

B



FINANSDEPARTEMENTET

Kan myndigheter utvärdera sig själva?

Rapport till

ESO

Expertgruppen för studier
i offentlig ekonomi

Ds 1996:36



FINANSDEPARTEMENTET

KAN MYNDIGHETER UTVÄRDERA SIG SJÄLVA?

99 06 14

Ex B

Rapport till expertgruppen för
studier i offentlig ekonomi

Av Catharina Barkman
och Stefan Fölster

Ds 1996:36

SOU och Ds kan köpas från Fritzes kundtjänst. För remissutsändningar av SOU och Ds svarar Fritzes, Offentliga Publikationer, på uppdrag av Regeringskansliets förvaltningskontor.

Beställningsadress: Fritzes kundtjänst
106 47 Stockholm
Fax: 08-20 50 21
Telefon: 08-690 91 90

Svara på remiss. Hur och Varför. Statsrådsberedningen, 1993.
– En liten broschyr som underlättar arbetet för den som skall svara på remiss.

Broschyren kan beställas hos:
Regeringskansliets förvaltningskontor
Distributionscentralen
103 33 Stockholm
Fax: 08-405 10 10
Telefon: 08-405 10 25

NORSTEDTS TRYCKERI AB
Stockholm 1996

ISBN 91-38-20292-1
ISSN 0284-6012

Förord

Enligt budgetordningen skall myndigheternas verksamhet utvärderas löpande; detta såsom ett självklart inslag i den budgetprocess som skall fastställa anslagen till verksamheten i fråga. Myndigheterna har själva en viktig uppgift i denna utvärdering; deras "självutvärdering" skall redovisa vad de presterat och vilka mätbara resultat verksamheten åstadkommit.

Någon systematisk granskning av dessa utvärderingar har dock hittills inte ägt rum. Vi har således inte vetat huruvida myndigheternas självutvärderingar verkligen förmått tillföra budgetprocessen den kunskap som förutsatts.

I denna rapport presenterar Catharina Barkman, RRV, och Stefan Fölster, Industriens Utredningsinstitut, en kartläggning och kvalitetsgranskning av myndigheternas självutvärderingar. Ett antal fristående experter – företrädesvis forskare eller utredare med egen erfarenhet av fältet – har genomfört granskningen av ett 20-tal myndigheters självutvärderingar. Resultatet är nedslående:

- Det finns många myndigheter som inte har genomfört några självutvärderingar alls.
- Bland de genomförda självutvärderingarna är kvaliteten ofta låg; det gäller såväl metodmässigt som resultatmässigt.
- Slående är att självutvärderingarna ofta helt saknar ett tillförlitligt resonemang om ändamålsenlighet och kostnadseffektivitet i den granskade verksamheten.

Visst finns några lysande undantag, och man kan också diskutera representativiteten i urvalet av självundersökningar i denna rapport. Bark-

mans och Fölsters slutsats är likväl att "de flesta av självvärderingarna är av så dålig kvalitet att de faktiskt inte bör användas som beslutsunderlag". Därmed saknas ett centralt inslag i den nya budgetprocessen – enligt Barkman/Fölster själva "motorn" i processen. Följaktligen är det svårt att hävda att resultatstyrningen fungerar som avsett.

Denna rapport är således såväl viktig som kontroversiell. Den reser en rad frågor rörande budgetprocessen, utvärderingsmekanismen och resultatstyrningen. Den reser också frågor om hur utvärderingsprocessen skall kunna förbättras och kvaliteten på självvärderingarna höjas. Författarna avslutar sin rapport med ett antal förslag om hur detta skall kunna åstadkommas.

Det är vår förhoppning att rapporten därmed skall bidra till en diskussion om hur budgetprocessen skall kunna bli mer stringent. Som vanligt i ESO-sammanhang är analysen och resultaten författarnas egna.

Mörksund i juni 1996

Klas Eklund
Ordförande för ESO

Innehållsförteckning

1	Sammanfattning	7
2	Inledning	11
3	Självvärderingarna och den nya budgetprocessen	15
4	Kartläggning av statliga myndigheters självvärderingar	23
4.1	Vad är en självvärdering?	23
4.2	Myndighetsurvalet	25
4.3	Självvärderingarna – några generella drag	29
4.3.1	AMS	31
4.3.2	Banverket	32
4.3.3	Datainspektionen	32
4.3.4	Glesbygsverket	33
4.3.5	Invandrarverket	33
4.3.6	Jordbruksverket	33
4.3.7	Konsumentverket	33
4.3.8	Kriminalvårdsstyrelsen	34
4.3.9	Livsmedelsverket	34
4.3.10	Naturvårdsverket	34
4.3.11	Närings- och teknikutvecklingsverket	35
4.3.12	Patent- och registreringsverket	35
4.3.13	Riksförsäkringsverket	36
4.3.14	Rikspolisstyrelsen	36
4.3.15	Riksskatteverket	37

4.3.16	Riksåklagaren	37
4.3.17	SIDA	38
4.3.18	Skogsstyrelsen	38
4.3.19	Skolverket	38
4.3.20	Socialstyrelsen	39
4.3.21	Tullverket	39
4.3.22	Vägverket	40
4.4	Några slutsatser	40
5	Metod för kvalitetsbedömning	41
5.1	Hur har självutvärderingarna valts ut?	42
5.2	Hur har de sakkunniga valts ut?	44
5.3	Vad skall de sakkunniga jämföra med?	46
6	Kvaliteten i statliga myndigheters självutvärderingar	49
6.1	Helhetsbilden	49
6.2	Dokumentation, data och metod	57
6.3	Slutsatser, kostnadseffektivitet och objektivitet	63
6.4	Några goda exempel på självutvärderingar	68
7	Slutsatser och rekommendationer	71
	Referenser	77
	Abstract	79
	<i>Bilagor:</i>	
Bilaga 1	Termen självutvärdering och projektets uppläggning	83
Bilaga 2	Instruktioner till sakkunnigbedömning	87
Bilaga 3	Förteckning över utvärderingar som bedömts av sakkunniga	93
Bilaga 4	Förteckning över de sakkunniga	95
Bilaga 5	Sakkunnigbedömningar	97
	Förteckning över ESO:s tidigare publicerade rapporter	317

1 Sammanfattning

Beslutet att införa den nya statliga budgetprocessen sedan 1988 har ibland kallats för "den största förvaltningsreformen sedan Axel Oxenstierna". Den skulle ge regering och riksdag möjlighet att styra de statliga myndigheterna på basis av kunskap om faktiskt uppnådda resultat. Tanken var att myndigheterna själva skulle utvärdera och redovisa vad de presterat och vilka mätbara effekter man åstadkommit.

Under ett antal år har myndigheterna nu presenterat sina resultat i årsredovisningar, fördjupade anslagsframställningar samt i separata utvärderingsrapporter. En fundamental fråga i detta sammanhang är om myndigheternas utvärderingar av den egna verksamheten *kan* användas. I vilken omfattning genomför myndigheter utvärderingar som verkligen belyser relevanta frågor om kärnverksamheten? Är utvärderingarna pålitliga och objektiva? Är kvaliteten i metodval och framställning godtagbar?

I denna rapport redovisas en kartläggning av 22 myndigheters självutvärderingar, dvs. utvärderingar av den egna verksamheten. Omfattningen och i viss mån karaktären av myndigheters självutvärderingsarbete under åren 1990–1995 bygger på RRV:s definition av begreppet utvärdering. Denna definition utgår från att utvärderingar bl.a. "klargör orsaksförhållanden när det gäller myndigheternas prestationer", vilket exkluderar en del enklare beskrivningar av verksamhetens utveckling som av myndigheterna själva ibland betraktas som självutvärderingar. Hälften av de utvalda myndigheterna tillhör de största, hälften är ett slumpmässigt urval av mindre och medelstora myndigheter.

Som ett led i arbetet har ett antal fristående sakkunniga personer anlitats för att bedöma kvaliteten i ett tvärsnitt av självvärderingarna. Totalt har 23 utvärderingar granskats. De utvalda utvärderingarna är i huvudsak sådana som myndigheterna själva framhåller bland sina bästa eller metodmässigt mest sofistikerade. För varje utvärdering har vi låtit två fristående och mot varandra anonyma sakkunniga göra en bedömning. Sammanlagt har 40 sakkunniga anlitats. De sakkunniga är forskare från olika discipliner och utredare med mångårig erfarenhet från utvärderingar. Myndigheterna har själva deltagit i urvalet av sakkunniga, och har kunnat välja bort sakkunniga.

Kartläggningen visar att av 22 myndigheter har 10 över en femårsperiod gjort färre än 1 självvärdering per 1 000 anställda. Sex av dessa, eller en dryg fjärdedel, har inte gjort några självvärderingar alls. Bland de övriga 14 dominerar attitydundersökningar och enkla bedömningar av olika delverksamheter eller organisatoriska förändringar. Relativt sällan har myndigheterna tagit sig an uppgiften att mäta effekter eller effektiviteten som uppnåts. Det bör här särskilt understrykas att analysen i rapporten vare sig handlar om hur väl en myndighet fungerar eller om att ge en bild av myndigheternas totala utvärderingsverksamhet. Rapporten behandlar enbart självvärderingsarbetet.

När det gäller kvaliteten på utvärderingarna har de sakkunniga i allmänhet lämnat mycket detaljerade motiveringar till sina bedömningar. En majoritet av de sakkunniga är mycket kritiska till kvaliteten i fråga om dokumentation, metodval, frågeställning m.m. I flertalet fall anser de att slutsatser dras som det inte finns stöd för i det empiriska materialet och att det hade gått att göra en väsentligt bättre utvärdering med samma arbetsinsats. Ofta har de sakkunniga anmärkt på att utvärderingar med tystnad förbigått de viktigaste frågorna angående ändamålsenlighet och kostnadseffektivitet i den utvärderade verksamheten.

Några fåtal goda exempel eller pärlor finns som har fått höga betyg. Dessa utvärderingar kännetecknas av att fokus ligger på en central frågeställning, att vetenskapligt acceptabla metoder kommit till användning och att man endast dragit sådana slutsatser som materialet tillåter.

Slutsatsen är att de flesta självvärderingarna är av så dålig kvalitet att de faktiskt inte bör användas som beslutsunderlag. I den meningen kan sägas att budgetprocessens centrala ambition att möjliggöra resultatstyrning ännu inte kommit att uppnås, även om den i flera avseen-

den varit framgångsrik. Utan en fungerande resultatmätning av godtagbar kvalitet saknas den motor i den nya budgetprocessen som krävs för att resultatstyrning skall kunna bli en realitet.

För att förverkliga budgetprocessens ambitioner krävs sannolikt en betydande uppstramning av självvärderingsprocessen. Ett avgörande inslag i detta bör vara en kvalitetssäkring. Till en del måste ansvaret för kvalitetssäkringen självfallet ligga hos de berörda myndigheterna. Men till en del bör även en oberoende instans som ligger utanför myndigheterna utöva en kvalitetssäkrande roll. Det förefaller t.ex. rimligt att RRV i högre grad än tidigare kommenterar och granskar kvaliteten i myndigheters självvärderingar.

Som metod för kvalitetssäkring av utvärderingar förefaller en väg vida överlägsen och även mest använd i olika utvärderingssammanhang. Det är samma metod som vi också har använt för denna rapport. Oberoende och anonyma (om de så önskar) sakkunniga anlitas för att bedöma utvärderingarna. Av särskild vikt är att dessa sakkunniga väljs ut av andra än av den som genomfört utvärderingen, den som utvärderas, eller närstående till dessa. Vi föreslår därför att någon instans ges uppgiften att agera som "Kvalitetssäkringsgrupp" med uppgift främst att förmedla sakkunniga. På begäran skulle Kvalitetssäkringsgruppen välja sakkunniga som passar till en viss självvärdering, och återföra de sakkunnigas bedömning till de berörda, under bevarandet av de sakkunnigas anonymitet om de så önskar. Kvalitetssäkringsgruppen skulle kunna anlitas av myndigheter som utför självvärderingar. Detta kan samtidigt fungera som en kvalitetsstämpel som ger läsaren en signal om pålitligheten. Men Kvalitetssäkringsgruppen skulle också kunna anlitas av t.ex. regeringskansliet och RRV i granskningen av självvärderingars kvalitet. Det är tänkbart, och även önskvärt, att myndigheter skaffar sig en kvalitetsstämpel på sina självvärderingar som ett sätt att förebygga den kritik som kan komma om någon annan i stället anlitar Kvalitetssäkringsgruppen. Självfallet är det också bra om myndigheter använder sig av kvalitetssäkringen i ett tidigt skede under en självvärderings framställning och lyckas arbeta in de sakkunnigas synpunkter.

En sådan kvalitetssäkringsprocess fungerar sannolikt också som katalysator som ökar utvärderingskunnandet inom myndigheterna. Våra resultat indikerar att en förbättring av självvärderingar handlar om kunnande, inte nödvändigtvis om högre utvärderingskostnader. En så-

dan kompetens är en förutsättning, inte bara för att utföra egna utvärderingar, utan även för att beställa och kontrollera utvärderingar gjorda utanför myndigheten av t.ex. utomstående konsulter.

Eftersom ökade kvalitetssäkringskrav kan medföra att myndigheterna gör färre självutvärderingar är det viktigt att kompletterande utvärderingar kan beställas av t.ex. revisionsmyndigheterna.

Först när kvaliteten i myndigheters självutvärderingar har förbättrats är det meningsfullt att återigen se på frågan om användningen av utvärderingarna. Det är mycket möjligt att en förbättrad kvalitet och tydligare information om enskilda utvärderingars kvalitet ökar användbarheten och därmed även användningen.

2 Inledning

Bakgrund

År 1988 fattade regering och riksdag beslut om en ny s.k. budgetprocess. Denna syftade till att låta regering och riksdag styra myndigheter utifrån kunskap om myndigheternas prestationer och effektivitet.

En viktig hörnsten i den nya budgetprocessen är att de statliga myndigheternas resultat skall mätas. I förordningen (1993:134) om myndigheters årsredovisning och anslagsframställan ställs t.ex. krav på att myndigheters resultatredovisning skall innehålla vad myndigheten "har presterat och mätbara effekter av prestationerna". Ansvar för resultatmätningen lades på de berörda myndigheterna själva.

Många myndigheter har som ett verksamhetsområde, ansvar för utvärderingar inom olika områden, dvs. utvärderingar som inte bedömer just den egna verksamheten. Utvärdering av den egna verksamheten, eller – som vi valt att kalla det i denna rapport – självutvärdering, utgör i normalfallet endast en bråkdel av myndigheters utvärderingsverksamhet. En bråkdel som dock är betydelsefull eftersom den utgör det fundament som regeringens styrning av de statliga myndigheterna vilar på.

Att myndigheter skall utvärdera sig själva har aldrig varit helt okontroversiellt. Trots detta har diskussionen kring hur självutvärderingar fungerar upptagit ett mycket begränsat utrymme, både i de utredningar som ledde fram till den nya budgetprocessen, och i de studier av budgetprocessen som har gjorts därefter. Mycket av arbetet har i stället fokuserat på omläggningen av de formella rutinerna. Uppföljningar som gjorts i efterhand har i korthet konstaterat att resultattänkandet hos myndigheter visserligen har ökat, men att regering och riksdag in-

te utnyttjar självutvärderingarna. Även i efterhand har diskussionen till stor del gällt de formella rutinerna samt regeringskansliets förmåga.

Syfte

Denna studie utgår från en mer grundläggande fråga. Kan myndigheterna utvärdera sin egen verksamhet? Hur många självutvärderingar görs egentligen? Hur mycket kostar utvärderingsarbetet? Undersöker de rätt saker? Är självutvärderingarna pålitliga? Kan potentiella användare bedöma tillförlitligheten?

Det har inte funnits någon samlad bild över i vilken utsträckning myndigheter utvärderar sin verksamhet eller om utvärderingarna håller tillräckligt hög kvalitet för att resultaten skall vara användbara. En del indikationer har dock funnits på att det kan finnas problem med myndigheters självutvärderingar. RRV:s granskning av NUTEK:s stöd till företag och forskning (RRV 1995:39) innehöll t.ex. kritik mot självutvärderingar som verket gjort.¹

En mer impressionistisk observation är att forskarvärlden till stor del förbigår myndigheters självutvärderingar med tystnad. Man bryr sig inte om att kritisera dem, men man refererar inte heller till dem.

Om det är så att myndigheters självutvärderingar ofta är behäftade med betydande kvalitetsbrister, kan det innebära att icke-användningen av självutvärderingar är en helt riktig reaktion. Att i ett sådant läge försöka tvinga fram en ökad användning av självutvärderingarna skulle enbart öka risken att beslut fattas på felaktiga grunder. Därmed skulle också den nya budgetprocessens legitimitet undergrävas. I den meningen kan sägas att en granskning av självutvärderingarnas kvalitet är en förutsättning för att framgångsrikt kunna ändra på icke-användningen av myndigheters resultatmätningar, och därmed även för att den nya budgetprocessen skall kunna fungera som tänkt.

¹ En av slutsatserna är att effektivitetsfrågorna är försummade och att bristfällig utvärderingsmetodik ibland använts.

Några definitioner

I denna rapport har vi valt att utgå från den definition av begreppet utvärdering som används av t.ex. RRV och Statskontoret. Enligt den definitionen skall en utvärdering bidra till kunskap om en åtgärds ändamålsenlighet och kostnadseffektivitet. Flera myndigheter har som en verksamhetsuppgift att göra utvärderingar inom olika områden, dvs. utvärderingar som inte bedömer just den egna verksamheten. För att påminna om att det handlar om myndigheternas utvärderingar av den *egna* verksamheten använder vi benämningen "*självutvärderingar*" (se även bilaga 1). Det spelar emellertid ingen roll om myndigheterna själva har gjort utvärderingarna eller om konsulter, anlitade av myndigheterna, har utfört utvärderingarna.

Inriktning och avgränsning

Rapporten bygger på en kartläggning av 22 myndigheters självutvärderingar under perioden 1990–1995. I detta urval ingår många av de största myndigheterna, mätt i antalet anställda. Dessutom ingår ett slumpmässigt urval av små och medelstora myndigheter. För varje myndighet har vi samlat in självutvärderingar eller listor på självutvärderingar. De flesta myndigheterna har hjälpt till med att ta fram detta underlag. Försvaret, universitet och högskolor – de s.k. stabsmyndigheterna samt statliga bolag och stiftelser ingår däremot inte.

Att granska kvaliteten i utvärderingar är inte lätt. En del skulle kanske hävda att utvärderingar har likheter med konstverk, dvs. det kan finnas många legitima uppfattningar om kvaliteten. Därtill kommer frågan om bedömarens objektivitet. Redan av den anledningen skulle inte vi, med någon trovärdighet, kunna utgå bara från våra egna bedömningar.

En grundsyn som de flesta utvärderare, liksom forskare, delar är dock att det finns vissa grundläggande kriterier som en utvärdering bör uppfylla för att räknas som trovärdig. På flera områden har dessa kriterier formulerats explicit (se t.ex. Sanders 1994), mer eller mindre omfattande. Kriterierna gäller bl.a. att dokumentation, metodval, data och slutsatser uppfyller vissa minimikrav. Kriterierna blir meningsfulla därför att de flesta initierade personer i många fall gör en likartad

bedömning av huruvida kriterierna är uppfyllda. Däremot kan vissa personer i enskilda fall göra avvikande bedömningar, och för vissa utvärderingar kan det uppstå en större spridning av bedömningar.

Därför har vi för vår kvalitetsgranskning valt ut ett stort antal självutvärderingar och flera sakkunniga som bedömt självutvärderingarna. Urvalet av självutvärderingar bygger i huvudsak på myndigheters förslag, och myndigheterna har även haft visst inflytande över valet av sakkunniga. En närmare diskussion av kvalitetsfrågan och vårt tillvägagångssätt förs i kapitel 5 (se även bilaga 1).

Vi för ingen diskussion om metodval i de olika utvärderingarna. Men de sakkunniga diskuterar i allmänhet metodfrågorna mycket grundligt och förtjänstfullt. Samtliga bedömningar finns i bilaga 5. Vidare finns i bilaga 4 en förteckning över de sakkunniga som granskat rapporterna och som inte har begärt att få vara anonyma.

Vi har inte gjort någon undersökning av i vilken mån självutvärderingar används internt eller av regeringskansli och riksdag.² Det är givetvis en intressant fråga i sig. Användningen av självutvärderingar är dock avhängig kvaliteten och relevansen vilket vi studerar i denna rapport.

² Exempel på studier om detta är Statskontoret, 1992:32, Att använda resultatmätt med gott resultat.

3 Självutvärderingarna och den nya budgetprocessen

Under det senaste decenniet har några statliga verk och myndigheter bantats, lagts ner, bolagiserats eller privatiserats. Andra myndigheter har varit föremål för omstruktureringar. De flesta arbetar emellertid vidare i en oförändrad myndighetsform medan deras styrning har ändrats framför allt genom den nya statliga budgetprocessen.

Den nya budgetprocessen infördes successivt från 1988.³ Den innebär i korthet att riksdag och regering skall besluta om inriktning av verksamheten genom att formulera mål. Samtidigt får myndigheterna friare händer att välja medel för att uppnå dessa mål. Resultaten skall följas upp och utvärderas. Tanken är att bristande resultatuppfyllelse skall leda till någon form av reaktion och utkrävande av ansvar.

Genom den förändrade budgetprocessen har myndigheterna fått ett betydligt större ansvar än tidigare för uppföljning och utvärdering av den egna verksamheten och de regelsystem man administrerar. I förordningen (1993:134) om myndigheters årsredovisning och anslagsframställan formaliseras kraven på att myndigheterna skall redovisa bland annat "prestationer och mätbara effekter av prestationerna".

Betydelsen av resultatmätning har också lyfts fram i senare uppföljningar av den nya budgetprocessen. I en rapport om åtgärder för att stärka resultatstyrningen (ELMA-gruppen Ds 1995:6) fastslås t.ex. att myndigheternas information bör vara den viktigaste basen för rege-

på olika nivåer i beslutskedjan... Utan sådana utvärderingar går det inte att göra medvetna prioriteringar."

Principen att myndigheterna skall utvärdera sig själva har varit och är kontroversiell. De argument som talar för självutvärderingar bygger på tanken att dessa inte bara fyller en *utåt* legitimerande funktion, utan även en viktig intern funktion. Myndigheter måste kunna utvärdera löpande som ett rutinmässigt led i styrningen av den egna verksamheten. Det är därför närmast en självklarhet att utvärderingsfunktionen inte uteslutande kan vara en uppgift för externa organisationer.

En väl fungerande intern utvärderingsfunktion ger också rent praktiskt ett försprång t.ex. i termer av detaljkunskap vid genomförande av utvärderingar. I bästa fall kan således myndigheten själv utföra utvärderingar billigare och bättre än någon extern utvärderare.

Kritiken av självutvärdering har ifrågasatt om myndigheter själva alltid har kompetens för att producera goda utvärderingar. Även beställning av utvärderingar kräver en hel del kompetens för att kunna välja en kunnig utförare och för att kunna formulera ett genomtänkt uppdrag. En annan viktig fråga är om myndigheter alltid har den objektivitet som krävs för att göra trovärdiga självutvärderingar.

Den formella styrningsprocessen

Mycket av det förberedande arbetet inför den nya budgetprocessen och uppföljningsarbetet har handlat om de formella procedurerna i budgetprocessen.

Åtminstone vart 6:e år gör regering och riksdag en fördjupad prövning av myndigheternas verksamhet. Myndigheterna förser regeringen med en fördjupad anslagsframställning som bl.a. skall ge svar på om de mål som uppställts för myndigheternas verksamhet har uppnåtts och vilka effekter de olika verksamheterna gett.

I förordningen (1993:134) om myndigheters årsredovisning och anslagsframställan finns allmänna föreskrifter om vilken information en fördjupad anslagsframställning skall innehålla. Ofta finns det emellertid skäl för regeringen att precisera kraven på den information myndigheten skall lämna. En sådan precisering görs med hjälp av myndighetsspecifika direktiv. I dessa begär regeringen in uppgifter som är

viktiga för att kunna fatta beslut om myndighetens verksamhet under nästa treårsperiod.

I förhållande till gällande förordning och föreskrifter innehåller direktiven ofta skärpta krav på resultatanalys inom områden som är under omprövning. På andra, mindre angelägna områden kan i stället direktiven lätta på kraven på information. När myndigheterna lämnat sina fördjupade anslagsframställningar granskas dessa i regeringskansliet.

Myndigheterna skall också varje år redovisa sina resultat för regeringen. I dessa årliga resultatredovisningar ges en fortlöpande återkoppling till verksamhetens resultat, kostnader, avvikelser från planer och ändrade förutsättningar. Årsredovisningarna innehåller förutom resultatredovisning även balans- och resultaträkning, anslagsredovisning och finansieringsanalys.

Årsredovisningen har utvecklats till ett av resultatstyrningens viktigaste dokument samtidigt som anslagsframställningens betydelse minskat. Ett annat styrningsdokument i budgetprocessen som har fått ökad vikt är regleringsbrevet som sedan 1991 innehåller både en verksamhetsdel och en finansieringsdel. I verksamhetsdelen anges mål och krav på resultat i verksamheten.

ELMA-gruppen diskuterar de olika styrningsdokumenten och menar att både dokumenten och själva styrningen i högre grad än tidigare bör verksamhetsanpassas. Det framhålls också att de myndighetsspecifika direktiven och de fördjupade anslagsframställningarna inte passar in i sin nuvarande form i resultatstyrningen och de centrala kraven på dessa dokument bör därför tas bort.

Erfarenheter av den nya budgetprocessen

Flera uppföljningar av den nya budgetprocessen konstaterar ett ökat resultatänkande både hos myndigheterna och inom regeringskansliet.⁴ Inom flera myndigheter har införandet av resultatstyrning följts av in-
terna förändringar. ELMA-gruppen skriver bl.a. att myndigheterna har

⁴ För en sammanfattning av studier se Barkman (1993), kapitel 9 i "Sveriges systemskifte i fara?", IUI.

varit aktivt kunskapssökande och engagerat sig i reformprocessen. Myndigheternas anslagsframställningar har i det långa perspektivet blivit mycket bättre.

En studie där RRV har jämfört anslagsframställningar över en 25-årsperiod visar att det har skett stora förbättringar. Anslagsframställningarna har utvecklats från att vara "synnerligen detaljerade tiggarebrev" till att bli "analytiska produkter" med högt informationsvärde. Myndigheternas resultatanalyser har utvecklats.

Samtidigt har en rad studier av den nya budgetprocessen emellertid lyft fram en del problem som den nya budgetprocessen inte lyckades lösa i enlighet med förhoppningarna. Riksdagens revisorer har gjort en studie, som bl.a. visade att myndigheternas resultatredovisningar inte gav det underlag för omprövningar av verksamheten som förutsatts. Underlag hämtades i stället från fristående utredningar. Viktiga beslut om omprövning av statlig verksamhet fattades helt eller delvis vid sidan av budgetprocessen.

Liknande synpunkter finns också i en uppföljning av den s.k. ELMA-gruppen (Ds 1995:6). I ELMA-gruppens rapport konstateras bl.a. att de nya "styrformerna inte lett till ett bättre underlag för departementens omprövningar av verksamheter genom den information som lämnas av myndigheterna". Gruppen menar också att större tonvikt bör läggas på redovisning av effekter och uppnådda resultat.

Bristande användning av resultatmått

Flera tidigare undersökningar har konstaterat att regeringskansliet inte har klarat av att hantera den informationsmängd som kom från myndigheterna, och därmed inte heller att agera och reagera på basis av myndigheternas redovisningar (t.ex. Arbetstagarkonsult, 1992; RRV, 1991; Statskontoret, 1990:24).

Dessa undersökningar har några år på nacken och vissa åtgärder för att förbättra regeringskansliet med avseende på resultatstyrning har sedan dess initierats, bland annat kompetensutveckling. Även ELMA-gruppen (Ds 1995:6) betonar emellertid vikten av att reagera på missförhållanden och att utkräva ansvar. RRV har en viktig roll att granska effektivitet och resultat i statlig verksamhet. "Så måste t.ex. revisionens anmärkningar leda till konkreta åtgärder från regeringen. Så-

dana åtgärder kan vara att ompröva myndighetsledningens uppdrag eller att återta delar av den omfattande delegation som givits myndigheterna." ELMA-gruppen fastslår också att det fortfarande finns brister främst vad gäller att formulera övergripande mål och verksamhetsmål.

Inom ramen för denna rapport har samtal förts med utvärderingsansvariga eller motsvarande på ett stort antal myndigheter. Det har då framkommit att det fortfarande finns en betydande skepsis mot att departementen är effektiva i sin roll som beställare och mottagare av resultat.

Flera myndighetsrepresentanter har också framfört att de upplever att de får olika signaler från skilda instanser. Departementen ger uttryck för en viss inriktning i regleringsbrev men anger en annan muntligen, mer informellt. RRV å andra sidan måste och bör i sin granskning av årsredovisningar (där även effekter skall redovisas) ha budgetförordningen (1993:134) och regleringsbrev som utgångspunkt. Myndighetsrepresentanter menar att detta kan skapa viss förvirring.

Ett annat problem som myndigheterna gett uttryck för är begreppsapparaten. Vissa myndigheter upplever att departementen använder andra definitioner, eller för allmänt hållna definitioner, på begrepp som t.ex. "effekter" än myndigheterna själva.

Varför används inte självutvärderingarna?

Mycket av diskussionen kring den nya budgetprocessen har handlat om vad som kan ändras inom regeringskansliet för att öka användningen. Det s.k. "non-utilization" problemet har också diskuterats mycket i utvärderingslitteraturen (t.ex. Tarschys 1982). Följande orsaker till att utvärderingar inte används betonas:

1. Politiker har svårt att bedöma trovärdigheten av olika utvärderingar.
2. Politikens handlingar styrs inte enbart av effektivitetstänkande utan även av hänsyn till särintressen, den allmänna opinionen, förhandlingsspelet mellan partierna och inte minst det pressade tidsschemat.

vitet emellertid inte heller riktigt entydiga signaler som kan användas i resultatstyrningen därför att det förblir okänt hur hög kostnadseffektivitet som skulle kunna uppnås med t.ex. en bättre ledning eller annorlunda arbetssätt.

Därför är ett ännu bättre resultatmått ett som ställer prestationen i relation till vad som är möjligt att prestera, eller vad andra presterar, eller vad som presteras med olika arbetssätt inom samma organisation. Detta fångas upp av den relativa kostnadseffektiviteten. Sådana resultatmått förutsätter inte bara att kvaliteten kan mätas utan även att det finns jämförbara verksamheter. De ger också alltid ett entydiga signaler till resultatstyrningen, nämligen att flytta över verksamhet till de som har den bästa kostnadseffektiviteten. Därmed är mått på relativ kostnadseffektivitet klart mest användbara för resultatstyrning.

Denna rangordning av resultatmåttens användbarhet är inte särskilt omstridd. I en mer systematisk undersökning har t.ex. Statskontoret (1990:24) visat att offentliga verksamheter sällan använder sig av resultatmått som mäter effektiviteten. I Statskontorets rapport visas också att resultatmått i den offentliga verksamheten sällan omsätts i åtgärder. Resultatmått används främst som allmän information och i bland som underlag för allokering av medel. I de företag som undersöktes i Statskontorets studie betonas däremot att resultatmätningar i allmänhet omfattade även effektiviteten och att de ledde till åtgärder.⁶

En möjlig orsak till resultatmåttens bristande användning kan därför vara att resultatmått i offentlig verksamhet oftast tillhör kategorin verksamhetsutveckling.

Sammanfattningsvis har av olika utredningar konstaterats att myndigheters självframtagna resultatmätningar inte används mycket. Diskussionen om detta problem har fokuserat på dels frågor om hur utvärderingsresultaten presenteras, dels olika blockeringar hos mottagarna. I denna rapport reses den mer grundläggande fråga om självutvärderingarna i nuvarande form går att använda för resultatstyrning. Om så inte är fallet, är innebörden att den kanske viktigaste länken i den nya budgetprocessen ännu inte fungerar.

⁶ Detta bekräftas av undersökningen "Bäst i världen" (IVA 1990) som visar också att de mest framgångsrika företagen utvärderar mer och omsätter resultaten i åtgärder mer än konkurrenterna.

4 Kartläggning av statliga myndigheters självutvärderingar

I detta kapitel redovisas en kartläggning av självutvärderingar i ett urval av myndigheter. Hur många utvärderingar som har genomförts under perioden 1991–1995 vad har utvärderats, vilka metoder har använts osv. Vi inleder med en diskussion om vad en självutvärdering är.

4.1 Vad är en självutvärdering?

Flera myndigheter har enligt sin instruktion uppgiften att göra utvärderingar inom olika områden, dvs. utvärderingar som inte bedömer just den egna verksamheten. För att understryka att det handlar om myndigheternas utvärderingar av den *egna* verksamheten använder vi benämningen "*självutvärderingar*". Utvärderingarna kan vara utförda av myndigheterna själva, eller av konsulter anlitade av myndigheterna. Vi bortser således från utvärderingar som har tillkommit på initiativ av externa aktörer.

Inom utvärderingslitteraturen florerar en rad olika definitioner av "utvärdering". Vedung (1994) t.ex. använder en vid definition: "Utvärdering är en noggrann efterhandsbedömning av utfall, slutprestationer och genomförande i offentlig politik, vilken avses spela en roll i praktiska handlingssituationer".

RRV:s definition av utvärdering inom ramen av den statliga budgetprocessen är att "bedöma och förklara prestationer och händelseförlopp" i en myndighetsverksamhet. Utvärderingen "skärskådar och klar-

gör orsaksförhållanden när det gäller myndighetens prestationer och/eller prestationernas effekter i samhället". (RRV 1994: Resultat, verksamhet, ekonomi, en handledning för myndigheterna).

En kommitté som bildats i USA (The Joint Committee on Standards for Educational Evaluation) har tagit fram standards eller kriterier (The Program Evaluation Standards) för hur en bra utvärdering bör se ut, både vad avser metodik, användning etc. Dessa standards har blivit accepterade i USA som nödvändiga för välgjorda och relevanta utvärderingar. Inom ramen för dess arbete finns en övergripande definition på utvärdering (evaluation) som är "The systematic investigation of the worth or merit of an object".

Visst kan det skapa förvirring med olika definitioner på utvärdering och det är givetvis eftersträvsvärt med en enhetlig begreppsapparat. Några utvärderingsansvariga inom olika myndigheter som vi diskuterat med inom ramen för denna studie har också anfört att det förekommer att myndigheterna använder sig av en definition samtidigt som regeringen (fackdepartementen) avser någonting annat än vad respektive myndighets definition rymmer.

Vi har i denna rapport valt att utgå ifrån RRV:s definition. Detta innebär att vi inte tar med enkla mått på verksamhetsutveckling, t.ex. redovisning av kostnadsutveckling, antal ärenden och enkla produktivitetsmått. Ofta redovisar myndigheter sådana mått i årsredovisningen, där de också fyller en viktig funktion att beskriva hur verksamheten har utvecklats. De uppfyller dock inte RRV:s definition av en utvärdering att "bedöma och förklara...", än mindre övriga definitioner. Enligt diskussionen i det föregående är enkla mått på verksamhetsutveckling inte heller tillräckliga eller i sig särskilt användbara för resultatstyrning.

Som självutvärderingar betraktar vi således utvärderingar enligt RRV:s definition som avser att bedöma eller mäta effekter av, eller effektiviteten i, åtgärder och verksamheter som en myndighet ansvarar för. I definitionen på självutvärderingar lägger vi inga krav på metoder. Det kan avse en enkel bedömning, en attitydundersökning eller en mer sofistikerad empirisk undersökning.

4.2 Myndighetsurvalet

Urvalet omfattar totalt 22 statliga myndigheter. Hälften av myndigheterna representerar de största mått i antalet anställda. Försvaret, högskolor och universitet, stabsmyndigheterna, statliga bolag och stiftelser ingår däremot inte. Övriga 11 myndigheter bygger på ett slumpmässigt urval av små och medelstora myndigheter.

I tabellerna 4.1 och 4.2 redovisas som bakgrund antalet årsarbetskrafter och ungefärligt anslag från regeringen. Vi har dessutom (tabell 4.2) även angett driftskostnad och transfereringar.

Myndigheterna skall i de flesta fall utvärdera den egna verksamheten, de bidrag man hanterar och de regelsystem man administrerar.

Tabell 4.1 Myndighetsurvalet. Årsarbetskrafter och anslag (mkr).

	Antal anställda 1995	Anslag (mkr) 1995/96
AMS	12 000	4 285
Banverket	6 532	0
Datainspektionen	40	22
Glesbygdsverket	17	20
Invandrarverket	2 850	589
Jordbruksverket	900	165
Konsumentverket	180	70
Kriminalvårdsstyrelsen	8 000	4 000
Livsmedelsverket	500	90
Naturvårdsverket	470	0
Närings- och Teknikutvecklingsverket	360	180
Patent- och registreringsverket	760	*
Riksförsäkringsverket	640	824
Rikspolisstyrelsen	27 000	11 000
Riksskatteverket	15 000	6 066
Riksåklagaren	1 370	582
SIDA	506	315
Skogsstyrelsen	1 500	279
Skolverket	250	243
Socialstyrelsen	390	347
Tullverket	3 500	1 200
Vägverket	7 000	13 000

* = s.k. 1 000-kronorsanslag

Källa: Myndighetskalender 1995, RRV. Siffrorna är ungefärliga. Exakta siffror återfinns i myndigheternas årsredovisningar.

Tabell 4.2 Myndigheternas driftskostnader och transfereringar (mkr)

	Drifts- kostnader	Transfe- reringar
AMS	4 285	75 645
Banverket	11 500	0
Datainspektionen	23	0
Glesbygdsverket	20	0
Invandrarverket	611	8 439
Jordbruksverket	650	9 000
Konsumentverket	111	0
Kriminalvårdsstyrelsen	4 000	0
Livsmedelsverket	260	0
Naturvårdsverket	203	105
Närings- och Teknikutvecklingsverket	180	3 000
Patent- och registreringsverket	380	0
Riksförsäkringsverket	924	167 000
Riksskatteverket	6 349	-613 000
Rikspolisstyrelsen	11 020	0
Riksåklagaren	651	0
SIDA	315	16 500
Skogsstyrelsen	1 189	0
Skolverket	246	2 169
Socialstyrelsen	365	3 728
Tullverket	1 200	-90 000
Vägverket	13 600	1 560

Källa: Myndighetskalender 1995, RRV. Siffrorna är ungefärliga. Exakta siffror återfinns i myndigheternas årsredovisningar.

För kartläggningen av myndigheternas självutvärderingar kontakades i de flesta fall den utvärderingsansvarige eller motsvarande på respektive myndighet och ombads att hjälpa till med att sammanställa en lista över självutvärderingarna. Vid ungefär en tredjedel av myndigheterna fanns redan färdiga listor. Flertalet myndigheter har varit mycket hjälpsamma och gjort sammanställningar av självutvärderingarna åt oss.

Detta förfarande är av naturliga skäl beroende av myndigheternas agerande. En risk är att den person vi har haft kontakt med på respektive myndighet inte känt till alla självutvärderingar. Detta kan vara en tänkbar felkälla. Dock kan man ifrågasätta i vilken mån självutvärderingar som inte gjorts kända inom myndigheten, för utvärderingsansvariga eller motsvarande, uppnår syftet med självutvärderingar. För att gardera mot sådana felkällor har vi dessutom brevlades frågat generaldirektören (motsv.) om våra uppgifter uppfattas som

korrekta. Några myndigheter har då kommit med kompletterande material som vi beaktat.

En annan risk kan vara att en myndighet valt att inte redovisa vissa självutvärderingar. I den mån sådant skulle ha förekommit, innebär det att vår sammanställning har med något färre, men förmodligen kvalitativt bättre, självutvärderingar än vad som faktiskt genomförts.

Med dessa reservationer redovisas i det följande *hur många* självutvärderingar respektive myndighet har genomfört, *vad dessa mäter* (produktivitet, effekter, kostnadseffektivitet, eller annat) samt vilken typ av *metod* som använts.

Tabell 4.3 Antal självutvärderingar per myndighet/1 000 anställd, 1990-1995.

	Antal själv- utvärderingar	Antal själv- utvärderingar per 1 000 anställd
AMS	19	1.58
Banverket	0	0
Datainspektionen	0	0
Glesbygdsverket	0	0
Invandrarverket	17	6
Jordbruksverket	7	8
Konsumentverket	23	127
Kriminalvårdsstyrelsen	3	0.4
Livsmedelsverket	16	32
Naturvårdsverket	2	4.2
Närings- och Teknikutvecklingsverket	25-49 ¹	69-136
Patent- och registreringsverket	0	0
Riksförsäkringsverket	6 ²	9
Rikspolisstyrelsen	5	0.19
Riksskatteverket	13	0.86
Riksåklagaren	0	0
SIDA	130 ³	257
Skogsstyrelsen	2	1.33
Skolverket	10	40
Socialstyrelsen	12	30
Tullverket	0	0
Vägverket	3	0.42

¹ För NUTEK faller 24 rapporter i en gråzon för vad som definieras som självutvärdering. Detta förklaras närmare längre fram.

² Alla sex avser dock Försäkringskassan snarare än RFV själv.

³ Uppskattat antal.

Beräkningen av antal utförda självutvärderingar måste tas med försiktighet. Vissa självutvärderingar är väldigt smala medan andra täcker en stor del av verksamheten. Att enbart se på antalet självutvärde-

ringar ger således ingen bra bild över i vilken utsträckning en myndighet verkligen utvärderar sig själv.

En viktig slutsats är emellertid att sex myndigheter överhuvudtaget inte redovisar några självutvärderingar. Tre av dessa tillhör dessutom kategorien stora myndigheter med mer än 1 000 anställda. Om vårt urval är representativt skulle det innebära att cirka en dryg fjärdedel av myndigheterna inte har gjort någon självutvärdering alls. Tio myndigheter, eller närmare hälften, redovisar under femårsperioden färre än en självutvärdering per 1 000 anställda.

På ett grovt sätt kan tabell 4.3 också användas för att få en bild av kostnaden för självutvärderingar. Vi har uppgifter om kostnaden för några självutvärderingar. Utifrån det kan man utgå ifrån att kostnaden för de flesta självutvärderingarna ligger mellan 30 000–150 000 kronor. Om man utgår från ett medelvärde på 70 000 kronor per självutvärdering, så utgör de totala kostnaderna för självutvärderingar 0,02 procent av myndigheters driftskostnader, eller en tusendels procent av myndigheters totala driftskostnad och transfereringar.

Det måste betonas att detta enbart avser kostnaden för självutvärdering. Det är sannolikt att kostnaden för relaterade uppgifter som verksamhetsbeskrivning, uppföljning och myndighetens spridning av information om sig själv vida överstiger kostnaden för självutvärdering.

4.3 Självutvärderingarna – några generella drag

I ett nästa steg delas självutvärderingarna in enligt vad de syftar till att mäta. Detta framgår av tabell 4.4. En svårighet här är att en grupp av självutvärderingar som i huvudsak består av en utvärderares bedömning inte tydligt mäter någonting specifikt. I stället gör de i allmänhet explicita eller implicita bedömningar av såväl produktivitet som effekter och effektivitet. Som framgår av tabellen är merparten av självutvärderingarna av denna karaktär.

Endast några få självutvärderingar mäter produktivitet. Vi påminner återigen om att vi inte räknar in enkla produktivetsmått, som ofta återfinns i årsredovisningarna, om det inte görs ett försök att "bedöma och förklara" varför produktiviteten har utvecklats på ett visst sätt.

Något vanligare är mätningar av effekter av åtgärder. Däremot förekommer mätningar av den relativa effektiviteten inte alls.

Tabell 4.4 Vad mäter självutvärderingarna?

	Bedömn/ uppfölj- ning	Produk- tivitet	Effektr/ effekti- vit	Relativ effekti- vit	Annat
AMS	15		3		1
Banverket					
Datainspektionen					
Glesbygdsverket			3		
Invandrarverket		14	2		
Jordbruksverket	5		5		
Konsumentverket	18				
Kriminalvårdsstyrelsen		6		1	
Livsmedelsverket	3		13		
Naturvårdsverket	2				4
NUTEK	13-33	2	11		
Patent- och registreringsverket					
Riksförsäkringsverket	2	1	3		
Rikspolisstyrelsen	4		1		
Riksskatteverket	5	5	3		
Riksåklagaren					
SIDA	121	3	6		
Skogsstyrelsen		1		1	
Skolverket	5		5		2
Socialstyrelsen		3		7	
Tullverket					
Vägverket	3				

Av tabell 4.5 framgår vilken metod som använts i självutvärderingarna. En metodkategori, professionella bedömningar, överensstämmer av naturliga skäl nästan helt med motsvarande kolumn i tabell 4.4.

I övrigt tycks intervjuundersökningar vara förhållandevis vanliga bland självutvärderingar, medan empiriska bearbetningar grundade på andra datakällor är mindre vanliga.

Tabell 4.5 Vilken metod används i självutvärderingarna?

	Profession- nella be- dömare	Intervjuer	Statistisk/ empiriskt	Annat
AMS	12	3	3	
Banverket				
Datainspektionen				
Glesbygdsverket				
Invandrarverket	8	6	8	
Jordbruksverket	5	1	1	
Konsumentverket	6	13	4	
Kriminalvårdsstyrelsen			1	6
Livsmedelsverket	4	7	5	
Naturvårdsverket		2		
NUTEK	13-33	13	5	3
Patent- och registreringsverket				
Riksförsäkringsverket	2	1	3	
Rikspolisstyrelsen	4		1	
Riksskatteverket	1	4	8	
Riksåklagaren				
SIDA	121	3	6	
Skogsstyrelsen		1	1	
Skolverket	3	7		
Socialstyrelsen		3	9	
Tullverket				
Vägverket	1	2		

Tabellerna 4.4 och 4.5 beskriver ganska schablonmässigt självutvärderingarnas omfattning. Därför följer en kort beskrivning av särskilt utmärkande drag i varje myndighets självutvärdering samt vår något subjektiva bedömning av hur väl självutvärderingarna täcker respektive myndighets verksamhet. Beskrivningarna gör inget anspråk på fullständighet.

4.3.1 AMS

Totalt har AMS redovisat ett 50-tal utvärderingar under tidsperioden. Av dessa kan 19 betecknas som självutvärderingar i vid bemärkelse. Tolv av dessa är beskrivningar/bedömningar av olika försöksverksamheter, samt uppföljningar där t.ex. redovisas hur stor andel av personer som genomgått en åtgärd är fortsatt arbetslösa efter en tidsperiod

utan att jämförelse görs med någon kontrollgrupp. Tre av självvärderingarna är s.k. servicemätningar där kunder på arbetsförmedlingarna ombeds att bedöma servicekvaliteten. En är en genomgång av empirisk litteratur på arbetsmarknadsområdet. Resterande tre är självvärderingar med ambitionen att mäta effekter av arbetsmarknadspolitiska åtgärder. Två av dessa har valts (se kapitel 6).

Ganska få självvärderingar belyser verksamhetens effekter. Till bilden hör dock att en hel del utvärderingar av arbetsmarknadspolitiska åtgärder har genomförts utanför AMS av fristående forskare och av EFA (Expertgruppen för arbetsmarknadspolitiska utvärderingsstudier). AMS har också byggt upp databaser som görs tillgängliga för utomstående forskare. Under undersökningsperioden syns en ökning av ambitionsnivån att göra självvärderingar. Detta reflekteras bl.a. i en förteckning över planerade och pågående utvärderingar som AMS sammanställt.

4.3.2 Banverket

Banverket har inte redovisat några självvärderingar alls. Några samhällsekonomiska kostnads-intäkts kalkyler över olika projekt har dock gjorts. Dessa kalkyler bygger emellertid i huvudsak på antagna eller faktiska förändringar i t.ex. antalet resande utan att utvärdera i vilken mån förändringen beror på en genomförd investering eller på externa faktorer. I vår (och två oberoende experters) bedömning bör kalkylerna därför inte räknas som fullgoda självvärderingar enligt RRV:s definition.

4.3.3 Datainspektionen

Datainspektionen har inte redovisat några självvärderingar.

4.3.4 Glesbygdsverket

Glesbygdsverket har inte redovisat några självvärderingar. Däremot har en omvärldsanalys enligt Delfitekniken gjorts.

4.3.5 Invandrarverket

Invandrarverket har redovisat 17 självvärderingar. Av dessa är tre interna granskningar av asylärenden. Tre är beskrivningar av produktivitet utveckling och handläggningstider. De övriga är bedömningar – vanligtvis grundade på intervjuer – av olika delar av verksamheten; ofta försöksverksamheter. Av dessa ingår två i denna studie. Täckningen av verksamheten genom utvärderingarna förefaller god.

4.3.6 Jordbruksverket

Statens jordbruksverk har redovisat sju självvärderingar av enskilda typer av åtgärder som verket administrerar. De flesta utgörs enbart av beskrivningar och bedömningar, i synnerhet de som avser mer centrala delar av livsmedels- och jordbrukspolitiken. Några använder sig dock av empiriska metoder, och mäter effekten av olika kampanjer, angående bekämpningsmedel. En av dessa ingår i denna studie. Täckningen av verksamhetsområdet genom självvärderingar förefaller något begränsad.

4.3.7 Konsumentverket

Konsumentverket har redovisat 23 självvärderingar eller uppföljningar av olika satsningar, informationskampanjer, Råd och Rön m.fl. Täckningen av verksamheten förefaller god. Många av utvärderingarna är förhållandevis enkla attitydundersökningar, men Konsumentverkets mål kan också sägas vara att påverka attityder. Därutöver finns också flera entydiga effektstudier där effekter på konsumentbeteende eller annat mäts. Två av dessa ingår i vår studie.

4.3.8 Kriminalvårdsstyrelsen

Kriminalvårdsstyrelsen har redovisat sju uppföljningar. En av dessa är en attitydundersökning. De resterande är oftast beskrivningar av antal återfall efter behandling. I vissa fall redovisas också antal återfall från olika behandlingar. I allmänhet görs dock inte några försök att förklara varför återfallsfrekvensen varierar, t.ex. genom att jämföra antal återfall med likartade grupper, eller att på annat sätt behandla selektionseffekter. Enbart tre av studierna har därför kunnat klassas som självutvärderingar. Mest ambitiös verkar en studie som Kriminalvårdsstyrelsen beställt av Väg- och Trafikinstitutet där ett mer systematiskt försök görs att undersöka effekten av en behandling. Denna studie har valts till sakkunnigbehandling.

4.3.9 Livsmedelsverket

Få självutvärderingar har redovisats för åren 1990–1993. Därefter sker en markant ökning, ungefär åtta stycken per år. Efter 1994 är täckningen av verkets verksamhet god. De flesta utvärderingarna kan betecknas som effektstudier. Självutvärderingarna utgör oftast intervjustudier med frågor om attityder, information och beteende. Två av dessa ingår i denna studie.

4.3.10 Naturvårdsverket

Naturvårdsverket har redovisat två självutvärderingar, båda intervjustudier bland användare an Naturvårdsverkets information. En genomgång av 9 miljömål har emellertid gjorts samt en beskrivning av vad olika aktörer gör på området. Verket redovisar också en sammanställning av regeringsuppdrag som verket fått under åren samt ett antal andra uppföljningar.

4.3.11 Närings- och teknikutvecklingsverket

En del av NUTEK:s uppföljningar och utvärderingar har tidigare granskats i en revision av RRV (RRV 1995:39). Där behandlades två verksamhetsområden, NUTEK Teknik och NUTEK företag. Därtill har vi här kartlagt NUTEK Regional och NUTEK Energi. Sammanlagt redovisar NUTEK ett 50-tal självutvärderingar för perioden. Drygt hälften av dessa är bedömningar gjorda av experter av projekt som får en del av sin finansiering av NUTEK. I vår mening är dessa endast i begränsad omfattning självutvärderingar. Experterna som gör bedömningen har i allmänhet inte information om hur stor andel av projektet som finansieras av NUTEK, och de uttalar sig endast ibland och då perifert om hur de bedömer NUTEK:s roll och betydelse för projektet.⁷

Bland de övriga förekommer en hel del uppföljningar som oftast med hjälp av intervjuer bland företag och forskare som fått stöd försöker skapa en bild av stödets effekter. De räknas således som effektstudier. I några studier ställs också effekterna i relation till kostnaderna och utgör således effektivitetsstudier. Täckningen av verksamheten med självutvärderingar är förhållandevis god.

I RRV:s revision påpekades metodproblem i en del utvärderingar. För att inte överlappa arbete som redan gjorts, har här valts en utvärdering av det regionalpolitiska stödet som NUTEK genomförde 1993 för sakkunnigbedömning. Utvärderingen består av en huvudrapport och ett antal delrapporter.

4.3.12 Patent- och registreringsverket

Patent- och registreringsverket har inte redovisat några självutvärderingar. Kvartalsvis görs dock en resultatuppföljning som också ger underlag till årsredovisningen. Där beräknas produktivitet och tidsåtgång för behandling av ärenden.

⁷ Jfr SIDA.

4.3.13 Riksförsäkringsverket

Riksförsäkringsverket redovisar 25 uppföljningar och 12 utvärderingar av socialförsäkringssystemet. Inga av dessa är emellertid entydiga självutvärderingar av verkets centrala aktiviteter. En typ av utvärdering som verket gör är att beräkna det s.k. ohälsotalet och att statistiskt försöka bestämma orsaksfaktorer bakom förändringar i detta mått. Antydning görs att förändringar som inte kan hänföras till de undersökta exogena faktorer (som t.ex. demografiska förändringar) kan tolkas som effekter av ändringar i socialförsäkringssystemet. I vår bedömning är dock detta en allt för indirekt härledning för att kunna räknas som en självutvärdering av verkets centrala aktiviteter.

RFV är tillsynsmyndighet över försäkringskassorna som dock har egen juridisk status. Jämfört med andra myndigheter med regional organisation har också kassorna en större självständighet. Försäkringskassorna tillhör inte myndigheten RFV. Med denna utgångspunkt är täckningen av Riksförsäkringsverkets aktiviteter begränsad. Även om kassorna räknas in i RFV:s aktiviteter är täckningen ändå inte överväldigande. Man kan däremot beskriva det som att RFV och kassorna tillhör socialförsäkringsadministrationen. Kassornas arbete är i nära anknytning till RFV:s aktiviteter. Med tanke på detta kan sex av ovan nämnda uppföljningar och utvärderingar räknas som självutvärderingar, varav flertalet berör olika försöksverksamheter. Två av dessa ingår i vår studie.

4.3.14 Rikspolisstyrelsen

Rikspolisstyrelsen har redovisat fem självutvärderingar, varav flertalet är bedömningar och intervjuer som rör försöksverksamheter. En av dessa har valts ut till sakkunniggranskning. Täckningen av den centrala verksamheten förefaller vara dålig.

4.3.15 Riksskatteverket

Riksskatteverket har redovisat 13 självutvärderingar under perioden. Av dessa är fyra attitydundersökningar där människor ombeds bedöma sina kontakter med Riksskatteverket. En självutvärdering är en expertbedömning av en del av verksamheten. Fem självutvärderingar är produktivetsstudier där behandling av olika typer av skatteärenden eller upptaxeringar till följd av granskning ställs i relation till kostnader. De tre övriga självutvärderingarna behandlar med olika metoder frågor som förekomsten av skattefusk och effekter av skattereformen. En av självutvärderingarna, som i vår bedömning tillhör de mest sofistikerade har valts till sakkunniggranskning.

Det är inte helt entydigt vad som skall betraktas som effekter av skattemyndigheters verksamhet. Upptäckt av fler fall av skatteundandragande behandlas i viss mån som en effekt, och i den mening är flera av självutvärderingarna också effektstudier, med en ganska god täckning av verksamheten. Men den vidare effekten på omfattning av försök till skatteundandragande behandlas i ganska liten omfattning. Detsamma gäller andra frågor som rör hur skattestrukturen och administrationen påverkar människors beteende. Sådana frågor undersöks dock även av utomstående forskare, bl.a. på uppdrag av Finansdepartementet.

4.3.16 Riksåklagaren

Riksåklagaren har inte genomfört några självutvärderingar.

4.3.17 SIDA

SIDA har genomfört ett 30-tal självvärderingar per år. Dessa utvärderingar avser i allmänhet ett biståndsprojekt, som oftast bedöms av någon eller flera experter.⁸

Av 20 SIDA-utvärderingar som vi har gått igenom fanns en som innehöll vissa mått på kostnadseffektivitet. I enskilda fall förekommer också produktivetsmått. Vanligen utgörs de dock huvudsakligen av verbala bedömningar. Sedan 1994 görs emellertid försök att i ökad utsträckning använda sig av kostnads-intäkts analyser.

I utvärderingarna beaktas oftast inte (och kan kanske inte beaktas) indirekta och makroekonomiska aspekter. En del studier tyder på att dessa effekter förtar positiva effekter av bistånd. En fråga som måste resas är emellertid om SIDA:s projektutvärderingarna skulle kunna ge sådan information om den är sann.

4.3.18 Skogsstyrelsen

Skogsstyrelsen har redovisat två självvärderingar. Den ena är i huvudsak en attitydundersökning bland skogsägare. Den andra är en effektstudie av särskilda skogsvårdsinsatser. Den senare ingår i vår studie.

4.3.19 Skolverket

Skolverket har redovisat tio självvärderingar. Tre av dessa är bedömningar av hur organisationsförändringar har påverkat den interna effektiviteten. Två är attitydundersökningar angående hur skolhuvudmän och andra såg på eller reagerat på insatser från Skolverket. Slut-

⁸ En jämförelse kan här göras med de utvärderingar som NUTEK låtit göra av projekt som finansieras. Även här är frågan om utvärderingarna kan ses som självvärderingar mer än i högst begränsad omfattning. I vår bedömning är emellertid SIDA:s operativa ansvar för utformning och genomförande av de flesta projekten större än t.ex. NUTEK Tekniks roll i projekt som får stöd av NUTEK. I den mån den är också utvärderingar av SIDA:s projekt mer att betrakta som självvärderingar än motsvarande hos NUTEK Teknik.

ligen har ett antal läsvärdesundersökningar av olika av Skolverkets skrifter och informationssatsningar gjorts.

Den mest avancerade undersökningen förefaller vara en angående genomslaget av den nationella utvärderingen av grundskolan. Den, samt en utvärdering av den nya fältorganisationen, har valts för sakkunnigbedömning.

Skolverket redovisar också en planerad fördjupning av utvärderingsverksamheten. Detta gäller bland annat effekter av tillsynsverksamhet.

4.3.20 Socialstyrelsen

Socialstyrelsen har redovisat 12 självvärderingar. Av dessa är två bedömningar gjorda av egen personal, och en är en bedömning gjord av en extern expert. Resterande nio är intervjuundersökningar. Av dessa behandlar sju i vilken omfattning Socialstyrelsens skrifter har lästs och har haft genomslag. En avser allmänhetens attityder gentemot Socialstyrelsen, och en är en pilotstudie om hur Socialstyrelsen uppfattas vid Lex Maria-anmälningar. En utvärdering som tar reda på om människor har läst en skrift kan i vår bedömning inte betecknas som effektutvärdering. Tre av utvärderingarna undersöker dock även effekter i andra termer. Två av dessa ingår i vår studie och en betecknas som "pilotstudie". Vi har valt den ändå, eftersom den metodmässigt förefaller mer gedigen än de andra utvärderingarna. Det faktum att den beskrivs som pilotstudie beaktas dock i kvalitetsbedömningen.

4.3.21 Tullverket

Tullverket har ett omfattande uppföljningssystem. Man har redovisat ett antal resultatmått som t.ex. utvecklingen av antal beslag, antal visitationer och mängden beslagtagna droger. Dessa relateras också till kostnader och utgör därför ett produktivetsmått. De relateras dock inte till arbetsmetoder och inte heller systematiskt till förändringar i omvärldsförhållanden. Det förekommer inte heller utvärderingar av hur tullens arbete påverkar omfattningen av den illegala handeln. Effekt- eller effektivitetsstudier görs således inte.

4.3.22 Vägverket

Vägverket redovisar en rad olika uppföljningar av nationella vägtrafikmål, där t.ex. utvecklingen av antal döda och skadade eller bärighet på vägarna beskrivs. I dessa uppföljningar görs emellertid i allmänhet inga försök att undersöka varför förändringar sker. Därför har endast tre kunnat klassas som självutvärderingar.⁹

4.4 Några slutsatser

Genomgången av myndigheternas självutvärdering tyder på en mycket ojämn fördelning i flera avseenden. Vissa myndigheter gör många självutvärderingar, andra inga alls. Bland självutvärderingarna använder sig de flesta av mycket enkla metoder, men några mer sofistikerade studier förekommer.

Till en del skulle denna ojämn fördelning kunna bero på olika förutsättningar på olika verksamhetsområden. Huvudintrycket är emellertid att ambitionsnivån oftast är avgörande.

⁹ Dessa tre redovisades först vid efterkontrollen. Ingen av dessa har därför valts ut för sakkunnigbedömning.

5 Metod för kvalitetsbedömning

En utvärderings kvalitet har många dimensioner. Den är också i viss mening subjektiv. Det som är viktigt för en läsare kan verka oväsentligt i en annans ögon. Olika utvärderare kan också göra olika bedömningar om lämpligheten i olika metoder eller vad rimliga tolkningar av resultat kan vara. Av dessa skäl kan det förefalla omöjligt att exakt mäta eller beskriva kvaliteten på en utvärdering.

Samtidigt finns ofta en betydande samstämmighet bland utvärderare och forskare om grundläggande principer för vad som utgör en bra utvärdering. Sådana principer har också formulerats på olika sätt. Ett är de s.k. "Program Evaluation Standards" (The Program Evaluation Standards. 2nd Edition. Sage Publications), en skrift som beskriver kriterier (standards) som man menar att en god utvärdering skall uppfylla. Ett sådant kriterium är t.ex. att dokumentationen skall vara så utförlig att läsaren i princip skulle kunna kontrollera de empiriska bearbetningarna. Därtill skall data och metod vara lämpliga och slutsatserna relevanta och användbara. Andra kriterier handlar om vikten av objektiva och trovärdiga utvärderare.

Kriterierna blir meningsfulla därför att de flesta initierade personer i de flesta fallen gör en likartad bedömning av huruvida de är uppfyllda. Däremot kan vissa personer i enskilda fall göra avvikande bedömningar, och för vissa utvärderingar kan det uppstå en större spridning av bedömningar.

Konsekvensen av denna grundsyn är att vi dels har valt att granska ett stort antal självutvärderingar, dels att två sakkunniga bedömt varje självutvärdering. Urvalet av självutvärderingar bygger i huvudsak på myndigheternas förslag om de bästa, och myndigheterna har dessutom haft visst inflytande över val av sakkunniga.

Det normala förfarandet i många utvärderings- och forskningssammanhang är att kvaliteten bedöms av oberoende experter eller utvärderare, med avseende på de uppställda kriterierna. Ett sådant förfarande har emellertid hittills aldrig tillämpats på de svenska statliga myndigheternas utvärderingar.

Inför vår tillämpning av en sådan sakkunniggranskning eller -bedömning aktualiserades flera frågor. Hur skall utvärderingar väljas ut? Hur skall sakkunniga väljas ut? Vilka kriterier skall de sakkunniga tillämpa, dvs. vad skall de jämföra med? Det är uppenbart att dessa frågor måste hanteras på ett sätt som inte snedvrider granskningen. Vi har därför lagt ner en hel del möda för att utveckla och tillämpa ett så rättvist förfarande som möjligt. Detta beskrivs i avsnitten 5.1–5.3.

5.1 Hur har självutvärderingarna valts ut?

En möjlighet hade varit att helt slumpmässigt göra ett urval av myndigheters självutvärderingar. Vår kartläggning visade emellertid uppenbart att ett sådant förfarande skulle kunna ha lett till felaktiga slutsatser. Ett problem är att fördelningen mellan myndigheter skulle ha blivit sned. De få myndigheter med många utvärderingar hade blivit överrepresenterade, medan flera med få utvärderingar inte hade kommit med alls. Ett än större problem är att ett slumpmässigt urval hade lett till att många enkla attitydundersökningar, som antalsmässigt dominerar, hade tagits ut till sakkunniggranskning, medan endast ett fåtal av de mer sofistikerade utvärderingarna hade kommit med.

Vi har här strävat efter att få med de mest kvalificerade självutvärderingarna. De mest sofistikerade utvärderingarna valdes ut. Konkret innebär detta att i första hand valdes utvärderingar som mätte effektivitet eller effekter på ett statistiskt/empiriskt sätt (enligt tabellerna 6.3 och 6.4).¹⁰ Om det fanns flera sådana för en myndighet så valdes två slumpmässigt. Om det inte fanns några sådana valdes två slumpmässigt av resterande utvärderingskategorier (enligt tabellerna

¹⁰ I de fall respektive myndighet framhöll att de själva betraktade vissa utvärderingar som bäst accepterade vid den bedömningen.

4.4 och 4.5). Ambitionen var således att välja två från varje myndighet.¹¹

Tabell 5.1 visar vilka utvärderingar som har valts ut för respektive myndighet (i bilaga 2 ges fullständiga referenser till respektive utvärdering). I två fall består utvärderingen av flera rapporter. Totalt har 23 utvärderingar valts ut för sakkunnigbedömning.

Tabell 5.1 Vad handlar de uttagna självutvärderingar om?

Självutvärderingens frågeställning	
Starta-eget bidrag	AMS
Rekryteringsstödet resultat	AMS
Bosnier hjälper Bosnier	Invandrarverket
Kontakt nät för invandrare	Invandrarverket
Minskade hälso-miljörisker av bekämpningsmedel	Jordbruksverket
Varudistribution	Konsumentverket
Värderad service	Konsumentverket
Kurs för fängelsedömda rattfyllerister	Kriminalvårdsstyrelsen
Infokampanj om fisk till gravida och ammande	Livsmedelsverket
Utvärdering av "Startpaket Nyckelhål"	Livsmedelsverket
Regionalpolitiska stödets effekter (Huvudrapport och 7 delstudier)	NUTEK
Försöksverksamhet inom rehabområden	Riksförsäkringsverket
Rehabarbetet på försäkringskassan	Riksförsäkringsverket
RAR-systemet (lokalt ADB-stöd)	Rikspolisstyrelsen
Kvalitetsutveckling inom skatteförvaltningen	Riksskatteverket
Importstöd till några u-länder:	
– Importstöd till Guinea-Bissau m.fl.	SIDA
– Importstöd till Tanzania	SIDA
– Importstöd till Cap Verde	SIDA
Särskilda skogsvårdsinsatser i norra Sverige	Skogsstyrelsen
Utvärdering av nya fältorganisationen	Skolverket
Nationell utvärdering av grundskolan (3 st. i en bok)	Skolverket
Lex Maria	Socialstyrelsen
Kvalitetssäkring inom vården	Socialstyrelsen

¹¹ I några fall valdes bara en, t.ex. Rikspolisstyrelsen.

5.2 Hur har de sakkunniga valts ut?

Bland experter inom olika områden finns i vissa fall sådana som är kända för att vara mer positiva eller mer negativa till en myndighet eller en myndighets verksamhet. Detta kan naturligtvis vara välgrundade bedömningar baserade på tidigare erfarenheter med respektive myndighet. Samtidigt kan det också finnas ett inslag av förutfattade meningar, eller personliga konflikter. Av dessa skäl har vi försökt välja sakkunniga som inte är kända för en positiv eller negativ inställning till respektive myndighet. I urval av sakkunniga har varje myndighet fått tillfälle att utöva ett inflytande.

Andra krav som ställs på de sakkunniga är att de dels har en gedigen kunskap om utvärderingsmetoder, dels har kunskap om sakområdet.

I en första omgång har vi sammanställt listor av mellan 5 och 10 tänkbara sakkunniga för varje myndighet. Listorna byggde till stor del på förslag från den referensgrupp som ESO knutit till vårt projekt.

I en andra omgång sändes listorna till respektive myndighet. Varje myndighet fick då tillfälle att stryka namn som de inte ville ha. Det fanns ingen explicit restriktion för hur många namn en myndighet fick stryka. Däremot gjordes klart att om myndigheten inte accepterade något av namnförslagen eller om den person som en myndighet helst vill ha inte ställer upp som sakkunnig så skulle nya sakkunniga väljas utan att myndigheten tillfrågades en gång till.

Tre myndigheter anförde klagomål mot samtliga sakkunniga eller mot urvalsförfarande för sakkunniga. I efterhand kan vi dock konstatera att det inte finns något samband mellan antal strykningar eller klagomål och utfallet. Självutvärderingar av myndigheter som var mer negativa till sakkunniga bedömdes inte i genomsnitt mer negativt av de sakkunniga.

Två myndigheter kommenterade också den stora andelen nationalekonomer bland sakkunniga. Dessa utgjorde 58 % av de som genomfört kvalitetsgranskning. Självutvärderingarna fokuserar emellertid ofta på begrepp som produktivitet eller effektivitet som i hög grad är ekonomiska begrepp. Det är därför försvarbart om det blir en viss övervikt av ekonomiskt orienterade personer bland sakkunniga. En

efterhandskontroll visar att nationalekonomerna bland de sakkunniga inte varit mer negativa än andra sakkunniga.

Sammanlagt valdes 40 sakkunniga. (En förteckning på sakkunniga finns i bilaga 3). I ett fall gjordes en bedömning av två forskare gemensamt. Tre av de sakkunniga har granskat två självutvärderingar var. Normalt valdes emellertid två sakkunniga per utvärdering. En sakkunnig har inte haft kännedom om vem mer som haft till uppgift att granska samma utvärdering. Detta gjorde att vi kunde göra en avstämning för att se om bedömningarna skilde sig åt i någon större utsträckning. Två av varandra oberoende utvärderare gör också bedömningarna i genomsnitt "säkrare". I efterhand har det visat sig att i de flesta fall var de båda sakkunnigas bedömningar av samma utvärdering i hög grad samstämmiga. Detta kommenteras mer längre fram.

Sakkunniga tilläts vara anonyma om de så önskade. Det diskuterades ingående i projektets referensgrupp¹² om anonymiteten skulle kunna leda sakkunniga till mindre gedigna granskningar. Mot det anfördes dock att en del sakkunniga inte skulle tordas skriva sin mening om deras namn offentliggjordes. Detta slutade i kompromissen att låta de sakkunniga avgöra frågan själv. Endast ett fåtal valde att förbli anonyma (fyra stycken). I efterhand kan sägas följande om farhågorna angående anonymiteten. Enligt vår bedömning kan endast några få bedömningar beskrivas som mindre genomarbetade. Någon enstaka av dessa är bland de anonyma. Det finns också sakkunnigbedömningar av de anonyma som tillhör de mest välskrivna och genomarbetade rapporterna. Nästan samtliga som har avgett positiva omdömen om en självutvärdering har gått med på att låta sitt namn offentliggöras. Å andra sidan finns sakkunniga som avgett kritiska bedömningar och har valt att inte vara anonyma. Med tanke på att de sakkunniga i allra flesta fall är mycket kunniga och trovärdiga är vårt subjektiva intryck att fördelarna med att tillåta anonymitet väger tyngre än risken för mindre gedigna utlåtanden som detta kan föra med sig.

De sakkunniga rekryterades från tre grupper med olika bakgrund. Den största gruppen bestod av forskare eller personer med forskaran-

¹² I referensgruppen ingick Göran Arvidsson, SNS, Peter Bohm, Stockholms universitet, Lars Hultkrantz, Högskolan Dalarna och Uppsala universitet, Marja Lemne, ESO, Ingrid Munck, SCB och Göran Schubert, ESO.

knytning inom ett område som låg nära det som utvärderingen fokuserade på. Den andra, och mycket mindre, gruppen bestod av utvärderingskunniga inom vissa utredande myndigheter t.ex. Statskontoret och SCB. Den tredje och minsta gruppen, bestod av konsulter med inriktning på utvärdering.

I vår bedömning har det generellt funnits en tydlig skillnad mellan konsultgruppen och de andra. I konsultgruppen tycktes frågor som layout och läsvänlighet ibland tillmätas större vikt än metodfrågorna. Det verkade inte heller finnas en erfarenhet av att skriva bedömningar om utvärderingar.

Vår slutsats av dessa observationer inför framtida granskningar är dels att anonymitet är viktigt, dels att urvalet av sakkunniga är så viktigt att det *inte* kan göras rutin- eller slumpmässigt.

5.3 Vad skall de sakkunniga jämföra med?

De sakkunnigas svar hade kunnat standardiserats med t.ex. en poängsättning. Vår referensgrupp gjorde emellertid den bedömningen att detta skulle styra in de sakkunniga alltför mycket. Därför har de sakkunniga haft stora möjligheter att formulera sin bedömning själv, styrd endast av några brett formulerade delfrågor.

När det gäller instruktioner till sakkunniga har vi utgått från "Program Evaluation Standards", men sedan ansett det nödvändigt att formulera om, slå ihop och lämna bort en del. Instruktionerna har gått genom två "pilotstudieomgångar" med sammanlagt sex testpersoner. Instruktionerna redovisas i *bilaga 2*.

De sakkunniga ombads att göra sin bedömning utifrån en jämförelse med "vad de bedömde som god praxis på området". Detta är ganska vagt hållet, och skulle kunna leda till en osäkerhet om vad de sakkunniga jämför med. Det har emellertid visat sig att de sakkunniga nästan alltid tydligt anger vad de jämför med. Samtidigt är de sakkunniga erfarna forskare och/eller utvärderare/utredare vilket gör att dom har en god grundkännedom om hur en godtagbar utvärdering bör se ut. Ofta anger de också hur de anser att en utvärdering förhåller sig till olika referenspunkter. En sakkunnig har t.ex. påpekat att rådande "praxis" (av utvärderingar inom området) inte är av god kvalitet och bedöm-

ningarna är därför baserade på vad den sakkunnige anser bör vara godtagbar kvalitet på en utvärdering. En annan sakkunnig skriver att utvärderingen bedöms vara bättre än de flesta av myndighetens utvärderingar, men ändå otillfredsställande jämfört med god praxis. En sakkunnig utgår från RRV:s definition av utvärdering för budgetprocessen nämligen att "bedöma och förklara prestationer och händelseförlopp i en myndighetsverksamhet" samt att "klargöra och skärskåda myndighetens prestationer och/eller prestationernas effekter i samhället". En annan sakkunnig utgår ifrån "en vägledande princip för en utvärdering av ett projekt är att undersöka om nytta av projektet överstiger uppoffringarna (kostnaderna) för att genomföra det". Den sakkunniga påpekar samtidigt svårigheten med att göra en perfekt utvärdering men har ändå angivit viktiga utgångspunkter.

I övrigt ombads de sakkunniga att kommentera följande punkter¹³:

¹³ Flera sakkunniga har påpekat att dessa punkter är heltäckande och genomtänkta.

A Dokumentation

Är verksamheten som utvärderas, tidigare utvärderingar, data som används, metoder, resultaten och annat av relevans beskrivet på ett sådant sätt att läsaren kan skapa sig en rimlig bild av utvärderingens tillförlitlighet?

B Data

Är de data som ligger till grund för utvärderingen lämpliga och tillräckligt tillförlitliga?

C Metod

Är utvärderingsmetoden lämplig? Är den konsekvent och riktigt tillämpad?

D Slutsatserna

Är de slutsatser som dras väl underbyggda i analysen? Finns andra rimliga tolkningar av analysen som inte beaktas?

E Kostnadseffektivitet

Anser du att det hade varit möjligt att göra en väsentligt mycket bättre utvärdering med en bättre uppläggning, men – såvitt du kan bedöma – ungefär samma resursinsats när det gäller t.ex. datainsamling, bearbetning, skrivande osv ?

F Objektivitet

Anser du att tillvägagångssättet i utvärderingen, val av frågeställning, tolkning av resultat m.m. ger uttryck för en godtagbar objektivitet hos utvärderaren?

6 Kvaliteten i statliga myndigheters självutvärderingar

6.1 Helhetsbilden

Eftersom de sakkunniga har bedömt självutvärderingarna med egna ord och utan någon standardiserad mall, så kan helhetsbilden inte återges i exakta siffror. Vi har därför valt att göra en enkel kodning av resultaten och sammanställt dessa i tabellform för att få en överskådlig bild. Vi har dessutom kommenterat och beskrivit viktiga dimensioner kvalitativt. Detta innebär givetvis en viss risk att bilden färgas av våra glasögon. För att minska den risken har vi dessutom låtit en oberoende expert koda de sakkunnigas bedömningar. Denna kontrollkodning gav nästan identiska resultat som vår kodning.¹⁴

Samtidigt kan dock vår bild kontrolleras genom att läsa sakkunnigutlåtanden (se bilaga 5). I många fall är de mycket välskrivna och utgör intressant läsning även för den som inte läst de underliggande självutvärderingarna. Vi har också angett utgångspunkter för den enkla kodningen och exemplifierat vårt tillvägagångssätt.

Vi har genomgående valt att koda bedömningar efter följande mall när det gäller de övergripande bedömningarna:

¹⁴ Kontrollkodningen genomfördes av ekonomie doktor Roger Svensson. För den övergripande bedömningen skildes inte kodningen från vår egen med mer än ett steg i något fall.

- (--) = mycket kritisk (mycket stora brister, uppfyller inte minimikraven, beskrivande rapport, någon utvärdering görs ej)
 (-) = kritisk (mindre bra, utvärderingen har en del brister)
 (0) = neutral (varken eller; rimlig nivå, finns både bra och dåliga punkter i utvärderingen)
 (+) = positiv (bra)
 (++) = mycket positiv (mycket bra, utmärkt, ett mycket viktigt bidrag)

I tabell 6.1 ger vi en bild av det övergripande resultatet. Vi har fått in totalt 46 olika utlåtanden (två per utvärdering). De övergripande bedömningarna fördelar sig på följande sätt.¹⁵

Tabell 6.1 Totalbild av de sakkunnigas helhetsbedömningar

(--)	(-)	(0)	(+)	(++)	Totalt
21	16	1	6	2	46

Av de utvärderingar vi har tagit ut har endast sex stycken av totalt 23 utvärderingar erhållit minst en positiv bedömning (+) eller (++) . Två utvärderingar (RFV, SIDA) har av de sakkunniga bedömts med (++) . Den totalt sett låga kvaliteten i utvärderingarna är något förvånande. De sakkunniga gav ett negativt slutomdöme till mer än två tredjedelar av självutvärderingarna.

¹⁵ Det är värt att notera att kontrollkodningen inte skiljer sig mer än ett steg i något fall från vår kodning.

I tabell 6.2 visas de övergripande bedömningarna av respektive utvärdering. Kodningen bygger på följande resonemang.

Det finns fall där de sakkunniga menar att de mest elementära kraven på en utvärdering saknas, avsaknad av försök att mäta resultat av en verksamhet, avsaknad av diskussion om i vilken utsträckning målen för verksamheten uppfylls, osv. Vi har då klassat den övervägande bedömningen som kritisk (-) eller (--).

Låt oss ge några exempel. En sakkunnig skriver i den övergripande bedömningen att:

"Den granskade "utvärderingen" är ett gott exempel på en ...relativt enkel uppföljningsinsats kring ett projekt...men svarar inte på de grundläggande frågor som ställs".

Här har vi gett bedömningen (--), dels för att den sakkunniga inte anser rapporten vara en utvärdering, dels för att den inte svarar på de frågor som ställs.

En annan bedömning som vi har tolkat som kritisk (-) men som kanske t.o.m. skulle kunna klassas som (-) har följande sammanfattning:

"Utvärderingens styrka är att det finns en intresseväckande beskrivning...Utvärderingens svaghet är att man kring de centrala frågeställningarna saknar en riktig datainsamling och en genomtänkt metod för att göra jämförelser. Därtill kan läggas att utvärderaren har en okritisk positiv hållning till hela projektet, vilket resulterar i att slutsatserna är lite distanslösa."

En sakkunnig skriver i den sammanfattande bedömningen att:

"...Då det dessutom finns mycket få mätbara kriterier för output i den organisation som utvärderas blir det utifrån det underlag jag erhållit omöjligt att bilda sig en välgrundad uppfattning om hur väl verksamheten fungerat".

Den bedömningen har vi tolkat som mycket kritisk (--) eftersom utvärderingen inte verkar bidra till kunskap om åtgärdens ändamålsenlighet (se kapitel 2 för en diskussion om begreppet utvärdering).

De sakkunniga som gett utvärderingen "väl godkänt" eller uttryckt att "huvudintrycket är mycket positivt" i sammanfattningen har vi klassat som (++).

I tolkningen av tabellerna skall man komma ihåg att den *inte* avser myndigheters *genomsnittliga* självutvärdering. Man bör därför inte övertolka utfallet för enskilda myndigheter.¹⁶

¹⁶ De sakkunnigas bedömningar återfinns i bilaga 5. Två granskningar per utvärdering har gjorts. I bilagan benämns dessa utlåtanden 1 och utlåtande 2. I tabellerna motsvarar kodningen benämningen i bilagan. Den första bedömningen i tabellerna kommer från utlåtande 1.

Tabell 6.2 Sakkunnigas övergripande bedömningar av självutvärderingarna¹

Starta eget-bidrag. (AMS)	(--)	(--)
Rekryteringsstödet resultat. 1995:1 (AMS)	(--)	(+)
Bosnier hjälper Bosnier (SIV)	(--)	(--)
Tusentals kontakter - nätverk för framtiden. (SIV)	(--)	(-)
Minskade hälso- och miljörisker vid användning av bekämpningsmedel (Jordbruksverket)	(-)	(+)
Varudistribution. (Konsumentverket)	(-)	(-)
Värderad service (Konsumentverket)	(-)	(-)
"Rattfällan." Utvärdering av kurs för fängelsedömda rattfyllerister (KVS)	(0)	(-)
"Goda råd om fisk" - utvärdering av information (Livsmedelsverket)	(--)	(-)
Utvärdering av "Startpaket Nyckelhål" (Livsmedelsverket)	(--)	(--)
"Värdet av regionalpolitiken" (NUTEK)	(--)	(-)
Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län (RFV)	(++)	(++)
Rehabiliteringsarbetet på försäkringskassan (RFV)	(--)	(-)
Utvärdering av RAR-systemet (RPS)	(--)	(-)
Kvalitetsutveckling inom skatteförvaltningen (RSV)	(-)	(--)
Import support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam. (SIDA)	(--)	(+)
Fungibility and Sustainability. Import Support to Tanzania (SIDA)	(+)	(+)
Swedish Support to Cap Verde. (SIDA)	(-)	(-)
Resultat av Särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	(--)	(--)
Utvärdering av den nya fältorganisationen (Skolverket)	(--)	(+)
Den nationella utvärderingen av grundskolan (Skolverket)	(--)	(-)
Pilotstudie av Lex Maria anmälningar och beslut. (Socialstyrelsen)	(-)	(--)
Rapport till Socialstyrelsen. 8 sept 1995 (Socialstyrelsen)	(-)	(--)

¹ I bilaga 2 återfinns en fullständig förteckning över utvärderingarna.

Det är intressant att se hur ofta två sakkunniga ger samma bedömning om samma utvärdering. Endast i 4 av 23 fall gjorde två sakkunniga bedömningar som skiljde sig mer än ett steg från varandra.

I flera fall är det givetvis förenat med svårigheter att göra en bra utvärdering som ger kunskap om verksamhetens ändamålsenlighet och kostnadseffektivitet. Vissa verksamheter är mer komplicerade att utvärdera än andra. Att göra en jämförelse mellan utfallen för en experimentgrupp, som utsätts för en åtgärd och en kontrollgrupp, som inte

deltar, kan t.ex. innebära praktiska problem. I flera granskningar lyfter också sakkunniga fram problemen med att göra goda utvärderingar. Men samtidigt anför många konkreta tillvägagångssätt för att höja kvaliteten på respektive utvärdering betydligt.

I en utvärdering av självhjälpsverksamhet – Bosnier hjälper Bosnier (Invandrarverket) – undersöks hur bosniska flyktingar i Sverige i egen regi sökt tillvarata den egna gruppens kompetens för att underlätta deras integration och möjliggöra eventuell återutvandring. Syftet med utvärderingen var bl.a. att ange vad Invandrarverkets stöd betytt för verksamheten. En sakkunnig skriver att själva utvärderingsmetoden främst går ut på att beskriva vad som utträttats i olika projekt. Trots det angivna syftet görs inte några försök att utröna vilken betydelse SIV:s stöd egentligen haft. I det sammanfattande omdömet anføres följande:

"Studien kan inte kallas för en utvärdering i gängse vetenskaplig mening. En sådan utvärdering är nämligen en jämförelse mellan utfallen för en experimentgrupp, som utsätts för åtgärden, och en kontrollgrupp, som inte deltar. De beskrivningar av utfallen i experimentgruppen som redovisas är därmed inte ointressanta, men man måste naturligtvis vara mycket försiktig med långtgående slutsatser. Ofta är det förenat med stora praktiska eller administrativa svårigheter att använda sig av en kontrollgrupp. I stödverksamhetsprojektet borde man dock ha kunnat använda sig av en sådan grupp, eftersom denna verksamhet inte kom till stånd på mer än två orter – Borås och Lund – och det finns andra orter med många bosnier som man ganska lätt kunnat jämföra med. Intervjuundersökningar skulle kunna ha genomförts med deltagare respektive icke-deltagare. En intressant fråga är t.ex. om icke-deltagare anlidade den reguljära sjukvården i större utsträckning än deltagare. I så fall skulle man ha kunnat peka på möjligheten till kostnadsbesparingar av självhjälpsprojektet. Enligt rapporten skall en intervjuundersökning genomföras, dock endast med bosnier på orter där verksamheten bedrivits."

I detta exempel på granskning framkommer en medvetenhet om metodsvårigheter, samtidigt som det på ett tämligen konkret sätt ges förslag på hur kvaliteten kan förbättras i utvärderingen.

I de olika granskningarna förekommer ofta hänvisningar till olika jämförelsenormer. I flera fall uttrycks en viss uppskattning av myndigheters försök, eller förbättringar i standarden trots att utvärderingen fortfarande beskrivs som otillfredställande. I en del fall sägs också att

en utvärdering skulle kunna vara acceptabel som en intern, "quick and dirty", uppföljning, men inte som en riktig utvärdering.

I en del självutvärderingar som vi tagit ut sägs i syftet att man vill "ge en bild" av någonting. Denna skrivning är viktig att diskutera. I syftet framgår att det är en kartläggning man vill göra vilket givetvis är intressant i olika sammanhang. Trots det försiktiga syftet kan det sedan komma slutsatser där kausala samband och effektiviteten av den undersökta åtgärden fastslås. Slutsatserna måste då givetvis prövas mot den metod som används i rapporten. Ett exempel är en AMS-rapport ("Starta eget-bidrag: En totalundersökning över de personer som avslutade bidragsperioden i oktober -93") som ingår i vår studie. Rapporten beskriver situationen för de personer och därigenom de företag som beviljats starta-eget bidrag (SEB). Dess fastslagna syfte är att kartlägga dessa personers arbetsmarknadssituation och deras uppfattningar om bidragshanteringen efter avslutad bidragsperiod. Sedan slås det emellertid fast i denna "kartläggande studie" att "Vart fjärde företag har mer än fem anställda. Eftersom antalet anställda i företagen överstiger antalet personer som ej längre bedriver verksamhet har SEB därmed resulterat i ett sysselsättningstillskott som överstiger antalet beviljade SEB". En av de sakkunniga menar att slutsatsen saknar grund och förutsätter en rad inte alltid rimlig antaganden. Detta diskuteras mer utförligt nedan där bedömningarna av utvärderingarnas slutsatser presenteras. Intressant är emellertid att studien är ett exempel där man i syftet "friskriver sig" från att göra en "riktig" utvärdering men ändå med bestämdhet slår fast olika resultat. I detta fall bör givetvis resultatet i studien bedömas utifrån om det i studien finns empiri och metod som håller för sådana slutsatser.

Just frågan om vad som kan räknas som en riktig utvärdering återkommer ofta i sakkunnigutlåtanden. Många konstaterar exempelvis att en utvärdering reser viktiga och relevanta frågor i inledningen, men sedan använder empiri som inte ger svar på dessa frågor utan på perifera, mindre relevanta frågor. En del anför att en utvärdering leds fel av att den inte utgår från någon klar frågeställning. Ett exempel är Skolverkets nationella utvärdering av grundskolan 1987-1994. En sakkunnig menar att flera olika syften dyker upp. Syftet som anges i utvärderingens inledning är inte detsamma som vad som används längre fram i samma utvärdering. En sakkunnig anför följande:

"Syftet har därför inte i första hand varit att bedöma Skolverkets sätt att genomföra den nationella utvärderingen (som de själva skriver är syftet). Det har därför varit oerhört frustrerande att försöka läsa denna text som en utvärdering av genomförandet".

Det finns sakkunniga som anser att utvärderingen över huvud taget inte nämner de i sammanhanget relevanta frågeställningarna. Ett exempel är en utvärdering av ett lokalt ADB-stöd (RAR – Rationell Anmälan Rutin) för brottsanmälan (Rikspolisstyrelsen). Av utvärderingen framgår inget annat än att valet stod mellan att sprida RAR till övriga polismyndigheter eller att behålla de manuella rutinerna. Båda sakkunniga påpekar (oberoende av varandra) att det är tveksamt om detta är den rätta frågeställningen. Den ena sakkunnige skriver:

"En fråga ... är om Rikspolisstyrelsen valt rätt huvudfrågeställning för utvärderingen. Stod Rikspolisstyrelsens val verkligen mellan att sprida just RAR till övriga polismyndigheter eller att behålla de manuella rutinerna för brottsanmälan?"

Den andra påpekar:

"Att man skall ha ett ADB-system för administrativa handläggningsrutiner i samband med brottsanmälan är en självklarhet i dagens IT-samhälle men att det skall utformas enligt RAR-systemet är en helt obesvarad fråga".

Ett annat exempel där frågeställningens betydelse diskuteras är en bedömning av Livsmedelsverkets utvärdering "Goda råd om fisk". Utvärderingen går ut på att belysa om verkets information i en viss fråga – kostråd om fisk till gravida och ammande kvinnor – nått ut till målgruppen, varit begriplig och lett till ändrade kostvanor. En sakkunnig har gett utvärderingen godkänt eller bra betyg för de enskilda aspekterna såsom dokumentation, data, metod, slutsatser osv. Men det övergripande omdömet blir ändå negativt. Den sakkunniga för en diskussion om utvärderingens nytta och åtgärdens kostnadseffektivitet och konstaterar att:

"Studien är så avgränsad att den inte ger något meningsfullt svar på frågan om den verksamhet som den utvärderar varit framgångsrik".

När frågeställningar diskuteras undrar en del sakkunniga om den utvärdering som de granskats skall ses som ett komplement till andra utvärderingar som undersöker de centrala frågorna. Utifrån vår kartläggning kan vi dock konstatera att det i inget av dessa fall funnits andra utvärderingar som har undersökt de frågor som sakkunniga lyft fram som de centrala.

En mycket tydlig slutsats av att så många sakkunniga reser dessa frågor är att urvalet av självutvärderingsfrågor är ett problem.

En annan aspekt som diskuteras i sakkunnigbedömningarna är avsaknaden av försök till kostnads- och intäktsvärdering. Ett stöd eller en ny verksamhet kan utvärderas utan att utvärderaren nämner den samhällsekonomiska kostnaden respektive förväntad intäkt för åtgärden. Detta problem påpekas av flera sakkunniga. Ett exempel är en utvärdering av rekryteringsstödet resultat (AMS) där en av de sakkunniga skriver att frågan om vad rekryteringsstödet kostar och om det är kostnadseffektivt relativt andra åtgärder helt har försumrats. Ett annat exempel är RFV:s studie av rehabiliteringsarbetet på försäkringskassan. Resultaten är av stort intresse menar en sakkunnig.

"Däremot ... är det en påtaglig brist att begrepp som kostnader och intäkter/nyttor helt lyser med sin frånvaro. Rapportens värde som beslutsunderlag skulle tillta dramatiskt om kostnaderna för olika delar av den diskuterade verksamheten redovisats."

6.2 Dokumentation, data och metod

De sakkunnigas bedömningar av dokumentation, data och metod framgår av tabell 6.3. Kodningen i tabellen bygger på följande resonemang. Om datamaterialet/underlaget bedöms som tillförlitligt och relevant har vi betraktat detta som positivt. En sakkunnig menar att

"...de data som används, metoden och resultaten är beskrivna på ett sätt så att läsaren kan skapa sig en rimlig bild av utvärderingens resultat".

Här har vi tolkat bedömningen av dokumentationen som positiv (+), dock inte mycket positiv (++) eftersom den sakkunniga samtidigt gör vissa anmärkningar på dokumentationen. Om bedömningen däremot

gett uttryck för att "materialet är för litet för att vara statistiskt tillförlitligt" har vi tolkat detta som negativ kritik (-). Om metoden bedöms som lämplig har vi angett detta som (+) eller (++) beroende på övriga kommentarer om metoden. Om den sakkunniga däremot beskriver metoden som den "enklast möjliga statistiska metod ... där "inga samband undersökts på ett tillfredsställande sätt" har detta tolkats som negativ kritik.¹⁷

¹⁷ De sakkunnigas bedömningar återfinns i bilaga 5. Två granskningar per utvärdering har gjorts. I bilagan benämns dessa bedömning 1 och bedömning 2. I tabellerna motsvarar kodningen benämningen i bilagan. Den första bedömningen i tabellen kommer från bedömning 1.

Tabell 6.3 Dokumentation, data och metod i utvärderingarna.

	Dokumen- tation	Data	Metod
Starta eget-bidrag. (AMS)	(-)(+)	(-)(-)	(-)(-)
Rekryteringsstödet resultat. 1995:1 (AMS)	(+)(++)	(++)(++)	(-)(+)
Bosnier hjälper Bosnier (SIV)	(-)(0)	(?)(0)	(+)(-)
Tusentals kontakter - nätverk för framtiden. (SIV)	(++)(-)	(-)(-)	(-)(-)
Minskade hälso- och miljörisker vid användning av bekämpningsmedel (Jordbruksverket)	(-)(+)	(?)(+)	(-)(0)
Varudistribution. (Konsumentverket)	(0)(-)	(-)(-)	(?)(-)
Värderad service (Konsumentverket)	(+)(0)	(+)(-)	(0)(-)
"Rattfällan." Utvärdering av kurs för fängelsedömda rattfyllerister (KVS)	(++)(+)	(+)(+)	(+)(0)
"Goda råd om fisk" - utvärdering av information (Livsmedelsverket)	(++)(+)	(+)(+)	(0)(+)
Utvärdering av "Startpaket Nyckelhål" (Livsmedelsverket)	(++)(0)	(?)(-)	(-)(-)
"Värdet av regionalpolitiken" (NUTEK)	(0)(+)	(-)(0)	(-)(0)
Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län (RFV)	i(+)(++)	(+)(++)	(+)(++)
Rehabiliteringsarbetet på försäkringskassan (RFV)	(0)(0)	(+)(0)	(?)(+)
Utvärdering av RAR-systemet (RPS)	(-)(0)	(-)(-)	(-)(0)
Kvalitetsutveckling inom skatteförvaltningen (RSV)	(-)(-)	(-)(0)	(-)(0)
Import support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam. (SIDA)	(+)(+)	(?)(?)	(?)(-)
Fungibility and Sustainability. Import Support to Tanzania (SIDA)	(++)(+)	(?)(?)	(+)(-)
Swedish Support to Cap Verde. (SIDA)	(-)(-)	(-)(-)	(-)(+)
Resultat av Särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	(-)(+)	(0)(+)	(-)(-)
Utvärdering av den nya fältorganisationen (Skolverket)	(-)(+)	(-)(-)	(-)(?)
Den nationella utvärderingen av grundskolan (Skolverket)	(-)(-)	(?)(-)	(-)(-)
Pilotstude av Lex Maria anmälningar och beslut (Socialstyrelsen)	(+)(-)	(-)(-)	(?)(+)
Rapport till Socialstyrelsen. 8 sept 1995 (Socialstyrelsen)	(-)(-)	(-)(-)	(+)(?)

¹ I bilaga 2 återfinns en fullständig förteckning över utvärderingarna.

De sakkunniga är i allmänhet mer positiva till enskilda aspekter såsom dokumentation, data, metod jämfört med de övergripande bedömningarna. I nästan hälften av granskningarna är de sakkunniga positiva till dokumentationen. det gäller dokumentationen är det betydligt fler sakkunniga som är positiva jämfört med deras övergripande bedömningar. I 17 fall har utvärderingarna bedömts positivt jämfört med 21 utlåtanden där de sakkunniga varit negativa. Många utvärderingar har varit lätta att följa och bilda sig en uppfattning om. Ett exempel på ett sakkunnigutlåtande om dokumentationen är en bedömning av Livsmedelsverkets utvärdering av "Startpaket Nyckelhål". En sakkunnig skriver:

"Rapporten ger en ganska tydlig bild av hur utvärderingen har gått till. Enkäterna finns med som bilaga, vilket ger möjligheter att se hur frågorna formulerades. Det hade dock varit ännu lättare att bedöma resultatens tillförlitlighet om frågorna, särskilt de öppna hade angivits i anslutning till resultatredovisningen."

Även Livsmedelsverkets andra utvärdering: "Goda råd om fisk" är till synes beskriven på ett tydligt sätt. En sakkunnig anför följande:

"Enkätundersökningen är dokumenterad på ett utmärkt sätt. Detta är utfört av SCB och innehåller så vitt jag kan se allt man behöver veta om population, urval, olika typer av bortfall och fel, uppräkningsförfaranden osv. Enkäten återges inte, men det förefaller som om enkätfrågornas exakta ordalydelse finns i resultattabellerna, vilket jag tycker är OK."

Det finns även kritik mot dokumentationen. En sakkunnig skriver följande i sitt utlåtande av en utvärdering som gjorts av Socialstyrelsen (Rapport till Socialstyrelsen):

"Mitt första intryck av rapporten är att den påminner om en omkullvänt bokhylla. Genom att i den löpande texten hänvisa till tabellerna längst bak i rapporten, tvingas jag att bläddra fram och tillbaka för att få en vettig överblick ..."

Fler bedömningar innehåller kritik mot dataunderlaget, jämfört med de positiva omdömena. Det finns emellertid några utvärderingar som har fått mycket beröm just avseende lämplig och tillförlitlig data. I

utlåtanden om RFV:s utvärdering av försöksverksamhet inom rehabiliteringsområdet i två län (1992:12) anføres följande:

"Tillgången till relevanta och tillförlitliga data synes ha varit imponerande god, och det är uppenbart att RFV satsat betydande resurser på att bygga upp tillförlitliga databaser som kan utnyttjas för en utvärdering av den egna verksamheten."

I ett mindre positivt utlåtande om AMS utvärdering av starta-eget-bidrag" påpekar en sakkunnig följande om datamaterialet:

"Undersökningen kallas totalundersökning vilket ger en felaktig bild eftersom materialet är ett urval av den underliggande populationen på 16 500 personer och urvalskriteriet är en månad ... Detta urvalskriterium diskuteras inte. Huruvida dessa är representativa för den underliggande populationen diskuteras inte heller. Man borde redovisa enkla statistiska mått för denna grupp och populationen för att kunna ge en bild av hur detta urval är relaterat till populationen ..."

Bedömningarna om metoden varierar. I ett utlåtande om Socialstyrelsens "Pilotstudie av Lex Maria-anmälningar och beslut" skriver en sakkunnig att:

"Metoden med personliga intervjuer bedömer jag som lämplig. Där emot kan så väl urvalet av ärenden ..., som urvalet av respondenter ifrågasättas."

RFV:s utvärdering av försöksverksamheten inom rehabiliteringsområdet får mycket beröm för metodval. En sakkunnig framhåller följande:

"Som tidigare angivits kan utvärderingen uppdelas i en statistiskt deskriptiv del, en ekonometrisk analys samt en mycket utförlig konkret beskrivning av hur arbetet i olika avseenden organiserats och förändrats vid försökskassorna under projektets gång. Enligt min uppfattning är detta en mycket lämplig metod för den aktuella utvärderingen. Min kritik gäller framför allt att den ekonometriska analysen redovisats så styvmoderligt i utvärderingsrapporten."

Ungefär hälften av granskningsrapporterna kritiserar den valda metoden. I en bedömning av NUTEK:s utvärdering (Värdet av regionalpolitiken) anføres att:

"Det är i första hand de företag som får eller har fått stöd som intervjuats, men även organisationer och statliga verk i stödområdena får säga sin mening. Utomstående företag/personer/organisationer utanför stödområdena får sällan eller aldrig komma till tals om de regionalpolitiska stöden. Det är inte otroligt att de företag som fått stöd eller kommuner i stödområdena, vars sysselsättning och skatteintäkter är beroende av dessa företag, kommer att säga att stöden har varit till större hjälp än vad de egentligen har varit ...".

Vidare påpekar en sakkunnig att:

"Statistiska simuleringar utförs för att estimerar effekterna av företagsstöd på företagets betalningsförmåga på kort och lång sikt (likviditet och soliditet). Man kommer fram till slutsatsen att ett positivt samband finns. Detta är dock inte förvånande, ty om man ger företag stöd i form av pengar så kommer naturligtvis deras finanser att förbättras."

I ett utlåtande om Skolverkets rapport om den nya fältorganisationen skriver den ena sakkunnige följande i sin bedömning av metoden:

"Utvärderingsmetoden är knappast lämplig, eftersom den är alltför journalistisk och impressionistisk. Utvärderarna har "plockat" aspekter ur bakgrundsmaterialet på ett så pass osystematiskt sätt att en utomstående näppeligen kan bedöma om en rättvisande bild".

I några fall framhålls att metoden är beskrivning. I ett utlåtande av Invandrarnätets utvärdering av "Tusentals kontakter – Nätverk för framtiden. Kontaktnätets första tid" framhålls följande om metoden:

"Metoden är en mycket noggrann beskrivning av genomförandet av Kontaktnätet. Man redogör för olika enkätundersökningar som gjorts samt lägger in ett stort antal intervjuer med guider och med flyktingar. Man beskriver också olika myndigheters roll och uppträdande och hur olika frivilligorganisationer har engagerats ... Detta är inte en utvärdering i egentlig mening för att bedöma effektiviteten av åtgärder."

Ett utlåtande om den metod som används i en utvärdering av självhjälpverksamhet framkommer liknande kommentarer om metoden:

"Utvärderingsmetoden går främst ut på att beskriva vad som utrettats i de olika projekten."

6.3 Slutsatser, kostnadseffektivitet och objektivitet

Vidare framgår av bedömningarna huruvida slutsatserna är välgrundade, om självutvärderingen varit kostnadseffektiv och om utvärderarna förefaller objektiva (tabell 6.4). Om slutsatserna anses väl underbyggda tolkas detta som en positiv bedömning (+) eller (++) beroende på övriga kommentarer om slutsatserna. Om slutsatserna däremot saknar grund eller är bristfälliga har detta tolkats som negativ kritik. I de fall sakkunniga menar att man kunde gjort en bättre utvärdering givet samma kostnad eller att man satsat onödigt stora resurser på utvärderingen (t.ex. onödigt stort antal intervjuer eller datamaterial) har detta tolkats som (-) eller (--). Om utvärderaren eller utvärderarens tillvägagångssätt anses ha godtagbar objektivitet har vi tolkat bedömningen av objektiviteten som positiv. Om den sakkunniga däremot skriver att "Trots att det lyser igenom att den negativa informationen överväger i bakgrundsunderlaget får ej detta fullt genomslag i texten..." har vi tolkat detta som negativ kritik. I de fall sakkunniga inte har bedömt eller kommenterat någon aspekt har vi markerat med (?) i tabellen.¹⁸

¹⁸ De sakkunnigas bedömningar återfinns i bilaga 5. Två granskningar per utvärdering har gjorts. I bilagan benämns dessa bedömning 1 och bedömning 2. I tabellerna motsvarar kodningen benämningen i bilagan. Den första bedömningen i tabellen kommer från bedömningen 1.

Tabell 6.4 Självutvärderingarnas slutsatser, kostnadseffektivitet och objektivitet.

	Slutsatser väl- grundande	Kostnads- effekti- vit	Objektivi- tet
Starta eget-bidrag. (AMS)	(-)(-)	(-)(-)	(-)(0)
Rekryteringsstödet resultat. 1995:1 (AMS)	(-)(+)	(?)(0)	(0)(++)
Bosnier hjälper Bosnier (SIV)	(-)(-)	(?)(-)	(-)(0)
Tusentals kontakter – nätverk för fram- tiden. (SIV)	(+)(-)	(-)(?)	(0)(-)
Minskade hälso- och miljörisker vid an- vändning av bekämpningsmedel (Jord- bruksverket)	(?)(+)	(-)(+)	(?)(+)
Varudistribution. (Konsumentverket)	(-)(-)	(-)(?)	(-)(-)
Värderad service (Konsumentverket)	(-)(?)	(0)(-)	(-)(-)
"Rattfällan." Utvärdering av kurs för fängelsedömda rattfyllerister (KVS)	(0)(+)	(?)(-)	(+)(+)
"Goda råd om fisk" – utvärdering av information (Livsmedelsverket)	(0)(+)	(-)(-)	(+)(+)
Utvärdering av "Startpaket Nyckelhål" (Livsmedelsverket)	(?)(-)	(?)(-)	(-)(+)
"Värdet av regionalpolitiken" (NUTEK)	(-)(0)	(-)(?)	(-)(0)
Försöksverksamhet inom rehabilite- ringsområdet i Malmöhus, Norr- bottens och Blekinge län (RFV)	(+)(++)	(+)(+)	(+)(+)
Rehabiliteringsarbetet på försäkrings- kassan (RFV)	(++)(+)	(-)(-)	(0)(+)
Utvärdering av RAR-systemet (RPS)	(-)(-)	(-)(-)	(-)(0)
Kvalitetsutveckling inom skatteför- valtningen (RSV)	(-)(?)	(-)(-)	(0)(?)
Import support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam. Fungibility and Sustainability.	(-)(-)	(?)(-)	(?)(0)
Import Support to Tanzania (SIDA)	(+)(-)	(?)(-)	(0)(0)
(SIDA)	(-)(-)	(?)(?)	(?)(?)
Swedish Support to Cap Verde. (SIDA)	(-)(-)	(-)(-)	(-)(-)
Resultat av Särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	(-)(-)	(-)(?)	(-)(0)
Utvärdering av den nya fältorganisationen (Skolverket)	(-)(-)	(?)(0)	(?)(+)
Den nationella utvärderingen av grund- skolan (Skolverket)	(-)(-)	(?)(0)	(?)(+)
Pilotstude av Lex Maria anmälningar och beslut (Socialstyrelsen)	(+)(-)	(?)(0)	(0)(-)
Rapport till Socialstyrelsen 8 sept 1995 (Socialstyrelsen)	(-)(-)	(-)(-)	(0)(0)

¹ I bilaga 2 återfinns en fullständig förteckning över utvärderingarna.

Precis som bedömningarna av de övriga enskilda aspekterna är även utlåtanden om slutsatser, kostnadseffektivitet och objektivitet i allmänhet mer positiva jämfört med de övergripande bedömningarna. Som framgår av tabell 6.4 anser många emellertid att de slutsatser som dras inte är välgrundade. Flera sakkunniga anser därtill att det hade gått att göra en bättre utvärdering med samma resursinsats, i flera fall t.o.m. med exakt samma data.

En del ger försiktiga skrivningar eller diskuterar problemet med objektivitet i självutvärderingar i allmänhet. I de fall de sakkunniga klagar på objektiviteten kan det handla om att utvärderarna har en alltför okritisk inställning till den åtgärd (t.ex. bidragssystem) som utvärderas eller att kritiska synpunkter inte getts samma tyngd som positiva omdömen. De sakkunnigas försiktighet angående objektiviteten står i viss motsats till deras bedömningar angående hur välgrundade slutsatserna är. I nästan samtliga fall där slutsatserna inte anses välgrundade har slutsatser trots detta dragits som är till myndighetens fördel.

Låt oss ge några exempel på vad de sakkunniga skriver om de slutsatser som dras i utvärderingarna. I en bedömning av Invandrarverkets utvärdering (av projektet Bosnier hjälper Bosnier) anför en sakkunnig att:

"Slutsatserna kring delprojekten är underbyggda i analysen men det saknas en direkt koppling till de tre grundfrågorna som anges som utvärderingens syften. De generella slutsatserna är dock sämre underbyggda och saknar även de en direkt koppling till grundfrågorna. Bl.a. saknar en analys av vad som skiljer den bosniska gruppen, som i utvärderingen beskrivs som ... relativt välutbildad ..., från andra invandrar- och flyktinggrupper. Denna analys borde vara nödvändig för att beskriva hur "majoritetssamhället mött gruppens mobilisering" eller för att "beskriva under vilka förutsättningar stöd till "etisk" mobilisering skall vara effektiv."

I en utvärdering av Lex Maria anmälningar och beslut (Socialstyrelsen) är syftet att ge underlag för konstruktion av en enkät avsedd för en mer omfattande undersökning och värdering av Lex Maria verksamheten. En sakkunnig anför följande om slutsatserna:

"De slutsatser som dras beträffande det sannolika utfallet av en större studie samt beträffande rekommendationen att vänta några år med en sådan studie känns inte tillräckligt väl underbyggda: "En fortsatt,

större och mer representativ undersökning i anslutning till pilotstudien skulle sannolikt ge ett liknande utslag. Utredaren rekommenderar SoS att avvakta något eller några år med en ny undersökning av effekterna av Lex Maria utredningar och annan kvalitetssäkringsarbete." Det kan inte uteslutas att ett annat urval av ärenden och andra respondenter hade givit en åtminstone delvis annan bild och andra slutsatser."

I en bedömning av KVS som avser en kurs för fängelsedömda rattfyllerister menar en sakkunnig:

"Givet den kvasi-experimentella ansatsen och givet den analysmetod som valts (i huvudsak signifikansprövningar av skillnader mellan deltagargrupp och jämförelsegrupp med hjälp av chi2-test) kan slutsatserna som dras i utvärderingarna betraktas som underbyggda."

Om objektivitet skriver en sakkunnig följande om Konsumentverkets utvärdering av varudistribution:

"I två avseenden brister rapporten i objektivitet: dels ifråga om urvalet av verksamhetsområden, där rapporten – visserligen öppet redovisat – begränsar sig till de områden där Konsumentverket anser sig ha goda resultat. Dels ifråga om rapportens upprepande slutsatser att insatserna lett till goda resultat utan att just denna rapport tydligt visar att så är fallet."

Om Skolverkets utvärdering av den nya fältorganisationen skriver en sakkunnig om objektiviteten:

"Trots att det lyser igenom att den negativa informationen överväger i bakgrundsunderlaget får ej detta fullt genomslag i texten genom att man i flera fall refererar enskilda positiva bedömningar och inslag oavsett om dessa är representativa."

Om Jordbruksverkets utvärdering av "minskade hälso- och miljörisiker vid användning av bekämpningsmedel" skriver en sakkunnig att:

"Jag har ingen anledning att ifrågasätta objektiviteten hos utvärderaren. P.g.a att det s.k. halveringsprogrammet är kvantitativt uttryckt finns små möjligheter att tolka resultaten på andra sätt."

I ett utlåtande om utvärderingen av kurs för fängelsedömda rattfyllerister anför en sakkunnig följande:

"Utvärderingen av Rattfällan visar prov på godtagbar objektivitet. I den första rapporten diskuteras t.ex. svagheter med en kvasi-experimentell ansats där matchning görs utifrån observerbara egenskaper. ... Även bortfallsanalys har genomförts. Utvärderaren sticker inte heller under stol med att de stora bortfallen kan ha påverkat de uppmätta skillnaderna mellan kursdeltagare och jämförelsegrupp, varför de erhållna resultaten bör tolkas med försiktighet."

När det gäller en utvärdering av särskilda skogsvårdsinsatser i norra Sverige anför en sakkunnig att:

"Utredarens objektivitet skulle kunna ifrågasättas eftersom huvudslutsatsen förefaller bristfälligt underbyggd. Man kan ju inte heller bortse från möjligheten att Skogsstyrelsen, för vilken utvärderingen genomförts, har ett intresse av att läsa om positiva effekter av subventioner till skogsbruk. Å andra sidan skall poängteras att utredaren lagt sig vinn om att uppmärksamma svagheter i utvärderingen, även om dessa senare inte vägts in i den heltäckande bedömningen. Det är inte heller helt lätt att komma runt de svårigheter som här påtalats."

Slutligen vill vi ge några exempel på vad de sakkunniga skriver om kostnadseffektiviteten, dvs. om självutvärderingen i sig verkar har varit kostnadseffektiv. AMS-rapporten om starta-eget-bidrag kommenteras på följande sätt avseende kostnadseffektiviteten:

... På basis av mina synpunkter ..., vill jag påstå att det är möjligt att åstadkomma bättre resultat med de resurser som rapporten tycks ha använt."

I en bedömning av RFV:s studie av rehabiliteringsarbetet på försäkringskassan skriver en sakkunnig att:

"Enligt ... uppgifter ... har RFV lagt ned betydande resurser på att samla in det datamaterial som ligger till grund för studien. Det är svårt att se att de slutsatser som dras i rapporten skulle motivera ett så stort kostnadsåtagande. Det bör dock beaktas i sammanhanget att datamaterialet antagligen skulle kunna användas för andra syften än den aktuella studien."

Livsmedelsverkets utvärdering av startpaket Nyckelhål (en informationskampanj) kommenteras av en sakkunnig på följande sätt:

"Det som har krävt mest resurser i utvärderingsarbetet är troligen att genomföra drygt 200 telefonintervjuer. I stället för att intervjua så många som möjligt borde kraft ha lagts ned på att få ett representativt urval och att utforma frågorna mer precist...Med ett bättre upplägg av undersökningen kunde troligen mer tillförlitliga resultat ha erhållits samtidigt som antalet intervjuer kunde ha minskats betydligt...Man borde således ha låtit kvaliteten gå för kvantiteten."

6.4 Några goda exempel på självutvärderingar

PärLAN bland utvärderingarna är en självutvärdering av *Riksförsäkringsverket* (RFV anser 1992:12. "Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län – en utvärdering"). Båda sakkunniga är mycket positiva till denna studie.

Särskilda medel tilldelades tre län för att vidareutveckla försäkringskassans roll i rehabiliteringsarbetet, dels för köp av rehabiliteringsinsatser. Syftet med försöksverksamheten var att utveckla nya arbetssätt och samarbetsformer i rehabiliteringsarbetet som på sikt kan leda till att människor snabbare än hittills får möjlighet att återgå i arbete efter sjukdom eller skada. Syftet med utvärderingen var att bedöma utvecklingen vid de aktuella försäkringskassorna samt uppnådda resultatet. I utvärderingen har man jämfört de områden där försäkringskassorna tilldelats extra medel med andra områden som inte tilldelats medel. Man har med andra ord gjort ett gediget försök att jämföra en ny situation (nya arbetssätt) med alternativsituationen, dvs. vad skulle ha hänt om man *inte* hade genomfört dessa åtgärder. Eftersom det är omöjligt att veta vad som skulle ha hänt i försöksområdena om man inte hade genomfört nya arbetssätt har andra likartade områden, där inga förändringar på försäkringsområdet inträffat, använts som jämförelse. Jämförelsen är sedan bl.a. baserad på regressionsanalys för ohälsotalet där eventuella områdesskillnader avseende kön, ålder, civilstånd, inkomstfördelning m.m. beaktats. En samhällsekonomisk kalkyl har gjorts av försöksverksamheten m.m. Dessutom har en kvalitativ ansats använts där intervjuundersökningar har genomförts med personalen på försökskassorna och deras medaktörer.

De sakkunniga menar att det är positivt att både kvalitativa och kvantitativa metoder har använts. Båda sakkunniga menar att slutsatserna av den genomförda analysen förefaller mycket väl underbyggda. En av de sakkunniga framhåller i sin sammanfattande bedömning att:

"Huvudintrycket av utvärderingen är mycket positivt. Detta beror framför allt på att utvärderingen innehåller en kärna av mera kvalificerad statistisk analys baserad på en mycket relevant databas".

Den andra ger utvärderingen "väl godkänd".

En annan utvärdering som fått positiva kommentarer är SIDA:s "Importstöd till Tanzania". De sakkunniga tog i det fallet ställning till tre utvärderingar som gällde importstöd till tre olika länder. Utvärderingen av importstöd till Tanzania fick positiva omdömen, medan de andra två fick mer negativ kritik. Syftet med utvärderingen är att kartlägga de makroekonomiska effekterna av svenskt importstöd till Tanzania. En fråga gäller hur stödet påverkat tillväxten och tillväxtförutsättningarna i Tanzania. En sakkunnig betonar vikten av ett långsiktigt perspektiv när man skall dra slutsatser om tillväxteffekter. Samtidigt bör man kontrollera för andra faktorer som kan tänkas påverka tillväxten t.ex. konjunkturreffekter, kapitalackumulering, den offentliga sektorns storlek, regleringar och handelshinder. Ett sådant långsiktigt perspektiv finns i utvärderingen av importstöd till Tanzania. I utvärderingen diskuteras hur betydelsen av importstöd empiriskt kan underbyggas genom att t.ex. studera hur pass känslig produktionen är för förändringar i tillgång på utländsk valuta. Importstödet effekter på makroekonomiska variabler behandlas. Båda sakkunniga är positiva till utvärderingen. Den ena sakkunniga skriver att:

"... utvärderingen till stora delar är mycket bra och relevant."

Den andra menar att:

"... detta är en väl genomarbetad utvärdering som uppfyller de mål som stipulerats i direktiven".

Det är intressant att notera att de övriga SIDA-utvärderingarna av importstöd till andra länder (Guinea-Bissau, Mozambique, Nicaragua, Cap Verde m.fl.) kommer till motsatt resultat. I båda dessa utvärde-

ringarna görs inga ambitiösa försök att bedöma de makroekonomiska implikationerna såsom det görs i utvärderingen av importstöd till Tanzania.

En tredje utvärdering som erhållit bra kritik är Kriminalvårdsstyrelsens utvärdering (Rattfällan. Utvärdering av kurs för fängelsedömda rattfyllerister). Syftet med studien var att utvärdera en frivillig kurs för fängelsedömda rattfyllerister. Dels gjordes en analys av i polisregistret registrerade återfall i trafiknykterhetsbrott. De personer som gick kursen jämfördes vad gäller återfall i registrerade trafiknykterhetsbrott inom en 2-årsperiod efter fängelsestraffet med en lika stor kontrollgrupp av individuellt matchade individer som, i likhet med kursdeltagarna, var dömda till fängelse. och där trafiknykterhetsbrott ingick i domen. Matchningsvariabler var ålder, intagningsmånad, kön, tidigare kriminalitet samt huvudbrott i aktuell dom. Dels gjordes en enkätstudie (hos båda grupper) där alkoholvanor samt attityder till, och kunskaper om alkohol och alkohol/trafik analyserades. Den ena sakkunniga bedömer det som försvarbart att utvärderingen inte lagts upp som ett experiment.

"Genom att välja en kontrollgruppsansats samt att utgå från det rimliga antagande att försöks- och kontrollgruppen inte skiljer sig med hänseende på upptäcktsrisk är det faktiskt möjligt att åtminstone delvis studera eventuella faktiska beteendeförändringar efter genomgången kurs."

Variablerna bedömdes som helt tillräckliga för att standardisera jämförelsegrupperna på ett tillfredsställande sätt. Båda sakkunniga påpekar samtidigt med enfaset problemet med selektion. Om det föreligger skillnader mellan grupperna avseende icke observerbara egenskaper, t.ex. skillnader i motivation (kursen är ju frivillig) uppstår ett selektionsproblem. En sakkunnig bedömer dock, trots selektionsproblemet, att slutsatserna som dras i utvärderingen kan betraktas som underbyggda.

7 Slutsatser och rekommendationer

Ungefär en fjärdedel av myndigheterna kan inte redovisa några självutvärderingar under den sex-årsperiod vi undersökt. Av de självutvärderingar som myndigheter genomfört under perioden 1990–95 har vi gjort ett urval av de självutvärderingar som myndigheterna redovisat/föreslagit eller som vi bedömt vara de mest sofistikerade. Dessa självutvärderingar har sedan granskats av flera sakkunniga. Flertalet av dessa självutvärderingar har fått mycket kritik av de sakkunniga.

Det har gått sju år sedan den nya budgetprocessen infördes. En bärande idé i denna reform var just införandet av resultatstyrningen som bygger på att resultat faktiskt mäts och redovisas. Många myndigheter redovisar visserligen enkla mått på verksamhetsutvecklingen i sina årsredovisningar. Dessa ger dock i sig inget underlag för resultatstyrning eftersom de oftast inte ger någon indikation som kan användas i styrningen, t.ex. om varför ändringar i verksamhetens utveckling har skett, hur betydelsefulla verksamhetens effekter eller effektivitet, och i vilken mån och hur det skulle gå att öka effektiviteten. Det är mot bakgrund av de sakkunnigbedömningar som redovisas i denna rapport svårt att undvika slutsatsen att det som borde vara en motor i den nya budgetprocessen ännu inte fungerar.

Olika utredningar har visat att riksdag och regering i liten utsträckning använder sig av myndigheternas självutvärderingar. Vår slutsats är att självutvärderingarnas fåtal och skiftande kvalitet faktiskt gör att de inte kan användas annat än av experter inom respektive område, vilka kan genomskåda dåligt underbyggda slutsatser.

Under den sex-årsperiod som vi granskat (1990–1995) finns visserligen tecken på en viss förbättring i antal självutvärderingar såväl som

i kvaliteten. Ändå förefaller förbättringstakten inte så hög att man kan utgå ifrån att problemet löser sig själv. Under tiden finns också en risk att självutvärderingar får dåligt rykte och att intresset för resultatmätning därmed avtar.

I själva verket finns redan tecken på att systemet med självutvärderingar håller på att undergrävas. En linje, som flera myndigheter driver, är att kraven på myndigheterna att mäta sina resultat bör sänkas. Ett argument som anförs för detta är att det i många fall skulle vara svårt eller omöjligt att mäta resultat och att resultatmätning då utgör en onödig utgift.

Dessa argument motsägs emellertid av de resultat som redovisas i denna rapport. De flesta sakkunniga som vi har anlitat anser att det skulle gå att göra väsentligt bättre självutvärderingar med samma resursinsats. Detta, i samband med att vi ändå valt utvärderingar som tillhör myndigheternas bättre, tyder på att myndigheterna befinner sig långt ifrån den punkt bortom vilken det inte går eller lönar sig att utvärdera längre. I själva verket tyder det på att det mycket väl skulle gå att sänka kostnaden för självutvärdering och ändå göra bättre resultatmätningar.

Vår egen bedömning mot bakgrund av kartläggningen och sakkunnigutlåtanden är att det framförallt är tre misstag som gör självutvärderingar onödigt dyra och minskar deras styrka:

- A) Utvärderingar läggs upp eller beställs i efterhand. Många självutvärderingar skulle bli väsentligt billigare om de planerades i förväg eller i samband med införandet av något program.¹⁹ Som ett exempel kan insamling av data då göras till en automatisk del av den normala administrativa hanteringen. Vidare kan jämförelsegrupper ofta definieras i förväg, men inte i efterhand.
- B) Den dyra delen av en utvärdering är ofta datainsamlingen. I självutvärderingarna görs datainsamlingen ofta med för lite tanke på

¹⁹ Ett exempel är uppläggningsen av ett stöd till biografer, det tidigare fraktstödet. Införandet av stödet gjordes på ett föredömligt sätt när det gäller att lägga upp verksamheter så att de blir utvärderingsbara. Stödet var vid införandet begränsat till vissa delar av landet, vilket gjorde det lätt att hitta en jämförelsesituation. Fraktstödet var utformat som en försöksverksamhet just för att effekterna skulle kunna analyseras på ett relativt enkelt sätt.

metoden. I många fall skulle det gå att nå en bättre utvärderingen med en mindre, men mer genomtänkt datainsamling.

- C) Val av konsulter som genomför utvärderingar på uppdrag verkar i flera fall ske för lättvindigt, vilket leder till höga arvoden och klena resultat.

Mot bakgrund av våra resultat och observationer föreslår vi ett antal reformer som skulle kunna fungera som katalysator för resultatstyrningen

1. Stärk den interna kvalitetssäkringen

Är det lämpligt att myndigheter utvärderar sig själva? I vår bedömning är det ofrånkomligt. Myndigheter måste kunna genomföra självutvärderingar som en del av det interna arbetet med att förbättra verksamheten. Till viss del kommer dessa självutvärderingar också att användas för att legitimera verksamheten utåt. I flera avseenden bör dock myndigheterna själva ta steg för att förbättra utvärderingskvaliteten.

I första hand kan detta göras genom att rutinmässigt genomföra kvalitetssäkring av självutvärderingar. En sådan kvalitetssäkring bör bygga på sakkunnigbedömningar på ett sätt som beskrivs närmare nedan under punkt 3. Därutöver bör utvärderingskompetens och utvärderingsfunktionen stärkas.

En god kompetens är nödvändig, inte bara för att utföra egna utvärderingar, utan även för att beställa och kontrollera utvärderingar gjorda utanför myndigheten t.ex. av utomstående konsulter.

En stark utvärderingsfunktion är också viktigt, t.ex. i form av ett aktivt stöd och satsning från ledningen. Inom ramen för denna rapport har vi haft telefonsamtal med utvärderingsansvariga inom många myndigheter. Inte sällan påpekar de en brist på uppbackning i ledningen vilket gör att de får föra en kamp på flera håll; dels att förankra själva idéerna till olika utvärderingar och dels att sedan få dessa riktigt utförda.

Man kan inte blunda för att utvärderare inom en myndighet befinner sig i en prekär situation. En utvärdering som avslöjar brister i en del av myndighetens verksamhet har kan leda till en svår sats för utvär-

deraren. Det finns åtskilliga exempel på utvärderare som har känt sig tvungna att sluta efter en sådan händelse. Även fristående konsulter som utför utvärdering på beställning av myndigheten kan självfallet hamna i en liknande konfliktsituation.

2. Förstärk den externa kvalitetssäkringen

Den interna kvalitetssäkringen kan emellertid inte lösa alla problem. I vår bedömning förekommer alltför många blockeringar inom myndigheterna för att man helt skall kunna lita på den interna kvalitetssäkringen. Objektiva utvärderingar uppfattas i många fall uppfattas som ett hot mot delar av en myndighets verksamhet. Det är därför att begära mycket om en myndighet förväntas både genomföra och kvalitetssäkra sådana utvärderingar helt själv.

En extern kvalitetssäkring kan stå på flera ben. Det ligger nära till hands att RRV i högre grad än tidigare kommenterar och bedömer självutvärderingars kvalitet och omfattning hos de myndigheter som granskas.

Därtill kan det vara rimligt att regeringskansliet i högre grad deltar i kvalitetssäkringen av myndigheternas utvärderingar.

Under nästa punkt utvecklas ett praktiskt och enkelt tillvägagångssätt på hur de olika instanser kan genomföra kvalitetsgranskningen.

3. Metod för kvalitetssäkring

Som metod för kvalitetssäkring av utvärderingar förefaller en väg vida överlägsen och även mest använd i olika utvärderingssammanhang. Det är samma metod som vi också har använt i denna rapport. Oberoende och anonyma (om de så önskar) sakkunniga anlitas för att kommentera utvärderingar. Av särskild vikt är att dessa sakkunniga inte väljs av den som genomfört utvärderingen, den som utvärderas, eller närstående till dessa. Vi föreslår därför att någon eller några tar på sig eller ges uppgiften att agera som "Kvalitetssäkringsgrupp" med uppgift främst att förmedla sakkunniga. På begäran skulle Kvalitetssäkringsgruppen välja sakkunniga som passar till en viss självutvärdering,

och återföra de sakkunnigas bedömning till de berörda, under bevarandet av de sakkunnigas anonymitet om de så önskar. Kvalitetssäkringsgruppen skulle kunna anlitas av myndigheter som utför självutvärderingar. Detta kan samtidigt fungera som en kvalitetsstämpel som ger läsaren en signal om tillförlitligheten. De sakkunnigbedömningar som kvalitetssäkringsgruppen förmedlar skall självfallet inte bestå av ett enkelt godkännande, utan kan innehålla krav på förbättringar som så småningom leder till ett godkännande.

Men Kvalitetssäkringsgruppen skulle också kunna anlitas t.ex. av ansvariga departement eller av RRV till hjälp i granskningen av självutvärderingars kvalitet. Det är tänkbart, och även önskvärt, att myndigheter skaffar sig en kvalitetsstämpel på sina självutvärderingar som ett sätt att förebygga den kritik som kan komma om någon annan istället anlitar Kvalitetssäkringsgruppen. Självfallet är det också bra om myndigheter använder sig av kvalitetssäkringen i ett tidigt skede under en självutvärderings framställning och lyckas arbeta in de sakkunnigas synpunkter.

En sådan kvalitetssäkringsprocess fungerar sannolikt också som katalysator som ökar utvärderingskunnandet inom myndigheterna. Vårt resultat indikerar att en förbättring av självutvärderingar handlar om kunnande, inte om högre utvärderingskostnader. En sådan kompetens är nödvändig, inte bara för att utföra egna utvärderingar, utan även för att beställa och kontrollera utvärderingar gjorda utanför myndigheten av t.ex. utomstående konsulter. En sådan kvalitetssäkringsprocess underlättar också arbetet betydligt för regeringskansliets tjänstemän som ofta är mottagare av självutvärderingarna. Finns det en kvalitetsstämpel på utvärderingarna kan mottagarna känna sig säkra på att resultaten är relevanta och korrekta.

4. Beställ kompletterande utvärderingar

Ökade kvalitetssäkringskrav kan medföra att myndigheterna gör färre självutvärderingar. Eftersom många myndigheter dessutom inte genomför några självutvärderingar alls, och andra inte har en bra täckning av sin verksamhet, förefaller det finnas behov av externa organ som kan beställa eller genomföra kompletterande utvärderingar på

eget initiativ. I princip finns sådana möjligheter redan idag. I praktiken fungerar de dock inte på ett tillfredsställande sätt.

De ansvariga fackdepartementen kan t.ex. beställa kompletterande utvärderingar. Det är dock ingen hemlighet att fackdepartementen alltför ofta "värnar" om sina myndigheter, anställer personal från sina myndigheter, osv. vilket gör att objektiviteten inte alltid finns. Andra instanser som RRV och Riksdagens Revisorer granskar visserligen myndigheter, men i en särskild form, som hittills inte innefattat att beställa eller genomföra kompletterande självutvärderingar. Slutligen finns även forskare som ibland utför utvärderingar. Ofta har emellertid forskare svårt att få tillgång till de data som myndigheten förfogar över, och även om de får det så styrs deras intresse ofta av frågeställningar i den akademiska forskningslitteraturen, vilket exkluderar många typiska självutvärderingsfrågor.

Förekomsten av kompletterande självutvärderingar kan stärkas på flera sätt. En möjlighet är att ge revisionsmyndigheterna ett tydligare mandat att beställa kompletterande självutvärderingar. Därtill skulle sektorsforskningsmyndigheter kunna få ett tydligare mandat att även satsa på projekt som utgör utvärdering av myndighetsarbete. Ett exempel på sådant är att Kommunikationsforskningsberedningen beställt en kostnads/intäktsanalys av Dennispaketet, som därmed blev den första oberoende utvärderingen av projektet.

Referenser

- Arbetstagarkonsult, 1992. En ny start för regeringskansliet.
Ds 1995:6 ELMA – Åtgärder för att stärka resultatstyrningen.
Finansdepartementet
- Ds 1995:14 Företagsstödet – Vad kostar det egentligen? Rapport till
ESO, Finansdepartementet
- Fölster Stefan m.fl., 1992. Sveriges systemskifte i fara? Erfarenheter
av privatisering, avreglering och decentralisering, Industriens
Utredningsinstitut
- IVA, 1991. Bäst i världen, Vad kan vi lära av världens ledande
företag?
- Riksdagens revisorer, 1991/92. Att ändra statsförvaltning: Styrnings-
problem i ljuset av fyra fallstudier.
- Riksdagens revisorer, 1991/92. Resultatorienterat budgetarbete
- RRV, 1991. Resultatredovisningar. Några erfarenheter från myndig-
hetsförsök.
- RRV, 1993. Resurser och resultat. En analys av innehållet i statliga
myndigheters anslagsframställningar under 25 år. Av Joachim
Schäfer
- RRV 1994:10 Resultat, Verksamhet, Ekonomi – en handledning för
myndigheterna.
- RRV 1995:39 NUTEK:s stöd till företag och forskning.
- Sanders, James R., 1994. The program evaluation standards. 2nd
Edition. The Joint Committee on Standards for Educational Evaluation.

- Seidel, A.D., 1993, Producing Usable Research: A Selected Review, Policy Studies Review, 3.
- Statskontoret, 1990:24. "Att mäta eller inte mäta resultat".
- Statskontoret 1991:32. Myndigheternas förnyelsearbete.
- Statskontoret, 1992:23. Att använda resultatmått med gott resultat.
- Tarschys, D. (1982) För döva öron? Politikern och utvärderaren, i Palmlund, I. (red.), Utvärdering av offentlig verksamhet, Liber.

Abstract

Can Government Agencies Evaluate themselves?

A major reform of the budgetary process in 1988 introduced an entirely new approach to governance of state agencies in Sweden. The central aim of the reform was to allow the government and Parliament to govern by performance management. Government agencies are now expected to evaluate and report their own performance and results.

The new budgetary process has itself been evaluated a number of times since its inception. A frequent finding has been that, while many agencies have become more performance-oriented, their performance evaluations are rarely utilized as a basis for decision-making by management, government or Parliament.

Non-utilization may have many explanations. One fundamental question is whether public agencies' self-evaluations should be used. Do self-evaluations examine relevant questions? Are they objective and trustworthy? Do the evaluations employ valid methods?

This report maps the extent and quality of self-evaluation in 22 government agencies. Self-evaluations are defined as attempts to establish causal links between an agency's efforts and results, that are initiated or conducted by the agency itself. Quality of self-evaluations was judged by independent referees. The self-evaluations chosen for refereeing were generally those the government agencies themselves considered to be among their best. Each self-evaluation was examined by two referees. In total, 40 referees were engaged. All of them were experts and researchers in the various fields with many years of evaluation experience. Public agencies participated in choosing referees, and could veto referees that they considered biased.

The mapping of self-evaluations showed that 10 out of 22 government agencies had conducted fewer than one self-evaluation per 1000 employees over the period 1990-1995. Six agencies had conducted no self-evaluations at all. Among the self-evaluations conducted, simple

surveys of attitudes and judgements dominate. Empirical measurement of effects or effectiveness is relatively rare.

As far as quality is concerned, referees generally supplied detailed motivations of their judgements, that are reprinted as an appendix to this report. A majority of referees are very critical of the quality in the examined self-evaluations. The criticism concerns documentation, choice of method, choice of evaluation question and other aspects. Most referees argued that conclusions drawn in the self-evaluations were not supported by the empirical evidence. Many also felt that much better self-evaluations could have been produced with the same effort and resource limit. Referees often pointed out that self-evaluations had entirely neglected central questions concerning how adequate and efficient the evaluated project might be.

A few self-evaluations were praised by the referees. These generally focused on a central question, employed acceptable empirical methods, and reported conclusions that were supported by the evidence.

In sum, this report indicated that most government agencies' self-evaluations in fact should not be utilized. This implies that the central aim of the new budgetary process – to allow management based on performance and results – has yet to be achieved.

A much improved self-evaluation process is probably necessary if performance management is to be realized. Quality control of self-evaluations ought to be a crucial ingredient. In part responsibility for quality control must lie with the government agencies themselves; in part organisations independent of the self-evaluating agency should participate more in quality control. It would, for example, appear reasonable that the National Audit Office to a greater extent comments and examines quality in government agencies' self-evaluations.

One approach to quality control is most frequently used, and probably most reliable. That is the approach used in preparing this report. Independent and anonymous (if they desire) referees examine self-evaluations. In this process it is particularly important that referees are not chosen by the government agencies themselves. It is therefore proposed that a "Quality Control Group" be organized with the main task of choosing and engaging referees. Upon request the Quality Control Group would choose referees for a certain self-

evaluation and forward the referees' report while maintaining the referees' anonymity.

The Quality Control Group could be engaged by government agencies themselves. This would provide a quality stamp that would signal to readers that an evaluation is reliable. The Quality Control Group could also be engaged for example by the ministries or the National Audit Office.

Finally, this report recommends that evaluations of government agencies' performance more often are initiated by external interests, such as the ministries or the National Audit Office.

When the quality of self-evaluations has been improved it will again be meaningful to consider the issue of non-utilization. Quite possibly improved quality, and improved information about quality, will prove sufficient to make self-evaluations usable.