kan komma till användning på olika sätt. Sven-Eric Liedman gör en intressant distinktion mellan vetenskapens praktiska-materiella användning: vetenskaplig kunskap engageras i utvecklingen av den sociologiska basen och den praktiska-ideella (eller ideologiska) användningen som innebär "vetenskapens användning för att be-

¹Berk & Rossi (1976:349) menar att utvärderarna "are always choosing sides even before empirical findings are produced."

²"One of the most appealing ideas of our century is the notion that science can be put to work to provide solutions to social problems." Zetterberg (1962:15). Här citerad efter Suchman (1967:1).

³Comte menade t.o.m. att vetenskapsmännen skulle bli den nya tidens makthavare.

⁴Se t.ex. Horowitz (1967). Om samhällsvetenskap i beslutsfattande se t.ex. C.H.Weiss (1977:1ff).

kräfta eller stärka eller förändra försanthållanden, värderingar och normer gällande värld, människor, religion och konst, politik och moral ..." (Liedman 1978:28f). När det gäller (i synnerhet) samhällsvetenskaperna skall detta kompletteras med användning "för att befästa, stärka, vidareutveckla eller förändra politiska institutioner och sedvänjor." Dvs. ett administrativt bruk av vetenskapen. (Liedman 1978:29f).

Utvärderings- eller evalueringsforskningens¹ framväxt i USA har understötts av nyttoargument. Administrativt och politiskt användbar kunskap har utlovats om reformers sätt att verka hur de kan effektiviseras². I realiteten tror jag att kampen om hur olika reformaktiviteter skall uppfattats, dvs. en slags ideologisk användning av evalueringsforskningen, spelat minst lika stor roll för dess utveckling. Evalueringsforskningen har också backats upp från "båda sidor": vetenskapsmän som vill och menar sig kunna göra en samhällsinsats³ och politiker/administratörer som av olika skäl efterfrågar utvärderingsverksamhet⁴.

Utvärderingsforskningen är således inte uttryck för något helt nytt förhållande mellan samhällsvetenskap och praktisk användning och detta varken på det idémässiga eller funktionella planet. Som en specifik genre är den en del eller vidareutveckling av tidigare moderiktningar i gränslandet mellan forskning och utredningsverksamhet, mellan universitet och förvaltning, som t.ex. policy science

¹För en diskussion om begreppet se nedan.

²"From their early beginnings, various fields of public service have been concerned with proving the effectiveness of their programs." Suchman (1967:12). "Evaluation research is a part of a broader effort to bring rationality into the policy-making process." Rossi & Williams (1972:xvii).

³"Sound evaluative research may allow social science to make a far greater contribution to society than it has in the past." Rossi & Williams (1972:xvii).

⁴"It is person of power and influence, not academicians, who are the evaluation research lobby and who are responsible primarily for the widespread growth of the evaluation endeavour." Freeman (1977:22).

(Lazarsfeld 1975), policy analysis¹ och i viss mån även den s.k. socialindikatorrörelsen². Utvärderingsforskningen kan också betraktas som ett uttryck för reformistisk samhällsuppfattning och strategi³.

Evaluerings- eller utvärderingsforskningens historia tar sig tämligen olika ut beroende på hur denna typ av forskningsaktivitet avgränsas från annan utrednings- eller samhällsvetenskaplig verksamhet. Det finns en relativt omfattande flora av definitioner⁴. Ansatserna till historieskrivning varierar också beroende på i vilken utsträckning författaren vill framställa evalueringsforskningen som något unikt och nytt eller om han/hon önskar understryka kontinuiteten och knyta an till tillämpad samhällsvetenskap överhuvudtaget⁵. Med dessa variationer är förbundna inte bara accentolikheter i historieskrivningen utan också den betydelsefulla skiljelinjen i diskussionerna om de metodologiska likheterna/skillnaderna mellan evalueringsforskning och tillämpad samhällsvetenskaplig forskning i allmänhet.

Det är klart att evalueringsforskningen och angränsande genrer på något sätt har med bedömningen av politiska/administrativa reformer eller program att göra. Suchman målar upp ett tvåhundraårigt perspektiv⁶, medan Bernstein och Freeman finner en

¹Policy analysis har enligt Wollmann (1979:16) misslyckats i sina ambitioner att fungera som beslutsteknik och -hjälp, därför att politiska/administrativa och byråkratiska maktrelationer och implementerbarhet i stor utsträckning ignoreras.

²För en delvis kritisk granskning och påtalandet av indikatorernas begränsningar se Sheldon & Freeman (1970).

³Bell (1975:25f). Jfr. också Weiss (1973:6): "Most of /the/ political implications of evaluation research have an 'establishment' orientation. They accept the world as it is defined in agency structure ... But the proclivity of evaluation research is reformist".

⁴Se t.ex. Bernstein & Freeman (1975:1), Caro (1971:2) eller Gordon & Morse (1975:339). Mig förefaller det som Lind (1979b:12) överdriver enhetligheten när han menar sig finna stor överensstämmelse mellan förekommande definitioner av utvärdering/evaluering.

⁵Freeman & Bernstein (1975:9) menar t.ex. att "/From the standpoint of the social science, evaluation research has a considerable history".

⁶"With the advent of the period of revolution and enlightenment in the eighteenth century, public service entered an era of research, experimentation, and extended program development. For a period of almost 200 years, beginning slowly but reaching a climax

kontinuerlig utveckling inom forskningsfältet de senaste 40 åren (Freeman 1977:21 och Freeman & Bernstein 1975:9). De hänvisar då till en av samtiden knappast uppmärksammas artikel, skriven 1935 av Stephen Stephan, som pläderar för en undersökning av New Deals effekter (Stephan 1935, även i Caro 1971:37ff). Som föregångare har också angivits hälsoundersökningar och utvärderingar av hälsoprogram i USA i början av 1900-talet (Suchman 1967:13f). Lester Ward som 1906 förutsåg användandet av vetenskapliga metoder i utvecklandet av sociallagstiftningen (Caro 1971:4), Stuart Chapins förslag att genomföra sociala experiment under 10-talet (Caro 1971:4 och Suchman 1967:14), experimentella undersökningar av produktivitet och moral hos arbetare från 20-talet och framåt⁷, Newcombs socialpsykologiska studier av utvärderande karaktär (Caro 1971:5), Kurt Lewins studier på 30- och 40-talet och den klassiska undersökningen "the American Soldier" av Stouffer räknas också in (Caro 1971:5 och Bernstein & Freeman 1975:ix). Listan är som synes lång och kan göras längre. Det är emellertid helt klart att evalueringsforskningen, som en speciell aktivitet med eget självmedvetande, egna tidskrifter och utvecklade i industriell skala växer fram först i och med Johnsonadministrationens omfattande socialpolitiska aktiviteter. Freeman argumenterar för att det både föreligger en kontinuerlig utveckling av evalueringsforskningen och att ett kvalitativt nytt moment utvecklats i densamma (Freeman 1977:21). Som en förelöpare till de mera specifika utvärderingsverksamheterna kan införandet av programbudgetering vid det amerikanska försvarsdepartementet 1961 ses (Schick 1971 och Cyr & de Leon 1975:377). Liknande system skulle på Johnsons order införas i den övriga federala byråkratin 1965 (Cyr & de Leon 1975:377). När så "the Great Society" och "the War on Poverty" proklamerades, engagerades samhällsvetare i en tidigare icke skådad omfattning för att genomföra och utvärdera projekten. Det är intressant att notera att samhällsvetenskapligt expert-

 in our present century, public service administration offered to the public a choice of preventive and service programs based increasingly upon 'scientific' evaluations of their effectiveness" Suchman (1967:13).

⁷Caro (1971:4) nämner Mayo och Roethlisberg & Dickson.

kunnande knappast alls var inblandat i den egentliga uppläggningsen av "the Great Society". Den stora boomen kom efter beslutet (Schick 1977:258).

Samhällsvetare engagerade sig sedermera livligt som agitatorer för den nya politiken¹. Det var helt uppenbart inte framsteg inom samhällsvetenskapen som ledde fram till dess ökande användning². Det var istället politiska och administrativa intressen som drev fram den stora expansionen av samhällsvetenskaplig aktivitet inom och i tjänst hos den federala statsmakten (Schick 1977:258, Poland 1974:333).

Åren som gått sedan 60-talets början har fört med sig avsevärda förändringar i användningen av samhällsvetenskaplig expertis i den amerikanska förvaltningen. Till en del har detta även inneburit förskjutningar mellan de akademiska disciplinernas involverande. Förloppet kan i korthet sammanfattas enligt följande³: Det sovjetiska uppsändandet av Sputnik I 1957 utlöste något av ett nationellt chocktillstånd i USA och ganska snart samlades nationen runt uppgiften att sända en man till månen. Febrila aktiviteter utvecklades på det rymdteknologiska området, men utbildningssystemets kapaciteter diskuterades också. Varför hade det inte lyckats för USA, jordens rikaste nation, att frambringa sådant kunnande och sådan teknik att landet kunde uppträda som den främste också i rymden? Var det fel på utbildningssystemet?⁴

Ur intresset att undersöka de sociala följdverkningarna som de stora federala rymdsatsningarna väntades ha på den sociala situationen (ökad välfärd) växte den s.k. socialindikatorrörelsen

¹Jfr. Schick (1971:154f) om samhällsvetare i rollen som socialpolitiska aktörer mer än som vetenskapsmän.

²Även om utvecklingen av computertekniken var en av förutsättningarna för "massproduktionen".

³Jag har haft stor nytta av Hellstern och Wollmanns mycket informativa och kunniga översikter som tillhör de bästa jag känner till.

⁴Sedan, när USA erövrat denna position i rymden, dök en annan mera jordnära fråga upp: "If we can land a man on the moon why can't we solve the problems of the ghetto?" Nelson (1974: 376).

fram. Den syftade till att utveckla mått på befolkningens levnadsstandard (Bauer 1966). Utvecklingen av och tilltron till generella planeringssystem - t.ex. programbudgetering - emanerade från rymd- och försvarsaktiviteterna. Kennedy- och Johnsonadministrationens liberaler anammade den rationalistisk-teknokratiska synen på politiken som utvecklats inom rymd/månprogrammen, vilka för övrigt kan ses som en komponent i reform-demokraternas¹ (den mer eller mindre levande traditionen från "New Deal") föreställningsvärld. Utmärkande för denna fas var att analysarbetet inriktades på att undersöka olika mer eller mindre realistiska handlingsalternativ i förväg genom att kalkylera deras respektive (tänkta) utfall - deras kostnader och nyttor - och utforma underlag för politiska beslut. Dessa ex-ante tekniker, dvs. utvärdering av (alternativa) planer i förväg kunde emellertid inte infria de högt ställda förväntningar som hade knutits till dem², vilket befordrade intresset för att undersöka hur redan genomförda eller under verkställighet varande projekt föll ut. Parentetiskt måste jag här inskjuta att de terminologiska frågorna ställer till en hel del trassel. Ex-anteundersökningar benämns överhuvudtaget ej som evalueringsstudier av t.ex. Schick som reserverar denna term för ex-poststudier. Cost-benefit och andra former av studier i förhand³ benämner han "analysis", i överensstämmelse med policy analysis - får man förmoda (Schick 1971). Bells användning av termen utvärdering svarar i stort mot Schicks evalueringsstudier (Bell 1975), medan hos Åhrén utvärdering i första hand står för ekonomiska förhandskalkyler (Åhrén 1975 och även Bohm 1972).

¹Bell (1975:195) pekar ut McNamaras krigsstrategiska tänkande som en av den sociala reformpolitikens iderötter.

²Se bland många andra som pekat på detta t.ex. Hellstern & Wollmann (1977:420).

³Cost-benefitanalyser förekommer även i ex-postskepnad. Se t.ex. Haverman (1972). Vad mig är bekant förekommer denna typ av studier sparsamt och är mycket modellcentrerade vilket betyder att empirin spelar en undaskymd roll. Williams (1972:6f) invänder mot en användning av termen "cost-benefit analysis" som innefattande både policy-analys ("a priori analys") och outcome evaluering ("ex-post analysis"). Han menar att cost-benefitanalyser bör skiljas från evaluering. Den förra gäller ex-ante, den senare ex-post.

Den engelska (eller amerikanska) termen "evaluation" kommer jag i det följande att översätta med "utvärdering". Med "evaluering - utvärdering" avser jag bedömningar av politiska reformer och program som faktiskt har genomförts eller är under implementering¹. Likväl kommer jag i det följande att något beröra även ex-anteanalyser eftersom de olika genrernas framväxt och utveckling är nära relaterade till varandra.

Den ovan omtalade förskjutningen från ex-ante till ex-postbedömningar beskriver Schick (1971) som menar att besvikelserna över the Great Society men också Vietnankrigets för USA-etablissermanget och den teknokratiska krigsledningen mindre lyckosamma förlopp var de faktorer som förklarar tillbakagången för reformansatser med vidlyftigare förändringsambitioner och överhuvudtaget beräkningar av programeffekter i förväg. Bedömningen av hur lyckosam the Great Society-politiken egentligen var är emellertid ett komplicerat kapitel. Det är klart att en besvikelse är motiverad om man jämför med de löften som gavs. De heta somrarna som utspelades mot slutet av 60-talet i svarta gettoområden i snart sagt varje amerikansk storstad förefaller till en del vara uttryck för en sådan besvikelse, dvs. till en del uttryck för ökade anspråk utvecklade bl.a. i de självorganiseringsprocesser i gettoområdena som fattigdomskrigets olika program, om inte skapade så i varje fall, bidrog till. De analyser som Piven och Cloward (1971) gjort av kriget, det krig som till övervägande del vände sig till de svarta områdena i storstäderna, tyder på att programmen ur en annan synvinkel inte var så helt misslyckade. Hela fattigdomskriget framstår som en reaktion inför hotande elektoral bakslag för demokraterna. Särskilt för Johnson måste valutslaget 1966 betecknas som mycket tillfredsställande. Den hastigt växande svarta befolkningen i storstäderna i norr - bortdriven från de alltmer mekaniserade jordbruken i södern - var en viktig grupp som han lyckades vinna. Även för John F. Kennedys valseger var de svarta väljarna i norr viktiga. Deras stöd tycks t.o.m. ha varit avgörande för utgången 1962. Frågan om i vilken utsträckning fattigdomskrigets misslyckande provocerade en

¹Detta överensstämmer med Hermerens (1980:9:19) språkbruk: "Jag skall därför föreslå att termen 'utvärdering' ... reserveras för uppföljning i efterhand (ex post) av hur de förslag eller alternativ man satsat på i praktiken visat sig fungera."

förändring av evalueringsforskningen från ex-ante till ex-postinriktning, eller om denna förändring bidrog till att skapa bilden av ett misslyckande, är därför svår att besvara på ett entydigt sätt, eftersom det inte finns några absoluta bedömningskriterier. Till en del är faktiskt undersökningarna av reformprogrammets effekter medvetna försök att stoppa dem, eller resultaten kan användas på ett sådant sätt. Den av Nixon och hans stab gjorda tolkningen av utvärderingen av det bekanta "Head Startprojektet" är ett typexempel (Rose 1976, Hellstern & Wollmann 1977:450, Lind 1979b:23ff). En del av sanningen tycks alltså vara att uppfattningen av the Great Society som ett stort misslyckande är ett resultat av framgångsrika försök bland reformmotståndare att misskreditera verksamheten¹. Å andra sidan är det uppenbart att programmen inte nått de resultat som utlovats, vilket - för det tredje - inte heller var den djupaste avsikten med programmen, eller annorlunda uttryckt: det politiska kraftfält inom vars ram de reformpolitiska löftena utvecklats rymde föga dynamik eller kraft för deras realiserande.

Hellstern & Wollmann (1977:102) förlägger övergången från ex-ante till ex-poststudier till ett tidigare datum. För dem är det en ökad känslighet för den inrikespolitiska konfliktnivån under 60-talet som öppnar fältet för forskning om hur politiken påverkar olika befolkningsgrupper och hur kostnad och nytta av olika åtgärder fördelas. De ser förskjutningen från frågan "What causes policy?" till "What policy causes?" som signifikativ för det ändrade politiska klimatet. Hela utvecklingen i tendensen från socialindikatorfasen² via förhandsanalyser (ex-anteutvärderingar) till genomförandestudier (implementering) och effektundersökningar avspeglar det amerikanska samhällsklimatets många och stegvisa förändringar från det tidiga 60-talet fram till idag. Detta betyder också att evalueringsforskningen förändrats i olika etapper. Så kan t.ex.

¹Rein & White (1977:253) menar att utvärderingen av Head Start lyckats sprida refrängen "nothing works" i vida kretsar, vilket var en indirekt men samtidigt mycket betydelsefull konsekvens av denna utvärdering.

²Socialindikatorundersökningar förmår som Sheldon & Freeman (1970:100) påpekar, inte åstadkomma kausala slutsatser. Ty även om förändringar av indikatorerna kan påvisas finns ingen metod att utrona om och i vilken utsträckning dessa härrör från i arbete varande reformprogram.

övergången från ex-ante till ex-poststudier inte återförs på ett enda moment i denna utveckling. Bakom programbudgetering och andra förhandsanalyser är tanken på en ökad styrning av verksamheterna lätt identifierbar. De olika lagverk som knåsatte the Great Society innehöll vanligen bestämmelser om att en viss procent - vanligen en - av anslagna medel skulle avsättas till evalueringar (Hellstern & Wollmann 1978:8). Tanken med detta synes ha varit att reformprogrammets effektivitet fortlöpande skulle kunna höjas. När så den sociala oron i storstädernas svarta områden tilltog låg naturligtvis frågan om vad som hände med reformprogrammen nära till hands och intresset riktades mot de socialpolitiska åtgärdernas effekter. Men varje undersökning av effekter innehåller inte bara en uppfordran till ökad effektivitet. Den innehåller också - eftersom reformprogram utan undantag misslyckas med att uppnå precis det utlovade resultatet eller kraven i mätförfarandet kan sättas så högt att tillräckliga utslag sällan uppnås - argument för verksamhetens avskaffande. Och det var detta som Nixonadministrationen tog fasta på i det ovan omnämnda exemplet Head Start.

Det är viktigt att observera att utvärderingsforskningen kunnat stå till tjänst oavsett om det gällt att uppamma tilltron till rationella styrsystem, legitimera omfattande socialpolitiska insatser eller motivera dessas avveckling (jfr. Hellstern & Wollmann 1977:450, Horowitz & Katz 1975:139 och Poland 1974:333). Den omfattande samhällsvetenskapliga forskning som använts i den federala byråkratin har alltså visat sig ha en stor anpassningsförmåga vid politiska klimatförändringar. Smidigheten härvidlag beror emellertid inte bara på förskjutningar av forskningsfokus efter politiska signaler utan också på att de flesta resultat lämnar stora möjligheter till tolkning för de politiska instanserna. Det finns också alltid möjligheter att mobilisera "motexpertis" om resultaten visar sig alltför besvärande, beroende på att ingen undersökning - allra minst av effektsammanhang - kan göras helt invändingsfri.

AKADEMISKA DISCIPLINER I UTVÄRDERING OCH UTVÄRDERINGENS GENRE-RIKEDOM

De förhandsanalytiska metoderna (programbudgetering, cost-benefit, etc.) stod i huvudsak ekonomer för (Nelson 1974:383). Den ekonomiska disciplinens väl utvecklade och långt drivna förmåga att bygga modeller kom här väl till pass. Det stora teoretiska problemet vid konstruktionen av (förment) rationellt beslutsunderlag är att reducera så många - helst alla - relevanta beslutsdimensioner som möjligt till en. De ekonomiska modellresonemangen lämpar sig härvidlag väl. Prisrelationer (marknadsbestämda eller fiktiva) blev den integrationsnyckel man opererade med. Förhandsanalyserna höll emellertid inte - oavsett vilka disciplinföreträdare som engagerade sig i deras utveckling eller tillämpning¹ - som jag redan konstaterat, vad de lovade².

I analyserna av de resultat som genomförda reformer uppnår har psykologer och pedagoger, men också sociologer intagit en ledande ställning³. Till en del beror detta säkerligen på att dessa discipliner är väl förtrogna med experimentell och kvasieexperimentell metodik - vilka dominerar resultatmätningarna, i varje fall som ideal - men också på att flera av de program som utvärderades ex-post gällde utbildningsreformer, behandling av kriminella och försök att påverka slumbefolkningens livsförhållanden, problem som innehållsmässigt ligger nära dessa discipliner. Övrigt, men väl omvitnat, är att den politiska vetenskapens företrädare, med få undantag⁴, mycket sent började intressera sig för resultatet av

¹Systemteoretiker och operationsanalytiker var också talrikt företrädare i detta arbete.

²"The loss of faith in the ability to predict on the basis of ex ante calculation what will work and not work and a new belief in pragmatic experimentalism and despair over the fact that few problems can actually be solved from the center, have all greatly eroded the influence of the policy analysis perspective". Nelson (1974:389).

³Rhoads (1978:114) konstaterar att den tidigare dominansen av ekonomer inom policy analysis har minskat på senare tid och detta i takt med att intresset förskjutits från analyser av föreslagna program till utvärdering av redan genomförda (ex-post).

⁴Standardexemplet är Laswell.

de politiska beslutsprocesser som man så noggrant följde fram till beslut fattats¹. Forskjutningen inom den politiska vetenskapen är, trots att den är sen och på intet sätt någon spjutspets, instruktiv och beskriven hos Hellstern och Wollmann (1977).

Lewis (1975) varning till den politologiska forskningen för dess försök att "modernisera" ämnet genom efterhärming av ekonomisk och sociologisk "mainstream" forskning är lätt att instämma i och bör tas på allvar, även bland ekonomer och sociologer som arbetar med utvärdering. Här skall jag använda några av de begrepp som lanserades inom politologin, men också inom den mer eller mindre "disciplinfria" evalueringsforskningen.

Under det inrikespolitiskt relativt lugna 50-talet och i början av 60-talet befann sig politikformuleringen eller politikens "input"-sida i centrum för politikvetenskapens studier (Wollmann 1979:10).

Denna systemteoretiska term syftar på de krav som av omgivningen riktas mot det politiska systemet. Frågor om hur politiken förverkligas eller vilka resultat den har lämnas åt sidan. Den normativa barlast som politikvetenskapen hade att släpa på i det kalla krigets dagar utslöt mer eller mindre fullständigt dessa typer av problem. Enligt den förhärskande demokratiuppfattningen var det otänkbart att beslut som fattats i demokratisk ordning annat än genomfördes. Politiken var ju den auktoritativa fördelningen av värden i samhället och politikvetenskapens ansträngningar riktades mot att undersöka hur politiken formulerades². Detta innebar t.ex. att den politologiska forskningen inte tog upp de inrikespolitiska kristecken som började visa sig under 50- och det tidiga 60-talet. Sent omsider bröts emellertid denna isolering och kraven på samhällsrelevans framfördes allt mera ihärdigt under

¹Recognition that the process does not necessarily stop with the policy decision has only recently begun to take hold. /In contemporary political science A.G./ and new questions are being asked, such as: 'What happens after the policy is made?' ..." Grumm (1975:439), "Many other social scientists were at work in this field and political scientists arrived at a very late stage". Dolbear (1974:108).

²Inputforskningen kritiserar av Rakoff & Schaefer (1970) som också visar att även inputsidan kan göras till föremål för kritiska och teoretiskt medvetna studier.

60-talets senare hälft. Dessa krav hade två källflöden. Ett var den federala byråkratins efterfrågan på "relevanta" kunskaper för de sociala reformprogrammen. Ett annat var en radikal samhällskritisk rörelse i första hand bland universitetsfolk och intellektuella vid denna tid.

Outputforskningen utvecklades under kort tid till en mycket livaktig forskningsgenre. Den syftade till att undersöka vad som kom ut av det politiska systemet. Vad fanns det för samband mellan input (socialekonomiska variabler/politiska majoritetsförhållanden) och output (utgiftsnivån hos delstater eller lokala förvaltningar)? Ett omfattande empiriskt material i form av räkenskaper från de olika amerikanska delstaterna eller från lokala förvaltningar och datateknikens utveckling var grunden för detta snabbväxande forskningsområde. Resultaten pekade mot att det är de socio-politiska variablerna som är avgörande för output. Politiska förhållanden som t.ex. konkurrenssituationen mellan partier och valdeltagande spelar mindre roll¹.

Output-forskningen kunde emellertid inte heller den svara mot de allt envisare kraven på samhällsrelevans. Visserligen undersöktes i viss mening resultatet av politiken (eller skall vi säga politikens uteblivna resultat), framförallt hur de delstatliga och lokala budgetposterna tog sig ut i dollar räknat och vissa andra "standardmått". Hur dessa förhållanden inverkar på olika grupper och hur kostnader och nyttor av olika offentliga verksamheter fördelades dem emellan förblev emellertid dolt i denna typ av

¹Rakoff & Schaefer (1970:51) och Jacob & Lipsky (1968). Här kan intressanta paralleller iakttas med behandlingsforskningen som ofta har svårigheter att påvisa skillnader i utfall beroende på olika behandlingar. De socioekonomiska bakgrundsvariablerna hos individerna tenderar att förklara hela variansen. Se t.ex. Börjesson (1966:62). Liknande resultat har också presenterats angående skolans effekter på elevernas prestationer. Christofferson (1979:105) menar att det enligt de flesta studierna på området ser ut "som alla samband mellan skolresultat (vad man än räknar dit) och insatta resurser (hur de än definieras) är svaga med undantag för social bakgrund som i allmänhet visar sig vara avgörande för skolframgången ..."

aggregerade data (Hellstern & Wollmann 1977:439). De politiska och administrativa variablerna återvinner sin betydelse om grupp-specifika undersökningar görs¹.

De forskningsgenrer som växer fram i reaktion mot inputforskningens ("what causes policy") inriktning och outputforskningens otillräckliga omorientering (dess oförmåga att ge relevanta svar på frågan "what policy causes?") har alla, liksom tidigare socialindikatorforskningen, cost-benefitanalysen, policy sciences etc., något klart modebetonat över sig. De proklamerades som sociala rörelser och tillväxtindustrier, vilket inte sker utan saklig grund. Det är svårt att värja sig mot intrycket att benämningar på forskningsområden lanseras för att etablera kommersiellt lönsamma och/eller akademiskt karriärskapande inmutningar. Intresseförskjutningen i riktning mot resultat- och genomförandeundersökningar är i alla fall positiv i den mån frågorna tas på allvar². Hos vissa författare har termen "outcome" införts för att markera en skillnad i förhållande till "output". För Easton t.ex. är "an output /.../ the stone tossed into the pond and its first splash, the outcomes are the ever widening and vanishing pattern of concentric ripples. The actual decisions and implementing actions are the outputs, the consequences traceable to them, however long the discernible chain of causation, are the outcomes"³. Medan t.ex. Dye inte gör någon distinktion och använder termen "outcome" (Dye 1966). Hans studie är att hänföra till outputgenren⁴. Eastons distinktion har påtagliga likheter med den skillnad som Cook och Scioli gör inom

¹Wollmann (1979:7). En nedbrytning av data inom skolforskningen visar parallella resultat. Christoffersson (1979:109) refererar till Summers & Wolfe (1977). De "har i en input-outputstudie undersökt hypotesen, att tidigare undersökningar i stort misslyckats påvisa positiva samband, därför att man använt allför aggregerade data för skol-inputs". Hypotesen kunde styrkas. När data aggregerades försvann i stort sett de blottlagda sambanden. Christoffersson (1979:110).

²Se min plädering för denna typ av studier i Gullberg (1981).

³Easton (1965:352). Roos beskriver i exemplet miljöpolitik skillnaden enligt följande: "Outputs are such things as the passage of laws and a forement efforts: outcomes are effects upon water quality, quality of the air, and so forth. The output of business-government interaction may not result in better outcomes". (Roos 1975:254).

⁴Detta gör t.ex. Jacob och Lipsky (1968:514).

en annan begreppsram, nämligen den mellan "primary" och "secondary impacts" som syftar på de omedelbara förändringar som resulterar från en aktivitet inom ramen för ett politiskt program, respektive sådana konsekvenser som gör sig gällande på lång sikt (Cook och Scioli 1975:101f).

"Impact-forskning" som alltså är den term Cook och Scioli använder för att karaktärisera resultatet av omorienteringen från politikens formuleringsprocess till dess effekter är en av de omtalade nya forskningsgenrerna. Som så ofta inom det samhällsvetenskapliga forskningsfältet tvingas vi här konstatera att detta modeord inte har någon mera allmänt accepterad och distinkt innebörd¹. Att det är uttryck för en önskan till nyorientering står emellertid klart. Lanserandet av termen har något av ritual över sig: att markera den egna forskningsinriktningen som originellare och nyare än vad den i själva verket är. Användningen av termerna "output", "outcome" och "impact" korsar varandra.

Dolbeares kritiska granskning av impactforskningen (Dolbeare 1974) ger vid handen att denna inte bidragit med mycket av värde. Impactstudier är för Dolbeare sådana studier som tar policy-innehållet som oberoende variabel och ställer frågan: "What difference /does/ it make /.../ to people and problems?" (Dolbeare 1974:93). Cook och Scioli förknippar impact-forskningen direkt till omorienteringen från frågeställningar om vad som bestämmer politiken till vad politiken bestämmer. För dem är impact-forskningen inriktad på frågor som har empiriskt innehåll (Cook och Scioli 1975:95ff). Grumms resonemang om impact går i samma riktning och han menar att: "Before impact, there is at least a two-stage process in which policy implementation can be considered an intervening variable between the policy and its impact." (Grumm 1975:422). Grumm som använder sig av ett lätt systemteoretiskt resonemangssätt - han hänvisar till Eastons begreppspar "output" och "outcome" - nämner åtta steg i

¹Eller som Dolbeare konstaterar: "/I/mpact studies ... is a rapidly developing subfield /in social science A.G./ with few shared definitions or boundaries. Depending on one's definition, there are between one and four thousand impact studies ... already published or stored in government warehouses". Dolbeare (1974:89).

processen från politikens formulering fram till det att dess impact eventuellt återverkar på politikens fortsatta output (Grumm 1975:444f) - dvs. feed-back i systemteoretisk jargong.

Implementeringsforskning är en annan gren inom den tillämpningsinriktade, sig till förvaltningen (politikerna) vändande samhällsvetenskapliga forskningen. "Implementation is fast becoming the newest shibboleth among academics and policy analysis concerned with complex social programs (Williams & Elmore 1976:xi). Den är av sent datum. "/W/hen the massive social programs of the Johnson Administration were being launched hardly anyone spoke of implementation. Five years ago, when the disappointing results of these policies and the frustration of Vietnam had dulled the high hopes of even the most optimistic reformers, several aouthers began to form an implementatism as the main problem. Now implementation is 'in'" (Berman 1978:157f). Implementering är stadiet mellan beslut och verksamhet (Williams & Elmore 1976:xi). Den kan ses som en klass av socio-politiska processer som stammar från och förutskickats av tidigare stadier i den politiska processen (Bunker 1972:72). Många forskare inom detta gebit har uttryckt sin förvåning över att genomförandet av en viss politik eller ett visst reformprogram inte undersöks närmare¹. Denna känsloreaktion kan jag dela. Det är naivt att tro att en politik kommer att genomföras bara för att "auktoritativa beslut fattats"². För min del väcks förvåningen också av det förhållandet att resultat av planering och politik knappast undersöks och då nästan bara på det sätt som output-forskningen gjort, dvs. utan att sätta resultatet i relation till de eventuella operationer som genomförts inom ramen för den studerade politiken.

Även implementationsforskningen är underkastad konceptuella mångtydligheter. V. Horn och v. Meter har funnit många, ofta motstridiga användningar av begreppet implementering i litteraturen (v Horn & v Meter 1976:45). Själva understryker de skill-

¹Se t.ex. Pressman och Wildavsky (1973), Williams (1976a:288) och Hargrove (1975). Den senare betecknar implementationen som "The missing link".

²There is an implicit assumption in most policy studies that once a policy has been formulated the policy will be implemented". Smith (1973:197).

naden mellan policy impact som de menar handlar om frågan "vad händer?" och policy implementering som gäller "varför händer det på detta sätt? (v Horn & v Meter 1976:45 som citerer Dolbear 1974). Enligt Berman utgår implementationsanalyser från att påståendet "om P (policy) så O (output)" skall ersättas med en kedja om minst två relationer: "om P (policy) så I (implementering av programmet", och "om I så O (outcomes)" (Berman 1978:163). Eftersom jag påstått att utvecklingen av dessa forskningsgenrer varit något av en reaktion mot ex-ante analyser, kan det vara av intresse att Williams menar att en implementationsanalys - innebärande att sannolikheterna för att a) den tilltänkta förändringen kan genomföras med tillgängliga resurser och att b) den kommer att accepteras bland dem som måste införa eller fullfölja det nya - bör göras i förhand. Naturligtvis kan implementationsundersökningar också handla om att uppskatta i vilken utsträckning en organisations prestationer kommer att kunna förändras i föreskriven riktning (Williams 1976a:280,282).

Evalueringsforskning får ofta som vi sett ovan stå som beteckning för bedömning av politiska program och reformer både ex-ante och ex-post. Denna terminologi är dock inte accepterad av alla. En viss användning av "evaluering" och "evalueringsforskning" ställer detta studieområde på jämställd fot med "impact" och "implementering". Det är framför allt när evaluering uppfattas som efterhandsbedömning av genomförda eller i kraft varande program så som jag föreslår ovan. Evalueringsforskning i denna betydelse har många likheter med de ovan nämnda forskningsområdena. För det första finns naturligtvis ingen allmänt godtagen definition (Bernstein & Freeman 1975:1,16ff). För det andra har en - icke otypisk - definition många beröringspunkter med definitioner av "impact" och "implementering": "the systematic study of the operation of social-action, treatment, or intervention programs, and of their impact" (Bernstein & Freeman 1975:1). Bernstein och Freeman skiljer mellan processevaluering och impactevaluering. Den förra undersöker huruvida ett visst program har implementerats i enlighet med uppdragna riktlinjer och den senare om det gör

någon skillnad med det aktuella programmet¹. Introduktionsartikeln till den 1977 startade tidskriften Evaluation Research innehåller en forskningsöversikt där både implementering och impactfrågor inkluderas i evalueringsforskningen².

Slutsatsen tycks given. De forskningsområden som seglar under följande flagg: implementering, impact, outcome och evaluering utgör i realiteten inga från varandra avskilda, med olika frågekomplex arbetande forskningsområden. Tvärtom går de in i varandra och kan t.o.m. behandla identiska problem. Det finns alltså ingen anledning att betrakta dessa forskningsområden som separata. Däremot pekar de olika begreppen i riktning mot olika typer av problemställningar. Det kan därför var motiverat att upprätthålla de begreppsmässiga distinktionerna, om dessa utarbetas en smula och man alltså undviker att låta begreppen äta upp varandra. Sjöblom har i en för övrigt mycket stimulerande och för politikvetenskapen sympatiskt (själv)prövande artikel föreslagit följande synsätt: "Through a series of processes some outputs are produced, in the form of decisions, with such and such a content, which are or are not implemented and which have or have not certain impacts on the environment" (Sjöblom 1977:19). Det finns i och för sig andra begreppsuppsättningar av liknande slag som skulle kunna refereras (t.ex. Grumm 1975). Jag ska dock nöja mig med att här komplettera Sjöblom med hjälp av ett konstaterande som Williams gör, nämligen att verbet implementera har två betydelser: dels att vidta åtgärder för att sätta i verket, dels sätta i verket (Williams 1976b:3). Denna språkundersökning bäddar för frågan om det är rimligt att tala om implementering i två olika bemärkelser. Jag menar att så är fallet. Om vi med output fortfarande menar de beslut och det programinnehåll som politiska processer leder fram till, så kan den första typen av implementering sägas bestå i att medlen för det beslutade programmet mobiliseras. Det kan t.ex. gälla organisatoriska förändringar, genomförande av personalutbildning, anvisning av

¹Freeman (1977:25ff) och Bernstein & Freeman (1975:18ff). De skriver också om "Comprehensive Evaluation" som omfattar båda de omnämnda undersökningstyperna.

²"Obviously ... the impact evaluation of a program should be related to the content of the program as implemented". Rossi & Wright (1977:26).

ekonomiska medel, anskaffande av information etc. Dessa omstruktureringar inom den offentliga apparaten eller i dess kontakter med "omgivningen" kan benämnas outcome. Det andra steget i implementeringen består då i att dessa mer eller mindre förändrade organisationer eller resurser (eventuellt) tas i bruk med hänvisning till och/eller med avsikt att uppnå de mål som formulerats i programmet, vilket alltså leder/inte leder till en påverkan av de (eller andra) samhällsliga processer varom målet handlade. Detta kan då benämnas "impact". I enlighet med mitt förslag ovan vill jag med evalueringsforskning mena sådan forskning som bedömer politiska program eller reformer oavsett om detta sker ex-post eller någonstans under resans gång efter start.

Den föreslagna nomenklaturen löser i sig knappast några problem men pekar på några olika aspekter som kan beaktas, eventuellt var för sig men helst i sina ömsesidiga relationer. Uppdelningen i steg kan i och för sig drivas utan slut. Pressman och Wildavsky talar t.ex. om kedjor av beslutspunkter i implementeringsprocessen, där varje punkt motsvarar varje gång ett instämmande måste registreras för att ett program skall löpa vidare (Pressman & Wildavsky 1973:69,xvi). Hur man än ställer sig till uppdelningen av vägen från den politiska beslutsprocessen till de sista (eventuella) utklingande effekterna (impacts) av politiken i olika steg måste man ha klart för sig att vi här har att göra med modellresonemang som i ringa grad avbildar verkligheten, utan fastmer erbjuder oss analytiska kategorier till hjälp /eller självp/1 när det gäller att avvinna denna verklighet dess (eller kanske betraktarens) hemligheter. Jag skulle t.ex. knappast ha några argument mot påståendet att det inte finns en väg - inte ens en krokig - som leder från intention till prestation och effekt. Däremot menar jag att sökandet efter sådana vägar bör kunna leda till intressanta upptäckter bara det inte drivs med sådan frenesi att det urartar till anläggningsarbete i enlighet med fixerade ritningar och utan känsla för terrängen.

 1Se t.ex. Majone och Wildavsky (1978) som pekar på de idealistiska och dualistiska misstag som uppstår då "policy" och "implementation" skiljas åt.

Tar man del av mera fordrande och seriösa bedömares betraktelser över utvärderingarnas aktuella läge och potentiella möjligheter fångas man i en viss förundran. De ger nämligen uttryck för en paradox, närmare bestämt för en paradox som är utvärderingsforskningens egen.

Snart sagt överallt uppstår utvärderingsforskning. "From the ambitions of the academic disciplines, from the convulsive reform of the educational system, from the battleground of the War on Poverty, from the ashes of the Great Society, from the reprisals of an indignant taxpaying public, there has emerged evaluation (Glass 1976:9). Specialtidskrifter ges ut och nya startar efter hand, handböcker trycks, symposier arrangeras och uppdragen strömmar in. Många samhällsvetare som med ganska stor sannolikhet eljest hade varit arbetslösa har sugits upp i utvärderingsarbete. I icke ringa grad förefaller efterfrågan ha influerat en stor del av forskningens inriktning även om forskarna själva visade sig ha god känsla för den nya "problemsituationen" (Hellstern & Wollmann 1977:422). Utvärdering har utvecklats till en robust aktivitet med beaktansvärd historia (Freeman & Bernstein 1975:9). Men likväl finns det ett djupt missnöje. Genren väcker, i all sin kommersiella glans, en tydlig olust¹. Ty även om det förvisso kan vara tillfredsställande att vara efterfrågad och göra pengar är detta inte tillräckligt för alla. Några vill dessutom eller ännu hellre göra ett bra jobb, och ett bra jobb för vetenskapsmän som effektivt socialiserats in i det vetenskapliga samhället betyder god kvalitet och rimliga relationer mellan undersökningarnas innehåll, deras slutsatser och eventuellt också deras användning. Vetenskapens uppgift är sanningen². Men detta är inte

¹Denna omfattas förvisso inte av alla. Guttentag & Saar (1977:11) menar att evalueringsforskningen snabbt nått mognad som disciplin, att det förekommer föga diskussion om vad evaluering är och inte är. "Considerable consensual validation now exists about the domain of evaluation ... Evaluation research is a discipline which now has immense spread throughout every sector of society. Jfr. också Freeman & Bernstein (1975:9). Utvärderingsforskningens utbredning är ovedersäglig. Dess mognad i vetenskapligt hänseende ligger dock fjärran.

²Eller som Berk och Rossi (1976:342) uttrycker sig: "Good social science is at least a demystifying activity".

vad den politiska/byråkratiska apparaten efterfrågar¹. Utvärderingsaktiviteter och -resultat utföres och avgives i de konflikter som där utvecklas och utspelas. Vilka olika funktioner utvärdering kan spela i dessa sammanhang när de nu uppenbarligen inte (bara) ingår i rationella kalkyler återkommer jag till nedan.

Utvärderingsefterfrågan har provocerat samhällsvetenskaperna till relevanta, men mycket svårbehandlade frågeställningar. Vi är helt enkelt med knapp nöd kapabla att angripa, än mindre att lösa dem. Det gäller, för att citera Walter Williams, ett av samhällsvetenskapernas svagaste områden: den samhälleliga dynamiken (Williams 1976a:271). Eller enligt en annan uppfattning det filosofiska tanke-handling problemet (Majone & Wildavsky 1978:103). Men trots detta flödar det av utvärderingsrapporter från universitetsinstitutioner, fristående icke kommersiella organisationer, profitinriktade konsultfirmor och byråkratiernas egna utrednings- och organisationsenheter. Som jag redan konstaterat är kvaliteten på utvärderingsstudierna generellt sett låg. Detta har emellertid inte så mycket att göra med att problemen är svåra, ty inte ens de metoder som finnes - även om de alltså inte riktigt förmår klara ut de frågor som ställs - kommer till användning (Gordon & Morse 1975).

Asplund kommer i sin stimulerande granskning av framtidsforskning, som ju till övervägande del vill uppträda som maktens rådgivare, fram till likartade resultat angående de så populära Delfiundersökningarna: Deras "uppseendeväckande låga kvalitet framstår ... på intet sätt som nödtvungen" (Asplund 1979:84).

Det paradoxala i att utvärderingsforskningen (implementeringsforskningen eller vilken skola man nu försvurit sig åt) på samma gång som den är en tillväxtindustri skapar olust hos sina utövare (t.ex. Majone & Wildavsky 1978) sammanhänger enligt min uppfattning med att de frågor som byråkratins och politikernas

¹Det är överhuvud osannolikt att "marknadsefterfrågan" befördrar god kvalitet i detta sammanhang. I fallet framtidsforskning konstaterar Asplund (1979:125): "En marknad ... höjer normalt profiten snarare än kvaliteten". Rossis slutsats: "/T/he state of the practised art /of evaluation A.G./ is improving response to the heightend level of demands for evaluation research" (1972:49) har på det hela taget visat sig hållbar.

hänvändelse med önskemål om utvärdering reser är mycket svåra och komplicerade och på samma gång oerhört relevanta för samhällsvetenskapernas teoretiska och metodologiska utveckling, medan däremot beställarna inte fäster något större avseende vid om de svar som producerats inom utvärderingsverksamheten lever upp till vetenskapliga kvalitetskrav eller, ännu mindre, förmår att bemästra tidigare olösta problem i sanningssökandets svåra konst.

Denna förmodan - att de frågor som förvaltning och politiska instanser ställer till utvärderingsforskningen mera reses för att svarets form (dess pretenderade vetenskaplighet) och innehåll skall kunna passas in i maktkamper än för att genuint undanröja ovetskap¹ - föranleder mig att diskutera utvärderingsaktiviteternas möjliga uppgifter/funktioner och samband med teoretiska/metodologiska problem: att se lite närmare på paradoxen.

¹Bernstein & Freeman (1975:x) "Currently it is mostly doing evaluation research studies, not their results, that provides the added impact to the complex political mosaic that surrounds important policy decisions".

UTVÄRDERINGSFORSKNINGEN SOM POLITIKENS REDSKAP OCH SOM SAN-
NINGENS APOSTEL - ERFARENHETER FRÅN DEN AMERIKANSKA FRONTEN.

Den idémässiga grundvalen för användningen av utvärderingsforskning (av vetenskap överhuvud) i politik är - om vi inte har att göra med ren cynism - tron på att "fakta" och rationell argumentation spelar roll. Detta är ett problematiskt antagande. Politik handlar inte primärt om sakskaäl utan om maktkamp och hävdande av intressen, där argument och analys understundom kan användas som maktmedel - as a tactic in the play of power (Weiss 1973 efter Lindblom 1968). Likväl hävdar jag här att vetenskapen har en potentiellt frigörande och kritisk förmåga. Förhoppningen om en rationalisering av politiken genom att vetenskapliga metoder och planeringstekniker introduceras i redan existerande institutioner under nuvarande förhållanden förefaller emellertid illusorisk. Vetenskap om politik och samhälle skulle i bästa fall kunna fungera avmystifierande.

I (själv)reflexionerna över den amerikanska utvärderingsforskningen har flera värdefulla synpunkter presenterats. Betydelsefulla slutsatser kan dras om denna typ av forskning - och om forskning överhuvud - i tjänst hos offentliga maktcentra. I sina huvuddrag är dessa enligt min uppfattning generaliserbara till andra västländer.

I föregående avsnitt har jag redan tjuvstartat med denna uppgift, dvs. att ur (diskussionen om) utvärderingsforskningens praktiska kunskaper om i synnerhet effektmätningens (mätning, uppskattningen av reformers och politiska programs effekter) problem men också möjligheter. Denna tjuvstart består däri att jag presenterat den amerikanska utvärderingsforskningen från ett speciellt perspektiv så att dess paradox blivit tydlig: den mellan att å ena sidan "lyckas" och å andra sidan alltid känna missnöje och olust.

Paradoxen blir möjlig att förstå, men inte att upphäva, mot bakgrund av den omnämnda idémässiga grundvalen för vetenskapens användning i politiken. Denna uppfattning av utvärderingsforskning som ett rationellt företag har givits många olika

formuleringar. "The assumption is that by providing 'the facts', evaluation assists decision-makers to make wise choices among future courses of action" konstaterar Carol Weiss (1973:1). Floden och Weiner menar också de att evaluering antas frambringa information som behövs vid rationellt beslutsfattande (Floden & Weiner 1978:9) och karakteriserar detta (efter Martin Rein) som en "decisionistic" modell av hur evaluering fungerar. Detta betyder i korthet att målen för politiska program antas vara stabila, ej föremål för konflikt och öppet deklarerade. Vidare anses evalueringen tjäna syftet att samla information om hur väl programmet svarar mot de formulerade målen och slutligen förutsätts beslutsfattarna arvända evalueringsinformationen i identifierbara, diskreta beslut i syfte att förbättra programmet (Floden & Weiner 1978:10). Rein och White talar för sin del om "a problem-solving image" inom policy-forskningen. "Belief in policy research, according to this image, appears to rest on the conviction that policy expresses a theory of action that includes: (1) a definition of the problems, (2) a set of possible courses of action, and (3) a goal or goals one seeks to achieve. The role of policy research is to a) provide data that clarify the problem b) conduct pilot projects or planned-variation studies to estimate whether the conceived courses of action approach the goals, c) evaluate implemented programs to determine if, and under what conditions, they are achieving goals (Rein & White 1977:249). Berk och Rossi finner att "evaluation findings are put forward as judgments above politics, based on scientific data, not organized interests. The implicit message is that if the 'truth' can be revealed decision makers will respond to it" (Berk & Rossi 1976:342). Några principiella problem att realisera förbättringarna antas inte föreligga inom den traditionella policy analysen eller som Nelson konstaterar: "Implicit in its basic characterization of problems and solutions is the image of someone or some collective mechanism that is actively involved in steering, making the policy decisions that guide the ship of state. The image involves both a steerman and a steering wheel well connected to the rudder" (Nelson 1974:381). Den tilltagande desillusionen om utvärderingsforskningens betydelse hos forskare sammanhänger säkerligen därmed

att "research findings were both sought and not used"¹, i varje fall inte på det sätt som en rationalistisk/decisionistisk modell förutsätter. Det är inte ovanligt, men en smula löjeväckande, att förslag till förbättringar i användningen av utvärderingsforskning som utgår från att detta synsätt kräver att politiken eller den politiska processen skall förändras för att bättre passa forskningen². Utvärderingsforskningen har varit föga framgångsrik om man betraktar den ur ett sådant perspektiv³. Men det finns andra skäl till varför utvärderingsverksamhet och -resultat efterfrågas och denna modell av utvärderingsforskningen är inte betydelselös för att denna "andra" efterfrågan verkligen skall kunna tillmötesgå. Rein och White skriver träffande att "the problem-solving image is a myth with the positive and negative properties of most myths" (Rein och White 1977:262). Positiv och negativ för de närmast berörda frestas man tillägga. Evaluerarna erhåller arbetsuppgifter och politiker/administratörer kan stärka sina positioner, vilket ur deras synvinkel kan betraktas som positivt. Negativ blir myten om den tas på allvar av utvärderarna genom besvikelse över hur resultaten används och av politiker/byråkrater om de handlar som om modellen vore verklighet.

Hellstern och Wollman har ställt samman en "katalog" över funktioner som utvärderingsforskningen (och annan uppdragsforskning) kan ha. De bygger på andra analytikers resultat (Hellstern och Wollmann 1977). Några former av närmast rituell evaluering (efter Suchman 1967:143) är 1) ögontvätt (eyewash): bara positiva programaspekter väljs ut. 2) vittvätt (whitewash): förnekande av objektiva resultat. 3) undervattensvariant (submarin): undersökningen syftar redan från början till att avskaffa det utvärderade programmet. 4) posering (posture): skenbart vetenskapligt före-

¹Rein och White (1977:243), som refererar resultat som Caplan et.al. (1975) kommit fram till.

²Jfr. t.ex. Rein och White (1977:245) som identifierar fyra strategier för förbättringar varav två innebär förändringar av politikens arbetssätt. Dels kravet att politiker skall formulera sina mål så att de blir möjliga att undersöka, dels att besluten skall centraliseras mera.

³"As a matter of record, relatively few evaluations had a noticeable effect on the making and remaking of public policy", Weiss (1973:4). "Federal program evaluations so far have been largely ineffective." Rein och White (1977:241).

havande utan att innehållsligt leva upp till skenet. 5) förhållning (postponement): forskningsuppdrag för att förhindra annan aktivitet. 6) ersättning (substitution): koncentration på undersökning av oviktiga men försvarbara detaljproblem. Vidare nämnes de fyra intressen som en förvaltning kan hysa för uppdragsforskning (efter Thompson 1975:16 och Bartholmæi 1977:286): 1. tidsvinst, 2. legitimitetsbehov, 3. systematisk framställning och vetenskaplig beskrivning, och 4. restkategori med t.ex. statussymbol, förbrukning av kontoöverskott. Slutligen anger de (efter Bandura & Gross 1976:249ff och Bandura 1976:15ff) fem funktioner hos vetenskapen: 1. upplysning, 2. diagnos, 3. terapi, 4. skiljedömande, och 5. manipulation. Floden och Weiner pekar bl.a. på utvärderingars betydelse som konfliktlösningsmekanism mellan eliter och dess förmåga att dämpa övermodet hos programföreträdare (Floden & Weiner 1978:15).

Ett av de minst subtila sätt genom vilket beslutsfattare avviker från den rationalistiska beslutsmodellen är att helt enkelt bara bry sig om sådana forskningsresultat som stämmer överens med deras egna förutfattade meningar. Tendenser i denna riktning omnämns av Weiss (1973:9). Men evaluering fungerar också som symboliska gester (i stället för handling), grundliga undersökningar vinner tid medan handlingskraftig image kan upprätthållas (Rein & White 1977:255). Men utvärdering kan också kastas in i mycket komplicerade policylägen där verkningarna kan återfinnas långt utanför det utvärderade programmet. Rein och White nämner den berömda utvärderingen av Head-Startprojektet som exempel på detta och menar att undersökningen skapade en opinion enligt slogan "nothing works" när det gäller sociala reformprogram överhuvudtaget (Rein & White 1977:253). Det aktuella programmet stoppades inte men anslagen ökades ej heller och effekterna av utvärderingen förefaller vara viktigare när det gäller reformpolitiken i stort än just beträffande det aktuella projektet¹.

¹Detta förhållande är av principiellt intresse för diskussionen om hur effekter skall identifieras, dvs. att de inte bara återfinns just där en åtgärd sätts in utan på olika sätt fortplantar sig i den sociala kontexten eller utgör skenmanövrer för att nå resultat på helt andra fronter.

Ett av de starkaste korten evalueringsforskningen har att komma med är sin vetenskapliga auktoritet - oavsett om undersökningarna i realiteten uppfyller rimliga krav på vetenskaplighet eller ej. Detta kort är eftertraktat på avnämarsidan och det har hävdats att det just är genom det evalueringen har sitt största inflytande¹. Denna auktoritet intar säkerligen en nyckelroll då evalueringen uppfyller sin rituella funktion så som Floden & Weiner har uppfattat den, nämligen att stilla medborgarnas ängslan och frammana en bild av ledningens rationalitet, effektivitet och pålitlighet. Evalueringen ger intrycket av att regeringen kan handskas med svåra situationer, vilket i sin tur - om medborgarna låter sig övertygas därom - faktiskt gör styrning och kommentering mera effektiv och detta helt oavsett om resultaten av evalueringen verkligen kommer till användning eller läggs i byrå-lådan efter att ha kungjorts inför allt folket.

Utvärdering som ett medel i taktiskt spel eller som (med eller utan någon agents avsikter) uppfyllande vissa funktioner (vid sidan av de som den rationalistiska beslutsmodellen utpekar) fungerar bäst om dessa funktioner inte uppdagas och dras fram i dagsljuset. Så kan man sammanfatta resultaten från de refererade granskningarna av genren. Här kan vi alltså se en aspekt av det rationella beslutsfattandet, närmare bestämt dess mytiska karaktär och med Floden och Weiner konstatera att "the ritual functions of evaluation are best fostered by not acknowledging their existence" (Floden & Weiner 1978:17).

Det blir nu lättare att förstå varför utvärdering och effektmätning så sällan använder sig av de, visserligen ej fulländade men dock, tillgängliga vetenskapliga metoderna och istället klart underskrider relativt elementära vetenskapliga krav. Avnämarna är helt enkelt inte (primärt) intresserade av att få tillförlitliga svar på de frågor som ställs till utvärderare. Det är innehållet - att det ligger i linje med beställarens intresse - och/eller formen (det - ofta förment - vetenskapliga) hos svaret som intresserar snarare än dess tillförlitlighet.

¹"Ironically, it is the scientific 'authority' of evaluation research that produces its greatest impact". Berk och Rossi (1976:342).

Mot denna bakgrund blir Colemans karaktäristik av tillämpad forskning som primärt - p.g.a. avnämarnas krav - intresserad av korrektheten i prediktioner och resultat mindre träffande eller rent av missvisande. Kriterier som sparsamhet och elegans sägs utmärka disciplinforskning till skillnad från policyinriktad forskning som gärna tar emot överskottsinformation (Coleman 1973, här citerad efter Gordon & Morse 1975:351). Coleman utvecklar sitt resonemang vid ett senare tillfälle och menar att det går en skarp gräns mellan policyinriktad och vetenskaplig utvärdering. Skillnaden består däri att den förstnämnda sägs vara inriktad på att tjäna som vägledning för handling och den senare på testning och utveckling av teorier (Coleman 1975:21). Detta förefaller bäst kunna liknas vid normativa utsagor eller önsketänkande.

Den tillämpade forskningen ställs knappast inför några stränga och påträngande krav om korrekthet hos prognoser och resultat. Inom disciplinforskningen - så långt det nu är möjligt att ta ställning till så svepande omdömen - förefaller vissa förfaringsätt vara utslagsgivande. Det är emellertid enkelt att påvisa förekomsten av en slags metodisk formalism även inom utvärderingsforskningen, i varje fall i den litteratur som diskuterar denna forskning och hur dess metoder skall kunna förbättras. Det är ingen överdrift att påstå att evalueringsforskningens metoddiskussion till stora stycken rekapitulerar vissa delar av samhällsvetenskaplig diskussion om undersökningsdesign och statistiska bearbetningsmetoder¹ delvis med annan terminologi². Gränsdragningen mellan tillämpad forskning och disciplinforskning blir därför svår i fallet utvärderingsforskning, i varje fall på det sätt som Coleman föreslår. Inte heller som en normativ rekommendation förefaller den vara välbetänkt. Om det verkligen ställdes effektiva krav på den tillämpade forskningen att utveckla tillförlitliga, korrekta prognoser och resultat kan jag knappast tänka mig någon effektivare hjälp i utvecklandet av teorier (som Coleman menar är disciplinforskningens uppgift).

¹Se T.ex. Wholy et. al. (1970:88) och de nästan obligatoriska rekapitulationerna hos Hellstern & Wollmann (1978), Bell (1975) och Dahlberg & Lindkvist (1976).

²Men som en redan refererad studie visat betyder detta inte att de genomförda studierna lever upp till konventionellt accepterad standard.

De många olika funktioner utöver eventuella "rationalistiska" som utvärderings- och överhuvud taget tillämpad forskning har/kan ha inom politik och förvaltning är att hänföra till den omständigheten att resultaten från denna forskning används i politiska sammanhang, i situationer av öppna eller övertäckta konflikter. Det går inte att bortse från den politiska kontext som resultaten förs in i. "Real conflicts or interests ... cannot be dissolved by sweet rationalism"¹. Resultatens användning politiseras (Weiss 1973:1, Rein & White 1977:261).

Vore det bara användningen av forskningsresultaten som politiserades skulle detta vara ett ringa bekymmer. Poängen med tillämpad forskning är ju just att den kommer till användning och resultat som en beställare inte finner tillfredsställande för den egna ståndpunktens befrämjande kan då användas av andra parter. Men det är inte bara resultat-användningen som politiseras. Även arbetet, dess uppläggning, metodval och utvärderarnas tolkningar av genomförda undersökningar - resultatpresentationen - "politiseras", påverkas eller får sin betydelse av det sammanhang utvärderaren och den utvärderade verksamheten befinner sig i. Konflikter och slitningar i samband med uppläggning och genomförande av utvärderingar är därför vanligt förekommande². En utvärdering som sker internt inom en organisation och på ledningens uppdrag uppfattas, kanske inte obefogat, som kontrollförsök av de undersökta (Gordon & Morse 1975, Bell 1975:63). En organisation som utsätts för utomstående granskning utan att utvärderarna är sympatiskt inställda till verksamheten är potentiellt hotad, särskilt om undersökningen utförs på kontrollerande och/eller anslagsgivande instansers uppdrag³. Genomförs däremot utvärderingen internt eller av personer med anknytning till verksamhetens ideologi är faran att resultaten skall kunna användas som förevändning för makt- eller resursinskränkning betyd-

¹Nelson 1974:410. Se också t.ex. Bell 1975).

²T.ex. det mångomtalade Head-Start projektet. Se Williams & Evans (1972). Redogörelser på svenska finns t.ex. hos Bell (1975:66ff), Lind (1979b:23ff).

³"Evaluation research is a signal that program under study is subject to further negotiation" Floden & Weiner (1978:13). "Administrators are not very enthusiastic about public evaluations." Gordon & Morse (1975:353).

ligt mindre. Inte heller resultaten är oberoende av utvärderingens institutionella inplacering. Det är detta man vet när det kämpas om utvärderingsaktiviteternas uppläggning och innehåll.

Vetenskaplig auktoritet förefaller alltså vara den hårdvaluta som politiker och administratörer vill komma i besittning av när utvärderingsforskning (inklusive dess systemgenrer) inkallas. Men hur klarar vetenskapen själv av att hävda sin auktoritet? Uppenbarligen mindre väl. Förslag har rests om att uppdrag endast borde accepteras när de yttre omständigheterna medger tillräckliga betingelser för vetenskapliga undersökningar av programs effekter, t.ex. att målen inte förskjuts under resans gång. Det framstår emellertid som bakvänt att låta tillgängliga metoder bestämma vad som skall undersökas, alternativt ställa krav på verkligheten att den skall förändras för att passa metoderna. Naturligare förefaller det vara att försöka utveckla metoder som kan användas för att möta uppkomna och formulerade problem. Att så inte skett sammanhänger med att effektmätningens problematik är mycket svårbehandlad och att forskarsamhället, egalt om det gäller disciplin- eller utvärderingsforskning, har så lite av vetenskapligt hållbar auktoritet att erbjuda när konkurrerande intressen gör sig gällande.

Men det är inte bara yttre influenser som hämmar utvärderingsforskningens potentiella avmystifierande insatser. Dess metodologiskt sett svaga position för nämligen inte bara med sig att resultaten blir vidöppna för tolkning och motståndskraften gentemot intressentpåtryckningar svag. Nej, själva de metoder och undersökningsdesigner som ofta används i utvärdering bär med sig ideologiska inslag som till vissa delar står i kontrast med genrens allmänt sett, i varje fall tidigare, reformpositiva uppfattning (Weiss 1973:vi). Det är metoder och designer som förnekar eller förminskar möjligheterna för världen och politiken att se annorlunda ut än vad den faktiskt gör. Detta känner vi från kritiken av positivism och empirism¹. "/M/ost common designs posses an

¹Se t.ex. Gullberg (1972). Se också hänvisningen hos Gordon & Morses (1975:350) till Barbers (1962:37f) karaktäristik av samhällsvetenskaperna som fortfarande befinnande sig i en helt empiristisk tradition.

inherent conservative bias" (Berk & Rossi 1976:339)¹. Berk och Rossi konstaterar att de vanliga utvärderingsmetoderna tenderar att bortse från historiska och sociala processer. Genom forskningsmetodik opereras det undersökta programmet ut ur den samhällsliga verklighet i vilken det ingår. Strävan efter s.k. intern validitet har drivit fram forskningsmetoder som gör sociala program ortogonala till andra sociala krafter som kan orsaka olika resultat hos experiment- och kontrollgrupp, enligt vad Berk och Rossi (1976:340) menar i en talande vändning.

De problem som en vetenskaplig verksamhet inriktad på att uppskatta eller mäta effekten av planering och politiska program har kan sammanfattas i två typer. För det första gäller det den institutionella inramning som utvärderingsforskningen ges: utvärderingsforskningens organisatoriska problem. För det andra rör det sig om teoretiska och metodologiska frågor, frågor som i första hand sammanhänger med svårigheterna att urskilja effekter av det undersökta programmet i det samhällsliga vimlet och att fixera vad programmet (som vanligen ändrar innehåll med tiden) faktiskt är². De två problemtyperna kan uttryckas i termer av kontexter, där organisationsfrågan har att göra med det undersökande subjektets (evaluerings-, evaluerarens) inplacering i politik, förvaltning, maktspel. Ur denna kontext kan det som Gordon och Morse kallar externa kriterier för evalueringsforskningen härledas. Metodfrågan däremot handlar om de samhällsliga sammanhang som den undersökta verksamheten a) griper/avser att gripa in i (impactsidan) och b) ger uttryck för, är en reaktion på (inputsidan). Överväganden om hur den undersökta verksamhetens relationer till och ställning i dessa båda kontexter skall kunna undersökas ger det som Gordon & Morse (1975:351) benämner interna kriterier.

¹Författarna fortsätter: "When they are imposed on existing programs with treatment content fully determined, the resulting findings are obviously constrained by empirical reality already in place. Alternative social interventions that have been rejected for implementation cannot be evaluated ..." Se också Nelson (1974:391) om konservatism.

²"Outcomes of the programs are obscured by the complexity of the social context in which they take place; furthermore, the public goals often change during the course of the evaluation". Floden & Weiner (1978:11).

De båda frågekomplexen - och försök som skall ha några utsikter att bemästra dem på ett fruktbart sätt - är nära förbundna med varandra (Berk & Rossi 1976:338,348). De krav som ställs på utvärderingsforskningen från uppdragsgivare och andra intressenter kan tillgodoses i olika utsträckning beroende på vilka metoder som används och vilka signifikanskrav som väljs. Gordon och Morse har t.ex. funnit att om mindre rigorösa metoder utnyttjas så tenderar resultaten att bli mer positiva till den utvärderade verksamhetens prestationer. Genom att ställa höga krav på statistisk signifikans kan sannolikheten för att ett program skall befinnas vara utan effekt ökas. De yttre kraven på utvärderingsforskningen har en relativt komplicerad relation till de inre. Vad behövs det för metodologiska ansatser för att utvärderingsforskningens uppdragsgivare skall bli nöjda? Gordon och Morse intar en mycket optimistisk hållning till de externa kriteriernas potentiellt positiva inverkan och refererar undersökningar av medicinens utveckling: "the major break-throughs occurred through problem-oriented research." (Gordon & Morse 1975:351). De refererar också Weinberg (1964) som är av åsikten att interna kriterier är en nödvändig förutsättning för vetenskaplig växt och att externa kriterier utgör tillräckliga betingelser för denna. Enligt min uppfattning är detta en överdriven idyll. Inte minst när det gäller utvärderingsforskning är situationen problematisk så tillvida att - som vi redan konstaterat - det mera är själva aktiviteten och mindre assistans att lösa vissa formulerade problem som efterfrågas¹. De externa kriterierna innehåller ett slags dubbelt budskap: frågor ställs till utvärderingsforskningen. Om dessa frågor gällande reformers och politiska programs effekter togs på allvar skulle det innebära en oerhörd utmaning till revolution inom samhällsvetenskaperna². I praktiken är dock, som vi konstaterat, frågorna av underordnat intresse. Trycket på utvärderarna att producera

¹Evalueringen som ursprungligen sågs som tekniskt hjälpmedel spelar i stället en betydande roll i formulerandet av den reformnegativa hållningen: "nothing works". Se Rein & White (1977:271). Evalueringsforskningen politiseras allt mer enligt Weiss (1973).

²"The moment one begins to take demystification seriously (vilket är liktydigt med att det svåra sanningssökandet kommer i första hand - A.G.), a number of methodological issues appear". Berk & Rossi (1976:344).

tillförlitliga resultat kan efter de undersökningar som gjorts av genren inte annat än betraktas som minimalt. Hur skall det annars förklaras att mindre än 20 % av undersökta studier når upp till minimala krav på tillförlitlighet vad gäller design, datainsamling eller analys? (Bernstein & Freeman 1975). Det finns snarare anledning misstänka att de ickerationalistiska funktioner utvärderingsgenren fyller inom politik och förvaltning motverkar snarare än stöder en sanningssökande aktivitet. Varför betala dyra och vetenskapligt djupt penetrerande, teori- och metodutvecklande undersökningar som verkligen försöker ge tillförlitliga svar på formulerade frågor och som dessutom riskerar att gå uppdragsgivaren emot, när en snabb konsultprodukt bättre kan fylla behoven av legitimitetsstöd i förhandlingar och konfliktsituationer?

De yttre kraven på utvärderingsforskningen är inte så entydigt produktiva som man med Gordon och Morse kunde hoppas. Deras negativa hållning till inre kriterier, att de förefaller förstärka status quo, är även den tvivelaktig. Sant är visserligen att gängse metoder inom utvärderingsforskningen har en konserverande, ohistorisk och ostrukturell slagsida (Berk & Rossi 1976). Det betyder emellertid inte att det skulle vara omöjligt att formulera interna kriterier som syftade till att övervinna dessa begränsningar, dvs. inrikta sig på att formulera sådana teoretiska och metodologiska strategier för bedömning av reformers effekter som inte helt lyfter dessa ut ur sina samhällliga kontexter. Utan hopp är inte en sådan strategi, ty den samhällsmodell som gängse utvärderingsmetoder implicerar, dvs. ett samhälle bestående av varandra ömsesidigt relativt oberoende delar stämmer föga överens med vad många samhällsvetare menar, bara det att deras hållning inte förts in i utvärderingsforskningens praxis (Berk & Rossi 1976:341). Utvärderingsforskningens misslyckande (om vi talar i andra termer än kommersiella) beror på att interna och externa kriterier inte samtidigt kunnat betonas och tillgodoses menar Gordon och Morse (1975:351)¹. Detta är sant. Utan yttre krav på sanningssökande och utan inre krav på teori och metod förmåga att tjäna detta syfte på effektmätningens svåra område får vi inte veta vad reformförsök och planeringsansträngningar förmår. Det betyder emellertid inte att det förbund på mytisk grund,

¹Utvärdering spelar ingen roll har det hävdats.

det "olustiga äktenskap" som politik och vetenskap ingått genom utvärderingsforskningen (Rein & White 1977:268), vare sig upplöses eller övergår i en harmonisk fas. Nej, denna genre kommer säkerligen även fortsättningsvis att frambringa "dilemmas that threaten continually to erode the responsibility and credibility of each partner in the match." (Rein & White 1977:268). Vad vi kan hoppas på och som jag kommer att försöka ge ett bidrag till, efter en genomgång av utvärderingsforskningen i Sverige i allmänhet och två effektmätningstudier i synnerhet, är att ge en något större plats för, vidga stödet - både externt och internt - för en utvärderings- (och då i synnerhet en effektbedömnings)forskning med avmystifierande, sanningsökande ambitioner¹. I politik och förvaltning borde denna typ av studier vara av särskilt intresse för en bredare allmänhet som inte har investerat existens och prestige i ett speciellt reformprogramms fortlevnad eller avskaffande.

Den andra sidan vetter mot vetenskapens utveckling. Inte utan skäl ser analytiker de stora teoretiska problemen skymta fram bakom evalueringsforskningen² och trots att jag inte instämmer i premisserna till Gordons och Morses påstående att "evaluation research ... is directly related to the development of the social sciences in general ..." ³ tror jag det blir en sanning om, och bara om, effektmätningens problem tas på allvar.

¹Det är att söka en väg som undviker de vanliga förslagen till reformer av "the research-policy alliance offered under the problemsolving image /, that/ generally have the effect either of taking research further away from policy relevance or of changing the basic nature of political processes". Rein & White (1977:268).

²Body-mindproblemet förs på tal av Rein och White (1977:261) och förhållandet mellan tanke och handling av Mayone och Wildavsky (1978).

³Gordon & Morse (1975:351). Lind (1979b:9) menar att "/u/tvärdering inte bara /är/ ett sätt att förbättra olika samhällsreformers verkningsgrad. Genom den ökade probleminriktningen och kraven på tvärvetenskaplighet kan också den vetenskapliga kunskapen om komplexa samhällsfenomen öka."

UTVÄRDERING I SVERIGE - INHEMSK MODELL OCH AMERIKAIMPORT

"I Sverige ligger vi som vanligt några år efter USA." (Lind 1979b:6). Så förhåller det sig när det gäller modedefenomen som hamburgerkedjor, jogging och travolteri. Utvärderingsvågen kan också inräknas bland modedefenomenen. Men även om metoden att undersöka den mest "avancerade" nationen för att bättre kunna föreställa sig de vägar andra länders invånare skall komma att fösas in på har använts av samhällsforskare av annan kaliber än den citerade observatören av utvärderingsforskningen "over there"¹, så bör man nog besinna sig en smula innan rollen som amerikansk modeattaché accepteras. Ty att Sverige ligger efter i t.ex. utvärderingssvängen, dvs. först något senare greps av den nya flugan, behöver ju inte betyda att det nya flygfäet, även om det har kromosombabyns växtkraft, i alla avseenden är överlägset de inhemska stammarna, att dessa är efterblivna. Kanske är de amerikanska inte heller så annorlunda, bara lite mer uppblåsta, självsäkra och modemässiga?

Varför saknas utvärderingsforskning och -tradition i Sverige frågar sig Bell (1975:156,16) och varför är, som ERU konstaterar (SOU 1978:47 s. 171) effektmätningar en sen företeelse i Sverige? Så sent som 1979 sände OECD en man till nationen för att utröna hur det kommer sig att Sverige är så fattigt på evalueringer². Var man än kommer att landa i sina beskrivnings- och förklaringsförsök när det gäller (bristen på) utvärdering i Sverige står ändå ett klart: någon utvärderingsrörelse med egna tidskrifter, konferenser, sammanslutningar³ och en egen yrkesidentitet har inte växt fram och som någon ny framstormande bransch på finanshimlen kan den

¹Marx' analys av den engelska kapitalismen motiverades som bekant med ett resonemang som har vissa likheter med det här anförda. För en kritik se t.ex. Giddens (1973).

²Muntlig information från forskningssekreterare Ingrid Lundahl, Statens råd för byggnadsforskning. Mannen ifråga var professor Jamous.

³Texten är från 1979. Numera finns ett embryonalt kontaktnät på nordisk basis.

inte heller fångas i kikaren¹. Möjligtvis kan den av allt att döma tilltagande förekomsten av konsultbolag på fickan hos forskare inom relevanta discipliner och hos högre tjänstemän dock tolkas som en första utveckling i riktning mot en kommersialisering av utvärderingsverksamheten (Bell 1975:16). Utvärdering syftar ibland /enbart/ på ex-anteundersökningar av övervägda handlings- eller planalternativ. Ibland talar man i detta sammanhang också om konsekvensanalys eller effektmodeller. Denna typ av utvärdering i form av förhandskalkyler är inte alls främmande i svenskt utredningsliv. Det är i synnerhet ekonomer som (liksom i USA) stått för denna genre med signaltermerna cost benefitanalyser, programbudgetering och - i Sverige också - positionsanalys². Ex-anteanalyserna har här, liksom i USA visat sig mindre fruktbara eftersom de inte förmått göra genomförandet till något problem i analysen³.

På senare tid har cost-benefitmetodernas begränsningar uppmärksamrats även bland dem som varit med om att utveckla deras kanske mest konsekventa användning i svensk förvaltning, dvs. Vägverket⁴. I ex-anteutvärderingarna är influenserna och inlånen

¹Uttalanden av mer eller mindre försiktig profetisk karaktär åt detta håll hos t.ex. Stevrin (1974) och Bell (1975) har inte riktigt slagit in.

²Se t.ex. Andersson (1978) med vidare referenser. Se också Bohm (1972), Mattsson (1970) och SOU 1967:11-13.

³Typiskt är Åhréns (1975:10) antagande att de planer som beslutas också kan genomföras.

⁴Ahlstrand (1977:53) menar att "/d/et räcker inte för tekniker, ekonomer och arkitekter att, som nu är vanligt, värdera vissa konsekvenser, kvantifiera andra och beskriva ytterligare andra ..." Detta är Ahlstrands personliga konklusion efter, som han själv skriver "tio års arbete med samhällsekonomiska bedömningar, fem års arbete med utveckling av utvärderingsmetodik för väginvesteringar vid vägverket och mångåriga kontakter med politiker ..." Det kritiska tonfallet är uppenbart. Sedan är det en annan sak att Ahlstrand tror att bot och bättring står att finna i att expandera metodernas pretentioner ytterligare. Utvärderarna skall inte "skjuta över ansvaret för att greppa helheten till politikerna." Detta skall utvärderarna själva göra. Besvikelserna efter sextitalets hybris i USA på den rationella beslutsanalysens område borde ge Ahlstrand anledning att göra helt om. Men det finns en rationalistisk planeringsoptimism som inga erfarenheter tycks bita på. Jfr. Gullberg (1981).

från den anglosaxiska världen tydliga.

Utvärderingar av pågående verksamheter eller i efterhand har - om man inte formulerar mycket snäva kriterier för vad som skall betraktas som utvärdering - en lång historia inom svensk förvaltning och politik. Jag tänker framför allt på det svenska utredningsväsendet med SOU-serien. En stor del av arbetet består i bedömningar av rådande förhållanden och pågående verksamheter som underlag för eventuella beslut om reformer eller justeringar av lagstiftning eller förvaltning.

Man kan till och med som Esping (1975:1f) hävda att "vi har ganska institutionaliserade inslag av utvärdering inom i varje fall hela statsförvaltningen och stora delar av den kommunala och landstingskommunala verksamheten". Denna sker i budgetprocessen, inom kommittéväsendet, genom departementens halvpermanenta expertgrupper som ERU och EFA, i riksdagsrevisorernas arbete och genom RRV:s granskningar och Statskontorets rationaliseringsråd. Det är denna institutionalisering av utvärderingen som inledningsvis föranledde mig till omdömet att den svenska modellen i vissa avseenden står framom den amerikanska. Uppräkningen ger ett imponerande intryck och ändå stämmer Brunsson och Jönssons omdöme "Uppföljning av vilka resultat verksamheter ger är notoriskt dålig inom den offentliga sektorn" (Brunssons & Jönsson 1979:40) väl överens med mina och många andra bedömares uppfattningar. ERU och EFA återkommer jag till i en särskild granskning av deras effektmättningsförsök. Men först till frågan om det inte inom de svenska universiteten utvecklats ambitioner och forskningsmetoder för att ge svar på frågor om reformers och planeringsansträngningars effekter eller med en något annan formulering: utvärdering i efterhand (ex-post). Först skall konstateras att sådana ambitioner existerar och att de ingår i argumentrepertoaren när ämnesföreträdare inom de olika samhällsvetenskaperna hävdar de egna verksamheternas samhällsnytta. Evalueringens historieskrivare i USA pekar som vi sett ovan på en lång tradition av tillämpad samhällsvetenskaplig forskning som förstadium till den nya forskningsgrenen. Samhällsvetenskaplig forskning i Sverige kan också ses i detta perspektiv även om efterhandsbedömningar inte kan sägas ha varit något centralt

tema. Men uttryckliga programformuleringar är inte omöjliga att leta fram, särskilt inte om man betraktar de senaste åren. Jag illustrerar detta med några relativt slumpmässigt valda exempel.

På den pedagogiska institutionen vid Stockholms universitet startades 1965 en utbildningslinje på försök. Den gick under benämningen SPUT-linjen (Socialpedagogisk utbildningsteknisk linje) och syftade till att utbilda personer kunniga i att undersöka pågående sociala reformprogram och utbildningsverksamheter samt att justera dem. Professor Arne Trankell introducerade forskningsmodellen action-cum-research som innebar att verksamheterna undersöktes mitt under pågående reformarbete och att forskaren aktivt borde gå in med rekommendationer under processens gång, därigenom undanryckande grunderna för sina egna negativa prognoser.

Ett annat exempel är doktorsavhandlingen "Om påföljders verkningar" från 1966 där Bengt Börjesson arbetar just med effektproblematiken. Han ställer frågan om på vilket sätt frihetsberövande påföljder påverkar återanpassningsutsikterna hos "i övrigt lika" individer. Hans metodik gick i korthet ut på att skilja ut personer som fått frihetsberövande påföljd från dem som ej fått sådan påföljd. Individerna grupperades i riskkategorier (för återfall) varefter återfallsfrekvensen för de olika grupperna jämfördes med konstanthållande av riskkategori¹. Börjessons problem gick ut på att isolera inflytandet från påföljdstyp (frihetsberövande - ej frihetsberövande) från all annan inverkan på återfallsbenägenheten.

"Studiet av effekter av politiska beslut är ingen nyhet i svensk statskunskap" skriver Sannerstedt (1979:22) och anför det stora kommunalforskningsprojektet under professor Jörgen Westerståhls ledning samt projektet "Statlig regional förvaltning", där han själv deltagit med en avhandling. De båda anförda projekten är bra exempel på att ambitionen att undersöka effekter av politiska

¹De statistiska metoder som kom till användning utsattes för hård kritik, vilket jag i detta sammanhang inte har någon anledning att gå in på.

beslut har funnits och finns inom statsvetenskaplig forskning i Sverige. Frågan om i vilken utsträckning detta verkligen lyckats diskuteras jag i Gullberg (1984:46-8).

Vi kan alltså dra följande slutsats. Bland universitetens företrädare förekommer i Sverige lika väl som i USA ambitioner att utveckla metoder och positioner inom samhällsvetenskaperna för att svara på frågor om reformers effekter¹. Skall vi tala om skillnader i analys- och utvärderingsverksamhet länderna emellan är det alltså varken förekomsten av ex-anteundersökningar eller ambitionerna inom de akademiska institutionerna att kunna stå till tjänst med svar på ex-postfrågor vi kan anföra, för såvitt vi inte talar i kvantitativa termer.

Avancerade metoder för beslutsförberedelser och för ledning och analys av myndigheter har utvecklats inom Riksrevisionsverket och Statskontoret. Detta utvecklingsarbete har resulterat i en rad publikationer². Uppskattning eller mätning av myndighetsutövandets effekter har emellertid inte stått i fokus. Problematiken berörs mera i förbigående (Statskontoret 1974:112) eller utan att ställas på sin spets³. En förskjutning mot ökat intresse för effektbedömningar kan dock registreras på senare tid⁴ (Riksrevisionsverket 1979). Man kan nog säga att den entusiasm som fanns hos reformatorer av offentlig förvaltning under senare delen av 60-talet som ett utflöde från förebilder i USA, manifesterande sig bl.a. i Programbudgetutredningen SOU 1967:11-13 och Budgetreform SOU 1973:43-46, har dämpats avsevärt. Ambitionerna hos RRV:s förvalt-

¹"Den traditionella akademiska forskningen är ... i allmänhet inte kopplad till beslutssituationer eller beslutsfattande varför effekten ofta blir ytterst indirekt eller uteblir helt" hävdar Esping (1977:18) för Sveriges del. Det är riktigt och strider inte mot mitt påstående att ambitionen formulerats. Se t.ex. Svensson (1973), Franke-Wikberg (1977) och Franke-Wikberg & Lundgren (1980,1981) när det gäller disciplinerna företagsekonomi och pedagogik.

²Riksrevisionsverket (1975, 1976a,b,c,d och 1979), Statskontoret (1971, 1973, 1974 och 1977). Också Esping (1977) och Gorpe (1978) kan på sätt och vis räknas dit.

³Statskontoret (1977:139ff) och Riksrevisionsverket (1975:6:18 och 7:5).

⁴Texten från 1979.

ningsrevision har fått sänkas p.g.a. kollisionsrisk med de politiska beslutsfattarna som uppstår vid en mera ambitiös granskning (Esping 1977) och Peter Gorpe (då) vid statskontoret är tämligen sval vid sin genomgång av vad de "nya" metoderna presterat (Gorpe 1978).

I Sverige finns ex-anteanalyser och utvärderingar ex-post. De finns inom förvaltning, utredningsväsende och akademisk forskning. Ändå känns Bells fråga om varför det inte finns någon utvärderingsforskning och -tradition inte orimlig. Den rörelse inom utvärderingen som Stevrin skriver om måste ha syftat på USA (Stevrin 1974:VI:3). Här i Sverige har tron på de rationella metoderna tagit sig något mera modesta former. Metoderna kastas inte heller in i den politiska kampen på samma sätt som i USA.

Till frågan om utvärderingsforskningens ställning i Sverige och vari skillnaderna jämfört med USA i detta avseende består kan möjligen förekommande presentationer av den amerikanska utvärderingsforskningen för svensk publik ge ett visst bidrag. Genomgången av dessa introduktioner ger mig också möjlighet att jämföra mina egna slutsatser från den amerikanska diskussionen: utvärderingsforskningens (särskilt efterhandsbedömningarnas) dilemma består både i metodologiska/teoretiska svårigheter att mäta effekter och institutionella/politiska bekymmer att placera iakttagaren/utvärderaren på en sådan observationspunkt att inte kortsiktiga politiska intressen skymmer synfältet.

DEN AMERIKANSKA EVALUERINGSFORSKNINGEN PRESETERAS I SVERIGE

Evaluerings- (eller utvärderings)forskningen ex-post i USA har uppmärksamats i Sverige med flera avhandlingar och även i andra publikationer. Hur har dessa introduktörer förhållit sig till denna rätt sammansatta forskningsinriktning? Hur har t.ex. nyhetsanspråken bedömts och eventuella jämförelser med svensk utvärderingspraxis utfallit, hur har de överdådiga löftena tagits upp? Har kritik riktats mot den "nya" genren och har det lyckats presentatörerna att avlocka evalueringsforskningen några hemligheter, i så fall vilka?

Ekonomiska utvärderingsmetoder i Samhällsplanering heter en avhandling från 1975 skriven av Per Åhrén vid Ekonomiska Forskningsinstitutet vid Handelshögskolan i Stockholm. Framställningen begränsas till ex-ante-metoder, dvs. utvärderingar av planer innan de beslutas, varför jag inte kommer att diskutera den vidare här.

I avhandlingen Att utvärdera sociala program ställer sig Robert Bell tre uppgifter: 1. Att presentera en del av den omfattande amerikanska litteraturen om hur man med beteende- och samhällsvetenskapliga redskap försöker bedöma verkningarna av sociala program. 2. Att återge egen försöksverksamhet och 3. Att jämföra litteraturen med egna erfarenheter i syfte att besvara frågan: Behövs formell utvärdering i Sverige?

Bell noterar det uppsving i utvärderingsforskningen som skedde i USA i samband med kriget mot fattigdomen, men utan att den nyhetsiver och de stora löften som understundom förekom bland de "nya" metodernas tillskyndare traderas. Bell kan stödja sig på en framväxande desillusion både över krigets resultat och utvärderingarnas betydelse. Han frågar sig om utvärderingsforskningen har någon framtid i USA och finner detta osäkert eftersom han menar att denna aktivitet förutsätter uppfattningen att det är möjligt och önskvärt att eftersträva en rationell politik. Som framgått ovan har emellertid evalueringsforskningen kunnat fungera utmärkt i ett helt annat ideologiskt klimat och faktiskt levererat

argument eller precisare uttryckt: kunnat användas som förövändning för uppfattningen att ingenting går. Detta var inte så lätt att se 1975 då Bell skrev sin avhandling.

Bell konstaterar också att det politiska systemet (särskilt det socialpolitiska som är det han i första hand inriktar sig på) fungerar annorlunda i USA jämfört med Sverige. Det gäller t.ex. det stora inslaget av frivilliga institutioner i USA, det utbyggda systemet med utredningsaktivitet och remissförfarande i Sverige som givetvis minskar behovet av en formaliserad utvärderingsverksamhet och det gäller också skillnaderna i ländernas storlek. I Sverige är det möjligt att besitta en överblick. Det är möjligt för alla experter att känna till varandras personer och åsikter.

Bells slutsatser om utvärderingsforskningens betydelse är inte överhövan optimistiska eller positiva. En av de viktigaste erfarenheterna från det amerikanska 60-talet är, menar han, att ambitionerna både vad gäller utvärdering och program måst skruvas ner. Men trots allt tror Bell att utvärderingsforskning och andra typer av rationella analyser äger sitt intresse som bidragsgivare till debatten om den offentliga sektorns användning som medel för samhällsförändring. När han diskuterar utvärderingsforskningen är det dock mest dess relationer till politiker och tjänstemän som stöts och blöts.

Problemen med utvärderingsforskningen redovisas så som de dryftades i USA under första delen av 70-talet. Den experimentella metoden betecknas som överlägsen andra metoder när det gäller att undersöka effekter men har nackdelen att den visat sig svår att tillämpa (Bell 1975:73). Det råder en något tragisk diskrepans mellan detta framhävande av den experimentella metoden och de utvärderingar som utförs (inklusive Bells egna). Bell gör inga försök att diskutera eventuella kunskapsmässiga begränsningar hos experimentförfarandet. De metoder som Bell använder - ett slags reflekterade iakttagelser - är rimliga och deras stringens och tillförlitlighet skulle kunna utvecklas betydligt om inte det gamla experimentella idealet blockerade sådana ansträngningar. Bell tar också fram den slutsats som flera författare i den amerikanska diskussionen drog, nämligen att utvärderingar sker i ett politiskt sammanhang, där manifesta och latenta konflikter är den omgivning

i vilken evalueringsforskaren utför sina undersökningar och redovisar sina resultat. Utvärderingsverksamheterna hotar att bli brickor i ett politiskt spel (Bell 1975:63).

Utvärderingens principiella problem definierar Bell som frågan om verkningar, om "impact" (Bell 1975:39). För att diskutera dessa frågor utvecklar Bell en verkningsmodell (Bell 1975:91). "Det problematiska är förstås frågan om hur man skall kunna komma fram till verkningar" (Bell 1975:105). Det gäller att bedöma projektets framgång när det gäller att uppnå givna mål (Bell 1975:136). Samtidigt som Bell här närmar sig en precisering av ex-postproblematiken uppenbarar sig dess traditionella bekymmer: Vilka är egentligen målen? De är nästan aldrig precist angivna. Vad skall man ta sig till för att mäta verkningar när experimentella undersökningar inte är möjliga? Angående den senare frågan pekar Bell på de framsteg som gjorts inom den samhällsvetenskapliga forskningen vad gäller kvasiexperimentella ansatser och manipulationer av data för prövning av kausala hypoteser. Temat om sociala programs positiva och negativa biverkningar diskuteras inte av Bell. S.k. programeffekter är dock ett exempel som nämns (Bell 1975:62).

Förtjänsten med Bells avhandling på det metodmässiga planet är att den formulerar frågan om effekter. Han för också en diskussion om vilka slutsatser som kan dras från olika registrerade utfall på målvariablerna (min terminologi A.G.). Men han når inte fram till en tillfredsställande formulering av relationerna mellan mål, aktiviteter och effekter. Dels blir behandlingen av de formulerade målen väl statisk. "Om man inte kan iakttaga den förväntade effekten, då kan man sluta sig till att programmet misslyckats på denna punkt" (Bell 1975:171). Detta konstaterande är i och för sig triviellt, men borde det inte finnas grader i eländet? Hur långt ifrån målet befinner man sig? Var det kanske lyckligt, t.ex. med hänsyn till nya bedömningar som gjorts sedan målet ställts upp, att det inte realiserades? Även i de fall där den registrerade utvecklingen varit motsatt den önskade kan det aktuella programmet inte utan vidare förkastas. Att det varit otillräckligt i förhållande till uppställda mål är ovedersägligt men att dess effekter ändå varit just i den riktning som målet anger är fullt möjligt. Det kan nämligen tänkas att situationen inom det

aktuella området varit ännu värre om programmet inte varit verksamt. Dels är Bells terminologi något olycklig just vad gäller effektresonemangen. Bell skriver att man "kan ... inte sluta sig till att projektet lyckats om man skulle iakttaga effekten /dvs. det mål som satts upp för verksamheten A.G./". Jag förstår inte det rimliga i att kalla detta effekt, ty effekt borde vara något som uppkommer som (kausalt) resultat av den manipulation som studeras eller av något annat specificerat förhållande. Det är enligt min uppfattning olämpligt att använda en terminologi där man menar att effekter (som helt rimligt för associationerna i riktning mot kausalresonemang) är möjliga att avläsa genom direkt observation. Jag föreslår termen utfall som bättre. Utfallet i ett socialt programs målvariabler behöver inte alls ha något med programmets (kausala) effekter att göra. Programmet kan ha varit verkningslöst just i detta avseende. Det är fullt möjligt att förändringar ändå inträffat i den aktuella variabeln under undersökningstiden t.ex. beroende på att pågående utvecklingsprocesser fortgår. Problemet med utvärderingsforskningen borde istället formuleras som (om jag övertar innehållet i Bells rekommendation) svårigheter att i det registrerade utfallet identifiera vad som är effekter av det genomförda programmet.

En annan förtjänst i Bells arbete består i redovisningen av de egna utvärderingserfarenheterna, där de praktiska problem som uppstår vid tillämpad forskning framställs på ett realistiskt sätt. Reflexionerna i anknytning till relationerna mellan forskare och beställare är upplysande. Presentationen av den amerikanska utvärderingsforskningen är välgjord men inte särskilt inträngande. Bells omdöme om densamma är balanserat avvaktande. Man kan emellertid spåra en viss tvehågsenhet hos Bell i det att han dels menar att utvärderingen har ett värde som en bland flera metoder inom en rationell och reformvänlig politik, dels att han i det närmaste brasklappsmässigt konstaterar att dessa metoder utvecklats som medel för maktutövning med McNamaras krigsplanering som exempel - något som Bell önskar distansera sig från. Bell stannar i sin avhandling vid det omedelbart synliga. Han är inte särskilt sofistikerad i de metodologiska resonemangen och lyckas inte avvinna utvärderingsproblemet någon av dess hemligheter.

Avhandlingen "Programorienterad utvärdering" från 1976 av Jörgen Dahlgren och Lars Lindkvist redovisar dels deras egen utvärdering av kommunal konsumentverksamhet, dels litteraturgenomgång och -diskussion samt försök att utveckla ett eget alternativt perspektiv på utvärdering. Alternativet kan ses som en reaktion mot utvärderingslitteraturens svagheter och svårigheterna författarna själva hade att komma tillrätta med utvärderingsproblematiken i den redovisade studien men också i tidigare egna försök. I utvärderingen av den kommunala konsumentverksamheten stötte Dahlgren och Lindkvist på många av genrens "klassiska" svårigheter. Det gällde t.ex. obefintliga eller vagt preciserade mål. Att genom intervjuer med ansvariga driva fram målpreciseringar visade sig inte möjligt (Dahlgren & Lindkvist 1976:II:126). Undersökarna hade alltså inga tillräckligt väl preciserade förväntade effekter av upplysningsverksamheten att inledningsvis rikta sitt intresse emot. Efter att på annat sätt bestämt på vilka områden effekter av upplysningsprogrammet var att vänta/önska och således värda att efterforska konstaterar författarna hur viktigt det är att kunna skilja ut effekterna av just det undersökta programmet från förändringar i den undersökta variabeln som härrör från andra omständigheter¹. Intresset för författarnas undersökningresultat var litet.

Dahlgren och Lindkvist fann i litteraturen att de inte var ensamma om dessa och andra problem. Men några lösningar stod där inte att få. Vidare föreföll det som om den alldeles övervägande delen av utvärderingsstudierna misslyckades i den meningen att de inte förde med sig några förändringar i verksamheterna. Resultaten användes ej². För att komma tillrätta med denna brydsamma situation gjorde författarna en fördjupad litteraturgenomgång med syfte att identifiera grundläggande föreställningar

¹Det var viktigt skriver Dahlgren & Lindkvist (1976:II:128) "inte bara att konstatera huruvida vissa effekter förekommit eller ej, utan också att dessa effekter entydigt kunde relateras till den kommunala konsumentverksamheten. Detta innebar att vi egentligen var tvungna att isolera effekterna av KKV från t.ex. effekterna av ..." Angående bruket av termen "effekter" i detta sammanhang jämför min kommentar till Bell ovan.

²Dahlgren & Lindkvist (1976:III:1). Detta är konstateranden som kan igenkännas från min forskningsöversikt ovan. Dvs. de kom inte till användning i en rationell beslutsprocess.

och på så sätt få avstamp för en omformulering av problematiken. Det som söktes var nya och bättre lösningsförslag (Dahlgren & Lindkvist 1976:III:2).

De disputerandes infallsvinkel till litteraturgenomgången är intressant. De tar diskussionen om designproblemet hos Alexander i termer av form och kontext som utgångspunkt och menar sig därigenom kunna undersöka överensstämmelsen mellan utvärderingens problem och dess lösningsförsök. De hänvisar också till form/kontext när de skall utveckla sitt eget alternativ. Beklagligtvis blir det inte så mycket av detta uppslag och litteraturen genomgås och sammanfattas i tre punkter, nämligen angående utvärderingens funktioner, de situationer den genomförs i och dess utformning. Slutsatsen angående "överensstämmelsen mellan kontext och utformning för den dominerande föreställningsramen som helhet /blir:/ utvärderingens utformning förefaller delvis vara oförenlig med upplevelsen av utvärderingens kontext" (Dahlgren & Lindkvist 1976:III:55). Inga av de principiellt spännande associationerna som hänvisningen till Alexander väcker tas tillvara. Slutsatsen kan uttryckas mycket enklare. I den utvärderingslitteratur som Dahlgren och Lindkvist gått igenom och refererat på ett förtjänstfullt sätt finns dels ett rationalistiskt ideal, dels en medvetenhet eller upplevelse av att utvärderingsforskningen och beslut angående utvärderade verksamheter inte försiggår i enlighet med rationalistiska principer. Evalueringen inte bara hotar att bli en ny bricka i det politiska spelet (Dahlgren & Lindkvist 1976:III:11), något som författarna betraktar som en av dess mindre legitima funktioner. Utvärdering är (bl.a.) just en sådan bricka.

Den lösning Dahlgren och Lindkvist söker på de identifierade problemen hoppas de finna i vad de själva benämner aktörssynsättet. Vägen dit går över en kritik av rationalistiska besluts- och organisationsteorier. I aktörsperspektivet får utvärderingen, menar författarna, en annan innebörd. "/E/nligt aktörssynsättet måste ett socialt program ses som ett resultat av en individbaserad situationstolkning" (Dahlgren & Lindkvist 1976:IV:34). Det blir inte särskilt klart vad detta innebär. Ett förhållande som dock pekas ut är den traditionella utvärderingsansatsens uppfattning om mål och andra utgångspunkter för programmet som oproblematiska. Med hänvisning till Esping (1975) konstaterar författarna att de

åtgärder som framstår som naturliga inom ett område beror på hur problemen uppfattas, hur situationen tolkas (Dahlgren & Lindkvist 1976:IV:34f). Dahlgren och Lindkvists (1976:IV:35) alternativa synsätt betyder att intresset inriktas på vad de kallar "de förutsättningar ett visst program baseras på" och benämner detta kritisk utvärdering. Goda inslag är ambitionen att inte ta målen så som de formuleras för givna och att undersöka effekter andra än de förväntade.

Ett aktörssynsätt på utvärdering betyder enligt Dahlgren och Lindkvist att utvärderingen bör utformas som dialog och med hjälp av en psykoanalytisk modell. De menar att utvärderingens kunskapsintresse i detta fall kan betraktas som hermeneutiskt-emancipatoriskt, vilket betyder att dess uppgift skulle vara att avslöja falska föreställningar eller myter. Detta är otvivelaktigt intressanta tankegångar. Dahlgren och Lindkvist lyckas emellertid inte göra något av sina uppslag. Den större kontext som undersökta program skall sättas in i är, för dem, individuella situationstolkningar och den analysand (jämförelsen med psykoanalytisk behandling görs av författarna) som evalueringsforskningen skall medvetandegöra om möjligheterna att se situationen på "nya" sätt (Dahlgren & Lindkvist 1976:IV:49) är ingen annan än beställaren/beslutsfattaren. Utvärderaren blir processkonsult, ibland en kritisk sådan i besluts- och genomförandeförloppen, menar Dahlgren och Lindkvist. Detta är ingen ny roll och författarna konstaterar helt riktigt att denna roll "förefaller vara den normala i de fall där en utvärdering beställes av en beslutsfattare ..." (Dahlgren & Lindkvist 1976:IV:49). Detta betyder varken mer eller mindre än att Dahlgren och Lindkvist med sitt alternativa synsätt på utvärdering givit en mera realistisk modell för hur utvärdering faktiskt går till. De redskap i form av dialogmetod och analogier till psykoanalytisk kommunikation och teknik som erbjuds utvärderingsverksamheten ger den ingen styrsel eller kritisk potens. Istället ger den alternativa modellen utvärderarna goda skäl att fortsätta som förut bara de ändrar sin självuppfattning. Anvisningar och förslag finns att hämta hos Dahlgren och Lindkvist. Utvärderare är processkonsulter som med hermeneutiskt-emancipatoriskt kunskapsintresse arbetar med att avslöja falska föreställningar och myter hos beslutsfattare och administratörer för sina uppdragsgivare, dvs. för beslutsfattare och administratörer.

Avslutningsvis kan vi konstatera att Dahlgren och Lindkvist genomfört en reflekterad och kritisk genomgång av utvärderingslitteraturen, att de identifierat väsentliga problem i denna lika väl som i sina egna arbeten. Men när de skall formulera ett eget program för hur problemen skall lösas kommer de ingen vart¹.

I ett projekt som behandlar beröringspunkterna mellan utvärderingsforskning och statlig förvaltningsrevision har Rolf Lind författat en översikt av "amerikansk evaluation research-litterature". Lind behandlar huvudsakligen den rationalistiskt orienterade utvärderingstraditionen som han, med fog - i varje fall volymmässigt - menar dominerar litteraturen. Lind har inga pretentioner på att fördjupa analysen av utvärderingsproblematiken, han önskar blott referera. I sin bakgrundsteckning pekar Lind på "besvikelsen över planeringsvägens relativa misslyckande" (Lind 1979b:7) och hur utvärderingsintresset, dvs. intresset att "kartlägga och värdera de faktiska effekterna av verksamheten", kommit i dess ställe². Förekomsten av nya professionella grupper - psykologer, pedagoger och sociologer - inom förvaltningen nämner Lind som en annan förklaring till utvärderingens framväxt (Lind 1979b:9) samt att intresset för utvärdering växer när expansionsutrymmet för offentlig verksamhet minskar (Lind 1979b:10). I den mest positiva vändningen i rapporten gentemot utvärderingsforskningen skriver Lind att "utvärdering är ... inte enbart ett sätt att förbättra olika samhällsreformers verkningsgrad. Genom den ökade probleminriktningen och kraven på tvärvetenskaplighet kan också den vetenskapliga kunskapen om komplexa samhällsfenomen öka. Utvärderingsforskningen kan på detta sätt ses som ett gemensamt intresse för reformistiskt orienterade politiker och samhällsengagerade forskare" (Lind 1979b:10). Men det är inte klart i vilken utsträckning detta är Linds egen åsikt eller bara ett referat. I förhållande till ambition och omfång måste man säga att Lind väl lyckas teckna en bild av den rationalistiska utvär-

¹De senare skrifter jag läst av författarna tyder inte på att de hittat rätt på den alternativstiga de försökt staka ut. Se Dahlgren & Lindkvist (u.å. och 1977).

²Detta har jag ovan beskrivit som en förskjutning av intresset från ex-ante- till ex-poststudier.

deringstraditionens problem. Svårigheter med diffusa mål, den experimentella metodiken som ett ouppnåligt ideal, program som med tiden ändrar inriktning, resultaten av utvärdering kommer ej till användning, den politiska miljö utvärderingarna företas i stället till bekymmer osv. Av Linds sätt att referera och citera får jag intrycket att han med reservation för de absurt höga förväntningar som till en början riktades mot och underblåstes av utvärderarna, ändå finner att utvärderingsforskning kan fungera som en del i ett omfattande beslutsunderlag (Lind 1979b:77). Lind intar emellertid en betydligt svalare hållning till utvärderingsforskningen i den sammanfattande rapporten där jämförelsen med förvaltningsrevisionen som utvärderingsförfarande genomförs. Omvänt är förvaltningsrevisionen tecknad i ljusa färger¹. Förvaltningsrevisionen har en fristående och oberoende ställning i förhållande till den granskade myndigheten stryker Lind under (Lind 1979a:10). Förvaltningsrevisionen verkar effektivitetshöjande genom sin problemidentifierande funktion. Den har en sanningssägarroll, en diagnosfunktion och den kan beskrivas som en form av konstruktiv kritik menar Lind (1979a:10). Några specifika metoder och tekniker för denna typ av revision finns inte². "Fristående och oberoende" är förvaltningsrevisionen i formell mening. Mitt förmodande, bl.a. grundat på studier i utvärderingslitteraturen, att dessa adjektiv när de används som en karaktäristik av hur RRV:s förvaltningsrevision faktiskt går till (för det var väl två utvärderingsansatser i deras tillämpning Lind ville jämföra?) är en skönmålning bestyrks av Esping i hans tankeväckande artikel om utvärdering av offentlig förvaltning. "Erfarenheterna från RRV:s förvaltningsrevision pekar på att utvärderingens problem inte är i strikt mening metodmässiga³ utan sammanhänger med dels analytikernas och besluts-

¹En elak fråga är ju om detta har något att göra med att uppdragsgivarna för Linds arbete är den statliga institution som genomför förvaltningsrevisioner, nämligen RRV. Det skulle i så fall stämma utmärkt överens med erfarenheter från andra utvärderingsstudier.

²Lind (1979a:11). Det att en specifik metod saknas utmärker enligt Gorpe (1978:37) istället utvärdering.

³Att de också är metodmässiga har framgått av min tidigare genomgång. Jag återkommer till metodproblemen.

fattarens förmåga att frigöra sig från de etablerade sanningarna¹ dels institutionella maktfaktorer i samhället" (Esping 1975:34). "Erfarenheterna från RRV visar att utvärderaren ... möts av repressiva attityder ..." (Esping 1975:34). Det kanske inte blir så helhjärtat med friståendet och oberoendet alla gånger.

Utgångspunkten för Linds arbete var, skriver han, "att utvärderingsforskningen skulle kunna ge värdefulla bidrag till metodutvecklingsarbete inom förvaltningsrevision" (Lind 1979a:26). Men Lind kommer fram till att själva utgångspunkten visat sig felaktig. Låt oss se hur denna slutsats vuxit fram. Lind begränsar sin undersökning av evalueringsforskningen till den rationalistiska traditionen. I förbigående (Lind 1979a:18) nämner han s.k. systemorienterad utvärdering och konstaterar att denna är av värde för förvaltningsrevisionen. Det som karakteriserar utvärderingsforskningen är dess inriktning på vetenskapliga metoder (läs: så experimentella och kvantitativa metoder som möjligt). Detta är förvisso sant för den rationalistiska "skolans" proklamationer. Med dess praktik förhåller det sig annorlunda². Andra "skolor" hyllar inte detta metodideal ens i festtalen. De speciella problem som har att göra med utvärderingsforskningens kvantitativa och hypotestestande³ inriktning är inte aktuella inom förvaltningsrevisionen. Denna arbetar nämligen enligt Lind med mjukare metoder (Lind 1979a:16). I detta läge kunde man förmoda att kritiken av den rationalistiska ansatsen och de mera kvalitativt inriktade, med fallstudier arbetande utvärderingsförsöken, skulle befinnas värda uppmärksamhet. Linds slutsatser om utvärderingsforskningen tycks alltså sammanhänga med att han begränsar sig till dess rationalistiska varianter och inte ens tar upp (själv)kritiken av dessa. Jag tror emellertid det finns andra skäl än en olycklig avgränsning i initialskedet av Linds undersökning som står i vägen för möjligheterna att avlocka utvärderingsforskningen några av dess lärorika

¹Vilket i viss mening just är en metodfråga av speciell vikt när det gäller att upptäcka icke förutsedda effekter.

²Som konstaterats ovan lever utvärderingsforskningen sällan upp till elementära kontventionella metodkrav.

³Validering av den "verkansmodell" man antagit "transporterar" det studerade programmets åtgärder via kausala mekanismer till effekter eller verkningar.

hemligheter. Lind anser nämligen inte att frågan om verksamheternas resultat är av något intresse för förvaltningsrevisionen. "Effektmätning är ... av begränsat intresse för förvaltningsrevisionen" (Lind 1979a:27). I ett kryptiskt försök att reda ut argumentationen meddelar Lind att "/förvaltningsrevisionen ska granska myndigheternas verksamhet, inte innehållet i den politik statsmakterna söker förverkliga genom att fackmyndigheterna ålägges olika arbetsuppgifter". Varpå man stilla undrar om inte myndigheternas verksamhet, i varje fall formellt, just är det som statsmakterna ålägger myndigheterna med syfte att /innehållet i/ politiken skall förverkligas? "Utvärderingsstudier inriktas däremot många gånger på åtgärder som är politiskt initierade. Syftet är att undersöka om dess åtgärder är verkningsfulla instrument att uppnå de mål som angetts", fortsätter Lind (1979a:28). Det bör alltså inte förvaltningsrevisionen göra. Den får på så sätt sig tilldelad ett slags restfält: sådana delar av myndigheternas verksamheter som inte ålagts dem av statsmakterna. En sådan prövning vill Lind med hänvisning till allmänna verksstadgan överlåta till myndigheterna själva.

Utvärdering kan som vi känner till upplevas som kontroll. Lind förmodar att denna upplevelse kan förstärkas om i exemplet förvaltningsrevision resultatet av myndigheternas verksamheter undersöktes (effektmätning) (Lind 1979a:28). Detta stämmer säkert och förmodligen är det här skon klämmer. Politiker vill inte så gärna bli upplysta om att av dem föreslagna verksamheter inte har medfört avsedda resultat och statliga myndigheter vill själva ha kontroll över bedömningen av den egna verksamhetens resultat. Så är det i USA, vilket utvärderingslitteraturen visar, och så är det med största sannolikhet även i Sverige. Att Lind avvisar effekt- och resultatproblematiken som ointressant för RRV:s förvaltningsrevision ser jag mera i ljuset av Espings konstaterande att institutionella faktorer och maktförhållanden i samhället spelar roll för utvärdering än som en övertygande argumentation om att RRV på grund av rationell arbetsdelning eller överskridande av någon rimlig politiseringsgräns borde avhålla sig från att bedöma resultat. Mäktiga intressen (t.ex. "tunga" ämbetsverk och ledande politiska grupperingar) försöker kontrollera sanningssägandet i den mån det inte kontrollerar sig självt. Det är föga överraskande. Vi

har sett det i evalueringsforskningen i USA och vi kan ana att det är anledningen till varför RRV ligger lågt. Kanske är det denna defensiv som skvimpat över och in i Linds slutsatser.

På svenska och av svenska forskare föreligger goda översikter av den amerikanska utvärderingsforskningen - i varje fall av dess rationalistiska huvudådra. De problem som är förknippade med denna utvärderingsstrategi blir tillfredsställande belysta och vissa - om än ofullgångna - försök görs att formulera strategier för att lösa dessa problem eller utveckla adekvata ansatser. Den övertro som utvärderingsforskningen till en början ackompanjerades av i USA har knappast fått så stort fotfäste i Sverige. Följdriktigt har inte några större besvikelser vädrats.

I Sverige finns ingen utvärderingstradition. Däremot har ett slags utvärderingar av tradition genomförts. Dessa har emellertid inte som i USA bildat en speciell forsknings- eller undersökningsgenre med professionsbildning etc. Påfallande är skillnaden i öppen politisering av utvärderingsstudierna i det svenska och amerikanska fallet. Utvärderingar har spelat en icke obetydlig roll i politiska stridigheter i den amerikanska inrikespolitiken. Konflikterna mellan presidentmakten, den federala och lokala förvaltningsnivån har bl.a. kommit till uttryck i utvärderingssammanhang. Kampen mellan en politik med reformatoriska förtecken (grovt: the Great Society) och med mera konservativ inriktning (grovt: the New Federalism) har också till en del förts med utvärdering som argument. Det svenska systemet är av tradition mera konfliktövertäckande. Genom en lång praktik av kommittéarbete har rutiner för konfliktdämpning och -övertäckning samt för utarbetande av kompromissalternativ utvecklats. Även metoder för kontroll av utredningsresultatens innehåll i stora drag och för att lägga obehagliga resultat åt sidan utan att särskilt stor uppståndelse väcks har haft goda möjligheter att utvecklas i ett litet land med korporativa tendenser.

En viktig skillnad mellan den svenska traditionen och det amerikanska sättet att organisera utvärdering är att i det senare fallet kopplas utvärderingen direkt till ett program, medan i Sverige den mera är inbäddad i aktiviteter av annat slag så som t.ex. en statlig utredning (jfr. t.ex. Bell 1975). Även om de båda

aspekter av utvärderingens (speciellt efterhandsbedömningarnas) problem som jag identifierat med hjälp av den amerikanska debatten kan spåras också i Sverige¹ leder den svenska modellen aldrig fram till så tydliga paradoxsituationer som den amerikanska utvärderingen gjort. Detta betyder att svensk utvärdering är besparad vissa obehag. Men det betyder också att chanserna att den svenska utvärderingen skall bli medveten om vad den gör är mindre. Ty den omständigheten att problemen inte träder i dagen så tydligt i Sverige innebär inte att de är lösta eller irrelevanta. Det innebär bara att de är så mycket svårare att upptäcka. Min tes är alltså att utvärderingsforskningens (effektbedömningens) paradox - att vara efterfrågad bara frågorna ej besvaras - och två problem - det institutionella/politiska och det teoretiska/metodologiska - är dolda i Sverige. Jag skall försöka visa detta genom en granskning av två av de mest uppmärksammade utvärderingsförsöken i Sverige. Det gäller EFA:s av arbetsmarknadspolitikens och ERU:s av regionalpolitikens effekter. Först när problemen blottläggs kan de angripas.

¹Esping (1975:34) skriver som jag redan påpekat att hindren för utvärdering ligger i institutionella förhållanden och maktfördelningen i samhället.

TVÅ HALVPERMANENTA EXPERTGRUPPERS FÖRSÖK ATT BEDÖMA EFFEKTER

Jag tror det finns en mycket lättbegriplig tendens hos personer vars dagliga gärning består i att fatta beslut eller att assistera dem som formellt lägger fast världens gång; en tendens att överskatta betydelsen hos det politiska/administrativa systemet och dess aktörer. Denna böjelse, att överdriva politiska åtgärder, planlagda reformer och administrativa ingrepp som samhällliga förändringsfaktorer skall jag här låta gå under beteckningen politikcentrism¹.

Vi kan också finna detta fenomen i den officiösa historieskrivningen om 30-talskrisens upphävande i Sverige. Det var den socialdemokratiska krispolitiken som förde in svensk ekonomi på rätt kurs får vi lära oss. Mycket talar emellertid för att det var den internationella rustningsekonomin som drog med sig det svenska näringslivet.

¹ Detta är en släkting till den decisionistiska uppfattningen som jag definierat ovan. Ett uttryck för denna hållning - men i realiteten med långt mera djupgående implikationer - är det språkbruk som substituerar 'staten' med 'samhället'. När staten, dvs. den politiska instansen inte längre betraktas som en separat institution utan som en representant för (eventuellt t.o.m. som identisk med) samhället som helhet, den allmänna viljan och det gemensamma förnuftet, ja när staten benämns samhället utan att vara bortvittrad, dvs. åternedstigen från sin maktutövande upphöjdhet tillbaka till institutioner av fria och lika individer, då ligger det nära till hands att se politiken (alias staten, alias samhället) som samtliga sociala förändringars källa. Inte ens det distributiva flertalets mekanism ryms i denna hypotetiska värld. (jfr. Elster 1972).

En lika naturlig tendens tror jag vi kan finna hos dem som är ansvariga för en politisk reform eller ett planeringsförsök (och för dem som sympatiserar med den förda politiken och/eller är beroende av den eller dess tillskyndare) att förtröstansfullt lita till att gjorda ingrepp varit av godo¹.

De två halvpermanenta expertgrupperna EFA (expertgruppen för utredningsverksamhet i arbetsmarknadsfrågor) och ERU (Expertgruppen för regional utredningsverksamhet) är knutna till för respektivepolitik ansvarigt departement. Dessa expertgrupper har bl.a. till uppgift att utvärdera den förda politiken, men bidrar också till idéer och strategiskt tänkande. Jag menar - med hänvisning till amerikanska erfarenheter - att det är rimligt anta att den institutionella inplaceringen av utvärderingsverksamheten påverkar dess inriktning, resultat och - kanske i första hand - tolkning. Jag skall leverera indicier för detta.

Utmärkande för svenskt kommittéväsande (av vilket de här diskuterade grupperna är att betrakta som en strategisk kunskapsproducerande och -förmedlande del) är de breda kompromisserna, de konfliktöverslätande politiska uppgörelserna på ett relativt tidigt stadium i den politiska beslutsprocessen. Det hör till ovanligheterna att utvärderingsverksamheten politiseras öppet i Sverige, till skillnad från t.ex. USA. Med detta menar jag att utvärderingsresultat och deras tolkningar knappast blir till medel i en öppen politisk strid i Sverige. Samtliga mera inflytelserika politiska grupperingar har på något sätt delansvar för kompromissresultaten. Vindkantringarna i den svenska inrikespolitiken har heller inte varit särskilt tvära under modern tid - möjligen har nu² sedan några år en skärpning skett. Utredningsväsende och utvärderingsverksamhet blir därför i Sverige, mycket mer än i USA, aktiviteter som på ett generellt plan går i god för den

¹Under kriget mot fattigdomen i USA genomfördes utvärderingar av politiken i en anda av reformoptimism och -tillit. Effektivisering och förbättring var dessas ändamål. Först senare blev utvärderingen en källa till kritik och ifrågasättande av reformerna. Ett resultat som pekar på att reformeffekterna inte varit riktigt tillfredsställande kan ju användas som argument i båda riktningarna, dvs. som stöd för uppfattningen att reformen borde avvecklas eller i en plädering för förstärkt reformering.

²Nuet är från 1979.

förda politikens förnuft. Det är när utvärderingsresultat används som stöd för än den ena än den rakt motsatta politiken som verksamhetens ideologiska funktioner spricker fram och betydelsen av utvärderingens institutionella inplacering blir uppenbar.

För EFA:s och ERU:s del förväntar jag mig således en tendens att överskatta de positiva effekterna av politiken, att när flera tolkningar är möjliga välja den för den förda politiken mest gynnsamma och att när resultaten inte direkt motsäger en positiv tolkning framhäva det troliga i att just den förda politiken spelat en stor roll.

I utredningsspråket finns som bekant en rad språkliga vändningar som i en tämligen komplicerad problemsituation förmår skala bort svårigheter och hävda ståndpunkter framsprungna ur politiska bedömningar och/eller kompromisser som naturliga slutsatser utan att - och det är viktigt - egentligen binda sig för dessa. Verb som "kan", "kan sägas" och "torde" är vanliga i dessa konstruktioner. Detta känner vi från flerfaldiga utredningar som syftat till policyformulering.

En särskilt intressant fråga att söka svar på är om denna resonemangstyp återfinns i EFA:s och ERU:s utvärderingar och hur den i så fall förhåller sig till de metoddiskussioner som förs angående mätning eller uppskattning av effekter. Jag inleder genomgången av utredningsmaterialet med en resumé av de allmänna metodresonemangen, redogör därpå för vilken typ av undersökningar som genomförts och vad de innehållit varpå de generella slutsatserna granskas. Slutligen ställer jag frågan om uppläggning och tolkning varit helt sunda. Min analys av ERU:s utvärderingar har publicerats i andra sammanhang¹ varför den här endast sammanfattats. Granskningen av EFA har presenterats vid en konferens² men är tidigare inte publicerad.

¹Den ursprungliga versionen författades 1979/80. Denna har redovisats i en BFR-rapport, Gullberg (1981:91-101), och i en utvidgad, bearbetad och uppdaterad version i Gullberg (1985b).

²Svenska sociologförbundets årsmöte i Stockholm 1983, Gullberg (1983).

Metodresonemang

EFA har omnämnts som ett undantag från den svenska bristen på utvärderingsstudier (SOU 1978:47 s. 171). Statskontoret riktade 1966 i ett utlåtande över arbetsmarknadspolitiken uppmärksamheten på behovet av kunskaper om hur denna politik fungerade. Sålunda påtalades den bristande kännedomen både vad gäller de förhållanden politiken avsåg att påverka (dvs. arbetsmarknaden) och effekterna av vidtagna åtgärder (dvs. effekterna av arbetsmarknadspolitiken). EFA tilldelades (bl.a.) uppgiften att utvärdera arbetsmarknadspolitiken.

EFA formulerar redan i sitt första betänkande - SOU 1974:29 - effektmättnings- eller bedömningsproblemet på ett mycket stringent sätt. Utvärderingarnas principiella uppgift att försöka belysa effekterna av ett visst program eller av vissa åtgärder "måste uppenbarligen innebära en jämförelse mellan två eller flera alternativ som skiljer sig åt i fråga om" den förda politiken, konstaterar EFA (SOU 1974:29 s. 78). "Man försöker alltså jämföra vad som faktiskt har hänt med vad som skulle ha hänt vid en något annorlunda politik."¹

Men det är stora problem förknippade med undersökningar av en politiks verkningar - eller som EFA formulerar sig, med att "prestera entydiga utvärderingar" - det betonar utredningen redan i inledningen (SOU 1974:29 s. 16).

Det som EFA i de ovanstående passagerna givit sig in på är att formulera problemet om effekter i termer av jämförelser mellan vad som faktiskt inträffat och vad som skulle ha hänt om något som faktiskt har ägt rum inte hade skett. Denna typ av resonemang går under beteckningen kontrafaktiska (eftersom de labore-

¹"Antag att man ger en forskare uppgiften att försöka belysa arbetsmarknadsutbildningens effekter. Det forskningsprojekt som blir följderna måste uppenbarligen innebära en jämförelse mellan två eller flera alternativ, som skiljer sig åt i frågan om arbetsmarknadsutbildning ... Ett av alternativen identifieras med den under de senaste åren faktiskt bedrivna arbetsmarknadsutbildningen. Övriga alternativ konstrueras genom att man föreställer sig att delar av den faktiskt bedrivna arbetsmarknadsutbildningen ej hade kommit till stånd." (SOU 1974:29 s. 78).

rar med ett tänkt händelseförlopp i förfluten tid, dvs. en historia som inte har inträffat och som därför inte är möjlig att undersöka med gängse empiriska metoder). Den är välbekant inom ekonomisk historia, där en hel skolbildning utvecklats med kontrafaktiska resonemang och förklaringar som gemensam nämnare. Detta har skett särskilt i USA och under benämningen "The New Economic History". Vidare är den kontrafaktiska problematiken väl känd inom analytisk filosofi där i synnerhet legitimiteten av kontrafaktiska förklaringar och dessas relation till gängse förklaringsstyper varit föremål för livliga meningsutbyten¹. Att formulera effektfrågan i kontrafaktiska termer tycker jag är en utmärkt idé. Det är nämligen en tillspetsning av frågeställningen. Lyckas man lösa eller behandla de problem denna resonemangstyp ger upphov till på ett godtagbart sätt innebär det ett stort steg framåt inte bara för bedömningen av politikeffekter utan för samhällsvetenskapen som sådan. Men jag vill genast understryka att den kontrafaktiska formuleringen av effektproblemet på intet sätt hjälper EFA att lösa de utvärderingsuppgifter gruppen ålagts. Istället är det ett ovanligt konsekvent och uppfordrande sätt att staka ut de uppgifter som expertgruppen hade att ta itu med. Även EFA:s andra betänkande - SOU 1978:60 - innehåller programmatiskt kontrafaktiska problemformuleringar².

ERU å sin sida hänvisar till EFA:s utvecklingsarbete på effektmätningens område (SOU 1978:46 s. 11). ERU resonerar också i kontrafaktisk riktning och konstaterar att man inte vet vad som skulle ha hänt med de företag som erhållit stöd om detta ej utgått. Därför kan man inte - så argumenterar ERU helt rimligt -

¹De kontrafaktiska förklaringarnas och resonemangens problem diskuteras på ett klagörande sätt av Elster (1978). Jag har diskuterat den kontrafaktiska problematiken i samband med utvärderingens teoretiska och metodologiska svårigheter i Gullberg (1984).

²"Den faktiskt observerade arbetsmarknadsutvecklingen utgör ett resultat både av den förda arbetsmarknadspolitiken och av andra faktorer ... Det är alltså svårt att fastställa hur arbetsmarknadsutvecklingen skulle ha sett ut i frånvaro av den förda politiken och därmed hur effekterna av arbetsmarknadspolitiken sett ut." SOU 1978:60 s. 20. Och vidare: "Ett centralt problem vid utvärderingsstudier kring effekter av en viss åtgärd är svårigheten att isolera effekterna av den vidtagna åtgärden från inverkan av andra faktorer." (s. 25).

utgå från att all sysselsättningsökning inom dessa företag är effekter av regionalpolitiken (SOU 1978:46 s. 43). I ett försök att studera effekterna av samtliga regionalpolitiska insatser på tillverkningsindustrins sysselsättning utför ERU en kalkyl där den faktiska sysselsättningen jämförs med "den som skulle ha uppstått vid oförändrad politik". (SOU 1978:46 s.43). Begreppet effekt används ... här, när en ... kvantitativ uppskattning av ett alternativförlopp utan stöd gjorts." skriver ERU¹.

Som kommentar till det kontrafaktiska sättet att formulera frågan om effekter skall jag här bara påpeka att detta står och faller med de utgångspunkter (eller kanske snarare teorier) och metoder som kommer till användning vid beräkning eller uppskattning av vad som hade hänt i det fall ingen politikförändring skett. Samtidigt är just denna formulering av effektfrågan så stringent och krävande att det förefaller svårt att negligera dess teoretiska och metodologiska implikationer. Har kanske EFA och ERU genom denna problemformulering lyckats skapa en sanningssökandets hemorts rätt i statligt utredningsväsende, ett skydd mot alltför flagranta och direkta influenser från eller eftergifter i förhållande till maktens intressen?

I både EFA:s och ERU:s regi har under deras mer än decennielånga² verksamhet ett mycket omfattande och värdefullt material av uppsatser, rapporter, diskussionsböcker och utredningar producerats. Det är här inte tal om att åstadkomma något som tillnärmelsevis skulle kunna liknas vid en utvärdering av expertgruppernas totala kunskapsproduktion, än mindre av de eventuella effekter denna produktion kan ha haft på beslutsfattande, politikinnehåll eller samhällelig förändring. Jag begränsar mig istället snävt till att syna de SOU-publikationer som grupperna avgivit och då endast de som med några större pretentioner behandlar frågan om bedömning av politikeffekter ex-post, s.k. efterhandsanalyser.

¹SOU 1978:47 s. 63. Det är alternativförloppet som enligt ERU är utan (regionalpolitiskt) stöd, inte beräkningarna.

²Texten är från 1979.

Att utvärdera arbetsmarknadspolitik (SOU 1974:29)

EFA har redovisat sina resultat i två digra SOU - SOU 1974:29 "Att värdera arbetsmarknadspolitik" och SOU 1978:80 "Arbetsmarknadspolitik i förändring". Båda dessa volymer innehåller material och resonemang syftande till att belysa arbetsmarknadspolitikens effekter. I det första betänkandet är det tre studier i genren efteranalys (SOU 1974:29 s. 77 och 42) som intresserar. Visserligen förs ett allmänt resonemang, ett slags försvarstal för arbetsmarknadspolitikerna som sägs tjäna syftet att utveckla en teoretisk grundval för infogande och tolkning av specialstudierna (SOU 1974:29 s. 20 och 44). Detta skulle kunna tolkas som en tvåvägsstrategi vid studiet av politikens effekter: en kombination av generella, (makro)teoretiskt grundade resonemang å den ena sidan och intensiva fallstudier å den andra. Jag kommer senare att ta upp denna tanke och diskutera hur relationerna mellan de två nivåerna bör se ut för att de skall kunna gripa in i varandra.

Men försvarstalet utvecklar sig aldrig till någon riktig arbetsmarknadspolitik teori och de genomförda mera detaljerade efterhandsanalyserna löper så vitt jag kan se utan influens från EFA:s mera vidlyftiga resonemang.

De tre områden som är föremål för efteranalyser är delar av arbetsmarknadsutbildningen (SOU 1974:29 kapitel 6), den geografiska rörligheten (SOU 1974:29 kapitel 7) och immigrationen (SOU 1974:29 kapitel 8). EFA karaktäriserar de i dessa studier använda metoderna som i stor utsträckning baserade på jämförelser mellan de personer som berörts av en viss politik och andra personer (SOU 1974:29 s. 42).

Tar vi vår utgångspunkt i EFA:s eget sätt att resonera om effekter av politik - det kontrafaktiska - blir vi nödsakade att rikta några grava invändningar mot de genomförda studierna, i varje fall enligt vad som går att utläsa om dem från de presentationer som återfinns i "Att utvärdera arbetsmarknadspolitik". Det visar sig nämligen att varken rörlighetsstudien eller immigrationsstudien undersöker effekterna av förord politik. Rörlighetsstudien syftar till "att belysa vilka konsekvenser flyttningar /som sker enligt arbetsmarknadskungörelsen A.G./ får för flyttarna

själva och för samhällsekonomin" (SOU 1974:29 s. 35). Detta är uppenbarligen något annat än att undersöka vad som skulle ha hänt om inte flyttningsbidrag utgått, vilket är den problemformulering man borde göra om det är politikens effekter som skall utvärderas. Att det föreligger en principiellt viktig skillnad mellan att undersöka effekten av flyttnings- och starthjälpsstödet å ena sidan och effekterna av de flyttningar som skett med stöd av dessa arbetsmarknadspolitiska insatser framgår av EFA:s eget påpekande: "En hel del av dem /som flyttat med starthjälpsstöd A.G./ skulle antagligen flyttat till dessa erbjudna arbetstillfällen även om arbetsförmedlingen ej haft möjlighet att erbjuda flyttningsbidrag." (SOU 1974:29 s. 330). Förbindelsen mellan förd politik och forskningsresultaten etableras avslutningsvis genom påståendet att "man kan utgå från att statsmakterna på olika sätt har möjlighet att påverka flyttningsströmmarna i samhället". (SOU 1974:29 s. 375). Efter utvärdering av arbetsmarknadspolitikens flyttningsstimulerande åtgärder är denna möjlighet inte något som man skulle behöva utgå från. Istället har man anledning vänta sig att sådana studier just skulle kunna svara på frågan om och, i så fall, i vilken utsträckning denna del av arbetsmarknadspolitiken ger påverkningsmöjligheter. Detta behöver inte - gör det inte heller - betyda att EFA:s rörlighetsstudier är irrelevanta¹.

Inte heller immigrationsstudien besvarar några frågor om politikens effekter. Istället försöker den "beräkna vilka kostnader och intäkter en ändring av immigrationens omfattning under slutet av 1960-talet skulle ha medfört" (SOU 1974:29 s. 39). Något kontrollgruppsförfarande har dessa EFA-studier inte tagit upp. Det är

¹Tvärtom levererar rörlighetsstudien intressanta informationer om flyttarnas villkor på de nya orterna. Studierna är upplagda så att ett urval flyttare jämförs med en grupp kvarstannare. Dessutom jämförs flyttarnas situation före och efter flyttningen. Projektet beskrivs i SOU 1974:29 s. 345ff där också vidare hänvisningar ges. Kontrollgruppskonstruktionen innebär alltid stora problem som dryftas i utredningen bl.a. på s. 347 och 86f. Jag återkommer till detta. Till de intressantare resultaten som rörlighetsstudien redovisar hör: Flyttarnas sysselsättning ökar men knappast lönenivån (s. 36) och därav följer att bruttoinkomsten ökar, men den disponibla inkomsten stiger mycket beskedligt (s. 331). De samhällsekonomiska vinsterna förefaller vara betydligt större än de individuella (s. 37). Det är klart att dessa och andra resultat från rörlighetsstudien har relevans för arbetsmarknadspolitiska ställningstaganden, men svarar på frågan om den flyttningsstimulerande politikens effekter gör studien inte.

ganska förnuftigt eftersom immigrationens effekter på t.ex. svenska ekonomiska förhållanden svårligen låter sig belysas med hjälp av en sådan teknik. Däremot kan naturligtvis - på samma sätt som de inhemska flyttarnas förändrade situation jämfördes med kvarstannarna i Norrland - immigranternas livssituation jämföras med "relevanta" grupper i emigrantlandet. Arbetsmarknadspolitiken spelar i denna studie en ännu mera undanskymd roll än vi kunde konstatera att den gjorde i flyttningsstudien. Immigrationens orsaker är i första hand den svenska konjunkturutvecklingen. I viss mån har också utvandringslandets konjunkturutveckling betydelse. "Immigrationsutvecklingen beror naturligtvis också på den svenska immigrationspolitiken" tillfogar expertgruppen. Men genom den fria nordiska arbetsmarknaden är regleringsmöjligheterna begränsade. (SOU 1974:29 s. 39). I denna studie är det allt annat än de tänkta effekterna av en alternativ politik som den faktiskt föreliggande verkligheten jämförs med.

De två här genomgångna undersökningarna ger klart belägg för att en stringent frågeställning (se ovan och SOU 1974:29 s. 78) på inget sätt garanterar att frågan besvaras enligt lika kvalificerade principer om ens alls något svar ges. Med studien av arbetsmarknadsutbildningen förhåller det sig emellertid annorlunda. I denna görs försök att jämföra personer som genomgått arbetsmarknadsutbildningen med personer i liknande situation som ej följt utbildningen (s. 31). Studien är i första hand av metodprövande och metodutvecklande karaktär (s. 31 och 323), men skall också ge vissa besked om den bedrivna arbetsmarknadsutbildningen (s. 31). Den slutsats som EFA presenterar framstår som nog så entydig: "Enligt resultaten ... medförde den studerade arbetsmarknadsutbildningen en betydande sysselsättningsökning för de utbildade själva (s. 32). "Arbetsmarknadsutbildningen har en avsevärd, höjande effekt på sysselsättningsgraden" (s. 271).

Jag skall i det följande peka på några omständigheter som gör denna entydighet en smula mindre, vilket naturligtvis inte utesluter att slutsatserna ändå skulle kunna vara riktiga. Här gäller det ju frågan om huruvida denna undersökning ger EFA anledning förkasta nollhypotesen. Ändå är det fullt möjligt att andra belägg för slutsatserna kan mobiliseras. Dessutom kommer jag att ta upp vissa metodproblem som berörs.

Vad hade hänt utan (de undersökta delarna av) arbetsmarknadsutbildningen? (SOU 1974:29 s. 31). Frågan kan uppenbarligen gälla antingen en person eller grupper av personer som genomgår utbildningen eller avse hela det undersökta kursutbudets inverkan på arbetsmarknaden. Urvalet till arbetsmarknadsutbildningen sker inte slumpmässigt. Därför är uppdelningen av personer som får och inte får utbildning otillgänglig för strikt experimentell kontroll (s. 259). Den kontrollgrupp som de arbetsmarknadsutbildade jämförs med är ett utval av arbetslöshetsregistrerade vid en tidpunkt och på platser som svarar mot undersökningsurvalet (s. 255).

Begränsar vi alltså diskussionen om utbildningseffekter till att gälla individuella differenser i sysselsättning och inkomst är det relevant att ställa frågan om (de eventuella) skillnader som undersökningsgruppens utveckling uppvisar i förhållande till kontrollgruppens beror på utbildningen eller på att grupperna redan från början var olika. För att komma tillrätta med detta problem genomfördes en utbildningsurvalsundersökning (s. 256f och 262ff). Denna studie undersöker skillnader i vissa variabler mellan huvudgruppen och kontrollgruppen. Någon studie av vad som bestämmer urvalet till utbildningen företas ej. De största skillnaderna uppmäts i variablerna (i nämnd ordning): intresse, ålder, förväntningar om förbättringar i nytt arbete, yrkesområde, sjukledighet och tidigare utbildning (s. 268). Undersökningen förklarar inte urvalsprocessen, den anger bara vissa samband mellan att följa/ej följa utbildningen å ena sidan och några andra variabler å den andra.

Ett resultat ger vid handen att sysselsättningsskillnaden mellan de två grupperna uppgick till 20% under tvåårsperioden närmast efter utbildningen. När hänsyn tagits till gruppernas olika sammanställning (enligt den ovan refererade undersökningen) reducerades skillnaden till 10.7% i båda fallen till förmån för de omskolade (s. 272). Någon klar skillnad i lönenivå mellan grupperna har inte kunnat konstateras (s. 277). Inkomstskillnaderna mellan utbildningsgrupp och kontrollgrupp på 35.6% av utbildningsgruppens inkomstnivå (s. 297) är nämligen att hänföra till den ökade sysselsätt-

ningsgraden. Denna skillnad minskar f.ö. till 19,8 om beräkningarna justeras för de observerade skillnaderna i gruppens sammansättning (s. 297).

För min del anser jag det mycket rimligt att anta att det föreligger skillnader mellan de båda grupperna i andra avseenden av betydelse för en framtida yrkeskarriär än de som fångats upp i utbildningsurvalsundersökningen. En spekulatio n är att personer som påbörjar arbetsmarknadsutbildningen är mer målinriktade och handlingskraftiga än den grupp av slumpmässigt valda arbetslösa som utgör kontrollgruppen, utöver de spår detta avsatt i de frågesvar om intresse för utbildning som ingick i undersökningen. Kanske kunde "det förhållandet att personer, som ej fullföljt utbildningen, och personer, som genomgått utbildningen, men ej placerats i utbildningsyrket hävdad sig nästan lika bra på arbetsmarknaden som personer, som fullföljt utbildningen" (s. 34) ge en tankeställare i denna riktning. Eftersom skillnaderna mellan utbildnings- och kontrollgruppen reduceras påtagligt i de variabler som EFA kontrollerat för håller jag det inte för otroligt att reduktionen skulle gå ännu längre om urvalsprocessen till utbildningen studerades närmare.

Ytterligare en skillnad som EFA pekar på, i något annat sammanhang, är att personer som genomgått arbetsmarknadsutbildningen tycks bli föemål för en intensifierad insats från arbetsförmedlingen (s. 312). Kan möjligen denna omständighet spela någon roll?

Nu är emellertid den typ av individuella och direkta effekter som möjligen kan undersökas med hjälp av jämförelser med kontrollgrupper av relativt begränsat intresse för en helhetsbedömning av arbetsmarknadsutbildningens effekter. Relevant är dock denna informationstyp om det gäller en enskild persons val att gå igenom arbetsmarknadsutbildningen eller ej. Utan att hemfalla åt klassificeringsraseri¹ finns det skäl att ägna frågan om olika typer av effekter en viss uppmärksamhet. En fråga om utbildningseffekter kan naturligtvis inte bara handla om vad som skiljer elevens öde efter utbildningen från vad han/hon hade fått uppleva utbild-

¹Så sker i ERU:s betänkande SOU 1978:47 s. 69, där inte mindre än 18 x 4 olika typer av effekter anges.

ningen förutan. Vi har minst lika stor anledning att intressera oss för den påverkan utbildningen kan tänkas utöva på strukturförhållanden i samhället, i detta fall kanske särskilt på arbetsmarknadens område. Utbildningens effekter stannar givetvis inte hos de individer som tagit del av densamma. Den påverkar även företagen - dvs. de som köper arbetskraft med olika kompetens - och andra arbetssökande, med andra ord: arbetsmarknaden¹.

EFA benämner denna typ av effekter indirekta². Ett exempel på denna effekttyp är den förändrade konkurrenssituation om lediga arbeten som kan uppstå genom att den utbildade gruppen ökar sina kvalifikationer. Personer som inte följt utbildningen kan då tänkas få sina möjligheter försämrade³. Möjligheten att det föreligger en konkurrens effekt gör det uppenbart att de direkta och individuella effekterna av arbetsmarknadsutbildningen inte kan undersökas utan att de indirekta effekterna beaktas. Förekomsten av konkurrens effekter innebär ju att inte bara undersökningsgruppen utan också kontrollgruppen påverkas av utbildningen⁴.

¹Hur kan jag vara så tvärsäker på det? Mitt resonemang är egentligen hypotetiskt. Om de individer som genomgår arbetsmarknadsutbildningen påverkas i sin situation på arbetsmarknaden finns det anledning anta att denna marknad och de aktörer som opererar på denna marknad också påverkas. Eventuella effekters storlek må undersökas empiriskt. En omständighet, vars betydelse inte heller den kan anges aprioristiskt, men som genast blir synlig är att de personer som är under utbildning för en tid befinner sig utanför arbetsmarknaden.

²Den definierar indirekta effekter som "effekter på sysselsättningsgrad och produktivitet hos personer och företag, som inte direkt berörs av kursdeltagandet" (s. 259).

³EFA refererar på s. 258 till en undersökning av Åke Dahlberg där fem olika slag av indirekta effekter omnämns, bl.a. konkurrens effekter. Dahlbergs resonemang är emellertid anpassat till ett ekonomiskt modelltänkande med kalkylkombinationer, vilket betyder att "hans" konkurrens effekt inte har riktigt samma innebörd som här.

⁴En tentativ bedömning av eventuella konkurrens effekter i den här refererade utvärderingen skulle kunna nås genom en undersökning av de utbildades antal i relation till stock och flöde på relevanta delarbetsmarknader. Kan det tänkas att de arbetsmarknadsutbildade påverkat kontrollgruppens situation på arbetsmarknaden? Jag går här inte in på sådana beräkningar eftersom det är det principiella förhållandet att individeffekter inte kan bedömas oberoende av indirekta effekter som jag är angelägen att fästa läsarnas uppmärksamhet på.

Det är alltså i princip fullt möjligt att skillnaderna i sysselsättning och inkomst mellan utbildningsgruppen och kontrollgruppen beror på att förstnämnda stärkt sin ställning på arbetsmarknaden på bekostnad av den senare. För det individuella valet att gå eller inte gå på utbildningen är skillnaden grupperna emellan naturligtvis fortfarande relevant eftersom vi i detta fall utgår från att utbildningen ges oberoende av hur en viss enskild person ställer sig till den. Men för en bedömning av arbetsmarknadsutbildningen som sådan är givetvis registreringen av denna typ av individuella effekter av begränsad relevans.

EFA uppmärksammar den typ av problem jag påtalat ovan. De indirekta effekterna kan möjligtvis vara av stor omfattning och de kan teoretiskt sett gå både i positiv och negativ riktning skriver gruppen (s. 303). De indirekta effekterna kan inte belysas genom det material som insamlats (s. 259) och det EFA åstadkommer är helt baserat "på antaganden och har inte kunnat grundas på empiriska observationer," (s. 308), men som en ursäkt konstateras att "för studiet av de indirekta effekterna har överhuvudtaget mycket litet gjorts i något land" (s. 326).

Mitt senaste resonemang kullkastar i och för sig inte slutsatsen att de som genomgått utbildningen nått högre sysselsättning och inkomst än jämförelsegruppen, men det reducerar högst avsevärt detta påståendes relevans för en bedömning av arbetsmarknadsutbildningens effekter.

Det primära syftet med studien var - som jag redan konstaterat - att utveckla och pröva metoder. Därför skall jag inte driva kritiken mot de formulerade slutsatserna särskilt långt. De konklusioner som görs angående de direkta individuella utbildningseffekterna är dessutom omgivna av reservationer. Likväl menar jag att dessa garderingar bort vara än mer långtgående, att vändningar av typen "det är troligt att just utbildningen bidragit till stabilare sysselsättning och färre arbetsbyten bland de omskolade än i kontrollgruppen" (s. 276f) har ett bräckligt stöd i under

sökningen och att dess politikaffirmativa karaktär är värd att uppmärksamma. Den politikcentrerade tendensen finns här men är inte särskilt långt driven.

När det gäller de metodologiska aspekterna menar jag att kontrollgruppsförfarandet har visat sig vara av ringa värde och på den punkten är jag strängare än EFA. För det första begränsas de hypotetiska alternativ som den faktiskt förda politiken ställs mot i de kontrafaktiska effektresonemangen till sådana som det är möjligt att finna kontrollgrupper för¹.

För det andra ger den icke-experimentella uppläggnings stora problem för den effektisoleringsstrategi som jämförelserna på individualnivå är uttryck för. Hur skall man kunna veta att det är utbildningen och inte skillnader i gruppmedlemmarnas egenskaper och deras behandling vid sidan av utbildningen som faller utslaget på arbetsmarknaden? Vissa statistiska metoder kan användas för att motverka sådana felkällor - t.ex. multipla regressionsanalyser (SOU 1974:29 s. 25) - men riktigt tillfredsställande blir dessa metoder aldrig. För det tredje kan de indirekta effekterna inte friläggas med hjälp av kontrollgruppsförfarandet, vilket också EFA, som vi redan sett, slagit fast. Min slutsats blir att arbetet med kontrollgrupper inte kan vara den huvudfråga som bedömningar av politiska reform- och planeringsansträngningars effekter skall försöka ta sig fram i.

De observationer som bör göras med anledning av undersökningen om arbetsmarknadsutbildningen kan enligt min uppfattning sammanfattas i fyra punkter:

1. Den kontrafaktiska frågeställning som en hel del energi satsas på att besvara - den om individuella effekter - visar sig ha begränsad relevans för en bedömning av utbildningens effekter. Dessutom är de slutsatser som dras minst sagt osäkra och en politikaffirmativ tendens kan spåras.

¹EFA pekar på denna begränsning (s. 87).

2. Kontrollgruppsförfarandet har visat sig otillfredsställande som metod att komma tillrätta med effektproblematiken. Det är visserligen möjligt att använda för bedömning av individuella effekter, men även på detta begränsade område är metoden behäftad med brister och begränsningar.

3. Frågan om hur de s.k. indirekta effekterna skall kunna bedömas rycker fram som ett av (ex-post)utvärderingarnas metodologiska kardinalproblem.

4. Trots att huvudsaken i undersökningen sägs vara metodprövning och -utveckling drar EFA inga särskilt långtgående slutsatser angående metodspörsmål från det intressanta material som levererats¹.

Arbetsmarknadspolitik i förändring (SOU 1978:60)

Det intressanta tillkottet i EFA:s andra betänkande, jämfört med det första, utgörs av ett försök att bedöma de sammantagna effekterna av arbetsmarknadspolitiken. EFA återknyter till tvåvägsstrategin vid studiet av effekter, där försöket att närma sig dessa samlade effekter är ena vägen. Den andra utgörs av "ett antal ingående studier kring effekterna av specifika arbetsmarknadspolitiska medel" (SOU 1978:60 s. 168). Fallstudierna skall jag inte gå närmare in på här² utan istället rikta mitt intresse mot makrobedömningen av arbetsmarknadspolitiken och relationen mellan denna studie och fallundersökningarna.

Frågan om effekterna av den förda arbetsmarknadspolitiken kan naturligtvis inte besvaras genom att en undersökning av arbetsmarknaden och en annan om arbetsmarknadspolitiken bara läggs bredvid varandra. Samband mellan politik- och marknadsförändringar behöver inte alls vara kausala. "Den faktiskt observerade arbetsmarknadsutvecklingen utgör ett resultat både av den förda arbetsmarknadspolitiken och av andra faktorer" fastslår EFA (SOU

¹Jfr. punkterna a-e SOU 1974:29 s. 328f.

²Se dock appendix I för en koncentrerad redogörelse och kritik.

1978:60 s. 20). Den granskning av "utvecklingen över tid för ekonomin som helhet vad gäller vissa målvariabler som bör ha påverkats av den förda arbetsmarknadspolitiken" (SOU 1978:60 s. 168. Min kursivering A.G.) kan man inte räkna med ger besked om vilket inflytande politiken haft¹. Hur har EFA gått till väga?

Det är målvariablernas utveckling som bildar utgångspunkten för EFA:s tentativa utvärdering. Situationen för tre olika grupper på arbetsmarknaden observeras. Det gäller kvinnor, ungdomar och äldre män. De tolkningar av politikens effekter som förs i dessa tre fall är alla mycket välvilliga i sin hållning till politiken. Det underförstådda antagandet är att arbetsmarknadspolitik är bra för arbetsmarknaden och att mera arbetsmarknadspolitik är ännu bättre.

"Även om kvinnors ökade arbetskraftsdeltagande underlättas av efterfrågeutvecklingen under perioden torde de arbetsmarknadspolitiska åtgärderna ha spelat en betydelsefull roll när det gäller att underlätta kvinnornas inträde på arbetsmarknaden." (s. 168). "Torde" skriver EFA, och det är inte lätt att finna handfast stöd för påståendet om den betydelsefulla roll som arbetsmarknadspolitiken förmodas ha spelat. "De arbetsmarknadspolitiska insatserna har inte varit tillräckliga för att motverka detta /ökade sysselsättningsproblem för äldre män. A.G./" (s. 169). Här underförstås att åtgärderna likväl har verkat eller att om de varit större de skulle ha kunnat balansera försämringen för denna åldersgrupp. "De arbetsmarknadspolitiska insatserna för ungdomar har ökat under perioden, men inte tillräckligt för att förhindra en uppgång i arbetslösheten." (s. 169). Detta uttalande implicerar att ökade resurser - dvs. tillräckliga - verkligen hade kunnat förhindra uppgången i ungdomsarbetslösheten.

"Den förda politiken har ... inte kunnat förhindra den samtidiga ökningen av arbetslöshetstiderna /som antalet arbetslösa under ett år har minskat A.G./. För detta hade det behövts en förstärkning

¹Ty det är som EFA skriver "svårt att fastställa i vilken utsträckning de arbetsmarknadspolitiska insatser som gjorts under perioden bidragit till den observerade utvecklingen." (SOU 1978:60 s. 147).

av åtgärder inriktade på att stimulera nyanställningar av arbetslösa." En smula pedantiskt kan om detta sägas att det ju inte är säkert att en politik med denna inriktning verkligen hade lyckats.

Det mest generella omdömet om arbetsmarknaden och arbetsmarknadspolitiken får man nog säga är synnerligen dystert för den förda politiken: "Vad gäller arbetsmarknadsutvecklingen har vi funnit tecken på negativa utvecklingstendenser, vilka man ej lyckats motverka med den förda arbetsmarknadspolitiken, ..." och det trots att insatserna både i fasta priser och i procent av BNP har en klart tillväxande profil¹.

I samtliga här citerade passager registreras en utveckling på arbetsmarknaden, den må betecknas som positiv som i fallet med den ökade kvinnliga förvärvsfrekvensen, eller negativ som i exemplen ungdomsarbetslöshetens ökning och de växande svårigheterna för äldre män på arbetsmarknaden. I den positiva utvecklingen påstås arbetsmarknadspolitiken ha spelat en betydelsefull roll och för negativa utvecklingstendenser gäller att politiken inte varit tillräckligt omfattande, underförstått att mindre av arbetsmarknadspolitik hade gjort situationen ännu värre och mera av samma vara kunde ha undanröjt eller mildrat problemen. Det intressanta med dessa resonemang är att de dels saknar substantiellt stöd i texten, några försök att pröva påståendena enligt den problemstruktur EFA givit (dvs. kontrafaktiskt) görs ej². Dels är sådana slutsatser som citerats ovan just vad anhängare av den förda och - enligt de samma principerna - utvidgade arbetsmarknadspolitiken behöver. På de punkter jag citerat ovan har politikcentrismen framträtt på ett betydligt mera tydligt sätt än i EFA:s första betänkande. Detta tar sig bl.a. uttryck i ett språk som suggererar tolkningar i en riktning utan att binda det konturlösa skrivande subjektet till några bestämda slutsatser eller till något ansvar.

¹Citat från s. 181. Om arbetsmarknadspolitiken kostnadsutveckling se s. 90.

²Detta utesluter givetvis inte att det kan förhålla sig just så som EFA påstår, men några belegg levereras ej.

Försöket att närma sig politikens effekter från två sidor, en makro- och en mikrosida har av EFA inte utvecklats i någon fruktbar riktning. De mera samlade bedömningarna av arbetsmarknadspolitikens effekter korresponderar inte med de genomförda fallstudierna. Därför uppstår ingen samverkan mellan de två studienivåerna, en samverkan som idealt sett skulle innebära att den okända effektstorheten trängs från två sidor - ju mera hopträngningen ökar desto klarare konturer från effekterna. De indirekta effekternas speciella problematik förs inte heller de närmare sin lösning eller rimliga behandling i EFA:s andra betänkande.

När en EFA-ledamot nyligen¹ påstod: "att vi inte på länge än kan åstadkomma annat än ytterst ofullständiga utvärderingar och, inte ens i princip, har tillgång till någon så att säga överordnad utvärderingsmetod" kan jag bara instämma².

Att främja regional utveckling

Man kan knappast hävda att regionalpolitiken skapat regional balans mellan landets olika delar. Norrland och i synnerhet Norrbotten befinner sig idag - 1980 - fortfarande i en mycket svår situation. När det gäller att utvärdera den förda politiken är det alltså lätt att konstatera att den inte löst de så kallade balansproblemen. Men har den kanske mildrat dem, ja kanske till och med lagt grunden till en ekonomisk utveckling i de eftersatta landsdelarna?

I den sammanfattning som ERU gjort av sina bedömningar av regionalpolitikens effekter råder ingen tvekan på den punkten. "Politiken har givit klart positiva effekter beträffande sysselsättning, produktion och lönebildning. ... I framtiden kan det krävas ytterligare öknings av stödinsatserna om en fortsatt

¹Texten är från 1979.

²Niklasson (1979:76) som menar att detta förhållande inte bör avhålla någon från att göra partiella eller begränsade utvärderingar, något som jag utan vidare instämmer i, dock med tillägget att förutsättningarna för studierna klart måste framgå.

utjämning mellan de nuvarande stödområdena och övriga landet skall kunna uppnås." (SOU 1978:46 s. 67). Nettoeffekten av det regionalpolitiska stödet under perioden 1965 till 1975 ligger enligt expertgruppens sammanfattning på mellan 10 000 och 20 000 arbetstillfällen (SOU 1978:46 s. 8).

ERU:s redovisning av sina utrednings- och forskningsrön inför riksdagens och regeringens ställningstagande 1978/79 om regionalpolitikens fortsatta utformning sker för det första i en huvudtext (SOU 1978:46) varur de ovan citerade slutsatserna är hämtade. Denna skrift består av en resonerande och redovisande del och av en sammanfattande. För det andra presenteras, ävenledes i SOU-serien (1978:47), en bilagedel med sammanfattning av forskningsrön syftande till att underbygga huvudtexten. För det tredje - vändande sig till en fackpublik - publicerar ERU en serie med "underlagsmaterial" som i vissa fall också fungerar som underlag för bilagedelens resonemang.

Jag har i annat sammanhang publicerat min analys av dessa utredningsprodukter (Gullberg 1981)¹, utvidgat denna analys med granskning av ett senare ERU-arbete och kompletterat med en genomgång av internationell litteratur om utvärdering av regionalpolitiska effekter (Gullberg 1985b) och dessutom utformat strategier för regionalpolitisk utvärderingsverksamhet (Gullberg 1985a). Här följer en kort sammanfattning av den ursprungliga granskningen. ERU formulerar problematiken kontrafaktiskt. Antaganden om en konstant eller konstant förändring av industrisysselsättningen i stödområden för det fall regionalpolitiken inte införts och intensifierats är emellertid otillräckliga för att ge presenterade slutsatser trovärdighet. Alternativa tolkningar är minst lika rimliga. ERU:s material varierar på ett intressant sätt i entydighet. Från bakgrundsmaterialet är det svårt att dra några bestämda slutsatser om regionalpolitikens effekter. Slutsatserna blir så allt mera entydiga - utan att nytt material som motiverar detta tillföres - ju närmare den korta sammanfattningen i början av SOU 1978:46 man kommer.

¹Till en del är följande passager oförändrade avsnitt från denna text.

De besynnerligheter som möter oss när slutsatserna om regionalpolitikens sysselsättningseffekter konfronteras med metod och belägg skall kort sammanfattas.

1. Slutsatsen talar om nettoeffekter men undersökningsuppläggningsen har inget att säga om kausala relationer - inte ens resonemangsmässigt, t.ex. genom hänvisning till någon teori, rättfärdigas kausalitetspåståendet. En fullt legitim alternativ slutsats är att det regionalpolitiska stödet i tiden sammanfaller med en ökning av industrissysselsättningen inom stödområdena (korrelation i stället för kausalitet)¹.

2. Slutsatsen talar om 10 000 till 20 000 nya jobb. Dessa tal bygger på antaganden om att stödområdenas industriandelar antingen skulle förblivit oförändrade på 1965 års nivå eller minskat i samma takt som under åren 1960-65 under förutsättning att stöd ej utgått. Några argument ges varken för dessa antaganden eller mot ett annat och alls ej orimligt, innebärande att en ökning av industriandelarna skulle kommit till även utan stödinsatser. Med utgångspunkt från det av ERU redovisade materialet skulle helt andra och lika rimliga siffror kunna räknas fram.

3. Genom att använda andra metoder än de som redovisas i ERU:s båda SOU-skrifter har två forskare kommit fram till slutsatser angående regionalpolitiken som inte stämmer överens med ERU:s. Dessa resultat redovisas ej.

¹Hade slutsatsen varit formulerad i samma försiktiga anda som när ERU presenterat sina studier som "undersökningar av hur regionalpolitiken kan ha påverkat stödområdenas sysselsättning ... under perioden 1965-75" (SOU 1978:47 s. 57, min kursivering A.G.) hade inget varit att invända mot den. Men då hade den också varit ointressant. I samma försiktiga anda konstaterar ERU "/a/tt den sjunkande sysselsättningsandelen 1960-65 förbyttes i en stigande trend efter 1965 är ... förenligt med ett antagande om att regionalpolitiken fått mätbara effekter på industrissysselsättningens regionala fördelning 1965-1975." (SOU 1978:47 s. 66, min kursivering A.G.). Det är för att råda bot på den osäkerheten som ERU presenterar sina egna undersökningar. Men som jag visat i Gullberg (1985b) blir vi genom dem inte så mycket klokare, varför citatet är en mera träffande slutsats än den ERU ger: "Politiken har givit klart positiva effekter beträffande sysselsättning, produktion och lönebildning." (SOU 1978:46 s. 67).

Det är svårt att tro att tillfälligheter ligger bakom den klara tendens att läsa in positiva effekter av den förda regionalpolitiken i resultaten. Vid en närmare granskning löser effekterna upp sig i osäkra antaganden. I stället skymtar jag det fenomen jag inledningsvis benämnde politikcentrism bakom de tystnader, luckor och djärva slutledningar jag redovisat ovan och som också visat sig vara förutsättningarna för det resultat riksdag och regering förelades av expertgruppen. Denna tendens att vilja tro på att en politik man har medverkat i utformningen av faktiskt har positiv betydelse t.o.m. utöver vad en svårtolkad och motsägelsefull verklighet vid sträng granskning riktigt ger anledning förmoda är som jag redan konstaterade i inledningen till detta avsnitt inget märkvärdigt. Men tendensen måste erkännas. Då blir det också lättare att förstå och bedöma utvärderingsresultat. Utvärderarna väljer alltid - har det hävdats - sida redan innan de empiriska undersökningarna gjorts (Berk & Rossi 1976:349).

Hur långt har den offentliga effektmätningen nått?

Genomgången av EFA:s och ERU:s utvärderingar¹ ger vid handen att rigorösa problemformuleringar och mångåriga ansträngningar inte utgör någon garanti för att en - ur sanningssökandets perspektiv - framgångsrik bedömning av politikens resultat skall komma till stånd. Trots att effektbedömningarnas principiella svårigheter formulerades klart redan på ett tidigt stadium och en potentiellt lovande - men väl ändå i sig otillräcklig - strategi antytts har metodproblemen och deras hantering inte givit frågeställningarna tillräcklig kraft att bana sig väg fram till lämpliga förfaringsätt som samtidigt måste vara så robusta att de förmår att motstå önsketänkande och påtryckningar och så känsliga att de kan fånga eventuella kausala utflöden från politiken.

Kontrollgruppsförfarandet har ägnats stor uppmärksamhet av EFA. Därigenom har dess begränsade värde blivit uppenbart och betydelsen av politikens indirekta effekter tydlig. Hur dessa

¹För undvikande av missförstånd kan det här vara på sin plats att understryka att de båda expertgruppernas insatser endast bedömts just vad gäller utvärdering (av effekter ex-post).

senare skall kunna ringas in förblir dock ett obehandlat problem. ERU ägnar stor energi åt omfattande räkneövningar som emellertid till sin grundstruktur visar sig vara schablonmässiga makrouppskattningar byggda på godtyckliga antaganden.

Metoden fångar korrelationer mellan politik och sysselsättning snarare än effekten av den förra på den senare. En strategi som till någon del skulle kunna motverka detta hypotetiska bollande med stora tal vore ett sökande efter de mekanismer som kan tänkas producera den antagna sysselsättningsökningen. Är det t.ex. möjligt att genom intensiva fallstudier rekonstruera kausala kedjor från stödformer till investeringsbeslut och ökande sysselsättning?

Någon fast mark att stå på har ansträngningarna att utveckla utvärderingsmetoder alltså inte lett fram till. Tolkningsutrymmet är vidsträckt, ja kanske hålls det vidöppet. I varje fall drar tolkningar och sammanfattningar av resultat iväg långt i en riktning som framhäver den förda politikens förtjänster i det förflutna och önskvärda fortlevnad och t.o.m. expansion i framtiden. Politikcentrismen är påtaglig. Man har sagt att ERU:s uppgift borde vara att bidra med kunskap som ger makten fantasi (Wirén 1974:320). Kanske har det lyckats. Makten har i varje fall fått hjälp att tala om att det den gjort är bra och att den nog bör få göra mera. Men vad händer med sanningen när maktens fantasi växer?

Ett ständigt återkommande missnöje sägs utmärka den amerikanska utvärderingsverksamheten. Så inte i Sverige. Uppenbarligen kan detta inte bero på att metodproblemen är lösta i vårt land. Inte heller verkar det som om en fristående och oberoende ställning för utvärderingsverksamheten skulle borga för ett lidelsefullt sanningssökande. Har vi i det svenska samhället en tillräckligt korporativ organisation, i synnerhet en tillräckligt formbar och modulerande, lagom politiserande och kunskapsinriktad statlig utredningsapparat för att denna utvärderingsforskningens brydsamma situation mellan maktens och sanningens rättesnören aldrig skall komma upp till ytan? Jag tror att så är fallet. Kan det också bero på att den politiska styrningen är lagom stark. Den finns, men inte så att utvärderarnas känsla av att vara fackmän - dvs. leverantörer av kunskapsunderlag för rationellt besluts-

fattande - rubbas? Kan det bero på att utredningsväsendet är tillräckligt nära politik och byråkrati för att göra det möjligt för en uppfattning om utvärderingarnas praktiska betydelse att leva vidare, samtidigt som avståndet är tillräckligt stort för att utredningar skall kunna leva ett eget liv när så är lämpligt, skall kunna föras åt sidan när så erfordras och lyftas fram när så finnes passligt? Är kanske utredningsväsendet och utredarrollen just institutionen som förmår blanda makthänsyn och kunskaps-sökande i till situationen väl anpassade proportioner? Ett är dock säkert. Att ge svar på frågan vad politiken åstadkommit leder till stora teoretiska/metodologiska och institutionella/politiska problem och Sverige är här inget undantag. Vad vi har kommit särskilt långt med här i landet är kanske att gömma undan svårigheterna. Det "olustiga äktenskap" (Gordon & Morse 1975:351) som politik och vetenskap ingått genom utvärderingsforskningen hålls här samman genom utredningsväsendets familjeterapi. Vem som lider mest, den teknokratiska politiken eller den anpassliga vetenskapen skall jag låta vara osagt, men tjänar på affären, det gör maktutövningen och inte sanningssökandet. Och vi får aldrig veta vad politik och planering varit och är goda för. Vill vi det?

DUBBELSTRATEGI FÖR EN VETENSKAPLIG UTVÄRDERING

Utvärderingens paradox¹ som omintetgör den sanningssökande och avmystifierande roll vetenskapen här borde spela måste angripas på två fronter samtidigt. Annars kan den inte mildras eller upplösas. Dels måste kulturvetenskaplig teori och metod utvecklas dithän att effektfrågorna kan besvaras på ett trovärdigt sätt, dels måste den vetenskapliga utvärderingsverksamheten arrangeras så att ett sökande efter de genuina svaren inte förhindras utan tvärtom stimuleras. Framryckning på den ena fronten blir verkninglös om inte avancemang samtidigt sker på den andra.

Institutionella arrangemang

Det är ett välkänt faktum att den institutionella placeringen av ett utvärderingsprojekt påverkar uppläggning, resultat och tolkning. Med institutionell placering menar jag då inte bara vid vilken institution projektet bedrivs utan också med vilka lojaliteter och aspirationer arbetet utföres.

Betydelsen av detta faktum är emellertid vanligen missförstådd. I vissa sammanhang överskattas den grovt som t.ex. när undersökningar utförda av en viss part alltid stämplas som vilseledande. I enskilda fall är det studiens uppläggning och genomförande som är avgörande för resultatens bärkraft. Starka parter kan i vissa situationer och på vissa områden vara intresserade av både att producera avancerade och rättframma utvärderingsstudier och att sprida de ofrisonerade resultaten.

¹Paradoxen växer fram ur en situation där utvärderaren uttryckligen ombeds att med vetenskaplig auktoritet besvara frågor om reformresultat och politikeffekter. Samtidigt är det den alltid outtalade eller i varje fall hemlighållna (hos beställaren inte sällan omedvetna) efterfrågan på vetenskapligt (sken)legitimerade argument som är den verkligt effektiva. En utvärderare som oreflekterat och efter bästa förmåga försöker besvara ställd fråga löper stor risk att misslyckas i branschen och den beställare som offentligt ger uttryck för den effektiva efterfrågefunktionen äventyrar utvärderingens användbarhet.

För utvärderingsverksamheten som sådan underskattas tvärtom vanligen betydelsen av de institutionella förhållandena. Resultat som ur kunskapssynpunkt inte bort accepteras läggs till grund eller används som argument för beslut. Vetenskapens mytiska kraft (House 1980:112) gör sig gällande och efterfrågas därför också. De kraftiga politiseringstendenserna förhindrar inte att enstaka demystifierande utvärderingar genomförs. För genren i sin helhet är dock detta inflytande katastrofalt. Det existerar nämligen knappast någon institutionaliserad kraft som verkar i riktning mot sanningssägande eller metod- och teoriutveckling. Utvärderingsmetodik och då i synnerhet effektmätning är synnerligen svåra och förvånansvärt outvecklade områden. Den stora efterfrågan på utvärderingsstudier motverkar t.o.m. ett seriöst arbete med teori- och metodutveckling då tillförlitliga och realistiska resultat inte är de sökta storheterna.

Nu är det inget fel i att parter söker argument för sina intressen och då bl.a. använder utvärderingar¹. Det kan istället vara till fördel att dylika mer eller mindre partsbundna utvärderingar genomförs. Problemet är dock att ingen motvikt förekommer, att inga institutioner företar angelägna utvärderingar av samhällsreformer och politiska program på tillräckligt avstånd från intressenter för att sanningen skall kunna sökas på ett mera reservationslöst sätt och att förbättringar i utvärderandets teori och metod skulle kunna komma till stånd.

House som menar att utvärderingar inte kan vara något mer än övertalningshandlingar (a.a.:72), att de till sin natur är politisk aktivitet (a.a.:121) och att de olyckligtvis alltför ofta tenderar att tjäna de styrandes intressen och inte den stora allmänhetens (a.a.:147f) förordar en strategi i vilken jag identifierar två intressanta komponenter. Enligt House skall utvärderingar inte bara vara sanna. De skall också vara rättvisa (a.a.:121). Utvärderaren bör därför sträva efter att involvera relevanta intressenter i utvärderingen (a.a.:187). Utvärderingar bör vidare företas på basis av en utvärderingsöverenskommelse. House diskuterar vilka punkter

¹Mera tveivelaktigt är det när partsförhållandet är dolt som ibland i offentliga utredningar. Jfr. Gullberg (1985b).

som bör ingå i en sådan överenskommelse (a.a.:159ff) och under vilka yttre omständigheter den skall träffas för att bedömas som juste (a.a.:170ff).

Om syftet med utvärderingsverksamhet är att sträva efter sanning och demystifikation framstår de från House hämtade rekommendationerna som otillräckliga men anpassade till rådande situation. En bred intresserepresentation i utvärderingar motverkar måhända ensidighet och partiskhet. Men någon pådrivande kraft i sanningsökandet är knappast sådana organisatoriska konstruktioner. Frågan om vad som är sant tenderar nämligen att bli en förhandlingsfråga. Utrymmet för vad som kan sägas minskar. Sanningen kan vara - och är det ofta - partisk i förhållande till existerande intressen. En formell överenskommelse om utvärderingsvillkoren kan ha sina fördelar men på samma sätt som själva utvärderingsaktiviteterna snabbt blev omspunna av dolda och öppna budskapsparadoxer hotar naturligtvis själva utvärderingsöverenskommelsen att bli det.

Som komplement till de utvärderingar som kommer att företas av parter förespråkar jag en från parterna formellt och reellt fristående utvärderingsverksamhet. En sådan verksamhet kan med fördel härbärgeras i en (reformerad och förstärkt) universitetsorganisation. House menar att en opartisk och fristående utvärdering blir blodfattig därför att den är irrelevant genom att ej representera de djupast berörda parterna (a.a.:187). Opartiskheten kan ha olika innebörder. Att i en utvärdering eftersträva resultat och skrivningar som inte gynnar någon part mer än någon annan leder givetvis till anemi. Om med opartiskhet istället avses systematiska försök att inte ta hänsyn till vilka intressen som inblandade parter har utan istället med största möjliga sanningslidelse besvara resta frågor tenderar utvärderingarna att bli brännande stoff. En röst i utvärderingskören som inte i första hand sneglar på intressentnyttan utan efter hur undersökta verksamheter faktiskt fungerar bör kunna spela en central roll i det allmänna demokratiska samtalet. Politiska program och reformer är faktiskt något som angår flera än de inom området organiserade intressenterna.

Fristående och gärna universitetsförankrade utvärderingsinstitutioner garanterar på intet sätt en produktion av tillförlitliga utvärderingar och ett utvecklingsarbete gällande utvärderingsteori och -metod. Men ett sådant arrangemang ökar åtminstone chanserna för att dylika aktiviteter skall kunna utvecklas. Utan organisatoriska nyordningar är något sådant högst osannolikt.

Metodologiska överväganden

Termen utvärderingsstudie täcker ett, som det framgått, vitt område både när det gäller behandlade frågeställningar, ambitionsgrad och institutionell förankring. Därför har också en klassifikatorisk verksamhet kommit att utövas.¹ Själv förespråkar jag en tredelning som i första hand syftar till att skilja perifera problem från det som är utvärderandets verkliga metodologiska stöttesten. Den ser ut som följer: 1. Utfallsprognos eller utvärderingar ex-ante, 2. Uppföljningar eller jämförelser mellan någon form av förhandsföreställning och hur det sedan faktiskt blev och 3. Utvärdering (i egentlig mening) eller studier av resultat och effekter.² Det är i studier av typ tre som vi möter de verkligt centrala metodfrågorna. Vilka effekter har den undersökta reformen, politiken (haft)? Detta problem är, liksom utvärderingens institutionella dito, välkänt men på det hela taget missförstått. Det uppfattas nämligen ofta som löst på ett principiellt plan genom en experimentell ansats. Av praktiska och moraliska skäl är emellertid experimentet inte alltid eller rent av sällan möjligt att tillämpa i sociala system varför den utvärderande effektstudien

¹Så skiljer t.ex. House (1980) mellan åtta eller egentligen nio olika typer av utvärdering (a.a.:21ff) på ett sätt som ur analytisk synvinkel innebär en otillfredsställande blandning av metodologiska och institutionella kriterier. Houses förslag står sig bäst som en ad hoc klassificering av förekommande utvärderingar. Nilstun (1980) presenterar med utgångspunkt från ett analytiskt schema fem olika typer av studier. Hans klassifikation har pedagogiska kvaliteter.

²Jfr. här Gullberg (1985a:22f).

tvingas retirera till s.k. kvasiexperiment eller naturliga experiment¹. Dessa framstår enligt detta synsätt som näst och tredje bästa ansatserna - ej perfekta men ändå fungerande.

Den springande punkten är att kunna skilja mellan korrelation (tidsmässig överensstämmelse t.ex. mellan politiskt beslut och viss relevant förändring) och kausalitet (att det politiska beslutet åstadkommer den relevanta förändringen). Hjälpa på vägen ger en s.k. kontrafaktisk formulering, dock utan att lösa problemet (jfr. Gullberg 1986:91ff). Formuleringen lyder: effekten av en reform, politik etc. är skillnaden mellan världen sådan den ser ut med reformen, etc. genomförd och sådan den skulle ha sett ut reformen förutan. Experimentet och dess "släktingar" är alla metoder som försöker producera dessa kontrafaktiska jämförelser och det sker genom att på olika sätt försöka stänga de system som studeras på ett sådant sätt att bara politiken varierar och utfallet därför restlöst kan hänföras till denna.

Det finns emellertid mycket starka argument för att samhällliga system inte låter sig stängas med mindre än att de förstörs eller förvrängs². Experimentet detroniseras därmed från att vara den excellentametoden inom det kulturvetenskapliga området. Den intressanta situationen uppstår att fall- och processtudier kan användas och utvecklas inte med en ursäkt för att de mera avancerade metoderna ej går att tillämpa utan i kraft av sina egna förtjänster. Det blir också möjligt att knyta mer direkt an till den intuitiva föreställningen om att kausalitet handlar om mekanismer varigenom något produceras. Det är det som ett politiskt program, en reform etc. producerar som bör vara den sökta storheten vid utvärdering.

I konventionell metodologi har fallstudier tilldelats en hypotesgenererande roll. Test av hypoteser anses endast de experimentella, kvasiexperimentella och naturliga experimenten lämpade för. Men genom att en stängning av de studerade samhällliga systemen inte låter sig göra finns det ingen entydig relation mellan empiriska mönster (korrelation) och effekter (kausalitet). De olika

¹Se Gullberg (1986:109ff) för en redogörelse för innebörden.

²För en redogörelse av argumentationen se Gullberg (1986:126ff).

"experimentmetoderna" kan därför inte fungera som dömande instans i effektfrågor men tjänar utmärkt i rollen som hypotes-generatorer. Dessutom kan olika manipulationer av empiriskt material i s.k. naturliga experiment bidra till detaljerad beskrivning av det som skall förklaras dvs. relationen mellan politik, genomförande och utfall¹.

Det alternativ till stängningsstrategin vid effektstudier som jag förordar kan sammanfattas i följande punkter²:

1. Effektundersökningar (ex-poststudier av en viss planerings kausala inverkan på ett visst område) kan modellmässigt formuleras gälla P - E; (E i O).
2. Både P och O undersöks som (fortlöpande) producerade och reproducerade, dvs. som processer.
3. Det eventuella mötet mellan P och O undersöks. När P, och det P influerar, fram till O och det som producerar O? Hur ser de eventuella effekterna ut - dvs. på vilket sätt påverkas O:s "produktionsförhållanden" och det faktiskt producerade O:et?
4. Steg 3. utföres, inte genom försök till stängning utan genom hänsynstagande till mångfalden av förekommande bestämningar i konkreta fall. Endast genom kontextuering av de processer som genererar P och O är detta möjligt att göra på ett rimligt sätt.
5. Experimentell verksamhet är ingalunda omöjlig att bedriva inom samhälls- och människovetenskaperna. Tolkningen av vad experiment innebär och av resultaten kan dock inte ske enligt gängse formler. Experiment skall normalt sett uppfattas på precis samma sätt som aktionsforskning. Någonting sätts igång (i detta sammanhang skulle det kallas P) och effekterna studeras. Både P och O får analyseras enligt principerna ovan eftersom stängning av systemet inte kan uppnås.
6. Det ovan sagda innebär ett förordande av teoretiskt förankrade fallstudier eller, bättre, realprocesstudier. Dessa kan genomföras utan ursäkter och inte därför att gängse "vetenskapliga" metoder är svåra eller omöjliga att använda. Dessa kontextu-

¹Om relationen mellan översiktlig empiri och fallstudier se t.ex. Gullberg (1985a:64ff).

²Jag hämtar dessa punkter från Gullberg (1986:146ff).

erande processtudiemetoder är vetenskapliga och en svårighet är istället att i studier som använder gängse kausalmetoder skilja substantiell kunskap från rena metodartefakter.

7. Mätningar i variabler som avbildar processerna P och O kan lämpligen företas vid flera tidpunkter. När det gäller O ger detta indikationer om hur det presumtiva effektområdet såg ut före, under och efter P och när det gäller P kan därigenom en första specifikation av P och dess eventuella förändringar över tid uppnås. Vid dylika beskrivningar är det emellertid viktigt att inte förväxla variabelvärdesförändringar med realprocesser. Vid storskaleundersökningar kan en lämplig strategi vara att med hjälp av massdata försöka utskilja några typfall som processtuderas, varefter dessa ånyo konfronteras med massdatabildden.

8. Generaliseringsproblemet har traditionellt varit en av de vanligaste invändningarna mot fallstudier. Man har kunnat konstatera att fallstudie har lagts till fallstudie utan att den generella förståelsen av studerade fenomen synes öka. Till en del beror detta på det sätt som fallstudier vanligen bedrivs. Fallstudier som på intet sätt inriktas mot det som kan vara generellt i det specifika, som inte söker pröva teoretiska antaganden och som inte heller relaterar sig till liknande studier kan inte förväntas bidra till en generell kunskapsstillväxt. Eller korrektare, överlämnar detta arbete till läsaren som får ikläda sig rollen som sekundäranalytiker. Men även bortsett från denna faktiska omständighet föreligger ett genuint problem: i vilka former kan de kunskaper som utvinnes ur realprocesstudier ackumuleras? Eller med andra ord: vilken typ av teoretiserande är lämpligt? Jag menar att kunskaps-kondensering bör ske i form av teorier/teorifragment gällande typer av agenter, typer av processer och typer av kontexter. Till en del kommer då generaliseringsproblemet att bli ett klassificeringsproblem.

Sammantaget innebär detta att metदानsträngningarna kan ges en mera realistisk inriktning. Vi behöver inte acceptera de uppenbarligen orimliga, humeanskt¹ baserade, undersökningsmetoderna som

¹Humeanskt efter David Hume. För en redogörelse se t.ex. Gullberg (1986).

bygger på möjligheten att stänga system¹ och vi är för den skull inte tvingade att ge upp vetenskapligheten, dvs. överge kausalitetsanspråken inom samhälls- och människovetenskaperna.

Så ser en metodstrategi ut som är betydligt mera lovande men också mera krävande än de vanliga metodrecepten. Den kan bara utvecklas under gynnsamma institutionella förhållanden.

Utvärdering i praktiken

För den vetenskapsman som åtar sig att göra utvärderingar är säkerligen den institutionella inplaceringen en långt mera betydelsefull fråga än de metodologiska frågorna². En gångbar utvärdering behöver alls inte innehålla de metodologiskt sett krävande studier som skulle ha fordrats för att möjliggöra dragna slutsatser. För framgång i utvärderingsbranschen är förmågan att bedöma beställares och andra intressenters nuvarande och framtida förväntningar av avgörande betydelse. Med skicklighet i detta värv kan en utvärderare nå långt och kanske t.o.m. tillskansa sig något, om än villkorligt, inflytande. Men det finns en fara med denna exploatering av vetenskapens good-will. Ju mer vetenskapens auktoritet åberopas och sprids desto mer urvattnat blir dess reella innehåll. Risker för rekyl är överhängande. Spelar vetenskapssamhället bort sin starkaste trumf - sanningssökandet - för en ringa plats i den offentliga maktens korridorer finns ingen räddning om bluffen synas. Problemet är allvarligt och kollektivt.

¹Nordenfelts rekommendation är välformulerad: "It is important to stress the primacy of detailed investigation. The reason for this is the complete openness of most historical (inklusive nutidshistoriska A.G.) systems including the system of a human individual. Many standard causal predictions break down, since the ceteris paribus clauses of their generalizations are not fulfilled." Nordenfelt (1981:54).

²Detta gäller givetvis också för utvärderingar som utföres utan vetenskaplig (sken)legitimitet.

Efterskrift och förslag

Denna rapport har en brokig tillkomsthistoria som på sätt och vis illustrerar utvärderingsforskningens institutionella dilemma. Mitt intresse för utvärderingsstudier emanerar ur ett arbete med planering. I vissa avseenden framstår nämligen utvärderingsansatser ex-post som en tänkbar utväg ur de paradoxer som planering och planeringstänkande befinner sig i (Gullberg 1981,1986).

Under en vistelse vid Wissenschaftszentrum i Berlin 1978/79 som BFR-stipendiat orienterade jag mig i den rikhaltiga amerikanska och tyska litteraturen om utvärdering. Detta resulterade dels i en första version av den här föreliggande texten, dels i ett forskningsprogram utarbetat tillsammans med Bjarne Lindström om Storstockholms utveckling efter andra världskriget där en utvärderingsansats skulle tillämpas. Litteraturstudien var avsedd som en förberedelse för Stockholmsprojektet som dock av finansiella skäl aldrig kom till utförande. Som surrogat för denna "praktiska demonstration" av en alternativ utvärderingsstrategi har olika metodologiska och sporadiskt finansierade kommentarer fått tillgripas (t.ex. Gullberg 1985a,b, 1986). Dessa har kortfattat refererats här. Sedan slutet av 1970-talet har flera satsningar gjorts på utvärderingsforskning. Vissa av dem har varit synnerligen omfattande. Utvärdering har också generellt sett ökat i popularitet. Utvärderingsverksamhetens grundläggande problem så som jag då, och i denna skrift, formulerat dem har emellertid därigenom inte förts närmare någon lösning eller rimlig behandling. Därför är denna snart decenniegamla analys av utvärderandets inneboende egenskaper och potentiella möjligheter nu t.o.m. mera aktuell än när den först formulerades.

Mina förslag till bättring lyder helt kort¹:

* Inrätta fristående forskningsorgan för utvärdering av politik-effekter.

¹Jag hämtar detta från Gullberg (1985a:1).

- * Stärk därvid den demokratiska processen genom att minimera intressenters inflytande över kunskapsproduktionen och förse medborgarna med politikgranskningens resultat.
- * Ge de gängse standardiseringsinriktade metoderna en roll som svarar mot deras förmåga; som hjälpmedel vid beskrivning.
- * Inrikta ansträngningarna mot att utveckla realistiska kausalundersökningsmetoder i form av flernivåbaserade, teoretiskt medvetna processtudier som utgår från samhälleliga förhållanden som skapade och återskapade.
- * Formulera effektfrågan som gällande: när politiken och det politiken initierar fram till det som skapar och återskapar relevanta samhällsförhållanden.

APPENDIX: Fallstudier i "Arbetsmarknadspolitik i förändring" SOU 1978:60

Denna text anknyter till not 2 sid. 76.

Fallstudierna är efter mitt sätt att se saken fem till antalet: 1. Försök med utökad arbetsförmedling för långtidsarbetslösa, 2. Införande av allmän platsanmälan, 3. Effekter av geografisk rörlighet, 4. Etablering av Algots Nord och 5. Marginella sysselsättningsstöd. Uppläggnings- och undersökningarna skall mycket kort beröras.

1 undersökning 1. har personer som erhållit utökade arbetsförmedlingsinsatser och personer som ingår i kontrollgruppen fördelats slumpmässigt, varigenom en av de invändningar jag ovan gjorde mot utbildningsstudien bortfaller. Vidare påpekas att de båda undersökta gruppernas resultat trots den slumpmässiga fördelningen inte med säkerhet är oberoende av varandra. De som får en utökad hjälp vid arbetsförmedlingen kan komma att tränga ut dem som fortsätter på normala villkor. Detta benämns trängseffekter och är exempel på en typ av indirekta effekter (SOU 1978:60 s. 185). Detta har jag tidigare kallat konkurrens-effekt. I undersökningen tror EFA ej att sådana effekter spelat in, eftersom under den aktuella perioden stor efterfrågan på arbetskraft rådde. Denna studie visar klara metodframsteg i förhållande till de studier som redovisats från EFA:s första betänkande. Dels är det verkligen en politikförändring som är föremål för undersökning, dels har försök gjorts att resonera om en av de tänkbara indirekta effekttyperna. Likväl kvarstår frågetecken som t.ex. vad var det i den utvidgade arbetsförmedlingen som orsakade förbättringen, vilka var de indirekta effekterna av försöksverksamheten (utöver tvångseffekterna), vilken betydelse hade det förhållandet att det verkligen rörde sig om en (fräsch?) försöksverksamhet för förmedlarnas entusiasm - kan resultaten generaliseras till en situation där "utvidgningen" är normal?

Undersökning 2. försöker bedöma effekterna av en lagstiftning som genomförts successivt för olika län. Här förefaller det svårare att få fatt på effekterna än i fallet utvidgad arbetsförmedling menar EFA eftersom målgruppen är svårare att peka ut. "Problemet med i samhällsekonomin vitt spridda återverkningar kommer ... till uttryck i en helt annan utsträckning när det gäller en informationsändring /vilket är vad den obligatoriska platsanmälan syftar till/ av kollektiv karaktär." (SOU 1978:60 s. 228). Egentligen förhåller det sig väl så att det är svårare att förbigå de indirekta effekterna när ingen klart avgränsad målgrupp utpekas för sökandet av de direkta effekterna. Inströmningen av lediga platser till förmedlingarna och förändringen i antalet kvarstående platser vid dessa förmedlingar jämfördes mellan de län som lagstiftningen först kom att omfatta och övriga Sverige. Jämförelser gjordes både före och efter lagens ikraftträdande (SOU 1978:60 s. 229). Inströmningen av lediga platser ökade i de tre berörda länen, men de insamlade uppgifterna om kvarstående platser var svårtolkade. Att säga något bestämt om reformernas effekter på arbetsmarknadens sätt att fungera visade sig därför svårt.

Studie 3. utgör i huvudsak en upprepning av och jämförelse med utvärderingen av den geografiska rörligheten i EFA:s första betänkande. De metodkommentarer jag gjorde till detta arbete är giltiga även här.

Undersökningen av Algots Nord - studie 4. - tar upp konsekvenserna av en lokalisering. Före/efterjämförelser har kompletterats med två kontrollgrupper. Dels en med kvinnor som sökt men inte fått/tagit arbete på de tre berörda orterna, dels kvinnor som slutat sina anställningar vid Algots Nord (SOU 1978:60 s. 317). Dessutom försöker man sig på en kostnads-intäktskalkyl på samhällsekonomisk nivå, dels ex-post för perioden 1972-76, dels ex-ante för perioden 1977-82.

Studie 5. av de marginella sysselsättningsstöden, dvs. stöd som täcker delar av företagets lönekostnader för en utökning eller inställd inskränkning av sysselsättningen pekar klarsynt på behovet av analys på såväl mikronivå (dvs. företagsnivå) och makronivå (i detta fall frågor om penningmängd och olika komponenter i den totala efterfrågan) samt att dessutom "det ömsesidiga beroendet mellan dessa nivåer /måste/ beaktas" (SOU 1978:60 s. 377). Denna studie ägnas åt olika modellresonemang för att belysa hur det marginella stödet kan tänkas påverka arbetsmarknaden och ekonomin i övrigt. Någon utvärdering i egentlig (empirisk) mening är det inte frågan om.

REFERENSER

- Ahlstrand, I. (1977) Metod för samhällsekonomisk utvärdering av kollektivtrafikinvesteringar, Statens vägverk, Stockholm.
- Andersson, R. (1978) "Samhällsekonomiska bedömningar av planeringsalternativ", Jämförelsemetoder i samhällsplaneringen, NU B 1978:3, ss. 15-70.
- Asplund, J. (1979) Teorier om framtiden, Stockholm.
- Bandura, B. & Gross, P. (1976) Sozialpolitische Perspektiven, München.
- Bandura, B. (hrsg.) (1976) Angewante Sozialforschung, Frankfurt a.M.
- Barber, B. (1962) Science and the Social Order, New York.
- Bartholomäi, R. (1977) "Ressortforschung: Aspekte der Vergabe und Forschungsbegleitung, in: Wissenschaftszentrum Berlin (hrsg.), Interaktion von Wissenschaft und Politik, Frankfurt a.M., ss. 285ff.
- Bauer, R.A. (ed.) (1966) Social Indicators, Cambridge, Mass.
- Bell, R. (1975) Att utvärdera sociala program, Stockholm.
- Berk, R. & Rossi, P. (1976) "Doing Good or Worse - Evaluation Research Politically Re-Examined", Social Problems, pp. 337-349.
- Berman, P. (1978) "The Study of Macro- and Micro-Implementation", Public Policy, pp. 157-184.
- Bernstein, I. & Freeman, H. (1975) Academic and Entrepreneurial Research, New York.
- Bohm, P. (1972) Samhällsekonomisk effektivitet, Stockholm.
- Brunsson, N. & Jönsson, S. (1979) Beslut och handling, Stockholm.
- Bunker, D. (1972) "Policy Sciences Perspectives on Implementation Processes", Policy Sciences, pp. 71-80.
- Börjesson, B. (1966) Om påföljders verkningar, Stockholm.
- Caplan, N. et al. (1975) The Use of Social Knowledge in Policy Decisions at the National Level, Ann Arbor.
- Caro, F. (1971) "Evaluation Research: An Overview", i densamme (ed.), Readings in Evaluation Research, New York, ss. 1-34.
- Christoffersson, N-O. (1979) "Är skolan effektiv?" i Riksrevisionsverket Är verksamheten effektiv?, ss. 98-112.
- Coleman, J.S. (1973) "The Principles Governing Policy Research", APA Monitor, p. 6.
- Coleman, J.S. (1975) "Problems of Conceptualization and Measurement in Studying Policy Impacts", in Dolbeare, K.M. (ed.) Public Policy Evaluation, pp. 19-40, Beverly Hills, Calif.
- Cook, T. & Scioli, F. (1975) "Impact Analysis in Public Policy Research", in Dolbeare, K.M. (ed.) Public Policy Evaluation, pp. 95-117, Beverly Hills, Calif.
- Cyr, A. & de Leon, P. (1975) "Comparative Policy Analysis", Policy Sciences, pp. 375-384.
- Dahlgren, J. & Lindkvist, L. (u.a.) Utvärdering - problem och utvecklingsmöjligheter, Ekonomiska institutionen, Linköpings universitet, Linköping.
- Dahlgren, J. & Lindkvist, L. (1976) Program-orienterad utvärdering, Lund.
- Dahlgren, J. & Lindkvist, L. (1977) Utvärderingsbehov och utvärderingssituationer, Ekonomiska institutionen, Linköpings universitet, Linköping.
- Dolbeare, K. (1974) "Impacts of Public Policy", i Cotter (ed.), Political Science Annual, Vol. 5, pp. 89-130, Indianapolis.
- Dye, T. (1966) Politics, Economics, and the Public Policy. Outcomes in the American States, Chicago.
- Easton, D. (1965) A Systems Analysis of Political Life, New York.
- Elster, J. (1972) Det distributive flertal, Oslo, (stencil).

- Elster, J. (1978) Logic and Society, Chichester.
- Esping, H. (1975) Utvärdering av offentlig förvaltning, Riksrevisionsverket, PM 1975-06-26.
- Esping, H. (1977) Förvaltning, förändring, framtid, Sekretariatet för framtidsstudier, Stockholm.
- Floden, R. & Weiner, S. (1978) "Rationality to Ritual: the Multiple Roles of Evaluation in Governmental Processes", Policy Sciences, pp. 8-18.
- Franke-Wikberg, S. (1977) Teoretisk referensram vid pedagogisk utvärdering - varför det? hur då?, Pedagogiska institutionen, Umeå universitet, Umeå.
- Franke-Wikberg, S. & Lundgren, U. (1980) Att utvärdera utbildning, Stockholm.
- Franke-Wikberg, S. & Lundgren, U. (1981) En antologi om pedagogisk utvärdering, Stockholm.
- Freeman, H. (1977) "The Present Status of Evaluation Research", In Guttentag, M. (ed.) Evaluation Studies Review Annual, Vol. 2, pp 17-51, Beverly Hills.
- Freeman, H. & Bernstein, I. (1975) "Evaluation Research and Public Policies", in Nagel, S. (ed.) Policy Studies and the Social Sciences, pp. 9-25, Lexington, Mass.
- Giddens, A. (1973) The Class Structure of the Advanced Societies, London.
- Glass, G. (1976) "Introduction", in Glass, G. (ed.) Evaluation Studies Review Annual, Vol. 1, pp. 9-12, Beverly Hills.
- Gordon, G. & Morse, E. (1975) "Evaluation Research", Annual Review of Sociology, pp. 339-361.
- Gorpe, P. (1978) Politikerna, byråkraterna och de nya styrformerna, Stockholm.
- Grumm, J. (1975) "The Analysis of Policy Impact" i Greenstein, F.I. & Polsby, N. (eds.), Policies and Policymaking - Handbook of Political Science, Vol. 6, pp. 439-473, Reading, Mass.
- Gullberg, A. (1972) Till den svenska sociologins historia, Stockholm.
- Gullberg, A. (1981) Det fängslande planeringstänkandet - planering och utvärdering granskade, BFR-rapport, 87:1981, Stockholm.
- Gullberg, A. (1983) Arbetsmarknadspolitiska effekter?, Papper till Svenska sociologförbundets årsmöte 1983, Stockholm.
- Gullberg, A. (1984) Det fängslande planeringstänkandet och sökandet efter en verklighetsutväg, Stockholm.
- Gullberg, A. (1985a) Att utvärdera regionalpolitikens effekter - en omöjlig uppgift?, NordREFO, Oslo.
- Gullberg, A. (1985b) "Lösliga effekter - om riter och rutiner inom regionalpolitisk utvärdering", Evaluering av regionalpolitik, Nord-REFO, Oslo.
- Gullberg, A. (1986) Det fängslande planeringstänkandet och sökandet efter en verklighetsutväg, andra utgåvan, Stockholm.
- Guttentag, M. & Saar, S. (1977) "Introduction", in Guttentag, M. (ed.) Evaluation Studies Review Annual, Vol. 2, pp. 11-14, Beverly Hills.
- Hargrove, E. (1975) The Missing Link, Washington, D.C.
- Haverman, R.H. (1972) The Economic Performance of Public Investments, Baltimore.
- Hellstern, G-M. & Wollmann, H. (1977) Sozialwissenschaftliche Untersuchungsregeln und Wirkungsforschungsregeln" i Peter Haugs (hrsg.) Res Publica, Festschrift für Dolf Sternberger, München.
- Hellstern, G-M. & Wollmann, H. (1978) Sanierungsmassnahmen, Schriftenreihe "Stadtentwicklung" des Bundesministers für Raumordnung, Bauwesen und Städtebau, Bonn.

- Hermerén, G. (1980) "Några logiskt distinkta steg i planeringsprocessen" i Att värdera byggd miljö, Stockholm.
- Horowitz, D. (ed.) (1967) The Rise and Fall of Project Camelot: Studies in the Relationship Between Social Science and Practical Politics, Cambridge, Mass.
- Horowitz, I. & Katz, J. (1975) Social Science and Public Policy in the United States, New York.
- House, E. (1980) Evaluating with Validity, Beverly Hills.
- Jacob, H. & Lipsky, M. (1968) "Outputs, Structure, and Power: an Assessment of Changes in the Study of State and Local Politics", The Journal of Politics, pp. 510-538.
- Lazarsfeld, P. (1975) "The Policy Science Movement" Policy Sciences, pp. 211-222.
- Liedman, S-E. (1978) "Humanistiska forskningstraditioner i Sverige. Kritiska och historiska perspektiv" i Forser, T. (red.) Humaniora på undantag?, Stockholm.
- Lind, R. (1979a) Förvaltningsrevision och utvärderingsforskning - en sammanfattande jämförelse, Research paper 6135, EFI, Stockholm.
- Lind, R. (1979b) Utvärderingsforskning - en litteraturoversikt, Research paper 6133, EFI, Stockholm.
- Lindblom, C. (1968) The Policy-making Process, New York.
- Lowi, T. (1975) "What Polical Scientists Don't Need to Ask About Policy Analysis", in Nagel, S. (ed.) Policy Studies and the Social Sciences, pp. 267-273, Lexington, Mass.
- Majone, G. & Wildavsky, A. (1978) "Implementation as Evolution", Policy Studies Review Annual, London.
- Mattsson, B. (1970) Samhällsekonomiska kalkyler, Stockholm.
- Nelson, R. (1974) "Intellectualizing about the Moon-Ghetto Metaphor: a Study of the Current Malaise of Rational Analysis of Social Problems", Policy Sciences, pp. 375-414.
- Niklasson, H. (1979) "Olika utvärderingsmetoder vid studiet av effekterna av arbetsmarknadspolitiska åtgärder", i Expertgruppen för utredningsverksamhet i arbetsmarknadsfrågor (ed.) Dokumentation från ett seminarium kring arbetsmarknadspolitisk forskning, Stockholm.
- Nilstun, T. (1980) Metoder för utvärdering av reformer i arbetslivet, Filosofiska Institutionen, Lunds universitet, Lund.
- Nordenfelt, L. (1981) Causation, Stockholm.
- Piven, F.F. & Cloward, R. (1971) Regulating the Poor, New York.
- Poland, O. (1974) "Program Evaluation and Administrative Theory", Public Administration Review, pp. 333-338.
- Pressman, J. & Wildavsky, A. (1973) Implementation, Berkely.
- Rakoff, S. & Schaefer, G. (1970) "Politics, Policy and Political Science. Theoretical Alternatives", Politics and Society, pp. 51ff.
- Rein, M. & White, S. (1977) "Policy research: Belief and Doubt", Policy Analysis, pp. 239-271.
- Rhoads, S. (1978) "Economists and Policy Analysis", Public Administration Review, pp. 112-120.
- Riksrevisionsverket (1975) Metoder för myndighetsanalyser, Stockholm.
- Riksrevisionsverket (1976a) Prestations- och effektbeskrivning, Stockholm.
- Riksrevisionsverket (1976b) Om metoder för att bedöma statliga myndigheter, Stockholm.
- Riksrevisionsverket (1976c) Mål & program i verksamhetsplanering, Stockholm.
- Riksrevisionsverket (1979) Är verksamheten effektiv?, Stockholm.
- Roos, L. (1975) "Quasi-Experiments and Environmental Policy", Policy Sciences, pp. 249-265.

- Rose, R. (1976) Managing Presidential Objectives, New York.
- Rossi, P. (1972) "Testing for Success and Failure in Social Action", i Rossi, P. & Williams W. (eds.) Evaluating Social Programs, New York.
- Rossi, P. & Wright, S. (1977) "Evaluation Research: an Assessment of Theory, Practice, and Politics", Evaluation Quarterly, pp. 5-52.
- Rossi, R. & Williams, W. (1972) "The Organization of the Volume and Some Key Definitions", i desamma (eds.) Evaluating Social Programs, New York.
- Sannerstedet, A. (1979) Fri konkurrens eller politisk styrning?, Lund.
- Schick, A. (1971) "From Analysis to Evaluation", Annals, pp. 57ff..
- Schick, A. (1977) "Beyond Analysis", Public Administration Review, pp. 258-263.
- Sheldon, E.B. & Freeman, H. (1970) "Notes on Social Indicators: Promises and Potential", Policy Sciences, pp. 97-111.
- Sjöblom, G. (1977) "The Cumulation Problem in Political Science. An Essay on Research Strategies", European Journal of Political Research, pp. 1-32.
- Smith, T. (1973) "The Policy Implementation Process", Policy Sciences, pp. 197-209.
- SOU (1967:11-13) Programbudgetering, Stockholm.
- SOU (1973:43-46) Budgetreform, Stockholm.
- SOU (1974:29) Att utvärdera arbetsmarknadspolitik, Stockholm.
- SOU (1978:46) Att främja regional utveckling, huvudtext, Stockholm.
- SOU (1978:47) Att främja regional utveckling, bilagedel, Stockholm.
- SOU (1978:60) Arbetsmarknadspolitik i förändring, Stockholm.
- Statskontoret (1971) Planerade förändringar, Stockholm.
- Statskontoret (1973) Organisation för långsiktspianering - åtta exempel från statsförvaltningen, Stockholm.
- Statskontoret (1974) Statsförvaltningen planerar, Stockholm.
- Statskontoret (1977) Om planering vid statliga myndigheter, Stockholm.
- Stephan, S. (1935) "Prospects and Possibilities: The New Deal and the New Social Research", Social Forces, pp. 515-521.
- Stevrin, P. (1974) Samhällsstyrning, Lund.
- Suchman, E. (1967) Evaluative Research, New York.
- Summers, A. & Wolfe, B. (1977) "Do schools make a Difference?", The American Economic Review, pp. 639-652.
- Svensson, C. (1973) Utvärdering - några arbetsmodeller, Företags-ekonomiska institutionen, Lunds universitet, Lund.
- Thompson, M.S. (1975) Evaluation for Decision in Social Programmes, Lexington, Mass.
- Van Horn, C. & Van Meter, D. (1976) "The Implementation of Intergovernmental Policy", in Jones, C. & Thomas, R. (eds.), Public policy making in a federal system, pp. 39-62, Beverly Hills, Calif.
- Weinberg, A. (1964) "Criteria for scientific choice", Phys. Today, 17:44.
- Weiss, C. (1973) "Where Politics and Evaluation Research Meet", Evaluation, nr. 3.
- Weiss, C. (ed.) (1977) Using Social Research in Public Policy Making, Lexington, Mass.
- Wholey, J. et. al. (1970) Federal Evaluation Policy, Washington, D.C.
- Williams, W. (1972) "The Capacity of Social Science Organizations to Perform Large-Scale Evaluative Research", i Rossi, P. & Williams, W. (eds.) Evaluating Social Programs, New York.

- Williams, W. (1976a) "Implementation Analysis and Assessment", in Williams, W. & Elmore, R. (eds.), Social Program Implementation, pp. 267-292, New York.
- Williams, W. (1976b) "Introduction", in Williams, W. & Elmore, R. (eds.), Social Program Implementation, pp. 3-14, New York.
- Williams, W. & Evans, J. (1972) "The Politics of Evaluation. The Case of Head Start", in Rossi, P. & Williams W. (eds.) Evaluating Social Programs, New York.
- Wirén, E. (1974) "Forskning som ger makten fantasi", Plan, nr. 6.
- Wollmann, H. (1979) "Implementationsforschung - eine Chance für kritische Verwaltungsforschung?", i densamme (hrsg.), Politik im Dickicht der Bürokratie, Leviathan, Sonderheft nr. 3.
- Zetterberg, H. (1962) Social Theory and Social Practice, Totowa, N.J.
- Åhren, P. (1975) Ekonomiska utvärderingsmetoder i samhällsplanering, Stockholm.



**Denna rapport hänför sig till forskningsanslag 850849-5
från Statens råd för byggnadsforskning till Hb SALG, Stockholm.**

R46: 1987

ISBN 91-540-4734-X

Statens råd för byggnadsforskning, Stockholm

Art.nr: 6707046

**Abonnemangsgrupp:
X. Samhällsplanering**

**Distribution:
Svensk Byggtjänst, Box 7853
103 99 Stockholm**

Cirka pris: 45 kr exkl moms