

UTVÄRDERING

Effekt- utvärdering

Att välja upplägg

esv

EKONOMISTYRNINGSVERKET



Innehållsförteckning

FÖRORD	5
INLEDNING	6
Syftet med effektutvärdering	6
Målgrupp för handledningen	6
Handledningens fokus	6
Handledningens disposition	7
Läsanvisningar	7
1. EFFEKTER OCH EFFEKTUTVÄRDERING	8
Definition av effekter	9
Effekter kan analyseras med hjälp av verksamhetslogik	9
Hur uppstår effekter?	11
Vad är en effektutvärdering?	11
Olika typer av effektutvärderingar	12
2. EXEMPEL PÅ UPPLÄGG	14
Vad är ett upplägg och varför behövs det olika upplägg?	15
Vad består ett upplägg av?	15
Översikt över basuppläggen	18
Presentation av basuppläggen	19
Upplägg A: Före-efterstudie utan kontrollgrupp	20
Kameraövervakning på Möllevångstorget i Malmö	
Upplägg B: Efterstudie utan kontrollgrupp	23
Introduktionsprogram för nyanställda handläggare i statlig tjänst	
Upplägg C: Före-efterstudie med slumpmässigt valda grupper	25
Rehabilitering av unga immigranter i primärvården	
Upplägg D: Efterstudie med slumpmässigt valda grupper	27
Prövning av kikhostevaccin	
Upplägg E: Före-efterstudie med matchade grupper	30
Effekter av lördagsöppna systembolagsbutiker	
Upplägg F: Efterstudie med matchade grupper	31
Anställningsstöd - en väg från arbetslöshet till arbete?	

Upplägg G: Teoretiskt upplägg	32
Informationskampanjen om HIV/AIDS - samhällsinformation som styrmedel	
Upplägg H: Flera upplägg inom samma utvärdering	34
Transportbidraget - en analys av motiv, effekter och alternativ	
Skolverkets arbete för att öka kvaliteten inom yrkesutbildningen	
3. ATT VÄLJA UPPLÄGG	38
Beskriv den åtgärd som ska utvärderas	39
Beskriv typen av effekter	39
Arbeta med verksamhetslogik	40
Bedöm vilka upplägg som kan vara lämpliga	40
Bedöm resursåtgången för alternativen	43
Lägg fast kravet på säkerställda resultat	44
Gör en sammanvägd bedömning och välj upplägg	45
BILAGA 1: ATT GENOMFÖRA EN EFFEKTUTVÄRDERING	
– SKOLVERKETS INSPEKTIONSVERKSAMHET	49
Utvärderingsuppdraget	49
Utvärderingens resultat	49
Valet av upplägg	49
Genomförande av utvärderingen	52
Skolverkets användning av resultaten	54
BILAGA 2: EXEMPEL PÅ OLIKA TYPER AV EFFEKTUTVÄRDERINGAR	55
BILAGA 3: FAKTORER SOM PÅVERKAR SÄKERHETEN I RESULTATEN	58
Faktorer specifika för upplägg utan kontrollgrupp (Upplägg A och B)	58
Faktorer specifika för upplägg med kontrollgrupp (Upplägg C - F)	58
Faktorer som är generella för alla upplägg	59
BILAGA 4: LITTERATURFÖRTECKNING	61
BILAGA 5: INDEX	63
FOTNOTER	64

Förord

EKONOMISTYRNINGSVERKET (ESV) arbetar för att öka och förbättra användningen av utvärdering i den svenska statsförvaltningen. ESV gör det främst genom att erbjuda utbildningar och producera handledningar inom utvärderingsområdet. ESV åtar sig även utvärderingsuppdrag från myndigheter och departement.

Den nya budgetprocess som infördes i den svenska statsförvaltningen i början av 1990-talet introducerade mål- och resultatstyrning. Den byggde på ökad långsiktighet, resultatorientering och en helhetssyn på verksamhet och ekonomi. Tanken var, och är fortfarande, att den ska bidra till en effektivare statsförvaltning. För att öka utrymmet för nya reformer ska existerande verksamheter kontinuerligt omprövas och de som fortfarande bedöms prioriterade ska effekti-

viseras. Detta ställer stora krav på att information tas fram som visar om verksamheten leder till önskade förändringar i samhället.

Utvecklingen i Sverige har kommit långt när det gäller styrning och redovisning av prestationer, men det finns behov av att fokusera mera på effekterna av verksamheterna. ESV:s förhoppning är att denna handledning ska medverka till detta.

ESV vill speciellt tacka de personer som med sina utvärderingar bidragit till att konkretisera handledningen. Tack också till ekon. dr Kim Forss, docent Patrick Ohlin och adjungerade professorn Johan Bring som kommenterat valda delar av slutmanuset.

Handledningen har utarbetats av Kajsa Nilsson och Lennart Widell.

Susanne Jansson
Avdelningen för statsmakternas styrning

Inledning

Syftet med effektutvärdering

Effektutvärdering är en del i statsmakternas mål- och resultatstyrning och en viktig faktor för att öka effektiviteten inom statsförvaltningen. En utvärdering kan till exempel användas som underlag för ett beslut som rör behovet av en verksamhet eller belysa om den är tillräckligt effektiv. Om en utvärdering pekar på att en verksamhet har en mycket begränsad effekt kan det leda till att verksamheten omprövas. Om till exempel en viss typ av trafikövervakning inte tycks förändra bilförarnas beteende, bör de ansvariga överväga en annan typ av övervakning. Om en informationskampanj mot rökning inte leder till att färre människor röker bör andra insatser övervägas.

Statsmakterna kan använda effektutvärderingar som underlag för propositioner och för sitt budgetarbete men också som utgångspunkt för dialogen med och styrningen av myndigheterna. Myndigheterna kan i sin tur använda effektutvärderingar för sin egen styrning och kunskapsuppbyggnad.

Det finns vissa krav formulerade för statsförvaltningen när det gäller utvärdering och redovisning av effekter. Kraven är dock mycket allmänt hållna. Lagen om statsbudgeten¹ ställer krav på regeringen att redovisa prestationer och effekter till riksdagen och riksdagen styrs av riksdagsordningen² som kräver att utskotten utvärderar riksdagsbesluten. Förordningen för årsredovisning och budgetunderlag³ ställer krav på myndigheterna att redovisa prestationer och effekter i förhållande till mål som regeringen ställt upp. Enligt verkförordningen⁴ ska styrel-

sen pröva om myndighetens verksamhet bedrivs effektivt och i överensstämmelse med syftet med verksamheten och behöver därför underlag för detta. Myndighetschefen ska fortlöpande följa upp och pröva den egna verksamheten och konsekvenserna av de föreskrifter och särskilda beslut som rör verksamheten. Myndighetschefen ska även se till att regeringen får det underlag som den behöver för att ta ställning till omfattningen och inriktningen av verksamheten.

Målgrupp för handledningen

Handledningen riktar sig till personer inom statsförvaltningen som ska genomföra en effektutvärdering. Den vänder sig också till personer som ska ta ställning till resultatet av effektutvärderingar eller beställa en effektutvärdering.⁵

Handledningens fokus

Faserna i en effektutvärdering är i princip desamma som i andra typer av utvärderingar eller i utredningar. De olika faserna visas i figur 1.

Vid en effektutvärdering är det viktigt att kunna samla in information, inte bara efter det att en verksamhet avslutats, utan även innan den startats. Planeringen av utvärderingen är därför avgörande och bör helst göras innan den verksamhet som ska utvärderas inleds. Bäst är om utvärderingen ses som en integrerad del av hela verksamheten.

Valet av upplägg sker inom planeringsfasen. Även om det finns tillfällen där valet är självklart och där utvärderaren istället kan ägna mer energi åt andra faser av utvärderingsprocessen är upp-



Figur 1: Utvärderingsprocessen

lägget strategiskt för att genomföra effektutvärderingar. Därför har ESV koncentrerat handledningen till detta moment och denna fas.

I teorin finns många olika upplägg för effektutvärderingar. I denna handledning presenterar vi ett antal vanliga upplägg, basupplägg, och diskuterar vilka som är lämpliga för olika situationer.

Valet av upplägg har konsekvenser för vilka datainsamlingsmetoder som ska användas och hur materialet ska analyseras. Vi har dock valt att inte explicit behandla datainsamlings- och analysmetoder i den här handledningen. Exempelen som presenteras är dock rika på detaljer om hur utvärderaren gått till väga för att samla in och analysera data. Vi har inte heller fokuserat på värdering av effekterna. Den som vill läsa mer om hur man i allmänhet genomför en utvärdering eller om metoder för datainsamling och värdering av effekter hänvisar vi till allmänna utvärderingshandböcker.⁶ I bilaga 1 ges också ett exempel på en effektutvärdering där alla faser i effektutvärderingen beskrivs.

Handledningens disposition


I det första kapitlet sätter vi in effektutvärdering i sitt sammanhang. I det andra beskriver vi ett antal basupplägg. Varje upplägg illustreras med ett praktiskt exempel.

I det tredje och sista kapitlet beskriver vi hur utvärderaren kan avgöra vilket upplägg han eller hon bör använda för att bedöma effekterna. Vi diskuterar hur olika förutsättningar påverkar valet av upplägg för utvärderingen och använder våra exempel från kapitel 2 för att illustrera detta.

I bilaga 1 finns en presentation av hur man praktiskt kan genomföra en effektutvärdering och i bilaga 2 ges ett antal exempel på effektutvärderingar som genomförts inom den offentliga förvaltningen. I bilaga 3 finns exempel på metodologiska problem vid effektutvärderingar.

Läsanvisningar

Handledningen kan läsas på flera olika sätt. Den som i allmänhet vill lära sig om effektutvärderingar läser hela handledningen. Den som främst vill få en överblick av vad effektutvärderingar handlar om kan nöja sig med att läsa kapitel 1. Den som vill läsa mer om ett speciellt upplägg för att utvärdera effekter (till exempel före-efterstudie utan kontrollgrupp) och praktiska tillämpningar hittar dessa i kapitel 2. Den som framför allt är intresserad av själva valet av upplägg finner detta i kapitel 3. Till sist, med hjälp av indexet i bilaga 5, är det enkelt att hitta förklaringar till olika begrepp som används inom effektutvärdering. ●



Effekter och effektutvärdering

Definition av effekter

**Effekter kan analyseras med hjälp
av verksamhetslogik**

Hur uppstår effekter?

Vad är effektutvärdering?

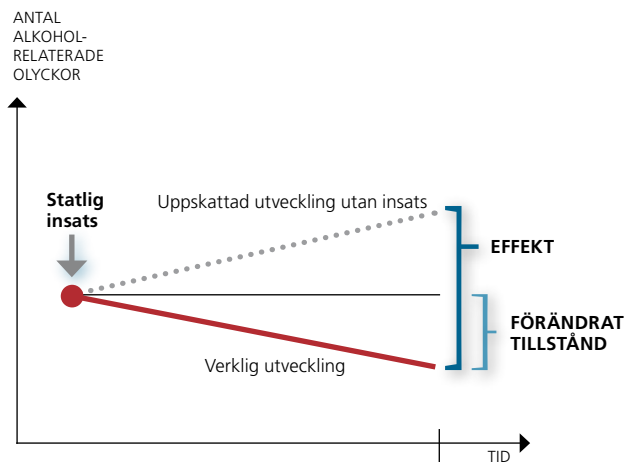
Olika typer av effektutvärderingar

Definition av effekter

Effekt är tillsammans med begreppet prestation centrala begrepp i den statliga ekonomistyrningen. ESV har valt att kalla både prestationer och effekter för resultat.⁷ Med *prestationer* menas tjänster eller produkter som lämnar myndigheten, det vill säga som tas emot av en målgrupp utanför myndigheten. Målgruppen kan vara enskilda individer, en grupp individer eller hela allmänheten men kan också bestå av olika organisationer eller företag. Exempel på en prestation är en inspektion, trafikövervakning, ett domstolsbeslut, en informationskampanj, arbetsmarknadsutbildning, polispatrullering och bidrag till företagare.⁸

ESV definierar en *effekt* som ”en förändring som inträffat som en följd av en vidtagen åtgärd och som annars inte skulle ha inträffat”.⁹ I ett konkret fall kan åtgärden eller prestationen bestå i att polisen ökar antalet alkoholutandningsprov som tas på vägarna. Om rattonykerheten minskar efter detta och därmed också de alkoholrelaterade olyckorna utgör minskningen ett förändrat tillstånd. Därmed inte sagt att ett förändrat tillstånd är detsamma som att en effekt uppstått. För att kunna tala om effekt av en insats måste det finnas ett orsak-verkan-samband mellan insatsen och det förändrade tillståndet. I det här fallet måste man kunna påvisa att en minskad trafikonykerhet är en följd av polisens ökade provtagning.¹⁰

Effekten av en enskild organisations verksamhet kan illustreras genom att jämföra vad som hypotetiskt skulle ha inträffat utan organisationens insatser, det vill säga *kontrafaktiskt*, och vad som faktiskt inträffade. Detta brukar ibland kallas för ”om-inte-problematiken” och visas i figur 2. I det här fallet beräknas effekten som skillnaden mellan det verkliga antalet olyckor som inträffat



Figur 2: Effekt och förändrat tillstånd

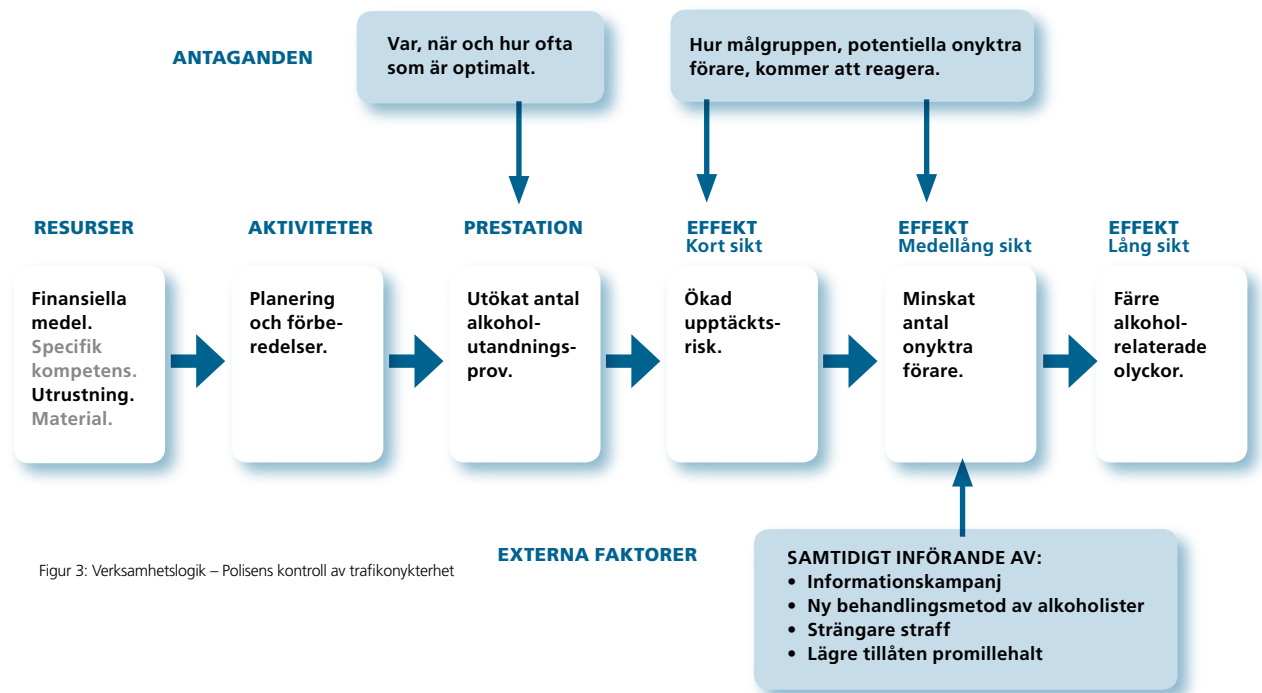
efter insatsen och det uppskattade antalet olyckor som skulle ha inträffat om inte insatsen genomförts.

Ibland är det, i likhet med fallet med den ökade alkoholkontrollen, relativt enkelt att förstå vad det kontrafaktiska resonemanget innebär. I andra fall kan det vara betydligt svårare och kräva stor fantasi och analytisk förmåga. Exempel på detta är frågan om effekten av att det finns statliga myndigheter eller effekten av Internet.

Det är lätt att förväxla effekt och förändrat tillstånd, vilket kan leda till att effektutvärderingen i själva verket enbart består av en beskrivning av hur tillståndet har förändrats.

Effekter kan analyseras med hjälp av verksamhetslogik

Effekter kan ibland uppstå i flera led och detta kallas vanligtvis *mål-medel-kedja* eller *effektkedja*, där en insats leder till en effekt som i sin tur leder till en ny effekt och så vidare.



Figur 3: Verksamhetslogik – Polisens kontroll av trafikonykterhet

Effektkedjan kan ingå i en beskrivningsmodell som länge har använts i utvärderingssammanhang, så kallad *verksamhetslogik* eller programteori.¹¹ Här beskrivs förväntade relationer mellan resursinsats, aktiviteter, prestationer och effekter. Figur 3 illustrerar detta.

Pilarna i bilden representerar ett förväntat orsak-verkan-samband. Effekterna representerar uppställda mål, det vill säga det man förväntar sig ska hända på grund av åtgärderna. I modellen ingår också beskrivningar av vilka antaganden som ligger bakom de förväntade orsakssambanden, vilka målgrupperna är i olika delar i kedjan och vilka externa aktörer och faktorer som också kan tänkas påverka utfallet.

När effekten av en myndighets verksamhet ska utvärderas i förhållande till ett mål kan verksam-

hetslogiken utgöra en grund för bedömning av vilka effekter i kedjan som är möjliga att mäta och hur långt myndighetens ansvar kan sträcka sig med hänsyn till andra aktörers agerande och faktorer i omvärlden. Det är då viktigt att det finns en beskrivning av antaganden om vilka mekanismer som ligger bakom de förväntade orsakssambanden. Om önskade effekter inte inträffar kan det till exempel bero på att åtgärden är byggd på fel antaganden.

Ett av de mål som polisen ska bidra till är ett minskat antal alkoholrelaterade trafikolyckor (se figur 3). Polisen kan välja att öka antalet kontroller med alkoholutandningsprov för att öka upptäcktsrisken med syfte att det ska påverka potentiella trafikonyktra att inte köra bil berusade. I vilken grad detta bidrar till minskad trafikonyk-

terhet och minskade alkoholrelaterade olyckor kan vara svårt att säkerställa. Andra händelser som kanske inträffar samtidigt kan också påverka utfallet, som till exempel en samtidigt pågående informationskampanj, en ny behandlingsmetod för alkoholister, strängare straff för att köra onyktert eller lägre accepterad promillegräns och så vidare. I verksamhetslogiken försöker man identifiera dessa händelser.

Om ökade kontroller inte tycks bidra till en positiv förändring kan det bero på att kontroller inte är en effektiv metod, det vill säga att förarna inte reagerar på ökad upptäcktsrisk eller att kontrollerna borde utföras vid andra tidpunkter eller på andra platser för att vara effektiva. Det kan alltså vara själva åtgärden eller implementeringen som inte fungerar. Det är därför viktigt att undersöka vilka antaganden om lämpliga tidpunkter och platser för kontroll som gjorts och också förarnas sätt att reagera på ökad upptäcktsrisk. En effektutvärdering kan hjälpa till att belysa detta.

Verksamhetslogik underlättar identifiering av tänkbara effekter, inte bara på målgruppen utan också på andra grupper i samhället. Verksamhetslogiken kan därför bidra till att utvärderaren redan från början kan bedöma om det finns behov av att studera effekter på andra grupper i samhället än målgruppen.

Effekterna påverkas för det mesta av flera aktörer eller faktorer i omvärlden. När statsmak-

terna fastställer mål för olika statliga insatser blir de därför sällan avgränsade till vad en enskild myndighet kan åstadkomma. I stället handlar målen ofta om effekter som flera aktörer och faktorer i samhället bidrar till. Målet för ett helt politikområde kan sägas vara den önskade effekten av olika insatser i statsförvaltningen. Målen för en enskild myndighet uttrycks av den anledningen ofta i termer av att myndigheten ska bidra till en önskad förändring.¹²

Hur uppstår effekter?

Syftet med olika typer av statliga åtgärder är att statsmakterna vill påverka grupper eller individer i samhället i en viss riktning eller skapa förutsättningar för positiva förändringar i samhället. Åtgärderna kan bestå av mindre insatser eller stora reformer, vara helt nya eller också en del av den ordinarie verksamheten.

En inventering av effektutvärderingar inom offentlig förvaltning som ESV gjort visar exempel på åtgärder och önskade effekter av dem. Några typexempel redovisas i tabellen på nästa sida. De rör åtgärder på olika nivåer i statsförvaltningen, från riksdagen till myndigheterna. I bilaga 2 finns en förteckning över ett antal utvärderingar som vi har sorterat efter vilken typ av upplägg som använts och typ av åtgärd.

Vad är en effektutvärdering?

Bedömningen av effekter kan göras med hjälp av såväl uppföljning¹³ som utvärdering.

Vid *uppföljning* mäts vanligen det vi kallar förändringar av tillstånd i samhället och dessa mått kombineras ofta med mått på de prestationer som var tänkta att påverka tillstånden. En prestation kan till exempel handla om förbättrade vägar och det motsvarande förändrade tillståndet kan vara ett minskat antal trafikolyckor.

Exempel på mål

Mål för politikområdet rättsväsendet:
Den enskildes rättstrygghet och rättssäkerhet.

Ett av målen för ekobrottsmyndigheten:
Risken för att brott begås ska minskas.

Exempel på statliga åtgärder och deras förväntade effekter

Typ av åtgärd	Exempel	Förväntad effekt
Skatt	Sänkning av momsen	Ökad konsumtion
Investering	Vägförbättring	Färre trafikolyckor och bättre framkomlighet
Stöd	Stöd till handikappade	Bättre levnadsförhållanden för handikappade
Information	Antirökkampanj	Minskat antal personer som får lungcancer
Utbildning	Utbildning i offentlighet och sekretess	Ökad rättssäkerhet
Omorganisation	Sammanslagning av två myndigheter	Ekonomiska stordriftsfördelar
Arbetsform	Nya behandlingsformer för narkomaner	Färre återfall bland narkotikamissbrukare
Administrativa rutiner	Förbättrade IT-system för handläggning av bidrag	Snabbare och säkrare handläggning
Tillsyn	Inspektion av skolor	Minskad mobbning

Uppföljning med mått på förändrade tillstånd ger inte svar på varför en förändring har inträffat och om och i vilken utsträckning en statlig verksamhet har påverkat förändringen i tillståndet. Detta mått visar bara i vilken riktning utvecklingen går. Uppföljningen ger därmed enbart en indikation på möjliga effekter. Det minskade antalet trafikolyckor kan bero på många andra faktorer än förbättrade vägar. Det kan till och med vara så att trots förbättrade vägar ökar trafikolyckorna och detta på grund av till exempel ökad trafik eller minskad trafikövervakning.

Ofta får beslutsfattare nöja sig med en uppföljning som fokuserar på ett förändrat tillstånd som utgångspunkt för sina beslut. Men för att få ett bättre underlag kan man behöva gå vidare med en *effektutvärdering*.

Begreppet utvärdering kan definieras på en rad olika sätt. Med utvärdering menar ESV *”en noggrann efterhandbedömning, dvs. en bedömning som görs efter det att en åtgärd har vidtagits”*. Utvärderingar kan vara inriktade på olika

faktorer som till exempel processer, prestationer, kvalitet eller effekter. Effektutvärderingar är som namnet anger inriktade på effekter. I praktiken genomförs dock ofta effektutvärderingar som ett led i en utvärdering som även fokuserar på prestationer eller processer inom myndigheten.

Olika typer av effektutvärderingar

I en effektutvärdering försöker utvärderaren fastställa ett orsakssamband mellan en insats och en förändring av ett tillstånd som organisationen i fråga försökt påverka. Effekterna bedöms ofta utifrån mål som statsmakterna eller organisationen själv satt upp. I det fallet talar vi om effektutvärdering gentemot mål, målutvärdering.

Myndigheter får ofta i uppdrag av regeringen att utvärdera effekter i förhållande till politiskt satta mål. I andra fall gör myndigheten utvärderingen på eget initiativ för att kunna redovisa resultaten till regeringen. Om regeringens mål inte är tillräckligt tydliga behöver myndigheten ta fram *indikatorer* på måluppfyllelsen.

Det går att använda andra normer än mål som utgångspunkt för en utvärdering. Exempel på alternativ är förväntningar hos olika målgrupper eller normer som utgår från forskningsresultat. Ett alternativ är att jämföra utvecklingen under en viss tidsperiod eller jämföra med andra organisationer – så kallad *benchmarking*. I vissa sammanhang får utvärderaren i uppdrag att även utvärdera positiva och negativa *bieffekter* och dessa kan vara både förutsedda och oförutsedda. Om en sänkt hastighetsgräns är tänkt att leda till ett minskat antal trafikolyckor innebär det sannolikt också att framkomligheten minskar. En positiv bieffekt kan vara att hastighetsgränsen inverkar positivt på miljön. I en *målfri* utvärdering behöver utvärderaren inte utgå från officiella mål för att bedöma vilka effekter som en insats lett till.

I en utvärdering kan effekterna jämföras med de resurser som åtgärderna kostat. Detta görs vanligen i de fall där utvärderingen utgår från givna mål men kan även användas i en målfri utvärdering. Om effekterna anses ”värda” mer än insatserna kostat brukar man tala om att verksamheten är *kostnadseffektiv* eller ger ”värde för pengarna”. I vissa fall, till exempel med hjälp av *cost-benefit-analysis*, kan effekterna värderas i pengar.¹⁴

En effektutvärdering kan också fokusera på att identifiera de mekanismer som gör att effekterna uppstår. Det kan röra sig om att förstå varför en viss information ger effekt eller vad det är som gör att en reform som syftar till att utveckla socialvården inte ger några effekter.

I vissa fall kan det vara viktigt att skilja mellan effekten av själva åtgärden och effekten av genomförandet av samma åtgärd.¹⁵ Det gäller i de flesta fall där själva mötet mellan utföraren och målgruppen är viktig. Om utvärderingen till

exempel avser två metoder inom socialvården och metod A visar sig vara mer effektiv än metod B kan det bero på att metod A i sig är effektivare. Men resultatet kan också bero på att genomföraren av metod A är skickligare än den som arbetar med metod B. I detta fall kan det därför vara intressant att undersöka hur metod B genomfördes för att eventuellt finna orsaken till det sämre resultatet. Slutsatsen kan bli att metoden B är den mest effektiva eller är lika effektiv som metod A om den genomförs på rätt sätt.

Effekter kan presenteras såväl kvantitativt som kvalitativt. Kvantitativt kan de uttryckas numeriskt (antal, volym) eller som procent av något. Kvalitativt kan effekterna uttryckas i form av citat från intervjupersoner till exempel rörande attityder, uppfattningar, motiv eller beteende.

Effekterna av en kurs kan till exempel uttryckas kvantitativt genom ett test av typen flervalsfrågor där det är lätt att se hur mycket deltagarna lärt sig. Kvalitativt kan effekterna redovisas genom citat som visar vad och hur mycket deltagarna lärt sig. Även dessa kvalitativa data kan bearbetas genom kvantitativa metoder. I detta fall kan utvärderaren till exempel räkna antalet gånger som intervjupersonerna anger att de använt sig av kursens innehåll.

Hur effekterna presenteras och bearbetas beror bland annat på syftet med utvärderingen. I många effektutvärderingar behöver utvärderaren uttrycka effekterna såväl kvantitativt som kvalitativt. ●



Exempel på upplägg

Vad är ett upplägg och varför behövs det olika upplägg?

Vad består ett upplägg av?

Översikt över basuppläggen

Presentation av basuppläggen

Vad är ett upplägg och varför behövs det olika upplägg?

Vi definierar ett upplägg för effektutvärdering som en ”plan som tar ställning till vilka grupper som ska studeras och jämföras, på vilket sätt dessa grupper ska väljas ut och när jämförelserna ska göras”.¹⁶ Vilka upplägg som är tänkbara vid en given utvärderingssituation beror delvis på när utvärderaren blir inkopplad men även syftet med utvärderingen och planerna för hur resultatet ska användas påverkar valet.

Ibland testas idéerna bakom en planerad statlig reform i mindre skala, till exempel i några utvalda län, innan den genomförs fullt ut i hela landet. Man brukar tala om pilotstudier eller försöksverksamheter. I utvärderingslitteraturen talar man om *experiment*. Oftast är då utvärderingen med som en del av försöket och utvärderaren kan tidigt påverka hur utvärderingen ska läggas upp.

Innan en utvärderare bestämmer sig för att göra en utvärdering behöver han eller hon bedöma förutsättningarna för utvärderingen. På så sätt går det att avgöra vilket upplägg som passar bäst. Det finns tre centrala frågeställningar som behöver belysas i det sammanhanget:

- Vad karakteriserar *utvärderingssituationen*?
- Vilka *resurser* finns till förfogande för utvärderingen?
- Vilka krav ställs på *säkerställda resultat*?¹⁷

I detta kapitel presenterar ESV ett antal basupplägg och analyserar när de är lämpliga att använda. Vi utgår då ifrån utvärderingssituationen, resursläget och kraven på säkerställda resultat, det vill säga med vilken säkerhet utvärderaren ska kunna uttala sig om de uppmätta effekterna. Uppläggen illustreras med exempel på utvärderingar där också valet av datainsamlingmetod

Säkerställda resultat

I utvärderingslitteraturen talar man om att en utvärdering bör uppfylla bestämda krav på hur pass säkerställda resultaten ska vara. Begreppet säkerställd brukar omfatta både *reliabilitet* och *validitet*.¹⁸ Begreppet innebär också att antalet observationer är tillräckligt stort.

Reliabilitet, ibland även kallad tillförlitlighet, anger precisionen i mätningen.¹⁹ Begreppet innebär att en upprepning av mätningen ska ge samma resultat.

Validitet står för ”att man mäter det man avser att mäta”. Frekvensen arresteringar bland ungdomar är en valid indikator på upptäckta brott bland ungdomsbrottslingar. Som indikator på ungdomsbrottsligheten är validiteten lägre eftersom det kan finnas många brott som inte upptäcks. Ofta är det lämpligt att använda en kombination av flera indikatorer på effekter för att få fram ett säkerställt resultat.

redovisas. I kapitel 3 sammanfattas och jämförs erfarenheterna från de olika uppläggen.

Vad består ett upplägg av?

Som vi visat på sidan 9 gäller det för utvärderaren att i en effektutvärdering kunna jämföra en faktisk utveckling med en kontrafaktisk, som indikerar vad som skulle ha hänt om inte den studerade åtgärden hade genomförts. Det finns olika metoder för att göra sådana jämförelser. Gemensamt för våra basupplägg är att de utgår ifrån följande:

- Vilka individer eller enheter ska studeras?
- Hur ska de väljas ut?
- När ska mätningar och jämförelser göras?

Vilka individer eller enheter ska studeras?

Ett sätt att kategorisera uppläggen är att dela in dem i studier med eller utan kontrollgrupp.

UTAN KONTROLLGRUPP

Orsakerna till att utvärderaren väljer att inte använda en kontrollgrupp kan vara flera. En orsak kan vara att utvärderaren kommer in i efterhand och att det inte går att etablera en kontrollgrupp. Men även om utvärderaren är med från början kan hon eller han välja att enbart studera gruppen som utsätts för åtgärden (den kallar vi då undersökningsgrupp). Detta kan bero på att en studie med kontrollgrupp inte skulle bidra till ett mer säkerställt resultat. Eller också kan det bero på att det är billigare att bara studera en grupp.

Utvärderaren kan i sådana fall välja att komplettera sin analys med andra typer av jämförelser för att bedöma påverkan av andra faktorer.²⁰ Två sådana metoder anges nedan:

Generisk kontroll

Generisk kontroll innebär att resultatet för undersökningsgruppen jämförs med vad som är normalt hos en större grupp, som anses vara allmängiltig och som inte utsatts för försöket eller insatsen. Om exempelvis en vägsträcka förbättrats för att minska trafikolyckorna på den vägen kan man använda sig av statistik över trafikolyckorna i samma län eller i hela Sverige som jämförelse.

Skuggkontroll

Skuggkontroll innebär att utvärderaren låter personer med särskilda insikter uppskatta insatsens effekter genom att bedöma vad som skulle ha hänt om åtgärden inte vidtagits. Personer som

tillfrågas kan vara experter inom området, någon i ansvarig position för insatsen eller personer som utsatts för insatsen.

MED KONTROLLGRUPP

Studier med kontrollgrupp bygger på att utvärderaren kan välja ut likvärdiga grupper där den ena, experimentgruppen, utsätts för en insats och den andra, kontrollgruppen, inte utsätts för insatsen. Ibland används flera experimentgrupper som utsätts för olika typer av åtgärder, till exempel olika mediciner mot samma sjukdom.

Beroende på hur grupperna tas ut skiljer man på kontrollerat experiment och kvasiexperiment. Gemensamt är att experimentgrupp och kontrollgrupp ska vara så lika varandra i alla väsentliga hänseenden som kan påverka resultatet och att bara åtgärden eller behandlingen de exponeras för ska skilja dem åt.

Kontrollerat experiment

Om individerna eller enheterna som ska ingå i ett experiment fördelas slumpmässigt mellan en experimentgrupp och en kontrollgrupp kallar vi det för ett *kontrollerat experiment*.²¹ Den slumpmässiga fördelningen gör att grupperna inte på ett systematiskt sätt skiljer sig åt, vilket annars skulle

Slumpmässigt urval

Slumpmässigt urval innebär att alla möjliga urval av en viss storlek har samma sannolikhet att bli valda ur en population (undersökningsgrupp).

Slumpmässigt urval kan genomföras med hjälp av slumptalsgeneratorer som till exempel en slumptalstabell eller ett datorprogram (programmet Excel innehåller till exempel en urvalsfunktion).

påverka resultatet. Skillnaden i resultat mellan experimentgruppen och kontrollgruppen kan signifikant testas. Detta test mäter om skillnaden är tillräckligt stor för att inte ha uppkommit av en slump. Det kontrollerade experimentet bygger också på att åtgärdens storlek och karaktär kan påverkas för experimentgruppen. För att öka säkerheten i resultatet utsätts kontrollgruppen ibland för en verkningslös åtgärd, så kallad placebo, för att gruppens vetskap om att den inte utsätts för åtgärden inte ska påverka utfallet. Detta benämns *blindning*. Om inte heller den person som genomför försöket vet vad de olika grupperna utsätts för talar man om dubbelblindning, vilket ytterligare ökar säkerheten i resultatet.

Vid ett kontrollerat experiment görs urvalet i två steg. I det första steget gäller det att utifrån vissa kriterier välja individer eller enheter som ska ingå i experimentet. I det andra steget fördelas urvalet slumpmässigt i experimentgrupp och kontrollgrupp.

I vissa situationer krävs att utvärderingen läggs upp som ett kontrollerat experiment. Det gäller när utvärderaren måste kunna uttala sig med stor säkerhet om de uppmätta effekterna, som fallet ofta är inom den medicinska sektorn.

Kvasiexperiment

Om det inte är möjligt att ta ut individer eller enheter slumpmässigt kan utvärderaren utse en kontrollgrupp som i väsentliga delar liknar experimentgruppen så mycket som möjligt, så kallad matchning. Ibland kan matchade grupper vara att föredra, till exempel om det handlar om mindre grupper. Experiment med matchad kontrollgrupp kallas för *kvasiexperiment*. Om till exempel en insats genomförs i en region kan en annan region få fungera som kontrollgrupp.

Det gäller då att denna region liknar experimentregionen så mycket som möjligt med avseende på de faktorer som kan påverka effekterna av åtgärden. I ett kvasiexperiment kan urvalet göras antingen före eller efter det att åtgärden vidtagits. Det går alltså att konstruera ett kvasiexperiment i efterhand om det finns data om individerna eller enheterna före insatsen.

Population, urval och datainsamling

Det är inte självklart att utvärderaren måste samla in data för alla enheter/individer som omfattas av åtgärden. Utvärderaren har tre huvudalternativ att välja mellan: att undersöka samtliga i den *population* som omfattas av åtgärden (totalundersökning) respektive att göra en urvalsundersökning i form av ett *slumpmässigt urval* eller ett *strategiskt urval*. Med strategiskt urval menas att urvalet görs utifrån specifika kriterier, för att till exempel få en storleksmässig eller geografisk spridning i urvalet, om det är faktorer som kan påverka resultatet. Anledningen är att man vill eliminera risken för att det inte ska gå att uttala sig om effekten av själva åtgärden. Olika urvalsmetoder är lämpliga i olika sammanhang. De utvärderingar som vi använder som exempel visar på olika urvalsmetoder.

Urvalets storlek påverkar i hög grad kostnaderna för en utvärdering. Om en population är mycket stor kan det av ekonomiska men även tidsmässiga skäl vara lämpligt att välja ut ett mindre urval ur populationen och enbart samla in data för denna grupp. En urvalsundersökning kan ibland även ge säkrare resultat än en totalundersökning. Om datainsamlingen sker med hjälp av intervjuer innebär ett färre antal intervjuer att det är lättare att använda mer kvalificerad personal. Datainsamlingen kan då dessutom göras mera noggrant.

När ska mätningar och jämförelser göras?

Upplägget av en utvärdering påverkas av om det går att få fram data om tillståndet innan åtgärden genomförs. En viktig faktor är om utvärderaren kan vara med från inledningen av en insats, eller ännu hellre redan före start.

Om utvärderaren är med från början

Om utvärderaren är med från början och det går att få fram data före åtgärden är det möjligt att göra ett kontrollerat experiment eller ett kvasiexperiment. I dessa kontrollsituationer kan ibland experiment- och kontrollgrupp antas vara lika före åtgärden och då behövs ingen jämförelse med situationen före åtgärden. Effekten utläses som skillnaden mellan grupperna efter åtgärden, vilket vi kallar *efterstudie*.

Om det finns risk för att experiment- och kontrollgrupp skiljer sig åt före åtgärden behöver utvärderaren jämföra skillnaden mellan förändringen i experimentgruppen och i kontrollgruppen. Utvärderaren justerar på så sätt effekten för skillnader i utgångsläget. Vi talar då om en *före-efterstudie*.²²

Även om utvärderaren är med från början väljer hon eller han ibland av olika anledningar att inte arbeta med en kontrollgrupp utan gör i stället en före-efterstudie av gruppen som utsätts för åtgärden. Helst bör utvärderingen då kompletteras med andra jämförelser, som till exempel med vad som är normalt i en större grupp, så kallad generisk kontroll.

Om utvärderaren kopplas in i efterhand

Om utvärderaren kopplas in i efterhand går det ofta inte att göra en före-efterstudie. Utvärderaren får då göra en efterstudie med till exempel skuggkontroll, där deltagarna själva får göra en

uppskattning av situationen före åtgärden eller där en expert gör uppskattningen. Även om utvärderaren kopplas in i efterhand kan det vara möjligt att genomföra en före-efterstudie. Det gäller om det redan finns data tillgängliga som har insamlats för andra ändamål eller som en del av en allmän uppföljning.

Vad betyder före och efter?

Begreppen ”före” och ”efter” kan låta enkla men det kan i praktiken vara svårt att exakt bestämma när mätningarna ska göras för att bli så rättvisande som möjligt. I vissa fall kan resultatet variera kraftigt med avseende på till exempel veckodag och årstid. I andra fall kan resultatet påverkas av hur lång tid efter det att åtgärden satts in som mätningarna görs. Ofta är det viktigt att följa upp ett försök efter det att en tid har gått för att avläsa om insatsen har effekt även senare eller om effekten avtar.

I vissa fall kan det vara intressant att göra mätningar under tiden då åtgärden sätts in. På det sättet kan utvärderaren inte bara bedöma om åtgärden haft effekt utan också få idéer om varför den haft effekt.

Översikt över basuppläggen

I detta avsnitt beskriver vi principerna för ett antal basupplägg, som är vanliga då man genomför effektutvärderingar. Sex av uppläggen grundar sig på insamling av data, så kallade *empiriska* upplägg, och de redovisas i tabellen på sidan 19.²³

Upplägg A-F: Empiriska upplägg

Uppläggen A och B bygger inte på någon kontrollgrupp. De fyra andra uppläggen grundar sig på användning av kontrollgrupp, C, D, E och F. Alternativen C och D är så kallade kontrollerade

Upplägg A-F: Empiriska upplägg

VILKA SKA INGÅ I STUDIEN?	UTAN KONTROLLGRUPP	MED KONTROLLGRUPP	
NÄR SKA MÄTNING GÖRAS?	Hela populationen eller ett urval	Med slumpmässigt fördelade grupper	Med matchade grupper
Före-efterstudie	A	C	E
Efterstudie	B	D	F

experiment och alternativen E och F är kvasiexperiment utan slumpmässigt urval.

Upplägg G-H: Övriga upplägg

De empiriska uppläggen bygger på ett antal specifika förutsättningar för att de ska fungera väl. De förutsätter bland annat att åtgärden är väldefinierad, att jämförelsen inte gäller ett flertal olika undersökningsgrupper och att det finns tillräckligt med tid och resurser. Tidsbrist, resursbrist eller svårigheter att få fram data kan innebära att utvärderaren föredrar ett teoretiskt upplägg (upplägg G). Om utvärderingen till exempel gäller flera undersökningsgrupper kan en kombination av upplägg behöva användas (upplägg H).

Upplägg G: Teoretiskt upplägg

Ett teoretiskt upplägg utgår inte ifrån empiri utan framför allt ifrån logiska resonemang byggda till exempel på verksamhetslogiken, statistiska modeller och litteraturstudier.

Upplägg H: Flera upplägg inom samma utvärdering

Utvärderaren kan i praktiska situationer kombinera olika upplägg som är lämpliga för olika

delar i utvärderingen. Upplägg H är därför inte ett upplägg utan består av kombinationer av uppläggen A–G.

Förekomst av olika upplägg

I statsförvaltningen är i dagsläget utvärderingar av typen A, B och H vanligast förekommande. Likvärdiga kontrollgrupper är ofta svåra att hitta, vilket gör att uppläggen C - F är ganska sällsynta. Kontrollerade experiment förekommer idag främst inom den medicinska sektorn, där de ofta är ett krav, och i viss mån inom arbetsmarknadssektorn och den sociala sektorn.

Presentation av basuppläggen

Vi har valt att redovisa uppläggen i en viss turordning. Vi börjar med de empiriska uppläggen, som i dag är vanligast förekommande i statsförvaltningen (A–B).²⁴ Turordningen ska i övrigt inte ses som en rangordning i någon bemärkelse.

Presentationen av uppläggen redovisas i form av ett antal verkliga exempel. Som en inledning till exemplen presenteras huvudprinciperna för upplägget med hjälp av samma exempel och en figur som visar hur effekterna kan utvärderas.

Utvärderingssituationen och upplägget

För att förklara hur de olika uppläggen förhåller sig till varandra tar vi ett exempel från utbildningsområdet. Exemplet är av pedagogiska skäl starkt förenklat.

Ibland är efterstudie utan kontrollgrupp tillräckligt

Antag att vi vill utvärdera effekterna i form av ökade kunskaper av en tvåveckors kurs i förvaltningskunskap för nyanställda vid statliga myndigheter och departement. Utvärderingssituationen är nu relativt enkel. Förmodligen finns det enbart en faktor som påverkar deltagarnas kunskaper och det är själva kursen. Utvärderaren kan därför anta att om deltagarna inte hade gått kursen hade deras kunskaper om förvaltning varit lika stora före som efter kursen. Därför behövs ingen kontrollgrupp.

I den här situationen räcker det med en före-efterstudie. En möjlighet är att testa deltagarna på deras kunskaper före respektive efter kursen (Upplägg A). Ett alternativ är att i efterhand fråga deltagarna hur deras kunskaper förändrats till följd av kursen (efterstudie, Upplägg B). Om utvärderaren misstänker att det trots allt finns andra faktorer som kan ha påverkat deltagarnas kunskaper kan han eller hon låta de inblandade själva bedöma denna påverkan (skuggkontroll).

Ibland krävs kontrollgrupper

Låt oss nu anta att staten vill introducera en helt ny pedagogisk uppläggning av kursen i förvaltningskunskap, från föreläsningar till interaktiv pedagogik. Innan den nya typen av kurs lanseras vill man i en provomgång

bedöma om den ger bättre resultat än den gamla. Utvärderingssituationen blir då svårare. Före-efterstudien och efterstudien ger även i detta fall indikationer på effekterna men nu måste utvärderaren jämföra hur deltagarnas kunskaper för de två kurserna förändras. Detta kan man göra genom att parallellt genomföra de två kurstyperna. Deltagarna på den nya kursen bildar nu experimentgruppen medan deltagarna på den existerande kurstypen utgör kontrollgruppen. En möjlighet är att slumpmässigt ta ut personer som får gå den gamla respektive den nya typen av kurs. Om antalet deltagare är stort är det troligt att kunskaperna för deltagarna i de två grupperna inte skiljer sig åt före utbildningens start. I det fallet är det rimligt att enbart mäta resultatet efteråt (Upplägg D). Om deltagarantalet är litet finns det risk för att deltagarnas kunskaper före utbildningen skiljer sig åt. I så fall måste utvärderaren ta hänsyn till detta och bedöma kunskaperna före såväl som efter kursen (Upplägg C).

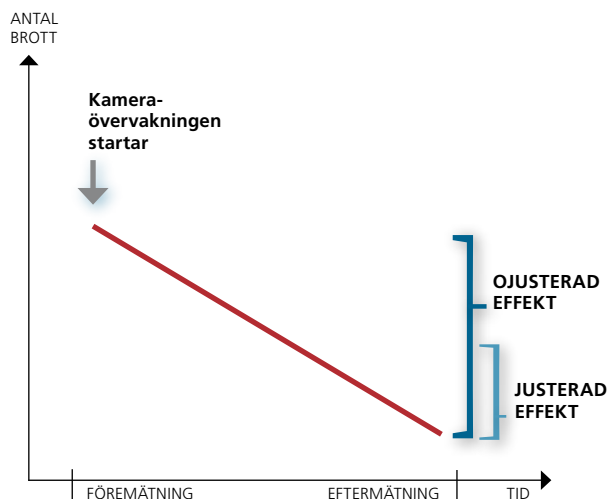
Troligen går det inte att slumpmässigt ta ut personer som får respektive inte får delta i en utbildning. I stället får utvärderaren utgå ifrån de personer som anmäler sig och försöka att matcha dem med varandra så att de har liknande bakgrund och kan tänkas veta lika mycket om förvaltningskunskap och vara lika motiverade att studera. Det mest säkerställda resultatet fås även här om mätningen görs såväl före som efter insatsen (Upplägg E). Ibland kan detta vara omöjligt och utvärderaren får då nöja sig med att göra en eftermätning (Upplägg F).

Upplägg A: Före-efterstudie utan kontrollgrupp

Utvärderaren utvärderar i detta fall effekten av insatsen genom att jämföra tillståndet före respektive efter insatsen hos den studerade gruppen. I figuren visar vi hur en utvärdering av effekterna

på brottsligheten av kameraövervakning av en speciell plats, till exempel ett torg, kan gå till.²⁷

Den ojusterade effekten fås genom att mäta brottsligheten före och efter det att kameraövervakningen satts in. Den är inte korrigerad för påverkan från eventuella andra faktorer. För att



Figur 4: Upplägg A – Kameraövervakning

ta hänsyn till att även dessa faktorer kan ha bidragit till en förändring behöver utvärderaren komplettera analysen med andra jämförelser. På detta sätt får vi en justerad effekt.

I avsaknad av kontrollgrupp kan utvärderaren till exempel jämföra resultaten med genomsnittliga data för en större grupp (generisk kontroll, se sidan 16). I det följande exemplet med kameraövervakning för att minska brottsligheten på ett torg går det inte att få fram ett likvärdigt torg som kan fungera som kontrollgrupp. Därför används i stället brottsstatistiken för hela kommunen som jämförelse. Genom att ta hänsyn till den allmänna trenden i kommunen kan utvärderaren få fram en justerad effekt som man bedömer kan hänföras till kameraövervakningen.

Ibland kan det vara lämpligt att hantera inverkan från andra faktorer genom att göra flera mätningar både före och efter det att åtgärden satts in. Mätningarna kan bilda en tidserie som analyseras manuellt (så kallad okulärbesiktning) eller med hjälp av olika statistiska modeller, till exempel *tidserieanalys*.²⁸

Evidensbaserade metoder (Evidence Based Practice)

Under senare år har det i utvärderingskretsar förts en intensiv diskussion om begreppet *evidens*.²⁵ Evidens kommer från engelskans *evidence* som betyder bevis. Evidens står för att ta fram kunskap på ett systematiskt, gärna vetenskapligt, sätt. Evidensbaserade beslut bygger på dessa kunskaper. Förespråkarna för evidens förordar upplägg med experimentgrupp och kontrollgrupp och grupperna ska helst väljas ut på ett slumpmässigt sätt (detta motsvarar uppläggen C - D och i viss mån även E - F).

Länder där denna diskussion varit mycket livlig är USA och England. En indikation på att det experimentella och kvantitativa upplägget har ett starkt grepp, åtminstone i USA, är att det amerikanska utbildningsdepartementet numera prioriterar ansökningar om medel för utvärderingsprojekt som använder experimentupplägg.²⁶

Vid den amerikanska utvärderingsorganisationens (AEA) kongress år 2004 debatterades evidensbaserade studier. Organisationen var kritisk mot departementets prioriteringar och påpekade att experimentella utvärderingar inte är de enda som kan påvisa effektsamband och att de inte lämpar sig för alla utvärderingssituationer.

Inom vissa sektorer, till exempel den medicinska, är den experimentella uppläggningsen vanlig. Inom detta område finns numera en typ av faktabank för evidensbaserade studier, Cochrane Collaboration. En motsvarande organisation inom beteendevetenskaperna har också bildats med samma inriktning, Campbell Collaboration. Båda organen gör systematiska studier, så kallade meta-analyser för att bedöma effekter inom respektive område. Se vidare <http://www.cochrane.org> och <http://www.campbellcollaboration.org>

Samtidigt kan utvärderaren kartlägga andra faktorer som kan ha påverkat tillståndet. För en utvärdering av ett kameraövervakat torg ger vi exempel på faktorer som skulle kunna vara av intresse för analysen:

- Ökade/minskade polisinsatser i området
- Storm och regn (som gör att få personer passerar)
- Torget avstängt
- Semestertider
- Stor festival vid torget
- TV-kamerorna slutar fungera

Utifrån denna analys kan utvärderaren intervjua personer med särskilda insikter inom området och be dem att utifrån det insamlade materialet ringa in åtgärdens betydelse (skuggkontroll, se sidan 16).

KAMERAÖVERVAKNING PÅ MÖLLEVÅNGSTORGET I MALMÖ ²⁹

Utvärderad effekt

Sommaren 2001 fick polisen i Malmö tillstånd att övervaka och spela in händelser vid Möllevångstorget i Malmö med syftet att förebygga brott. Utvärderingen studerade övervakningens effekter på brottsligheten och konsekvenser för tryggheten och den personliga integriteten.

Utvärderingssituationen och val av upplägg

I detta fall var det svårt att hitta en kontrollgrupp i form av ett likvärdigt torg. I stället jämförde utvärderarna brottsutvecklingen vid Möllevångstorget med den i Malmös centrala stadsdelar och kommunen som helhet, så kallad generisk kontroll (se sidan 16). För att undersöka eventuella omfördelningseffekter studerades också brottsutvecklingen på platser dit brotten kunde tänkas ha flyttat.

Konsekvenser för tryggheten och den personliga integriteten men även upplevelsen av brottsligheten studerades med hjälp av intervjuundersökningar. Dessa riktades till befolkningen i kommunen, till besökare på torget samt till boende vid platsen (skuggkontroll). Dessutom genomfördes längre intervjuer med de boende, näringsidkare, representanter för olika nätverk och brottsförebyggande aktörer.

Datainsamlingen gällde de polisanmälda brott som kunde antas förebyggas av kameraövervakningen på det övervakade torget. Uppgifterna om anmälda brott samlades in för de tre åren före kameraövervakningens start till drygt ett år efter starten. I analysen ingick endast de brott som hade begåtts utomhus på Möllevångstorget. Insamlingen genomfördes genom olika typer av intervjuer samt brottsstatistik före och efter kameraövervakningen.

Resultat

De polisanmälda brotten på Möllevångstorget minskade tydligt efter det att kameraövervakningen infördes. Året efter det att kamerorna satts upp minskade antalet brott på torget med 40 procent. Störst var minskningen av brott mot person, vilka halverades jämfört med de föregående åren. Även stöldbrotten minskade tydligt. Känslan av trygghet bland besökare, boende och näringsidkare på torget tycks ha ökat efter det att kamerorna satts upp. Utvärderingen visade också att de flesta besökare, näringsidkare och boende runt torget inte ansåg att kameraövervakningen innebar en kränkning av integriteten.

Upplägg B: Efterstudie utan kontrollgrupp

I många fall är det bara möjligt att arbeta med en efterstudie utan kontrollgrupp. Om utvärderingen till exempel påbörjas efter det att insatserna satts in kan det vara svårt att genomföra en föremätning, speciellt om det inte finns statistik tillgänglig. Av samma anledning kan det vara svårt att identifiera en kontrollgrupp. Ibland kan dessutom, som i vårt exempel nedan, värdet av en kontrollgrupp ifrågasättas.

I detta upplägg frågar utvärderaren i efterhand personerna i undersökningsgruppen om de själva upplever en förändring. På så sätt kan den ojusterade effekten tas fram. Samtidigt får gruppen ange i vilken mån den specifika insatsen bidragit till förändringen (skuggkontroll, se sidan 16)³⁰ och på så sätt kan den justerade effekten bedömas.

Det följande exemplet handlar om effekterna på arbetet för deltagarna av ett introduktionsprogram för nyanställda i statsförvaltningen. Figuren visar hur en eftermätning kan gå till. I detta fall får deltagarna via en enkät uppskatta sin arbetsef-

ektivitet, före respektive efter programmet.³¹

Det finns ett stort problem med att i efterhand fråga vad en person kommer ihåg av det som hände för en tid sedan, personen kan ha glömt vad som hänt. Ett annat problem gäller själva frågetekniken. Det är lätt att frågorna enbart handlar om den uppskattade förändringen och att de inte ringar in den uppskattade effekten som har att göra med insatsen.

Exempel på frågor i en efterstudie

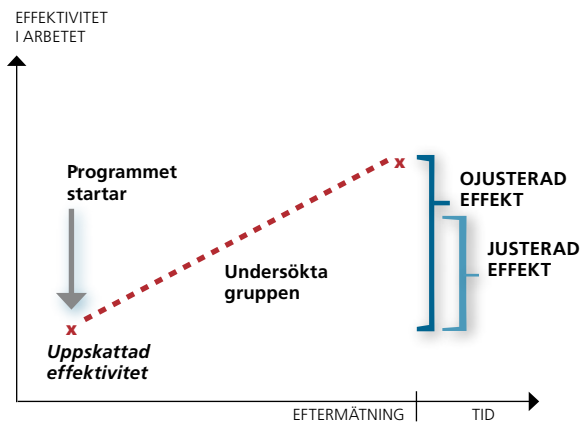
- På vilket sätt har utbildningen påverkat ditt sätt att hantera dina arbetsuppgifter?
- Finns det andra faktorer som du tror har påverkat ditt sätt att arbeta?
- Hur mycket tror du att förändringen i ditt sätt att hantera dina arbetsuppgifter har att göra med utbildningen?

INTRODUKTION FÖR NYANSTÄLLDA HANDLÄGGARE I STATLIG TJÄNST ³²

Utvärderade effekter

Statens kvalitets- och kompetensråd (KKR) fick i samband med att regeringens förvaltningspolitiska handlingsprogram introducerades år 2000 i uppdrag att utveckla ett generellt program för nyanställda handläggare vid de statliga myndigheterna.

Utvärderingen gällde dels introduktionsprogrammets nytta eller effekter för deltagarna i deras arbete, dels om utbildningen hade gett effekter på verksamheten inom de myndigheter som deltagarna representerade. Alla de fyra parallella pilotomgångarna av programmet som startade hösten 2001 utvärderades.



Figur 5: Upplägg B – Introduktionsprogram för nyanställda

Utvärderingssituationen och val av upplägg

Utvärderingen skulle belysa utbildningens effekt för enskilda individer och för den organisation, som deltagarna kom från. Effekterna gällde hur deltagarna använt sina kunskaper och utvärderingen genomfördes därför cirka ett halvår efter det att de avslutat sin utbildning. Utvärderingen utgick främst från de övergripande målen för programmet, statsförvaltningens grundläggande värden samt några av de speciella kännetecknen för programmet som uttrycks i KKR:s beskrivningar av utbildningen.

Utvärderarna kom in i efterhand och det var därför inte möjligt att ta fram en kontrollgrupp. Men även om utvärderarna hade varit med från starten är det, enligt dem själva, inte troligt att de hade valt ett sådant upplägg. Orsaken var att de ansåg att det skulle vara svårt att göra en relevant jämförelse mellan gruppen som gått kursen och en tänkt grupp som inte hade gått kursen. Frågan var vad man skulle mäta hos gruppen som inte gått kursen.

Utvärderingen genomfördes som en efterstudie baserad på enkäter till deltagare och berörda myndigheter samt intervjuer med programansvariga (skuggkontroll, se sidan 16). Utvärderarna valde enkätupplägget dels på grund av begränsade resurser, dels på grund av att det var enklare att genom enkätsvaren göra jämförelser mellan kursomgångarna.

Utvärderarna belyste effekterna för den enskilda individen genom en enkätundersökning till deltagarna med frågor om vilken nytta de haft i det egna arbetet av de moment som ingått i programmet.

För att ta reda på effekterna för hela verksamheten använde utvärderarna en enkät till utvalda personer på de berörda myndigheterna. I den tillfrågades de om vilken nytta myndigheten

haft av att man sänt deltagare till programmet. Utvärderarna intervjuade även de ansvariga programledarna med hjälp av ett halvstrukturerat intervjuformulär, som sändes ut i förväg.

Av flera skäl gjorde utvärderarna ett riktat urval till de myndigheter som sänt deltagare till pilotprogrammen och urvalet var därför inte slumpmässigt. Man bedömde också vilka personer och myndigheter som rimligtvis kunde besvara den typ av frågor som ställdes i enkäten. I övrigt eftersträvade man en blandning av stora och små myndigheter och av olika typer av myndigheter. Ett rimligt antagande var därför, enligt utvärderarna, att de som besvarat myndighetsenkäten inte var representativa för samtliga myndigheter.

Resultat

Deltagarna lämnade många positiva omdömen om programmet. Några av de övergripande positiva effekter som deltagarna lyfte fram var:

- Större säkerhet i arbetet.
- Inspiration i arbetet.
- Utbytet av att deltagarna representerade många olika myndigheter.

Utvärderarna konstaterade att det var mycket begärt att de som besvarat myndighetsenkäten skulle kunna uttala sig om effekter på verksamhetsnivå. Utvärderarna ansåg dock att om deltagarna upplevde att utbildningen varit nyttig för deras arbete fanns förutsättningar för effekter även för hela verksamheten. Men med tanke på att endast en, eller ett fåtal deltagare, från respektive myndighet deltog i pilotprogrammen kunde man knappast förvänta sig att det skulle gå att iaktta några effekter på verksamhetsnivå.

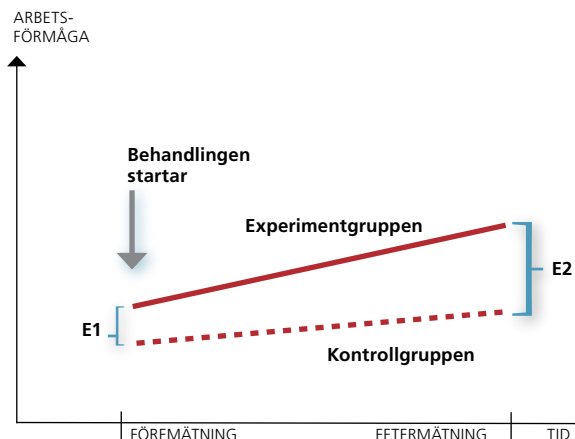
Upplägg C: Före-efterstudie med slumpmässigt valda grupper

När en utvärdering läggs upp som ett kontrollerat experiment väljer utvärderaren utifrån vissa kriterier ut en grupp individer eller enheter som ska ingå i studien. Därefter fördelas individerna eller enheterna slumpmässigt på en experimentgrupp och en kontrollgrupp. Medan experimentgruppen ”utsätts” för insatsen ska kontrollgruppen inte påverkas av densamma.

I det följande exemplet jämförs en ny behandlingsmetod för att rehabilitera sjukskrivna med en traditionell metod för att undersöka om den nya metoden leder till att arbetsförmågan ökar i större utsträckning jämfört med den traditionella metoden.

Om den undersökta populationen inte anses tillräckligt stor för att utvärderaren ska känna sig säker på att grupperna är lika i sin struktur från början bör utvärderaren göra en före-efterstudie. Effekten av verksamheten mäts då genom att minska skillnaden mellan kontrollgruppens och experimentgruppens genomsnittliga tillstånd efter insatsen med skillnaden före insatsen (E2 - E1). Ju mindre urvalet är, desto större är som regel behovet av att mäta tillståndet såväl före som efter insatsen.

Ibland finns det risk för att en föremätning i sig leder till en felaktig uppskattning av effekterna, eftersom de utvalda personerna då får veta att de deltar i ett experiment. I utvärderingslitteraturen kallas detta *Hawthorne-effekten* och gäller inte bara detta utan alla upplägg som innefattar en föremätning. I till exempel medicinska experiment använder man ofta så kallad blindning för att komma tillrätta med problemet. Det innebär att experimentgruppen får ta del av medicinen medan kontrollgruppen får en placebo, ”sockerpiller”, som är ett medel utan verkan. Blind-



Figur 6: Upplägg C – Rehabilitering av sjukskrivna

ningen blir ännu effektivare om man tillämpar så kallat dubbel-blindtest där inte heller den som genomför testet känner till vad de olika grupperna får.

REHABILITERING AV UNGA IMMIGRANTER I PRIMÄRVÅRDEN ³³

Utvärderad effekt

En utvärdering gjordes vid distriktsjukvården i Rinkeby, en stadsdel i Stockholm, av 92 personer i åldern 25–45, som varit sjukskrivna i mer än sex veckor på grund av långvarig, godartad smärta. Av Rinkebys invånare i åldersgruppen 16–45 år var 82 procent första eller andra generationens immigranter från mer än 70 länder.

Distriktssjukvårdens tidigare erfarenheter av rehabilitering av patienter med kronisk smärta var att den hade begränsade effekter. Detta kan ha berott på många faktorer, inte minst psykosocial stress och språksvårigheter. För immigranter

som lever under ogynnsamma omständigheter kan rehabilitering till och med motverka syftet.

Syftet med studien var att undersöka om strukturerade patient-doktor-samtal, som fokuserar på patientens smärtupplevelse, ökar välbefinnandet och arbetsförmågan i större utsträckning än traditionella metoder.

Utvärderingssituationen och val av upplägg

Eftersom utvärderingen gällde en ny behandlingsmetod krävdes enligt utvärderarna ett kontrollerat experiment. Deras största metodproblem var att bedöma den nödvändiga storleken på gruppen och att rekrytera tillräckligt många deltagare.

Undersökningen lades upp som ett fyra veckors experiment med två behandlingsgrupper. Alla personer mellan 24 och 45 år som varit sjukskrivna i mer än sex veckor blev erbjudna att ingå i ett rehabiliteringsprogram. De som antog erbjudandet delades sedan slumpmässigt upp i två grupper genom hemlig lottning. Den ena gruppen, grupp A, fick genomgå den nya behandlingsmetoden medan grupp B fick behandling på traditionellt sätt med läkarbesök utan inriktning på föreställningar om smärta.

Rehabiliteringsprogrammet bestod av två kliniska undersökningar som utfördes av läkare med 28 dagars mellanrum. Alla diagnoser och bedömningar gjordes av läkarna, baserat på deltagarnas presentation av sina problem och doktors undersökningsresultat.

Deltagarna fick peka ut de smärtsamma ställena på kroppen. Resultatet ritades in på en smärtteckning av en läkare. Smärtuppträdandet bedömdes sedan efter en bestämd skala.

Slutligen fick deltagarna och läkarna, oberoende av varandra, gradera funktionsförmågan i förhållande till arbetsuppgifterna. Graderingen

av den reducerade arbetsförmågan följde riktlinjerna för graderingen i sjukförsäkrans som gäller i Sverige (ingen, ringa, någon, måttlig, allvarlig). En skillnad på två steg i skalan eller mera bedömdes som en tydlig förändring.

Försöksgruppen fick komma tillbaka dagen efter den första undersökningen för en bandad intervju och för tre ytterligare veckobesök. Intervjun baserades på öppna frågor som koncentrerades på personens egen förklaring av hans eller hennes smärta. Under de följande mötena diskuterades den intervjuades uppfattning om smärta.

Studien innehöll också fysiska och psykiska undersökningar samt självbedömningar av arbetsförmåga och smärtupplevelse och de gjordes före och efter programmet. Data över sjukfrånvaron studerades vid tre- och åttamånadersuppföljningar och data hämtades från de lokala försäkringskassorna. När en deltagare hade börjat ett omskolningsprogram eller hade arbetat åtminstone deltid i 30 dagar eller mera räknades det som återgång till arbete.

Studien tog uppskattningsvis åtta månader att genomföra.

Resultat

Ett stort antal deltagare klev av programmet i förtid men totalt fullföljde 45 personer programmet, 23 i experimentgruppen och 22 i kontrollgruppen. Resultatet visade bland annat att smärtångesten hade minskat signifikant mer i experimentgruppen. En majoritet i båda grupperna rapporterade mycket mindre smärta efter programmet. Läkarnas bedömning av arbetsförmågan visade att fyra i experimentgruppen men ingen i kontrollgruppen hade tydligt förbättrad arbetsförmåga. Deltagarnas egna bedömningar visade att fem deltagare i experimentgruppen och två i kontrollgruppen hade tydligt förbätt-

rad arbetsförmåga. Två deltagare i kontrollgruppen hade försämrade subjektiv arbetsförmåga.

Sammantaget var antalet deltagare som förbättrades med avseende på en eller en kombination av variabler som depression, rapporterad smärtångest och självbedömd arbetsförmåga, arton stycken i experimentgruppen och sju i kontrollgruppen. Vid uppföljningen efter åtta månader visade det sig att det var färre personer i experimentgruppen som var sjukskrivna på fulltid jämfört med kontrollgruppen.

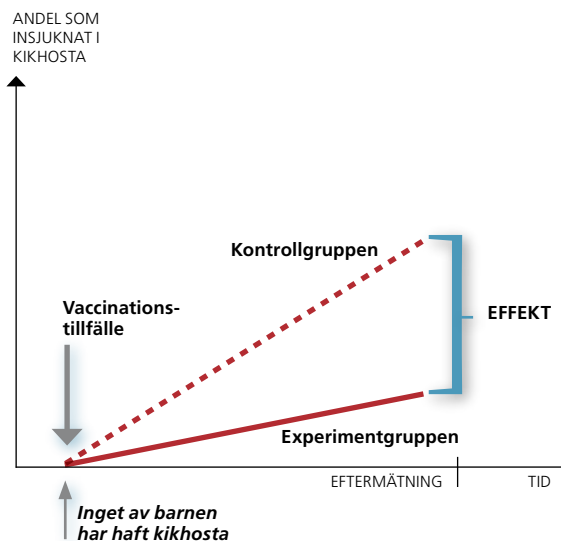
Utvärderarna anser att resultatet är uppmuntrande. Men vidare studier är nödvändiga för att säkerställa att effekterna var orsakade av den nya behandlingsmetoden.

Upplägg D: Efterstudie med slumpmässigt valda grupper

Upplägg D är, liksom upplägg C, ett kontrollerat experiment där utvärderaren utifrån vissa kriterier väljer ut en grupp individer eller enheter som ska ingå i studien. Därefter fördelas individerna eller enheterna slumpmässigt på en experimentgrupp och en kontrollgrupp.

Om urvalet är stort och slumpmässigt kan man anta att förutsättningarna för experimentrespektive kontrollgruppen är lika före insatsen och att det inte finns några systematiska skillnader mellan grupperna. Prövningen av kikhostevaccin används som exempel för det här upplägget. Där görs det slumpmässiga urvalet bland barn som inte haft kikhosta.

I detta fall får experimentgruppen kikhostevaccin medan kontrollgruppen får ett medel som är verkningslöst mot kikhosta. Effekten mäts sedan som skillnaden i andelen insjuknade barn i kikhosta för experimentgruppen respektive kontrollgruppen.



Figur 7: Upplägg D – Prövning av kikhostevaccin

PRÖVNING AV KIKHOSTEVACCIN ³⁴

Utvärderad effekt

Vaccination mot kikhosta ingick i det svenska vaccinationsprogrammet från 1950-talet till 1979, men upphörde för att det dåvarande kikhostevaccinet inte hade fullgod effekt och på grund av misstankar om biverkningar. Kikhostevaccinet bestod då av avdödade bakterier, så kallat helcellsvaccin. Kikhosta kom tillbaka under 1980-talet och det fanns därmed behov av att börja vaccinera igen. Nya så kallade acellulära kikhostevacciner, som består av en eller flera rena beståndsdelar från kikhostebakterien, togs fram och förväntades kunna ersätta de traditionella helcellsvaccinerna. Det var därför viktigt att testa och jämföra effekterna av de nya komponentvaccinerna med det gamla helcellsvaccinet.

Utvärderingssituationen och val av upplägg

Läkemedelsverket måste godkänna de vacciner som ska användas i det svenska vaccinationsprogrammet enligt författningarna om registrering av läkemedel. Kontrollerade experiment med slumpmässigt utvalda grupper är den standard som gäller. Därför valde utvärderaren ett så kallat dubbel-blind-placebo-kontrollerat experiment, där ett helcellsvaccin, ett femkomponentsvaccin och ett tvåkomponentsvaccin prövades. Dubbel-blind betyder att varken de som utsätts för eller de som genomför försöket vet vad individerna får i de olika grupperna. Placebo innebär att en kontrollgrupp får ett överksamt medel i stället för vaccinet.

Efter att ha fått tillstånd av föräldrarna gjordes ett slumpmässigt urval av 9 829 barn, som var två månader gamla och som inte tidigare haft kikhosta, och dessa fick ett av de tre vaccinerna eller placebo. Placebogruppen fick ett vaccin mot difteri och stelkramp. Vaccinet gavs vid tre tillfällen - vid två, fyra och sex månaders ålder.

Enligt utvärderarna var det ett etiskt problem att använda helcellsvaccin i studien eftersom man visste att det medförde större risk för biverkningar än de andra vaccinerna. Helcellsvaccinet betraktades dock internationellt som standard och kravet var att det skulle utvärderas.

Studien planerades under cirka ett år, 1991. Fältarbetet pågick från januari 1992 till januari 1995 och studien rapporterades i början av juli 1995.

Särskilt utbildade sjuksköterskor fick följa familjerna under två till tre år. Vaccinet levererades i identiska flaskor för att varken föräldrar eller sjuksköterskor skulle veta vilket vaccin varje barn fick. Biverkningar, som till exempel feber, lokala reaktioner och allvarliga symptom som kollaps, allergiska reaktioner eller kramper, undersöktes genom att sjuksköterskorna gjorde telefoninter-

vjuer en respektive fjorton dagar efter varje dos. Föräldrarna instruerades att ringa om barnet eller någon annan familjemedlem hostat i åtminstone sju dagar eller om föräldrarna misstänkte att någon i familjen fått kikhosta. I sådana fall besökte sjuksköterskan hemmet och tog prover.

Studien finansierades med ett anslag från National Institute of Health i USA på cirka 90 miljoner kronor. En andra prövning gjordes 1993–1996 med cirka 83 000 barn som ett komplement till den första och den prövningen blev billigare (40 miljoner kronor). Det berodde på att uppföljningen förenklades genom att utvärderarna via rapporter från landets laboratorier och enkäter till föräldrarna tog reda på vilka barn som fått kikhosta.

Resultat

Effektiviteten av ett vaccin mäts som minskningen av sjukdomen bland vaccinerade jämfört med ovaccinerade uttryckt i procent. Vid tidpunkten för utvärderingen hade Världshälsoorganisationen (WHO) skärpt kravet på effektiviteten av nya kikhostevacciner till minst 70 procent, vilket innebär att risken att få kikhosta för vaccinerade ska vara 70 procent lägre jämfört med ovaccinerade.

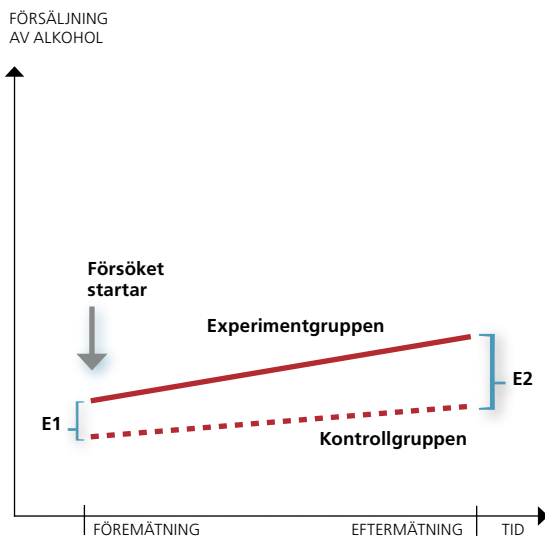
Femkomponentsvaccinets effektivitet beräknades till 85 procent, tvåkomponentsvaccinets till 58 procent och helcellsvaccinets till 48 procent. Både tvåkomponentsvaccinet och helcellsvaccinet hade alltså lägre effektivitet än den nya standarden krävde. Dessutom kunde utvärderarna konstatera att komponentvaccinerna hade mycket mindre bieffekter än helcellsvaccinet.

Baserat på resultaten från denna studie och en med den svenska studien samordnad italiensk studie beslutades att återinföra allmän kikhostevaccinering i Sverige i januari 1996.

Upplägg E: Före-efterstudie med matchade grupper

Ibland är det inte lämpligt att slumpmässigt fördela individerna eller enheterna i en experiment- respektive kontrollgrupp. I det följande exemplet, som gäller ett försök med lördagsöppna systembolag, fungerar det inte att slumpmässigt ta ut vissa län som får ta del av respektive inte får ta del av de nya öppettiderna. I detta fall får utvärderaren välja ut en grupp län som på försök får ha lördagsöppna systembolag (experimentgrupp) och en annan grupp län, som matchas med experimentgruppen och som inte får ha lördagsöppet (kontrollgrupp). Detta upplägg kallas i utvärderingslitteraturen för ”kvasiexperiment” (se sidan 17).³⁵

Med hjälp av statistik över försäljningen av alkohol går det att jämföra skillnader i försäljningen före och efter försöket. Effekten mäts som E2 - E1, precis som i upplägg C (se sidan 25).



Figur 8: Upplägg E – Lördagsöppna systembolag

EFFEKTER AV LÖRDAGSÖPPNA SYSTEMBOLAGSBUTIKER³⁶

Utvärderad effekt

Systembolaget startade försöksverksamhet med lördagsöppna butiker i en del av landet i februari 2000. En utvärdering gjordes av de tolv första månaderna och senare gjordes en uppdatering genom att ytterligare fem månader inkluderades. Lördagsöppna butiker infördes i hela landet i juli 2001.

Effekten som skulle utvärderas var i vilken mån lördagsöppet på systembolagen påverkar alkoholkonsumtionen mätt i termer av försäljning samt misshandel och rattfylleri.

Utvärderingssituationen och val av upplägg

Detta var ett experiment som utfördes i tre områden och utvärderarna använde sju kontrollområden. Deras uppfattning var att ett upplägg med kontrollområden skulle ge säkrare resultat än bara före-eftermätning i experimentområdena. Eftersom utvärderarna var inkopplade från början var det möjligt att lägga upp en sådan studie. De tre experimentområdena var Nord- sverige, Stockholms län och Skåne län och de sju kontrollområdena var Värmlands län, Örebro län, Västmanlands län, Östergötlands län, Kalmar län, Jönköpings län och Västra Götalands län. För att underlätta datainsamlingen valdes existerande administrativa enheter som experiment- och kontrollområden.

Det största metodproblemet gällde hur utvärderarna skulle hantera andra faktorer som kunde påverka försäljningen i kontrollområdena. Om experimentområdena var befolkningsmässigt små kunde den slumpmässiga variationen vara besvärande när de undersökta händelserna är få, som i det här fallet. Därför valdes få stora experiment- och kontrollområden som var be-

folkningsmässigt ungefär lika stora.

För att undvika att eventuella effekter skulle kunna bero på strukturella faktorer som till exempel urbaniseringsgrad och geografiskt läge valdes experimentområdena så att man fick variation i dessa avseenden. Om experimentområdena var lika skulle det vara svårt att bedöma om eventuella effekter berodde på dessa faktorer eller på att systembolagen hade lördagsöppet. Speciellt viktigt vid valet var att ta hänsyn till situationen i Danmark som redan hade lördagsöppna butiker.

Ytterligare en anledning till att utvärderarna valde få stora områden snarare än många små var att undvika den felkälla som ett eventuellt handelsläckage utgjorde. Med handelsläckage menas tillströmning av kunder som normalt handlar i något annat område. Av den anledningen såg man också till att inte välja experiment- och kontrollområden som gränsar till varandra.

Det kan vara svårt att avgöra om en förändring verkligen är en reell effekt eller om den ligger inom ramen för den normala variationen i form av långsiktiga trender, säsongvariation och mer tillfälliga, slumpmässiga fluktuationer. För att avgöra detta baserades analyserna på månadsdata för perioden fyra till fem år före och 17 månader efter försöket med lördagsöppna systembolag.

Effekten av ”lördagsöppet” skattades genom statistiska analyser av månadsdata som visade hur försäljning och skador utvecklats i experimentområdena jämfört med kontrollområdena under den nämnda perioden. När det gäller skador fokuserades analysen på utvecklingen under lördag-söndag, eftersom eventuella effekter antogs vara koncentrerade till dessa dagar.

Effekterna på misshandel och rattfylleri var

svåra att mäta. Det beror på att de visar på stora slumpmässiga variationer och påverkas av en rad andra faktorer. Det är också händelser som inte är så vanliga, vilket innebär att underlaget blir begränsat.

Studien tog cirka tre till fyra månader och kostade 200 000 kronor.

Resultat

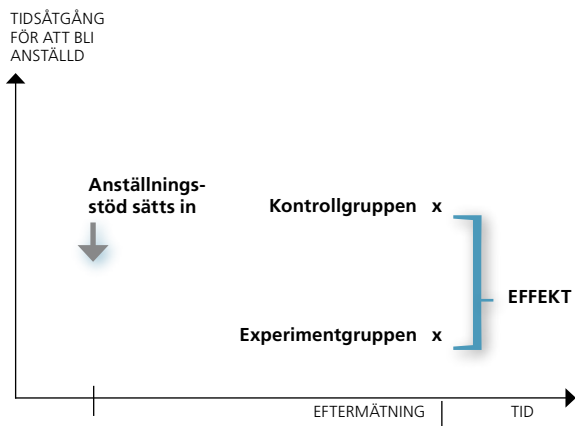
Försöket ledde till en statistiskt säkerställd försäljningsökning på drygt tre procent i experimentområdena. Effekterna på de undersökta skadeindikatorerna var mer osäkra och det beror på att dessa påverkas mer av andra faktorer och att det var svårt att isolera effekten av de lördagsöppna systembolagen. Detta i kombination med en kort observationsperiod gjorde, enligt utvärderarna, att de statistiska felmarginalerna var stora. Ökningen i rattfylleri i experimentområdena verkade, enligt dem själva, huvudsakligen bero på mer riktade poliskontroller.

Upplägg F: Efterstudie med matchade grupper

Ofta kopplas utvärderaren in efter det att insatsen har genomförts. Om inte statistik för tiden före insatsen finns tillgänglig är det då enbart möjligt att studera vad som hänt efter det att insatsen genomförts, det vill säga en efterstudie. Även i övrigt kan begränsningar i tid och resurser göra det nödvändigt för utvärderaren att begränsa utvärderingen till en efterstudie.

I detta upplägg bestämmer utvärderaren i efterhand vilka individer eller enheter som ska ingå i studien. Om det är möjligt kan en kontrollgrupp utses genom matchning.

I det följande exemplet undersöks effekterna av att långtidsarbetslösa får anställningsstöd för



Figur 9: Upplägg F – Anställningsstöd

att snabbare komma i arbete. En grupp långtidsarbetslösa som fått anställningsstöd väljs ut som experimentgrupp och denna grupp matchas sedan med långtidsarbetslösa som inte fått stöd men som i övrigt liknar experimentgruppen så mycket som möjligt. Båda grupperna är i utgångsläget arbetslösa och har därmed samma utgångsläge.

I detta fall mäts effekten genom att jämföra hur snabbt de två olika grupperna får anställning.

ANSTÄLLNINGSTÖD – EN VÄG FRÅN ARBETSLÖSHET TILL ARBETE? ³⁷

Utvärderad effekt

Anställningsstöd för långtidsarbetslösa har använts sedan januari 1998. Den som är långtidsarbetslös har en lägre sannolikhet att hitta ett arbete än en person som har varit arbetslös en kortare tid. En förklaring till detta kan vara att företagen uppfattar att det finns en risk att en långtidsarbetslös inte på ett fullgott sätt kan utföra sina arbetsuppgifter. Anställningsstöd kan ses som ett sätt att

subventionera företagen för den risk de anser sig ta genom att ändå anställa en långtidsarbetslös. Förhoppningen är att den som blir anställd med anställningsstöd senare ska ha lättare att hitta ett ordinarie jobb.

Den studerade effekten gällde hur mycket snabbare de långtidsarbetslösa personer som blivit anställda med anställningsstöd fick ett arbete jämfört med om de inte fick ta del av stödet.

Utvärderingssituationen och val av upplägg
Utvärderarna kom in i efterhand och utvärderingen måste därför av naturliga skäl bli en efterstudie. Trots detta var det möjligt att matcha individer som fått olika behandling mot varandra och undersöka hur snabbt de kom ut i arbete. Därför är detta ett exempel på ett i efterhand konstruerat kvasiexperiment.

Utvärderarna konstaterade att det fanns systematiska skillnader i ålder, kön etc. mellan de personer som fått anställningsstöd och arbetsökande som inte fått stöd. Eftersom grupperna inte kunde väljas slumpmässigt använde utvärderarna i stället matchningsmetoden. Detta var möjligt eftersom tillgängliga data om de arbetsökande innehöll relativt rik information.

Till kontrollgruppen valdes personer som var så lika dem som fick anställningsstöd som möjligt för att kunna representera vad som skulle ha hänt med deltagarna i stödet om de inte hade fått anställningsstöd. Varje enskild deltagare i anställningsstödet matchades med en icke-deltagare som i alla viktiga avseenden (till exempel ålder, kön, utbildning och inskrivningsperiod) var lika. Det innebar att 60 procent av deltagarna i anställningsstödet fick avlägsnas i underlaget eftersom ingen matchande individ fanns i kontrollgruppen.

Utvärderarna ansåg att det fanns risk för att

resultaten inte skulle gå att generalisera. Det berodde på att den grupp som matchats var liten och att det därför fanns risk för att skillnader i resultatet uppstått av en slump och inte på grund av att den ena gruppen utsatts för en åtgärd. För att undersöka betydelsen av detta användes också andra matchningsmetoder där den matchade gruppen gjordes ungefär dubbelt så stor. Resultaten blev i stort sett identiska med denna metod som med den första.

I utvärderingen ingick personer som registrerat sig som arbetssökande sedan augusti 1991. De uppgifter som användes sträcker sig från 1991 till och med oktober 2002.

Uppskattningsvis arbetade tre personer med utvärderingen i cirka sex månader utspridda över cirka två år.

Resultat

Resultatet visade att anställning med anställningsstöd påskyndade återgången i arbete. Resultatet visade också att anställningsstödet bidrog till en signifikant förkortad arbetslöshetstid för deltagarna, i genomsnitt åtta månader.

Upplägg G: Teoretiskt upplägg

Ibland kan det vara svårt för utvärderaren att empiriskt mäta effekterna av en statlig insats. Det kan bero på att utvärderaren kommer in för sent, att datainsamlingen skulle kosta för mycket eller att det inte är möjligt att ta fram data över effekter från tidpunkten för utvärderingen. Det senare kan bero på att effekterna förväntas uppstå först på längre sikt. Det kan också vara så att den statliga insatsen är alltför heterogen och består av för många olika typer av delinsatser för att uppläggen A–F ska fungera. Ett exempel på detta är lokala utvecklingsprojekt som brukar innefatta ett

stort antal prestationer. I dessa fall kan det vara intressant för utvärderaren att utifrån en verksamhetslogik teoretiskt bedöma om insatserna kan tänkas ha gett effekter.

Det teoretiska upplägget bygger inte på att det måste finnas en explicit nedskreven verksamhetslogik. Men för att en verksamhet ska bedömas ha gett effekter måste de personer som arbetar med verksamheten ha en idé om hur insatsen ska kunna tänkas ge effekt och den idén måste vara hållbar. Resultatet av en teorisk utvärdering har vanligen inte samma tyngd som en väl utförd empiriskt baserad utvärdering och resultatet bör därför presenteras i termer av till exempel att verksamheten ”troligen inte kan ha haft någon effekt”.

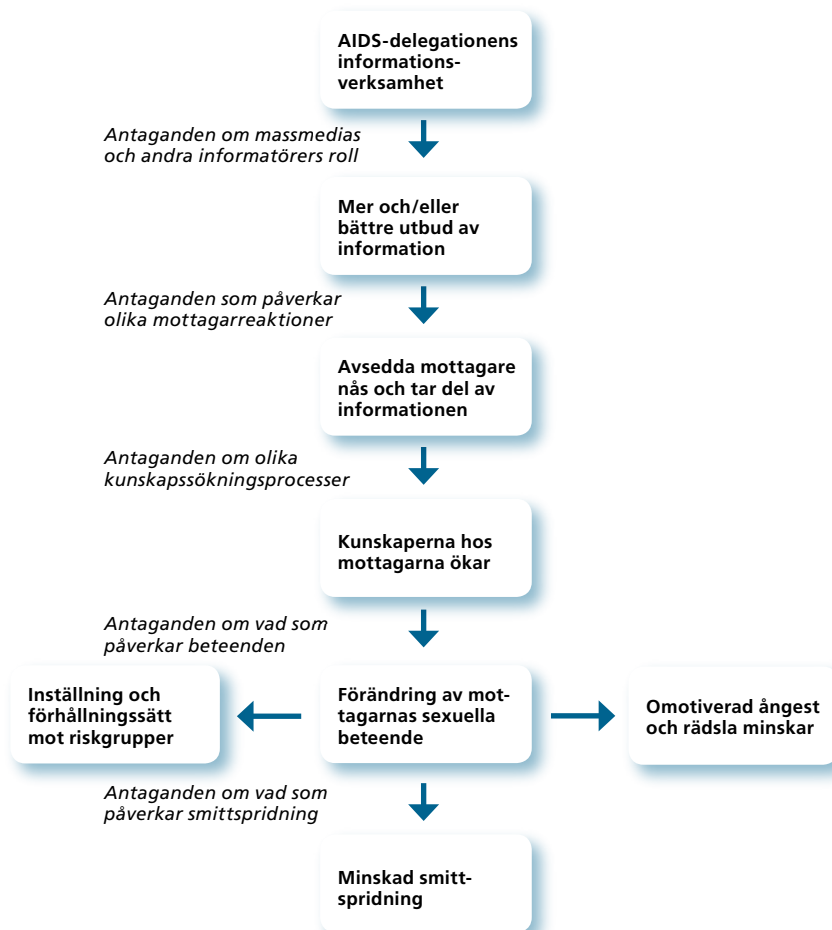
INFORMATIONSKAMPANJEN OM HIV/AIDS – SAMHÄLLSINFORMATION SOM STYRMEDEL ³⁸

Utvärderad effekt

Riksrevisionsverket (RRV) fick år 1988 regeringens uppdrag att utvärdera om de informationsinsatser i fråga om HIV/AIDS som fram till dess genomförts hade varit ändamålsenliga med hänsyn till målsättningen. RRV skulle också lägga fram ett förslag om en mer omfattande utvärdering av det samlade arbetet med informationen om HIV/AIDS.

Det primära målet för informationsverksamheten var att minska smittspridningen medan de sekundära målen var att minska fördomarna kring HIV/AIDS samt att dämpa ångesten för HIV/AIDS hos allmänheten.

Analysen gick ut på att undersöka om uppläggningsen av AIDS-delegationens informationsarbete byggde på bästa möjliga kunskap om informationskampanjer som fanns tillgänglig vid beslutstillfället.



Figur 10: Verksamhetslogik - HIV/AIDS-kampanj

Utvärderingssituationen och val av upplägg

Ur metodologisk synvinkel är det svårt att fastställa vilken effekt en informationssatsning haft på spridningen av HIV/AIDS. Det beror på flera faktorer. För det första är det många andra aktörer och faktorer som påverkar spridningen. För det andra är effekterna långsiktiga, det vill säga det kan ta lång tid innan en person som smittats får veta att han eller hon fått HIV och sedan utvecklat AIDS. Dessutom kan insatserna vara effektiva om de bedöms utifrån ett av målen men ineffektiva om de bedöms utifrån ett annat.

Utvärderingen lades upp som en teoretisk utvärdering där utvärderarna skulle pröva den teoretiska hållbarheten i den förda verksamheten.

För att göra detta började utvärderarna med att ta fram en verksamhetslogik för AIDS-delegationens informationsverksamhet (se figur 10).

Kedjan redovisar ett antal orsakssamband som måste vara till hands för att de olika effekterna skulle kunna uppträda och vilka antaganden som de bygger på. Som ett exempel nämns att antagandet att förändringar i det sexuella beteendet verkligen påverkar smittspridningen förutsätter att de bakomliggande antagandena om hur smittan sprids är riktiga.

För att kunna bedöma informationsverksamhetens tänkbara effekter utifrån den kunskap som då fanns om hur människor påverkas, gjorde utvärderarna en litteraturstudie. Med hjälp av

denna ville de få en bild av den senare samhällsvetenskapliga forskningens syn på förutsättningarna att lyckas med informationskampanjer. Utvärderarna tog också del av erfarenheter bland praktiskt verksamma samhällsinformatörer och personer som hade till uppgift att informera om sex- och samlevnadsfrågor.

Studien tog cirka tre till fyra månader att genomföra.

Resultat

RRV ansåg det rimligt att AIDS-delegationens masskommunikation kunde fylla en funktion när det gällde att påverka undersökningsgruppens beteende, fördomar och kanske också oro och ångest. När det gällde informationens innehåll ansåg RRV att endast en mycket liten del av informationen hade varit inriktad på konkreta beteenden som mer påtagligt ökar risken för smitta. Positivt var att delegationen undvikit skrämselfpropaganda som få forskare anser säkrar avskräckningseffekter.

Upplägg H: Flera upplägg inom samma utvärdering

I praktiken behöver utvärderaren ofta utnyttja olika upplägg för en utvärdering. Speciellt gäller detta då utvärderingssituationen är komplex:

- Insatsen är omfattande och producerar ett stort antal prestationer som dessutom är av olika kvalitet.
- Effekterna är uttryckta i många dimensioner.
- Effekterna gäller flera olika undersökningsgrupper.
- Många andra faktorer påverkar utfallet.

form av utvärdering med flera olika upplägg. De består av en litteraturgenomgång med sammanställning och bedömning av ett antal redan gjorda utvärderingar. Dessa kan, men behöver inte, vara genomförda utifrån olika upplägg. Metaanalyser är vanliga inom bland annat bistånds- och det medicinska området.³⁹

Nedan följer två exempel på utvärderingar med en kombination av flera upplägg. Det första exemplet bygger på att det fanns två undersökningsgrupper och det andra på att det fanns både flera prestationer och flera undersökningsgrupper.

TRANSPORTBIDRAGET – EN ANALYS AV MOTIV, EFFEKTER OCH ALTERNATIV ⁴⁰

Utvärderad effekt

Transportbidraget infördes 1971 och fördelas varje år till industriföretag i de fyra nordligaste länen med syfte att kompensera för företagens långa transportsträckor och främja ökad förädling inom produktionen.

Utvärderingen gällde transportbidragets betydelse för företagen respektive för regionen (effekter på vidareförädling, på transportköparnas prisnivåer, miljöeffekter, jämställdhetseffekter samt effekter på den regionala utvecklingen).

Utvärderingssituationen och val av upplägg

Utvärderingen arbetade med två utvärderingsfrågor, företagens respektive regionens utveckling. För att bedöma effekterna användes två olika upplägg, efterstudie utan kontrollgrupp respektive före-efterstudie.

Effekter för företagen

Eftersom det var svårt att hitta en kontrollgrupp av företag som inte fått bidrag, men som i övrigt

liknade dem som fått bidrag, använde utvärderarna i stället en efterstudie med skuggkontroll (upplägg B, se sidan 23) som grund för bedömning av transportbidragets effekter. Datainsamlingen genomfördes i form av tjugoåtta telefonintervjuer med industriföretag lokaliserade i bidragsområdet som fick bidrag. Urvalet av intervjupersoner och företag gjordes utifrån tre kriterier:

- Storlek på det erhållna bidraget.
- Företagets näringsgren.
- Företagets lokalisering.

Syftet var att få god spridning såväl geografiskt som vad gäller näringsgren.

För att öka säkerheten i resultatet intervjuade utvärderarna också fem transportföretag inom bidragsområdet. Dessutom intervjuades ett industriföretag i samma region som inte fått bidrag samt fyra industriföretag lokaliserade i andra regioner av Sverige som därför inte kunde få transportbidrag. Industriföretagen var delvis matchade i förhållande till de företag som fått bidrag.

Effekter för regionen

Effekterna för regionen är betydligt svårare att bedöma än effekterna för de enskilda företagen. Det beror på att många andra faktorer och aktörer påverkar den regionala utvecklingen. För att belysa transportbidragets betydelse för den regionala utvecklingen kompletterades intervjuerna ovan med analyser av statistik ur Nuteks databas över transportbidraget samt med regional statistik från Statistiska Centralbyrån (SCB). Statistiken gjorde det möjligt att följa utvecklingen från tiden innan företagen fick transportbidrag. Upplägget för att utvärdera effekterna för regionen var alltså en före-efterstudie (upplägg A, se sidan 20).

Utvärderingen byggde på statistik för bland annat produktions- och exportvärde för industrin i regionen och statistik om befolkning och sysselsättning före och efter det att bidraget införts. Utvärderarna jämförde också bidragsdata med industriomsättning, förädlingsvärde, befolkning och sysselsatta på en nedbruten geografisk nivå. Där så var möjligt och relevant konsuppdelades statistiken.

Resultat

Transportbidraget hade enligt utvärderingen stor betydelse för cirka hälften av de intervjuade företagen, särskilt de som hade långa transporter och som låg i inlandet. Dessa företag hade dessutom stor betydelse för de orter där de var verksamma. Bidraget hade därmed stor betydelse för den regionala utvecklingen, särskilt i inlandet.

SKOLVERKETS ARBETE FÖR ATT ÖKA KVALITETEN INOM UTBILDNINGEN ⁴¹

Utvärderade effekter

Skolverket fick år 1999 av regeringen i uppdrag att för 52 miljoner kronor under en treårsperiod speciellt arbeta för att öka kvaliteten inom yrkesutbildningen.

Effekten som skulle utvärderas var projektets bidrag till ”ökad kvalitet” inom yrkesutbildningen.

Utvärderingssituationen och val av upplägg

Uppdraget omfattade en rad olika prestationer. Framför allt bestod projektet av:

- Produktion av böcker om yrkesutbildning.
- Universitetskurser för lärare i yrkesutbildningspedagogik.
- Regionala konferenser inom yrkesutbildning.

- Stöd till forskarnätverk inom yrkesutbildning.
- Utveckling av speciella kursplaner inom yrkesutbildning.
- Stöd till lokala utvecklingsprojekt.
- Beskrivning av ett antal ”lärande exempel”.

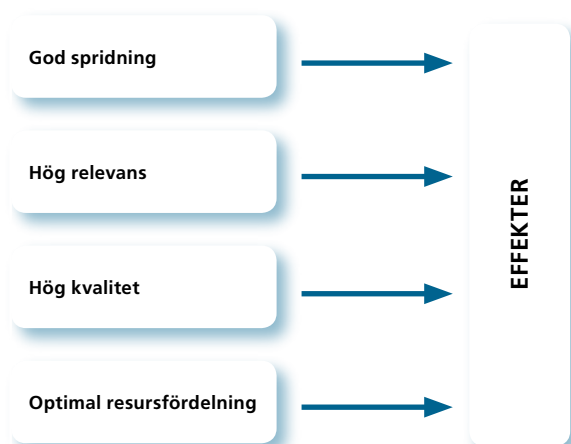
Prestationerna riktade sig till olika målgrupper: lärare/rektorer, forskare och elever inom gymnasiets yrkesinriktade program. Det var en av anledningarna till att det inte gick att etablera kontrollgrupper. Ett annat problem var att utvärderarna kom in i efterhand och inte hade möjlighet att planera utvärderingen i förtid och göra före-efterstudier. Sammanfattningsvis var utvärderingssituationen svår och gjorde det omöjligt att utgå från ett enda upplägg.

”Dataläget” var gott på så sätt att det fanns ett antal utvärderingar för de olika prestationerna som redan var gjorda av interna utvärderare. Dessa data hade dock samlats in på olika sätt, vilket till exempel innebar att enkätfrågor om effekterna hade formulerats på olika sätt i olika studier. Detta gjorde det svårt att sammanställa resultaten till en samlad bild. Dessutom var det svårt att nå ut till alla de olika målgrupper som nåtts av Skolverkets satsning om utvärderarna ville samla in mer data.

Upplägget byggde på en enklare typ av verksamhetslogik och beskrivs i fig. 11.

För att uppnå stora effekter måste Skolverket enligt modellen sprida sitt budskap effektivt, göra det med ett relevant budskap, vars kvalitet uppskattas av mottagarna. Resursfördelningen mellan olika aktiviteter borde dessutom ge prioritet till aktiviteter som bäst leder till stora effekter.

Baserat på det material som projektet redan samlat in och på enklare egen datainsamling fokuserade utvärderarna på specifika prestationer



Figur 11: Modell för utvärdering av Skolverkets regeringsuppdrag

och analyserade effekterna av dem.

För vissa utvecklingsprojekt gjordes en efterstudie utan kontrollgrupp (upplägg B, se sidan 23). I den tillfrågades de lokala projektledarna om hur verksamheten utvecklats och vilken betydelse som Skolverkets utvecklingsprojekt haft för den ordinarie verksamhetens utveckling (skuggkontroll, se sidan 16).

För de tio universitetskurser som anordnades om yrkesutbildningsfrågor gjordes en teoretisk effektutvärdering (upplägg G, se sidan 32). Utifrån en analys av kursplaner, litteraturlistor etc. bedömde utvärderaren tillsammans med en fackexpert inom yrkespedagogik om kurserna kunde tänkas ha givit de åsyftade effekterna.

De iakttagelser som utvärderarna gjorde om faktorerna ovan (spridning, kvalitet, relevans och resursfördelning) systematiserades och analyserades. Sammantaget kunde dessa observationer inte ge ett enkelt svar på frågan om effekterna av

Skolverkets verksamhet men de kunde användas för en diskussion och bedömning av effekterna av verksamheten.

Om utvärderarna hade varit med från början av projektet hade de, enligt dem själva, haft möjlighet att se till att alla projektets aktiviteter utvärderats på ett standardiserat sätt, till exempel med enkäter som bygger på exakt samma frågor om effekterna. På så sätt skulle utvärderingen ha kunnat läggas upp som en efterstudie där resultat från olika aktiviteter kunde ha sammanställts.

Studien kostade 600 000 kronor och tog 1,5 år. Den bestod av två delar: en del som handlade om effekter och en som handlade om Skolverkets handläggning. Effekttvärderingen kostade cirka 200 000 kronor och genomfördes på cirka två månader.

Resultat

Utvärderarna ansåg inte att det gick att sammanfatta effekterna i ett enda mått. De bedömde att Skolverket nått ut till ett relativt stort antal personer som arbetar med skolfrågor men ansåg att det varit svårt för projektet att nå viktiga målgrupper, till exempel skolledare och ”arbetslivet” samt vissa delar av inlandet och glesbygden. Målgrupperna uppskattade genomgående aktiviteternas kvalitet. Utvärderarna satte frågetecken för relevansen av ett antal av projektens aktiviteter och det gällde bland annat vissa högskolekurser och utvecklingsprojekt. Det var svårt att bedöma om resursfördelningen varit optimal men satsningen på högskolekurser skulle enligt utvärderarna möjligen ha kunnat begränsas.



Att välja upplägg

Beskriv den åtgärd som ska utvärderas

Beskriv typen av effekter

Arbeta med verksamhetslogik

Bedöm vilka upplägg som kan vara lämpliga

Bedöm resursbehovet för alternativen

Lägg fast kravet på säkerställda resultat

**Gör en sammanvägd bedömning
och välj upplägg**

SOM VI REDAN BESKREV i inledningen kan utvärderingsprocessen beskrivas i ett antal faser. En av dessa faser är planeringen av utvärderingen.

Planeringsfasen omfattar grundläggande överväganden kring utvärderingen som syfte och användning, metodologiska överväganden som urvalsmetoder, datainsamlingsmetoder och analysmetoder, men också administrativa planeringsfrågor som budget och tidplan. Valet av upplägg är en av de viktigaste frågorna i planeringsfasen för en effektutvärdering.

Valet av upplägg kan beskrivas i sju steg.⁴²

1. Beskriv den åtgärd som ska utvärderas.
2. Beskriv typen av effekter.
3. Arbeta fram en verksamhetslogik.
4. Bedöm vilka upplägg som kan vara lämpliga.
5. Bedöm resursåtgången för alternativen.
6. Lägg fast kravet på säkerställda resultat.
7. Gör en sammanvägd bedömning och välj upplägg.

Kapitlet bygger på de utvärderingar som använts som exempel i kapitel 2.

Beskriv den åtgärd som ska utvärderas

Det är avgörande för utvärderaren att få klart för sig vad insatsen som ska utvärderas består av. I detta steg beskriver därför utvärderaren först åtgärden eller åtgärderna (prestationerna) som ska utvärderas. En insats kan vara mer eller mindre omfattande och tydligt avgränsad. *Skolverkets kvalitetsarbete inom yrkesutbildningen* består till exempel av ett antal olika prestationer. Åtgärden kan också vara mer eller mindre standardiserad och utföras på ett eller många olika sätt. I exemplet om utvärdering av en ny typ av behandling för *rehabilitering av unga immigranter* krävs att

behandlingen utförs på ett standardiserat sätt för att det ska vara möjligt att utvärdera effekterna av den.

Målgruppen

Under detta steg identifierar utvärderaren också vilka som utgör målgruppen för den studerade åtgärden. Ibland rör det sig om en, ibland flera målgrupper och dessa kan bestå av till exempel individer eller organisationer. Generellt sett gäller att ju fler grupper åtgärden omfattar, desto mera tidskrävande är det att utvärdera effekterna och desto svårare blir det att använda ett enda upplägg för studien. I exemplet om *Skolverkets kvalitetsarbete inom yrkesutbildningen* riktades insatserna till ett flertal målgrupper (lärare, elever, forskare m.fl.) och det gjorde det svårare för utvärderaren att bedöma de totala effekterna av Skolverkets insatser. Även *informationskampanjen* om HIV/AIDS har olika målgrupper som måste adresseras på olika sätt i kampanjen.

Beskriv typen av effekter

När utvärderaren ska undersöka om en åtgärd har lett till uppställda mål talar man om en målutvärdering. Om målen inte är tydligt preciserade behöver utvärderaren operationalisera dem och bestämma indikatorer på måluppfyllelse. Exempel på detta är utvärderingarna av *transportbidraget* och *introduktionsprogram för nyanställda*.

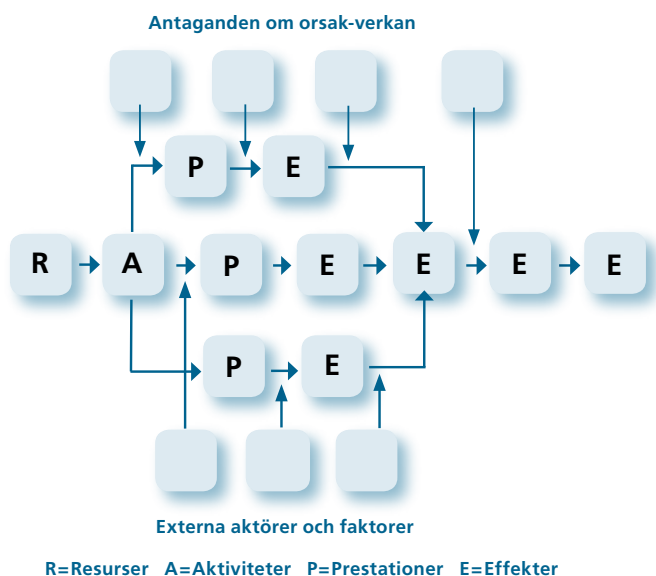
Det kan också vara viktigt att identifiera möjliga positiva och negativa bieffekter av åtgärden. Effekterna av transportstödet kan till exempel leda till, inte bara effekter för företagen som fått bidrag, utan också till en rad bieffekter i den berörda regionen. I det här fallet förväntades de vara positiva.

Effekterna kan vara en- eller flerdimensionella.

ENKEL



KOMPLICERAD



Figur 12: Enkel och komplicerad verksamhetslogik

Ju fler dimensioner som ska beaktas, desto mer komplex blir effektutvärderingen. Utvärderingen av *kameraövervakningen* blir enklare om enbart fokus ligger på antalet begångna brott och inte dessutom på medborgarnas trygghet.

En annan viktig aspekt är tidsdimensionen. Ju mer långsiktiga effekterna är desto längre tid måste utvärderaren vänta för att kunna avläsa dem. Utvärderingens upplägg är beroende av när utvärderaren måste lämna sin bedömning, innan, direkt efter eller långt efter det att den utvärderade verksamheten slutförts. Om

effekterna är långsiktiga och utvärderingen måste levereras i ett tidigt skede går det inte att bedöma effekterna. Ett exempel på detta är utvärderingen av *informationskampanjen om HIV/AIDS* som i stället fick inrikta sig på en teoretisk bedömning av om informationskampanjen var upplagd på ett ändamålsenligt sätt så att den kunde tänkas leda till någon effekt.

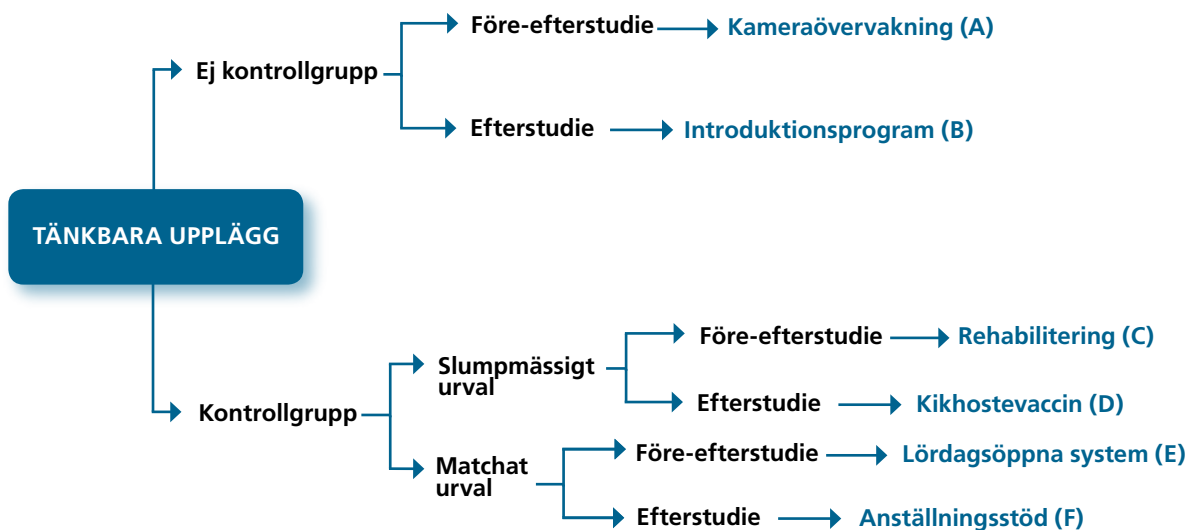
Arbeta med verksamhetslogik

Ett effektivt verktyg för att beskriva och analysera åtgärden, målgruppen och vilka effekter åtgärden kan leda till är verksamhetslogiken (se sidan 9). Den kan hjälpa till att identifiera vilka effekter som är möjliga att utvärdera, vilka målgrupperna är, vilka andra aktörer eller faktorer som kan tänkas påverka de önskade effekterna, negativa eller positiva bieffekter, indikatorer på effekterna och antaganden om mekanismerna bakom förväntade orsak-verkan-samband. Den tydliggör också utvärderingssituationens komplexitet.

En enkel tumregel är att ju fler boxar och pilar verksamhetslogiken omfattar, desto mer komplex är utvärderingssituationen och desto svårare är det att utvärdera effekterna (se figur 12). Ju fler prestationer, ju fler antaganden som måste göras och ju fler externa aktörer och faktorer som påverkar utfallet desto svårare blir utvärderingssituationen. I vissa fall kan till och med verksamhetslogiken visa på en så komplex situation att utvärderaren måste fundera på om det verkligen går att utvärdera effekterna.

Bedöm vilka upplägg som kan vara lämpliga

Under detta steg bedömer utvärderaren vilka upplägg som är möjliga att genomföra för att utvärdera effekterna.



Figur 13: Beslutsträd för utvärderare

Utifrån de upplägg som vi beskrivit i kapitel 2 kan valsituationen illustreras enligt figur 13. Utvärderaren kan välja mellan att arbeta med eller utan kontrollgrupp, i det senare fallet med slumpmässigt urval eller matchat urval samt med före-efterstudie eller bara efterstudie. Av pedagogiska skäl finns utvärderingarna som vi använder som exempel i kapitel 2 inlagda i figuren.

Kontrollgrupp eller inte

I vissa situationer krävs att en åtgärd först testas i ett kontrollerat experiment innan den tillämpas i full skala. Det gäller när utvärderaren med stor säkerhet måste kunna uttala sig om de uppmätta effekterna, som inom den medicinska sektorn. En förutsättning är att det går att fördela grupperna slumpmässigt och att utvärderaren kan

förändra åtgärdens storlek och karaktär. Studien avseende *rehabilitering av unga immigranter* och prövningen av *kikhostevaccin* illustrerar den här typen av utvärderingssituationer.

I utvärderingen av *anställningsstöd* var syftet att undersöka om anställningsstöd ledde till snabbare anställning. Här gick det i efterhand att matcha gruppen långtidsarbetslösa som beviljats anställningsstöd med sådana som inte fått det och på så sätt i efterhand konstruera ett kvasiexperiment.

I försöket med *lördagsöppna systembolag* var utgångspunkten att försöket skulle genomföras i ett antal län för att studera effekten på alkoholkonsumtionen och på förekomsten av misshandel och rattfylleri i länet. För att få ett säkrare resultat än vad enbart en före-eftermätning i

försökslänen skulle ge, valdes ett antal kontrollän ut genom matchning. En jämförelse kunde sedan göras mellan statistik över försäljningen och över förekomsten av misshandel och rattffylleri.

I utvärderingen av *kameraövervakningens* effekter på brottsligheten och konsekvenser för tryggheten och den personliga integriteten vid Möllevångstorget i Malmö hade det varit önskvärt, men var inte möjligt, att hitta en kontrollgrupp i form av ett likvärdigt torg. I stället använde utvärderarna generisk kontroll genom att jämföra brottsutvecklingen vid Möllevångstorget med den i Malmös centrala stadsdelar och kommunen som helhet.

I studien av *transportbidragets* effekter på företag och region var det inte möjligt att hitta en kontrollgrupp av företag som inte fått bidrag och som i övrigt liknade de som fått bidrag. I stället använde utvärderarna skuggkontroll för att undersöka effekterna för företagen och en före-efterstudie för effekterna på regionen.

Urval

Valet av vilka och hur många som ska ingå i en utvärdering är avgörande för hur säkerställda resultaten ska kunna bli. De urval som refereras i våra olika exempel går från ett mycket litet antal personer upp till 90 000 personer i testningen av kikhostevaccinerna.

Valet av vilka som ska ingå i en utvärdering görs ibland slumpmässigt för att undvika systematiska fel. Ibland krävs emellertid att utvärderaren gör ett strategiskt urval utifrån specifika kriterier. Det kan till exempel vara viktigt att få en storleksmässig eller geografisk spridning i urvalet, om det är faktorer som kan påverka resultatet. I utvärderingen av *transportbidraget* gjordes urvalet av företag som skulle intervjuas med syfte att få god spridning både geografiskt

och med avseende på näringsgren. Ett annat exempel på strategiskt urval är försöket med *lördagsöppna systembolag*. I det fallet såg utvärderarna till exempel till att inte välja experiment- och kontrollområden som gränsar till varandra. Anledningen var att man ville undvika så kallat handelsläckage, som innebär tillströmning av kunder som normalt handlar i något annat område.

Ibland väljs helt enkelt individer eller enheter som finns lätt till hands och som får representera hela populationen, så kallat bekvämlighetsurval⁴³. Exempel på ett sådant urval är valet av besökare på Möllevångstorget som skulle intervjuas i undersökningen av *kameraövervakning*. Där valde utvärderarna att intervju besökare som råkade finnas på torget under tiden som utvärderingen genomfördes.

Vid kontrollerade experiment görs urvalet i två steg. I det första steget gäller det att välja individer eller enheter som ska ingå i experimentet utifrån vissa kriterier. I det andra steget fördelas dessa slumpmässigt på en experimentgrupp och en kontrollgrupp.

Före-efterstudie eller inte

Upplägget av en utvärdering påverkas av om det går att få fram data om tillståndet innan åtgärden genomförs. En viktig faktor är om utvärderaren kan vara med från starten av en insats, eller ännu hellre redan före starten.

Om utvärderaren är med från början och det går att få fram data före åtgärden är det möjligt att göra ett kontrollerat experiment eller ett kvasiexperiment. Effekten kan då utläsas som skillnaden mellan grupperna efter åtgärden (efterstudie).

Om det finns risk för att experiment- och kontrollgruppen skiljer sig åt före åtgärden behöver

utvärderaren jämföra skillnaden mellan förändringen i experimentgruppen och i kontrollgruppen. Utvärderaren justerar på så sätt effekten för skillnader i utgångsläget. Exempel på experiment med före-efterstudie är *lördagsöppna systembolag* och *rehabilitering av unga immigranter*.

När det inte går att etablera en kontrollgrupp kan utvärderaren i stället göra en före-efterstudie av gruppen som utsätts för åtgärden och helst komplettera utvärderingen med andra jämförelser, som till exempel generisk kontroll. Ett exempel på en sådan studie är *kameraövervakningen*.

Före-efterstudier utan kontrollgrupp är ofta ett lämpligt upplägg när risken för att andra faktorer kan ha påverkat utfallet är liten, till exempel då effekten är nära knuten till åtgärden. Ett område där denna typ av upplägg är vanligt är trafikområdet och effekter av åtgärder på trafiksäkerheten.

Om utvärderaren kopplas in i efterhand går det ofta inte att göra en före-efterstudie. I utvärderingen av *introduktionsprogrammet för nyanställda* handläggare kom utvärderarna in i efterhand och kunde enbart göra en efterstudie. De använde sig då av skuggkontroll baserad på enkäter till deltagare och myndigheter samt intervjuer med programansvariga.

Även om utvärderaren kopplas in i efterhand kan det vara möjligt att genomföra en före-efterstudie. Det gäller om det redan finns data tillgängliga som har insamlats för andra ändamål eller som en del av en allmän uppföljning. I utvärderingen av *transportbidraget* kom utvärderarna in i efterhand. I den del som handlade om effekter på regionen kunde de ändå göra en före-efterstudie. Detta berodde på att det redan fanns regional näringslivsstatistik och statistik om befolkning och sysselsättning för tiden före det att insatsen genomfördes.

Bedöm resursåtgången för alternativen

Utvärderarens resurser kan mätas på olika sätt, till exempel i form av:

- Kompetens
- Tid
- Budget

Givet vilken frågeställning eller typ av effekter som ska utvärderas måste kompetens, tid och budget vägas in i bedömningen av vilket upplägg som är lämpligt. Det gäller att hitta den mest kostnadseffektiva lösningen.

Kompetens

Ju högre kompetens utvärderaren själv besitter eller har möjlighet att köpa in, desto mer avancerade metoder kan utvärderingen bygga på. Vid mer komplicerade utvärderingar är behovet av metodkunskap, analytisk förmåga och förmåga att tolka befintliga data extra viktigt. En generell erfarenhet är dock att de flesta effektutvärderingar framför allt kräver eftertanke och systematiskt tänkande men inte oerhört avancerade metodkunskaper.

Tid

Tiden som det tagit att producera de utvärderingar som presenteras här spänner från några månader upp till flera år. Skillnaden beror framför allt på effekternas karaktär och i vilket skede som utvärderingen genomförs. Att utvärdera kortsiktiga effekter med efterstudier där utvärderaren kommer in i efterhand kan gå snabbt. Att däremot, som i fallet med *prövningen av vacciner*, lägga upp ett experiment där effekterna är långsiktiga betyder att utvärderingen kommer att ta tid. Även i fallet med *lördagsöppna systembolag* var det viktigt att basera analysen på data för en lång period, både före och efter, för att

kunna avgöra om en förändring verkligen var en reell effekt eller om den låg inom ramen för den normala variationen i form av långsiktiga trender, säsongsvariation eller mer tillfälliga, slumpmässiga fluktuationer.

Budget

Våra exempel redovisar helt olika förutsättningar vad gäller kostnader. Budgetarna varierar från cirka 200 000 kr för utvärdering av *lördagsöppna systembolag* respektive *Skolverkets kvalitetsarbete* inom yrkesutbildningen med upp till 90 miljoner kr för utvärdering av *vaccinationsprogrammet*.⁴⁴

Begränsad budget och tidsram kan medföra att utvärderaren måste gå genvägar och använda sig av enklare och snabbare metoder. Det kan då påverka hur säkerställda resultaten blir. Ett annat alternativ är att begränsa utvärderingen. I utvärderingen av *rehabilitering av unga immigranter* menade utvärderarna att en kortare tidsram skulle ha medfört att de skulle ha studerat färre variabler, som till exempel endast förändringar i den subjektiva arbetsförmågan och hoppat över alla, eller de flesta, kliniska variabler.

I våra exempel använde utvärderarna sammantaget de flesta etablerade datainsamlingsmetoder: intervjuer, enkäter, mätningar, statistikinsamling, litteraturstudier etc. Kostnaderna för dessa moment upptar ofta en stor del av utvärderingens totala kostnader. Det betyder att utvärderaren redan på planeringsstadiet måste veta hur datainsamlingen ska gå till. Kostnaderna för en utvärdering kan också bero på hur mycket och vilken typ av data som finns tillgängliga för utvärderaren.

Hur skulle utvärderingen ha sett ut om utvärderarna hade haft mindre resurser? Vi frågade utvärderarna av *kameraövervakningen*. De menade att med en mera begränsad budget hade de fått avgränsa utvärderingen och fokuserat på brotts-

statistiken och inte studerat tryggheten och integriteten som omfattade intervjuer. Vi frågade också utvärderarna av *anställningsstödet*. De ansåg att de inte hade velat genomföra uppdraget med mindre resurser på grund av risken att resultaten då inte skulle bli tillräckligt säkerställda.

Hur skulle utvärderingen se ut om utvärderarna hade haft mer resurser? Om utvärderarna i studien av *introduktionsutbildningen* av nyanställda hade haft en större budget hade de behållit enkätupplägget. Men de hade kompletterat studien med intervjuer av ett antal deltagare. De hade då kunnat basera enkätfrågorna på resultatet från intervjuerna och därmed kunnat öka kvaliteten i frågorna.

Utvärderarna av *transportbidraget* skulle med en större budget ha gjort flera intervjuer av industriföretag. Anledningen var att de upplevde att det var intervjuerna som gav mest information när det gäller effekterna för företagen. Företagens svar spelade även en viktig roll vid bedömningen av effekterna för regionen. De kvalitetssäkrade analysen av den statistik som gällde regionen.

Lägg fast kravet på säkerställda resultat

Som vi redan påpekade i kapitel 2 talar man i utvärderingslitteraturen om att en utvärdering bör uppfylla kraven på säkerställda resultat, vilket brukar omfatta både reliabilitet och validitet. Reliabiliteten, det vill säga tillförlitligheten, anger precisionen i mätningen. Det innebär att en upprepning av mätningen ska ge samma resultat. Validiteten brukar stå för ”att man mäter det man avser att mäta”.

En viktig utgångspunkt för utvärderaren är att definiera med vilken säkerhet han eller hon ska kunna uttala sig om effekterna. I praktiken

kan det vara svårt att explicit uttrycka hur pass säkerställda resultaten måste vara.⁴⁵

Kravet på säkerställda slutsatser påverkas av syftet med utvärderingen och hur den ska användas. En kontrollutvärdering ställer normalt sett större krav på säkerhet än en främjandeutvärdering som ska användas som ett redskap för att förbättra verksamheten.⁴⁶ Utvärderaren måste ställa sig frågan i vilken grad effekterna måste vara säkerställda. Slutsatsen beror på hur viktigt det är att göra en korrekt och hur riskabelt det är att göra en felaktig bedömning. För att peka på några ytterligheter är det i medicinska sammanhang oerhört viktigt att veta om till exempel ett vaccin har effekt eller inte och det kräver att resultatet är statistiskt säkerställt. Om det däremot handlar om effekterna av en utbildningsåtgärkning är kraven troligen betydligt lägre.

Ett kontrollerat experiment anses, när det är realistiskt att utföra, vara det upplägg som är mest effektivt för att få fram säkerställda resultat. Det finns dock många undantag från denna regel. Ibland är det till exempel inte lämpligt att välja experiment- och kontrollgrupp slumpmässigt utan det krävs att grupperna matchas. Det kan bero på att populationen är för liten för att använda slumpmässigt urval. Det kan också vara etiskt tveksamt att vissa personer ska få pröva respektive inte pröva en ny typ av behandling.

Om det inte är möjligt eller relevant att arbeta med kontrollgrupp är den bästa lösningen för att säkerställa resultatet att göra en före-efterstudie, helst kompletterad med någon annan typ av kontroll som till exempel generisk kontroll. Om det inte är möjligt att ta fram data om situationen före åtgärden måste utvärderaren nöja sig med en efterstudie, men då bör

den kompletteras med andra kontroller som till exempel skuggkontroll.

En sammanställning över ett antal vanliga problem med de olika uppläggen ges i bilaga 3.

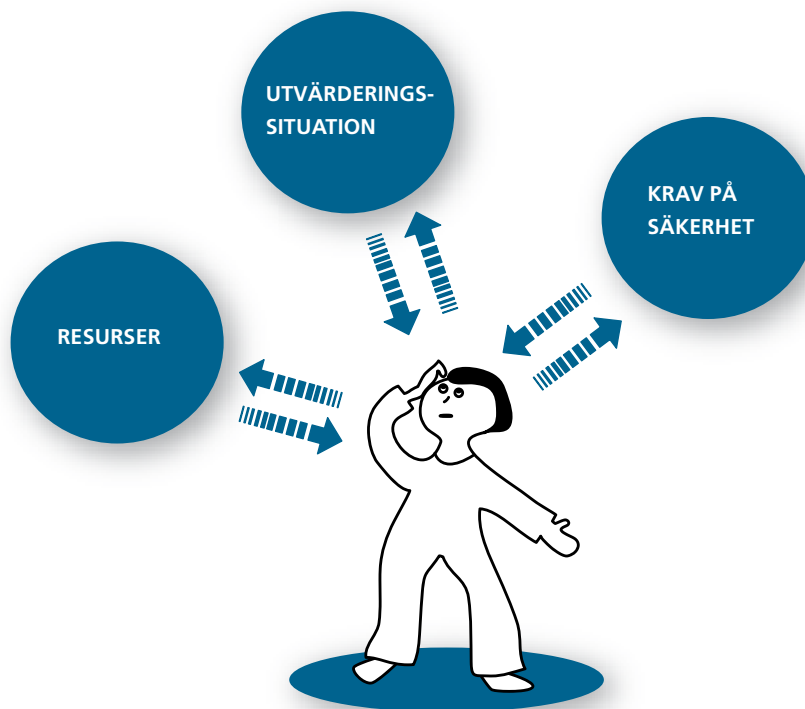
Gör en sammanvägd bedömning och välj upplägg

I detta steg väger utvärderaren samman tre faktorer:

- Utvärderingssituationen och de praktiska möjligheterna att genomföra olika upplägg.
- Resursläget.
- Krav på med vilken säkerhet det ska gå att uttala sig om resultatet.

När sammanvägningen är gjord kan utvärderarna välja upplägg för sin effektutvärdering. Utvärderaren får ofta gå en balansgång mellan kravet på säkerställda resultat och önskan om snabba och billiga studier. I de utvärderingar som vi beskrivit finns ett antal exempel på detta. Om studien ska bli billigare måste utvärderaren vidta åtgärder som till exempel att använda enklare analysmetoder, rikta sig till en mindre grupp för intervjuer, utesluta intervjuer och i stället använda enkäter eller bara göra en teoretisk utvärdering. Kravet på säkerställda resultat liksom de nödvändiga resurserna för att genomföra utvärderingen varierar mellan de olika uppläggen. Därför är uppläggen idealiska i en bestämd utvärderingssituation, men inget upplägg är optimalt för alla situationer.

Ibland tvingas utvärderaren inse att det inte är möjligt att komma fram till tillräckligt säkerställda resultat, även om resurserna är omfattande. Utvärderaren måste då ta ställning till om han eller hon verkligen bör genomföra utvärderingen.



Figur 14: Utvärderarens valsituation

Två situationer kan illustrera valsituationen. I prövningen av kikhostevaccin fanns höga krav på säkerställda resultat eftersom kontrollerat experiment är den standard som gäller vid prövning av vacciner. Resursutrymmet var också väl tilltaget och kraven på att kunna mäta skyddseffekten på ett tillförlitligt sätt krävde en viss tidsram. Helt annorlunda förutsättningar gäller vid till exempel en utvärdering av effekterna av en omorganisation. Kraven på exakta resultat är troligen mycket lägre, tiden för att genomföra utvärderingen kortare och budgeten mera begränsad än i exemplet om prövning av kikhostevaccin. Om omorganisationen enbart gäller delar av en myndighet kan det visserligen vara möjligt att hitta

kontrollgrupper inom organisationen, men det stora problemet är att hitta kontrollenheter som i sin struktur verkligen liknar de enheter som omorganiserats. I detta fall kanske effektutvärderingen i stället kommer att bestå av en efterstudie baserad på intervjuer med personalen där den får uttala sig om situationen före respektive efter omorganisationen (skuggkontroll).

Sammanfattningsvis går det inte att ge några generella råd om vilket upplägg som utvärderaren ska välja. Om förutsättningarna däremot är givna vad gäller hur utvärderingssituationen ser ut, resursläget och hur stora krav man har på säkerställda resultat är det däremot ofta lätt att fastställa vilket upplägg som passar bäst. ●

Råd från utvärderarna

Vi frågade utvärderarna som ansvarade för de utvärderingar vi använt som exempel vilka råd de skulle ge till utvärderare i färd med att göra liknande utvärderingar. Här är deras svar.

REHABILITERING AV UNGA IMMIGRANTER:

”Ge inte tappt! I det här fallet handlade det om sjuka människor och man måste hantera dem med deras eget bästa för ögonen och inte vara egoistisk.”

PRÖVNING AV KIKHOSTEVACCIN:

”Avsätt tillräckligt lång tid för planering. Konsultera all tillgänglig expertis. Visa stor öppenhet om studiernas uppläggning. Redovisa även negativa resultat.”

LÖRDAGSÖPPNA SYSTEMBOLAG:

”Var observant på effekter av andra faktorer.”

ANSTÄLLNINGSSTÖD:

”När man gör studier på observationsdata är man beroende av att kunna övertyga läsare om att de identifierade antagandena som görs är rimliga. Detta innebär att data-tillgång och kunskap om hur programmet fungerar är oundgängliga i varje effektutvärdering. Därför bör man gå igenom dessa aspekter noggrant innan man bestämmer sig för om en utvärdering överhuvudtaget är möjlig.”

KAMERAÖVERVAKNING:

”Var med redan från början och håll koll på processen, det vill säga när effekter kan börja förväntas. Använd om möjligt mer detaljerad information än årsdata, eftersom den här typen av åtgärder ofta ger tydliga effekter snabbt efter införandet.”

KVALITETSÖKNING INOM

YRKESUTBILDNINGEN:

”Försök att avgränsa utvärderingen. Arbeta systematiskt.”

TRANSPORTBIDRAGET:

”Involvera relevanta aktörer genom intervjuer av relevanta grupper samt avstämningar inom en brett sammansatt referensgrupp.”

INFORMATIONSKAMPANJ OM HIV/AIDS:

”Det är ibland möjligt att blottlägga antaganden som måste vara sanna för att en verksamhet ska leda till avsedda effekter. Dessa kan prövas mot bästa möjliga kunskap. En sådan prövning kan visa att en verksamhet inte kan fungera. Motsatsen gäller dock inte. Även om alla antaganden skulle stämma vet vi inte om verksamheten leder till avsedda effekter.”

INTRODUKTIONSPROGRAM

FÖR NYANSTÄLLDA:

”Man måste fråga sig vad som är rimligt att ta reda på och vad annat som påverkar.”

Bilagor

**Bilaga 1: Att genomföra en effektutvärdering –
Skolverkets inspektionsverksamhet**

**Bilaga 2: Exempel på olika typer av effektut-
värderingar**

**Bilaga 3: Faktorer som påverkar säkerheten
i resultaten**

Bilaga 4: Litteraturförteckning

Bilaga 5: Index

Bilaga 1

ATT GENOMFÖRA EN EFFEKTUTVÄRDERING – SKOLVERKETS INSPEKTIONSVERKSAMHET ⁴⁷

Utvärderingsuppdraget

Skolverket har sedan år 2003 som en viktig uppgift att utföra inspektion av landets skolor och förskolor. Tanken med inspektionen är bland annat att den ska upptäcka avvikelser, både positiva och negativa, och att bristerna ska åtgärdas så att de nationella målen för skolan uppfylls.

Skolverket gav år 2004 Ekonomistyrningsverket (ESV) i uppdrag att utvärdera effekterna av inspektionen.⁴⁸ Silja Jundin på Skolverket var ansvarig för upphandlingen och säger så här om hur uppdraget till ESV kom till:

”Det står i vårt regleringsbrev att vi ska utvärdera effekterna av vår verksamhet och eftersom inspektionen är så viktig för oss är det naturligt att vi utvärderar den. Samtidigt vill vi ju utveckla verksamheten. På så sätt finns det inslag av både kontroll och främjande i utvärderingen. Vi tyckte att det var viktigt att vi fick en oberoende bedömning av inspektionen. Därför var det naturligt att vi vände oss till en oberoende utvärderare. Att det blev just ESV beror på att vi vet att ESV har kompetens att göra den här typen av utvärderingar.”

Utvärderingen gjordes dels för att kontrollera och rapportera vidare om inspektionens effekter, dels för att kunna utveckla inspektionsverksamheten. I uppdraget angavs inte närmare vilka effekter som särskilt skulle studeras. Uppdraget har därmed likheter med det som brukar kallas målfri utvärdering.

Utvärderingens resultat

Det gick inte att ange effekterna av Skolverkets inspektion som ett värde utan utvärderarna fick ange effekter länkade till den verksamhetslogik som man tagit fram (se fig.15 på sidan 50/51).

Majoriteten av dem som besvarade ESV:s enkät var nöjda med skolinspektionens agerande. Därför bedömde ESV att det också fanns förutsättningar för att inspektionerna skulle

komma att leda till att skolorna förändrade sitt agerande. Mer än hälften av dem som besvarat ESV:s enkät uppgav också att de trodde att inspektionen skulle komma att leda till långsiktiga effekter, främst genom nya administrativa rutiner, omorganisationer samt förändrade arbetssätt.

Det kom också fram resultat länkade till de olika faserna av den verksamhetslogik som användes. Till exempel ansåg 30 procent av respondenterna att de redan kände till de problem som påpekades i inspektionsrapporten men att skolinspektionen i sig gjorde att man snabbare vidtagit åtgärder för att komma till rätta med bristerna.

Valet av upplägg

I kapitel 2 och 3 angavs att valet av upplägg bestäms av tre faktorer:

- Utvärderingssituationen och de tekniska möjligheterna att välja ett upplägg.
- Resurserna.
- Kravet på säkerställda resultat.

I detta exempel var det främst utvärderingssituationen och kravet på säkerställda resultat som bestämde valet av upplägg.

I kapitel 3 angav vi hur utvärderarna kan gå till väga för att besluta vilket upplägg de bör använda. Valsituationen beskrev vi i sju steg:

1. Beskriv den åtgärd som ska utvärderas.
2. Beskriv typen av effekter.
3. Arbeta fram en verksamhetslogik.
4. Bedöm vilka upplägg som kan vara lämpliga.
5. Bedöm resursåtgången för alternativen.
6. Lägg fast kravet på säkerställda resultat.
7. Gör en sammanvägd bedömning och välj upplägg.

Denna process används, något förenklad, för att analysera valsituationen i detta projekt.

Beskrivning av åtgärden

Skolverkets prestationer, inspektionerna, är i stort sett standardiserade. Ett metodproblem som ESV:s utvärderare måste hantera var dock att metoderna för inspektionen utvecklats sedan starten.

Beskrivning av typen av effekter

Skolans övergripande mål är att leva upp till de nationella målen för skolan.

Effekterna är mångdimensionella; en bra skola kan avläsas på många sätt; högre betyg, minskad mobbning etc. Detta gjorde utvärderarnas arbete svårare.

Rolf Sandahl var projektledare för utvärderingen och kommenterar detta så här: *”Det hade varit enklare om vi bara hade tittat på inspektionens betydelse för en speciell fråga, till exempel elever med särskilda behov. Ännu lättare hade det varit om vi till exempel hade fokuserat på andelen behöriga lärare. Då skulle det också ha gått att ta fram statistik som belyste situationen före och efter inspektionen.”*

Effekterna är långsiktiga, att utveckla en ”bra skola” tar tid och det kan vara svårt att snabbt avläsa effekterna. Det här är också en av anledningarna till att Skolverket ville vänta ett år med att utvärdera effekterna.

Verksamhetslogiken

Om skolan klarar att leva upp till de nationella målen beror på en stor mängd faktorer och aktörer, av vilka Skolverkets inspektion förmodligen bara är ”en liten kugge”. Det här gjorde det svårt att identifiera just skolinspektionens betydelse. En enkel analys med hjälp av verksamhetslogiken kan klargöra detta (se vidare figur 15).

Bedömning av vilka upplägg som är tänkbara

Går det att definiera en kontrollgrupp?

Frågan om att etablera en kontrollgrupp diskuterades noga av utvärderarna. För att undersöka vad som skulle ha hänt om skolinspektionen inte hade existerat, skulle det teoretiskt sett varit möjligt att arbeta med en kontrollgrupp. Den grupp som låg närmast till hands var skolor/kommuner som ännu



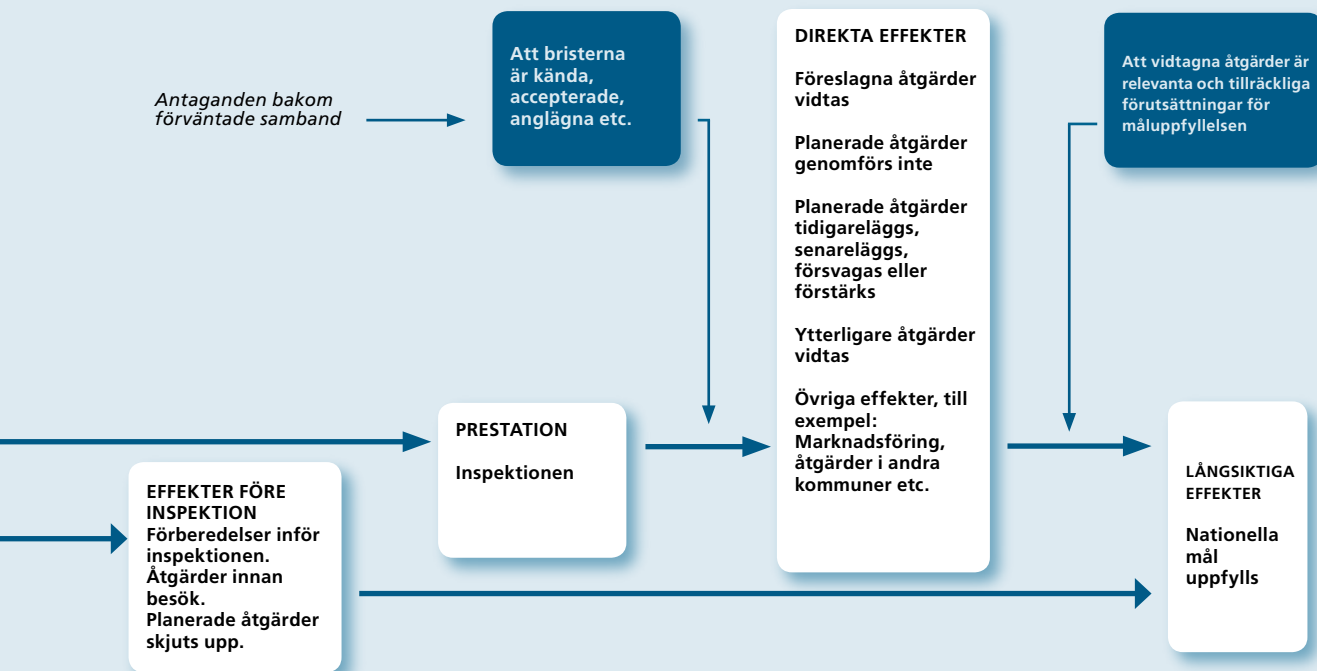
inte inspekterats. Svårigheten var att hitta jämförbara skolor. För varje enskild skola i den inspekterade kommunen behövde man då hitta en jämförbar skola i en ”icke-inspekterad kommun” med exakt samma brister. Detta var naturligtvis hart när omöjligt.

Går det att göra en före-efterstudie?

Ett tänkbart angreppssätt var att studera samma kommun/skola en tid före det att inspektionen ägde rum och sedan återkomma och göra besök då inspektionen ägt rum (före-efterstudie). Men i detta fall var det omöjligt, eftersom inspektionen redan hade ägt rum då utvärderarna fick sitt uppdrag.

Det valda upplägget

Med hänsyn tagen till resurserna och kravet på säkerställda resultat var utvärderarnas enda möjlighet att göra en efterstudie kring vad som skulle ha hänt i kommunerna och



Figur 15: Verksamhetslogik – Skolverkets inspektionsverksamhet

på skolorna om inte Skolverket hade genomfört sin inspektion. Metoden bygger på att utvärderarna i efterhand ställer frågor till personer på de berörda skolorna om vad som skulle ha hänt om inte Skolverket hade inspekterat. Kommunen/skolan jämför sig därför med sig själv och blir på så sätt sin egen kontrollgrupp, vilket brukar kallas skuggkontroll.

Utvärderingen var en totalundersökning som avsåg alla skolor som hade inspekterats av Skolverket.

Alternativa upplägg

Hur skulle utvärderingen ha lagts upp om utvärderarna haft mer eller mindre tid på sig?

Mer tid

Rolf Sandahl: "Även om vi hade haft mer tid på oss hade det nog inte påverkat metoden som sådan. Vi hade nog ändå gjort en enkät och den hade sett likadan ut, för det går inte att ha fler frågor. Kanske hade vi redan i denna studie gjort

fallstudier, nu kommer de i stället i nästa del av denna studie. En annan möjlighet hade varit att klassificera de olika inspektörernas kompetens och arbetssätt för att se om detta hade något inflytande på effekterna."

Mindre tid

Rolf Sandahl: "Om vi hade haft mindre tid, säg 20 arbetsdagar, hade vi fått avgränsa oss rejält. Vi hade fått minska antalet frågor rejält och kanske hade vi använt oss av telefonintervjuer. Eller också hade vi använt fokusgrupper och samlat ett antal rektorer och förvaltningschefer för en diskussion. Vi hade också fått dra ned på antalet frågor med öppna svar för de tog lång tid att bearbeta. En annan möjlighet hade varit att begränsa effekterna till några områden, till exempel effekterna på elever med särskilda behov."

Genomförande av utvärderingen

Tidplan

Utvärderingsuppdraget genomfördes av två erfarna utvärderare, Rolf Sandahl och Sara Bringle vid ESV. De använde under loppet av sex månader totalt cirka 100 persondagar för uppdraget. Följande aktiviteter genomfördes.

AKTIVITET/TID

- Inläsning av material + inledande intervjuer (Augusti–September)
- Formulering av verksamhetslogik samt förstudie i kommuner (Augusti+Oktober)
- Enkät:
Konstruktion (Oktober–December)
Genomförande (December–Januari)
Bearbetning (Januari)
- Tolkning och rapportskrivning (December–Januari)
- Rapportering till Skolverket (Januari)

Det är omöjligt att exakt ange hur stor del av tiden som användes för de enskilda aktiviteterna men en erfarenhet, enligt utvärderarna, var att konstruktionen av enkäten tog lång tid.

Inläsning och inledande intervjuer

ESV:s utvärderare startade med att göra en förstudie i några kommuner som blivit inspekterade där berörda personer intervjuades. Utvärderarna deltog också i några inspektioner. Syftet var att förstå hur skolinspektionen uppfattades och hur inspektionerna gick till. Samtidigt tog man del av gällande lagar och förordningstexter för inspektionen.

Formulering av verksamhetslogiken

Detta var förmodligen det viktigaste momentet i hela utvärderingen och det berodde på att datainsamlingen helt byggde på de resultat som kom fram via verksamhetslogiken.

Rolf Sandahl: *”Vi läste in oss på syftet med inspektionen och sedan satte vi oss helt enkelt ned och tänkte ut olika tänkbara effekter. Efteråt, när vi besökte några kommuner,*

la vi upp skissen på bordet och bad våra intervjupersoner att kommentera. På så sätt kunde vi lägga till några ytterligare exempel på effekter.”

Utvärderingen undersökte effekter vid två tidpunkter. Det första tillfället var efter skolornas vetskap om att en inspektion var förestående och det andra efter själva inspektionen. Figur 15 (se sidan 50/51) visar verksamhetslogiken över inspektionsverksamheten med tänkbara effekter och gjorda antaganden om förutsättningar för de tänkta orsak-verksamhetsbanden.

Som framgår av figuren är det yttersta syftet med inspektionen att den ska säkerställa att de nationella målen för skolan uppfylls. Ett första steg är att inspektionen upptäcker brister och sedan att inspektionen påtalar bristerna. Därefter krävs att bristerna åtgärdas. För att det sista ledet i kedjan ska påverkas krävs således att alla tidigare led är uppfyllda.

Effekter efter vetskap om en förestående inspektion

Vetskapen om en förestående inspektion kan skapa reaktioner hos dem som ska inspekteras. Skolan kan till exempel ta reda på vilka områden som granskas och därmed vidta planerade åtgärder innan inspektionen. Vetskapen om att en inspektion är på väg kan också få motsatt effekt, att skolan avvaktar för att se vilka synpunkter inspektionen kommer med. I direkt anslutning till besöket kan man även tänka sig att skolan genomför en viss ”uppsnygning” av sina lokaler.

Effekter efter inspektionen

Inspektionen kan leda till att skolan åtgärdar de förslag som inspektionen föreslår. Men skolan kan också genomföra åtgärder utöver dem som inspektionen föreslagit. Resultatet av inspektionen kan till exempel användas som ett styrinstrument inom kommunen mellan central ledning och rektorer. Inspektionens resultat kan också leda till effekter inom kommuner som Skolverket ännu inte inspekterat. Andra kommuner kan läsa om de kommuner som blivit inspekterade och som en följd av detta vidta åtgärder. Detta är en typ av icke-planerade effekter.

Enkät - konstruktion, bearbetning och tolkning

En svårighet med att mäta effekterna av inspektionen var att de kanske inte inträffade direkt efter att Skolverkets inspektion genomförts. Det innebar att utvärderarna måste vänta ett tag med att samla in data. Samtidigt kunde de inte vänta för länge, eftersom så kallade externa faktorer (till exempel neddragningar av resurser eller att en ny rektor tillsätts) kunde få för stort inflytande. I så fall skulle det bli svårare för intervjupersonerna att komma ihåg eller fastställa vad som orsakat en eventuell förändring på skolan. Tidpunkten för mätningen var därför inte självklar. ESV:s mätning via enkäten gjordes ett år efter det att inspektionerna ägt rum.

ESV:s enkät var Internetbaserad och gick ut till 378 personer i 21 kommuner som var berörda av Skolverkets

inspektion: lärare, rektorer, förvaltningstjänstemän och skolpolitiker.

Rolf Sandahl: *”Vi valde enkät framför fallstudier och intervjuer därför att vi bedömde att det skulle gå snabbare.”*

Utvärderarna lade ner mycket kraft på att utforma enkäten och använde sig av flera referenspersoner: kollegor på ESV, forskare, en representant från dåvarande kommunförbundet samt några personer från de kommuner utvärderarna besökt.

Utvärderarna stod inför två utmaningar vid enkätkonstruktionen. För det första gällde det att göra frågorna enkla att besvara och för det andra få frågorna att verkligen belysa de effekter som kom fram i arbetet med verksamhetslogiken. Tabellen visar exempel på hur utvärderarna använt effekter som identifierats via verksamhetslogiken som

KOPPLINGEN MELLAN TÄNKBARA EFFEKTER OCH ENKÄTFRÅGOR

Tänkbara effekter (enligt verksamhetslogiken)	Motsvarande fråga i enkäten	Svarsalternativ
Förberedelser på skolorna inför inspektionen	Vidtogs några åtgärder i Er verksamhet mellan information om att inspektion skulle göras och själva inspektionen?	<input type="checkbox"/> Ja <input type="checkbox"/> Nej <input type="checkbox"/> Vet ej
	Om ja, var det följande typer av åtgärder? (Ange ett eller flera alternativ)	<input type="checkbox"/> Kontaktade andra granskade kommuner. Läste på Skolverkets hemsida eller i andra dokument för att ta reda på hur inspektionen går till. <input type="checkbox"/> Sköt upp någon eller några planerade åtgärder i avvaktan på inspektionen. <input type="checkbox"/> Uppdaterade vissa dokument och/eller stämde av den egna verksamheten mot Skolverkets kriterier.
Åtgärder på grund av inspektionen	Har Din organisation påbörjat arbetet med de utvecklingsområden som inspektionen framfört i rapporten	<input type="checkbox"/> Ja, på alla områden. <input type="checkbox"/> Ja, på några områden. <input type="checkbox"/> Vi har inte börjat ännu. <input type="checkbox"/> Vet ej.
	Om Du kände till de brister/utvecklingsområden som inspektionen påtalade, tror Du att dessa skulle ha åtgärdats ändå?	<input type="checkbox"/> Ja, de var redan planerade att åtgärdas. <input type="checkbox"/> Ja, de skulle åtgärdas vid ett senare tillfälle. <input type="checkbox"/> Nej, de var inte så viktiga att de behövde åtgärdas. <input type="checkbox"/> Nej, det fanns inte tid/resurser/intresse. <input type="checkbox"/> Vet ej.

utgångspunkt för formulering av frågor i enkäten.

Enkätarbetet innehöll ett antal praktiska problem som tog mycket tid i anspråk. Dessa gällde allt från att ta fram e-postadresser till respondenterna till att hantera problemet med att inte alla respondenter förstod hur man rent praktiskt skulle hantera en webbenkät.

Tolkning och rapportskrivning

Den datainsamlingsmetod som utvärderarna valde rymmer ett antal metodproblem som är specifika för enkäter kring effekter. Det största problemet gällde hur man skulle komma åt den kontrafaktiska situationen, det vill säga vad respondenterna skulle ha gjort om inte inspektionen hade ägt rum. Problemet bestod i att det kunde vara svårt att komma ihåg hur läget såg ut när inspektionen utfördes och ännu svårare att föreställa sig vad som skulle ha hänt om inte inspektionen genomförts.

Rolf Sandahl: *”Resultaten grundar sig ju på vad intervju-personerna själva påstår att inspektionen kommer att leda till. Men vi måste tolka svaren med stor försiktighet. Det är mänskligt att i efterhand försöka ”frisera” svaren för att komma i bättre dager. I del II av utvärderingen kommer vi också att göra fallstudier i några kommuner och då blir det lättare att krypa in under pannan på våra intervju-personer.”*

Rapportering till Skolverket

Under utvärderingsarbetet höll utvärderarna löpande kontakter med uppdragsgivaren på Skolverket. I samband med publiceringen av rapporten presenterade de rapporten för generaldirektören, Skolverkets ledningsgrupp, inspektionens avdelningsledning samt enhetschefer och vid ett separat tillfälle för samtliga skolinspektörer.

Skolverkets användning av resultaten

Silja Jundin på Skolverket säger så här om hur Skolverket har använt och tänker använda resultaten från ESV:s utvärdering: *”Rapporten har använts i många sammanhang. Vi har tagit upp den på ett avdelningsmöte med alla inspektörer och dessutom på ett seminarium med verkets ledning. Sedan har vi haft den uppe i diskussionerna med departementet. Rapporten användes också vid återrapporteringen i årsredovisningen. Vi publicerar varje år en nationell rapport om inspektionen och i årets upplaga nämndes utvärderingen. Efter det är det många som frågat efter rapporten, även bland massmedia. Vi vill ju hela tiden utveckla inspektionens metoder. Rapporten gav oss en del synpunkter på hur våra inspektörer uppträder och det tar vi upp på våra träffar. Såväl det positiva som det negativa. Vi kommer också att använda den för att utveckla våra inspektionsmetoder.”*

Bilaga 2

EXEMPEL PÅ OLIKA TYPER AV EFFEKTUTVÄRDERINGAR

Upplägg	Typ av åtgärd	Titel	Beställare / Utvärderare	År
A	Information– miljö	Den svenska klimatkampanjen–en del av Sveriges klimatstrategi	Regeringen / Naturvårdsverket	2004
A	Kontroll– skatter	En ny kraft i ekobrottsbekämpningen– En utvärdering av skattebrottsenheterna, RSV Rapport 2001:7	Regeringen / Riksskatteverket	2001
A	Utvecklingsinsatser– HIV/AIDS	Committing to Results: Improving the Effective- ness of HIV/AIDS Assistance	Världsbanken	2005
A	Utvecklingsinsatser –kommunal infrastruktur	Drivkraft till förbättring - Betydelsen av begränsade kommunala insatser i spontan bebyggelse. Hai Si Salah, Oran, Algeriet	Birgitta Mekibes, Kungliga Tekniska Högskolan (KTH)	1988
A	Utvecklingsinsatser– narkotika	Polisens insatser mot narkotikabrottsligheten, Rapport 2003:12	Brottsförebyggande Rådet (BRÅ)	2003
A	Utvecklingsinsatser– trafik	Det nationella trafikprogrammet 1995–2000– ett misslyckande? Artikel till Svenska utvärde- ringsföreningens konferens, 2004	Matts-Åke Belin (Vägverket)	2004
A	Övervakning– brottsförebyggande	Kameraövervakning i brottsförebyggande syfte, Rapport 2003:11	Brottsförebyggande Rådet (BRÅ)	2003
B	Bidrag– datorer	Lånedatorsystemet–Gratis för arbetsgivare– dyrt för stat och kommun, Statskontoret 2005:14	Regeringen / Statskontoret	2005
B	Information– rökning	Våldtagen av en Prince, En Rökfri Generation	En Rökfri Generation / Petra Stenqvist & Lotta Thiringer	2005
B	Kontroll– skatter	Skattefel och skattefusk - en utvärdering av skattekontrollen 1992–1997, RSV Rapport 1998:3	Regeringen / Riksskatteverket	1998
B	Ny lag–kriminalvård	Besöksförbud–En utvärdering av lagen och dess tillämpning, Rapport 2003:2	Brottsförebyggande Rådet (BRÅ)	2003

Upplägg A = Före-efterstudie utan kontrollgrupp. Upplägg B = Efterstudie utan kontrollgrupp.

Upplägg	Typ av åtgärd	Titel	Beställare / Utvärderare	År
B	Styrning–miljöledning	Miljöledning i statliga myndigheter, Naturvårdsverket, Rapport 5346	Regeringen / Naturvårdsverket	2004
B	Policyverksamhet–bistånd	Utvärdering av Expertgruppen för studier av utvecklingsfrågor (EGDI)	Utrikesdepartementet / Sinova	2001
B	Skatteavdrag–byggnation	ROT-avdragets effekter, Rapport 2001/02:8	Riksdagens revisorer	2001
B	Tillsyn–skolan	Skolverkets utbildningsinspektion–ger den några effekter, slutrapport, 2006.	Skolverket / Ekonomistyrningsverket	2005
B	Utbildning–invandrare	Integration–var god dröj, Integrationsverket rapport 2004:01	Integrationsverket / Ramboll Management	2004
B	Utbildning–statsanställda	Utvärdering av KKR:s introduktionsprogram för nyanställda i statlig tjänst	Statens kvalitets- och kompetensråd (KKR) / IPF	2002
B	Utbildning–kvinnliga chefer	Uppföljning av ett utvecklingsprogram för kvinnliga chefer i staten, 2000–2001	Statens kvalitets- och kompetensråd / Svenska PA Forum et al.	2002
B	Utbildning–forskning	Effekter av det svenska kompetenscentrumprogrammet 1995-2003	Vinnova / Technopolis	2004
B	Utvecklingsinsatser–hatbrott	Hatbrott, Rapport 2002:9	Brottsförebyggande Rådet (BRÅ)	2002
B	Utvecklingsinsatser–regionalpolitik	Resurscentra för kvinnor–en kraft för hållbar regional utveckling?, NUTEK R 2004:14	NUTEK / IM-Gruppen	2004
B	Valfrihet–utbildning	Valfrihet och dess effekter inom skolområdet, 03:805	Skolverket	2003
C	Behandling–medicin	Rehabilitering av unga immigranter i primärvården, Scand J Prim Health Care 1997: 15	M. Lövander, A. Engström, H. Theander, A-K Furhoff	1997
D	Behandling–vaccin	Efficacy trial of acellular pertussis vaccines, Final report 1991-2003	Smittskyddsinstitutet	2004
D	Investeringsstöd–livsmedelsindustri	Utvärdering av ett investeringsstöd till livsmedelsindustrin, Rapport 2001:2	Regeringen / Livsmedelskonomiska institutet	2001
D	Kontroll–Sjukvård	Hur påverkar kravet på läkarintyg sjukfrånvaron? Erfarenheter från ett socialt experiment, ur Ekonomisk debatt 2/2006	Hesselius, Patrik; Johansson, Per; Larsson, Laura (IFAU)	2006
E	Arbets tid–militärer	Effekter av ständig nattpermission på sömn, prestationsförmåga, attityder och sjukanmälningsfrekvens Rapport A57002-HG	Försvarets forskningsanstalt (FOA)	1975
E	Kontroll–fiske	Den Svenska fiskerikontrollen–en utvärdering. Betänkande av Utredningen om fiskerikontrollen SOU 2005:27	Regeringen / SOU-utredning	2005

Upplägg	Typ av åtgärd	Titel	Beställare / Utvärderare	År
E	Servicehöjning–alkohol	Effekter av lördagsöppna systembolagsbutiker, uppföljning av de första 17 månaderna	Regeringen / Institutet för social forskning, Stockholms universitet och Det norske videnskapsakademi	2001
E	Utvecklingsinsatser–alkohol	Rattfylleriprogrammet i Stockholm, VTI meddelande 842 -1998	Väg- och transportforskningsinstitutet (VTI)	1998
E	Utvecklingsinsatser–socialt arbete	Med arbete som insats–Klienteffekter och samhällsekonomisk lönsamhet	Centrum för utvärdering av socialt arbete (CUS)	2002
F	Bidrag–anställning	Anställningsstöd–en väg från arbetslöshet till arbete?, Rapport 2004:17	Institutet för arbetsmarknadspolitisk utvärdering (IFAU)	2004
F	Bidrag–arbetsbefrämjande	Uppsalamodellen och socialbidragstagarna, CUS skrift 1998:1	Centrum för utvärdering av socialt arbete (CUS)	1998
F	Utvecklingsinsatser–narkotika	Kriminalvårdens särskilda narkotikasatsning, Rapport 2005:3	Brottsförebyggande Rådet (BRÅ)	2005
G	Delegering av ansvar–miljö	Hur påverkas beslut av delegering, Statskontoret 1992:28	Regeringen / Statskontoret	2005
G	Information–HIV/AIDS	Informationskampanjen om HIV/AIDS, RRV 1988:1836	Riksrevisionsverket (RRV)	1992
G	Styrning–intern organisation	Sida's Management Response System–Sida Studies in Evaluation 06/01	Sida / Centrum för Utvärderingsforskning (UCER)	1989
H (A+B)	Bidrag–transport	Transportbidraget, En analys av motiv, effekter och alternativ, NUTEK R 2004:15	NUTEK / Infraplan	2004
H (A+B)	Förändrade föreskrifter–anslag	Frysningen 2005–begränsning av myndigheternas anslagssparande och anslagskredit, ESV 2006:3	Ekonomistyrningsverket	2006
H (A+B)	Sektorspolitik–skog	Skogsvårdsorganisationens utvärdering av skogspolitikens effekter, SUS 2001	Regeringen / Skogsstyrelsen & Naturvårdsverket	2001
H (B+G)	Utbildning–mål 3	Samverkan för kompetens och konkurrenskraft, NUTEK 019-2004	NUTEK / IM-Gruppen	2004
H (B+G)	Utveckling–utbildning	Skolverkets arbete med att öka kvaliteten inom yrkesutbildningen	Skolverket / Eurofutures & L Widell Consulting	2003
H (Meta)	Behandling–fetma	Fetma–problem och åtgärder–en systematisk litteraturoversikt	Statens beredning för medicinsk utvärdering (SBU)	2002
H (Meta)	Insatser–trafik	Nybyggnad och förbättring–effektkatalog, Vägverket 2001:78:	Vägverket	2001
H (Meta)	Sektorspolitik–arbetsmarknad	Vad vet vi om den svenska arbetsmarknadspolitikens effekter?, Rapport 2002:8	Institutet för arbetsmarknadspolitisk utvärdering (IFAU)	2002

Bilaga 3

FAKTORER SOM PÅVERKAR SÄKERHETEN I RESULTATEN

De exempel vi presenterade i kapitel 2 visar upp ett antal faktorer som påverkar hur säkra resultaten kan bli. Faktorerna påverkar ofta samtliga empiriska upplägg men med olika styrka. Vi delar upp dem på sådana som främst berör upplägg utan kontrollgrupp (A och B), sådana som främst påverkar upplägg med kontrollgrupp (C – F) och sådana som är aktuella för alla typer av upplägg.⁴⁹

Faktorer specifika för upplägg utan kontrollgrupp (Upplägg A och B)

Mognad

En faktor som kan göra resultaten mindre säkra kallas vanligen för *mognad*. Det är en process som har med testpersonerna i sig att göra, det vill säga en biologisk eller psykologisk process. De flesta i en yrkesgrupp blir till exempel skickligare med tiden och patienter kan få bättre kunskap om sin sjukdom. Därför kan en del av förändringen i det uppmätta tillståndet bero på detta. Detta kan till exempel vara relevant i utvärderingen om *rehabilitering av unga immigranter*.

Detta problem är svårt att hantera. Det viktiga är att vara medveten om det i bedömningen av resultatet. Ett sätt att få en indikation på problemet är att göra ett flertal mätningar.

Tidpunkt för mätningarna

En viktig fråga är vid vilken tidpunkt som föremätningen ska göras. När det gäller verksamheter som fluktuerar, som olycksfrekvenser, brottsutveckling eller liknande, eller där man vet att värdena för tillfället är extrema, brukar dessa efter en upp- eller nedgång ofta sträva mot "mitten" igen, oavsett om man sätter in någon insats eller inte.

Ett sätt att komma tillrätta med problemet är att göra analyser av hur utfallet brukar fluktuerar och sedan bestämma tid för mätningen. Ett annat sätt är att göra ett antal mätningar i form av en tidserie. Generisk kontroll kan vara ytterligare ett sätt att fånga upp allmänna trender.

Varför är dessa faktorer speciella för upplägg utan kontrollgrupp (A och B)

Syftet med att använda en kontrollgrupp är att gardera sig mot de problem som vi tar upp ovan. Utgångspunkten i ett experiment med kontrollgrupp är att experiment- och kontrollgrupp är lika till sin karaktär och att problemen därför slår på samma sätt i de båda grupperna. Det gäller både kontrollerade experiment och kvasiexperiment. Vi kan naturligtvis aldrig helt gardera oss för att grupperna är olika. Men den experimentella uppläggningsen med kontrollgrupp är ett sätt att minimera systematiska fel.

Faktorer specifika för upplägg med kontrollgrupp (Upplägg C - F)

Även upplägg med kontrollgrupp innefattar ett antal faktorer som påverkar med vilken säkerhet man kan uttala sig om resultaten. Här är några av de viktigaste.

Selektion

En faktor rör urvalsprinciperna, speciellt för kvasiexperiment, där de deltagande inte är slumpmässigt valda. Om utvärderaren inte har förvissat sig om att experiment- och kontrollgrupp är lika i karaktär går det inte att utifrån en statistisk utgångspunkt hänföra eventuella effekter till åtgärden. Detta brukar kallas för att *selektionen* i gruppen inte är likvärdig. Även om grupperna kan verka lika i ett företest kan personerna i de båda grupperna utvecklas på olika sätt.

I exemplet med *anställningsstöd* anger utvärderarna att de jämfört olika sätt att matcha fram individerna mot varandra för att se om själva metoden för matchning spelade någon roll.

Låg yttre validitet

En experimentsituation kan vara så konstlad att resultatet inte kan tänkas gälla i det verkliga livet. Detta brukar kallas att den *yttre validiteten* är låg. I debatten kring ett svenskt experiment med förkortad arbetstid hävdades till exempel att situationen var alltför konstlad för att resultaten från försöket skulle kunna användas som beslutsunderlag. Kontrollgruppen var i det fallet arbetstagare som skulle arbeta utifrån nuvarande arbetsvillkor (åtta timmars arbetstid). Experimentgruppens deltagare skulle få arbeta sex timmars arbetsdag men med full lön, en konstruktion som ifrågasattes.⁵⁰

Om utvärderaren är med då experimentet startar finns det möjligheter att påverka projektets uppläggning så att det blir så verklighetsnära som möjligt.

Förändrat beteende

Individer i experiment- eller kontrollgrupper kan förändra sitt beteende under experimentet och detta kan göra resultatet ogiltigt eller svårtolkat. Om till exempel individer flyttar in och ut ur insatsområdet kan det också innebära att de flyttar in och ut ur experiment- och kontrollgrupperna.

Detta problem diskuteras bland annat i exemplet med *lördagsöppna systembolag*, där genomförarna medvetet valde experiment- och kontrollän som inte låg för nära varandra. På så sätt kunde man minska överströmningen mellan länen och problemet med att personer som bodde i kontrolläna åkte till experimentläna för att handla alkohol.

Imitationseffekten

Om det är svårt att skilja experiment- och kontrollgruppens medlemmar åt kan finnas risk att resultaten inte blir tillräckligt säkerställda. I ett utvecklingsprojekt kan det till exempel hända att experimentgruppens medlemmar delar med sig av sina nyförvärvade kunskaper till kontrollgruppens medlemmar. Detta kallas ibland för *imitationseffekt*.

En möjlighet är att göra ett strategiskt urval och försöka se till att personer som hamnar i de två grupperna inte har kontakt med varandra.

Faktorer som är generella för alla upplägg

Vetskap om testningen

Ett riskmoment för säkerheten i resultaten rör själva testningen. En person kan påverkas av själva vetskapen att han eller hon ingår i ett test och till exempel anstränga sig extra (detta brukar också kallas för *Hawthorne-effekten*). Själva testet kan också leda till ett förändrat beteende. I och med att en person till exempel ställer sig på en väg för att få ett ingångsvärde kan personen få incitament till ett förändrat beteende oberoende av vilka åtgärder som sätts in. Det finns även en risk att försökspersonerna kan få ett bättre resultat vid den andra testningen beroende på att de vant sig vid metoden, vilket är vanligt vid så kallade intelligenstagningar.

Det finns olika metoder för att komma åt problemet med att en testsituation i sig påverkar resultatet. En möjlighet är att inte göra någon föremätning, utan att i stället använda sig av endast ett eftertest. Vid kontrollgruppsupplägg kan man arbeta med blindning där grupperna inte vet om de får en verkningfull behandling eller inte. Inom det medicinska området kallas det verkningslösa medel som kontrollgruppen får för placebo.

konsekvenserna av bortfallet.

Instrumentering

Ytterligare ett problem som har betydelse för säkerheten i resultaten är *instrumenteringen*, det vill säga felaktigheter i mätinstrumenten. Dessa problem är oftast vanligare för så kallade sociala metoder (intervjuer, enkäter eller observationer) än så kallade fysiska metoder (till exempel fysiska mätningar). Det kan hända att en intervjuare uppträder olika i före- och eftermätningen, kanske därför att hon eller han vid den andra mätningen känner intervjupersonen. Om utvärderingen i stället använder olika intervjuare i före- och eftermätningen kan problemet i stället bestå i att intervjuarna upplever eller uppfattar saker och ting på olika sätt.

Om utvärderingen bygger på intervjuer bör intervjuarna rotera och utvärderarna dessutom bedöma om resultaten från de olika intervjuarna skiljer sig åt.

Om utvärderingen bygger på beteendeobservationer eller reaktioner på medicinska behandlingar bör observatörerna iakttä både experiment- och kontrollgruppen och då inte känna till vilka individer som hör till vilken grupp. Om inte heller deltagarna vet vilken grupp de tillhör kallas det för dubbel-blindning och ökar säkerheten i resultaten. Även här finns det möjlighet att bedöma om resultaten skiljer sig åt för de olika observatörerna.

Bortfall

Ett gemensamt problem för effektutvärderingar är *bortfallet*. Bortfallet kan i sig vara stort eller skilja sig mellan experiment- och kontrollgruppen. Båda förhållandena gör att resultatet kan bli mindre säkert. Detta behandlas till exempel i vårt exempel med *rehabilitering av unga immigranter*.

I vissa situationer är det möjligt att minska bortfallet. Det kan till exempel gälla för en enkätundersökning. En möjlighet är här att göra enkäten lättläst, inte för lång, snabb att svara på och se till att respondenterna inser värdet av att de svarar. Ett annat sätt att hantera bortfallsproblemet är att med hjälp av statistiska analyser bedöma

Bilaga 4

LITTERATURFÖRTECKNING:

Litteraturförteckningen består av såväl refererad litteratur som litteratur som på annat sätt påverkat handledningens innehåll.

Björklund, Anders: Experimentera mera!, ur Ekonomisk Debatt 1/2003.

Bohm, Peter (1996): Samhällsekonomisk effektivitet: SNS förlag.

Britton, Mona: Så graderas en studies vetenskapliga bevisvärde och slutsatsernas styrka, Läkartidningen, volym 97, nr 40.

Dahmström, Karin (2000): Från datainsamling till rapport: Studentlitteratur.

Ekonomistyrningsverket 2001:16: Verksamhetslogik.

Ekonomistyrningsverket 2004:14: ESV:s ordbok.

Ekonomistyrningsverket 2005:26: Att beställa utvärderingar - en vägledning.

Ekonomistyrningsverket 2006:7: Måluppfyllelseanalys.

Esaiasson, Peter; Gilljam, Mikael; Oscarsson, Henrik; Wängnerud, Lena (2003): Metodpraktikan - Konsten att studera samhälle, individ och marknad: Norstedts juridik.

Finansdepartementet Ds 2003:31: Politik på prov - en ESO-rapport om experimentell ekonomi.

Fredriksson, Peter (2005): Om bruk och missbruk av experiment: ur Ekonomisk Debatt 7/2005.

Förordning (2000:605) om årsredovisning och budgetunderlag.

Isaac, Stephen; Michael, William B. (1971): Handbook in Research and Evaluation: EdITS Publishers.

Karlsson, Ove (1995): Att utvärdera - mot vad? HLS Förlag.

Lag (1996:1059) om statsbudgeten.

Mathison, Sandra, Editor (2005): Encyclopedia of Evaluation: Sage.

Miller, Delbert C.; Salkind, Neil J. (2002): Handbook of Research Design and Social Measurement: Sage.

Morén, Stefan; Blom, Björn (2003): Insatser och resultat - Om utvärdering i socialt arbete: Umeå Universitet.

Morton, Sally C.; Rolph, John E., Editors (2000): Public Policy and Statistics - Case Studies from RAND: Springer.

Petrosino, Anthony; Boruch Robert F.; Soydan, Haluk; Duggan, Lorna; Sanchez-Meca, Julio (2001): Meeting the Challenges of Evidence-Based Policy: The Campbell Collaboration: Särtryck ur ANNALS, vol 578, November 2001.

Public Policy and Administration, Vol. 17, No. 3, 2002: Special Issue Evidence Based Policy.

Riksdagsordningen, SFS 1998:1450.

Riksrevisionsverket (1991): Resultatanalys.

Riksrevisionsverket 1996:6: Att mäta effekter.

Rossi P.; Lipsey M.; Freeman, H. (2004): Evaluation - A systematic Approach: Sage.

Shadish, William R.; Cook, Thomas D.; Campbell, Donald T. (2002): Experimental and Quasi-Experimental Designs: Houghton Mifflin Company.

Socialstyrelsen (2003): Perspektiv på kunskapsutveckling inom socialtjänsten - en antologi.

Statistiska Centralbyrån (1994): ABC i statistik.

Statskontoret 1993:4: Det går att mäta effekter - Metoder och exempel.

Statskontoret 2002:21: Utvärderingar och politik II - Hur använder regeringen utvärderingar?

Svensson, Lennart; Randle, Hanne; Eklund, Jörgen och Aronsson, Gunnar (utkommer preliminärt våren 2006): Projekt eller långsiktiga effekter?

Vedung, Evert (1998): Utvärdering i politik och förvaltning: Studentlitteratur.

Verksförordningen (1995:1322).

Weiss, Carol H. (1998): Evaluation: Upper Saddle River.

Vägverket 2001:78: Nybyggnad och förbättring, effektkatalog.

UTVÄRDERINGAR SOM ANVÄNDS SOM EXEMPEL I HANDLEDNINGEN:

BRÅ Rapport 2003:11: Kameraövervakning i brottsförebyggande syfte.

Eurofutures/L Widell Consulting (2003): Utvärdering av Skolverkets arbete med regeringsuppdraget Kvalitetsförbättrande insatser för gymnasieskolans yrkesutbildningar.

Institutet för arbetsmarknadspolitisk utvärdering (IFAU) (2004): Anställningsstöd - en väg från arbetslöshet till arbete?, Rapport 2004:17.

Institutet för personal- & företagsutveckling (IPF) (2002): Utvärdering av KKR:s introduktionsprogram för nyanställda handläggare i statlig tjänst.

Institutet för social forskning, Stockholms universitet och Senter for høyere studier, Det Norske videnskapsakademi (2001): Effekter av lördagsöppna systembolagsbutiker - uppföljning av de första 17 månaderna.

Löfvander, Monica; Engström, Alf; Theander, Hans; Furhoff, Anna-Karin (1997): Rehabilitation of young immigrants in primary care - a comparison between two treatment models: Scand J Prim Health Care 1997:15.

NUTEK R 2004:15: Transportbidraget - en analys av motiv, effekter och alternativ.

Ohlin, Patrick (2004): Efficacy Trial of acellular pertussis vaccines - final report 1991 - 2003: Smittskyddsinstitutet.

Riksrevisionsverket 1988:1836: Informationskampanjen om HIV/AIDS - samhällsinformation som styrmedel. Regeringsuppdrag.

Skolverket (2006): Skolverkets utbildningsinspektion - ger den några effekter?

Bilaga 5

INDEX	Sidhänvisning		
Bekvämlighetsurval	42	Målgrupp	9, 37, 39
Benchmarking	13	Mål-medel-kedja	9
Bieffekt	13	Måluppfyllelse	39
Blindning	16–17, 25, 60	Målutvärdering	12, 39
Bortfall	60	Om-inte-problematik	9
Budget	39, 43–44, 65 (not 44)	Orsakssamband	12
Cost-benefit-analys	13, 64 (not 16)	Pilotstudie	15
Dubbel-blindning	25, 60	Placebo(effekt)	16, 25, 28, 59
Efterstudie	18, 23–24, 27–28, 30–32, 35, 36, 41, 42–43, 47, 50, 55–57	Population	17
Effekt	9 (definition), 10–13	Prestation	9–10, 50–51
Effektkedja	9	Process	12
Effektutvärdering	12 (definition)	Programteori	10
Empiriska upplägg	18	Reliabilitet	15, 44
Evidens	21	Regressionsanalys	48 (not16)
Evidensbaserade studier	21	Resurser	15, 39, 43–44, 52, 65 (not 44)
Experiment	15, 18, 20, 25–32, 41–42, 46, 56–57, 58–59	Selektion	58
Experimentgrupp	16–20, 25–32, 41–42	Signifikanstest	17
Externa aktörer/faktorer	10, 40, 53	Skuggkontroll	16, 20, 22, 23, 35, 36, 42, 43, 45, 51
Fallstudie	64 (not16)	Slumptalgenerator	16
Felkälla	58–60	Slumptalstabell	16
Försöksverksamhet	15	Slumpmässigt urval	17, 19, 25, 27, 42
Före-efterstudie	18–22, 25–27, 29–30, 35, 41, 42–43, 50, 55–57	Strategiskt urval	17, 42
Förändrat beteende	59	Systematiska fel	58
Generisk kontroll	16, 18, 21, 22, 42, 45, 58	Säkerställda resultat	15, 39, 44–45, 49
Genomförande	13, 52	Teoretiskt upplägg	19, 32–34
Hawthorne-effekt	25, 59	Tidserieanalys	21
Imitationseffekt	59	Tillförlitlighet	15
Indikator	12	Totalundersökning	17, 51
Instrumentering	60	Undersökningsgrupp	16
Kompetens (hos utvärderaren)	43	Uppföljning	11
Kontrafaktisk	9	Upplägg	15 (definition)
Kontrollerat experiment	16, 18–19, 20, 25–28, 41–42, 45, 46, 56, 58–59	Urvalsundersökning	17, 42, 58
Kontrollgrupp	16, 19, 20, 25, 27, 29, 30, 41–42, 50, 55–57, 58–59	Utvärdering	12 (definition)
Korrelationsanalys	64 (not 16)	Utvärderingssituation	15, 20, 40, 45
Kostnadseffektiv	13	Validitet	15, 44, 59
Kvalitativt	13	Verksamhetslogik	10, 32–34, 36–37, 40, 50, 52–53
Kvantitativt	13	Vetskap om testning	59
Kvasiexperiment	17, 29–32, 41, 42, 56–57, 58–59	Yttre validitet	59
Metaanalys	34		
Mognad	58		
Målfri utvärdering	13, 49		

FOTNOTER

- 1 Lag (1996: 1059) om statsbudgeteten, 1 §.
- 2 SFS 2003:1057, 4 kap. 18 §.
- 3 Förordning (2000: 605) om årsredovisning och budget underlag, 3 kap. 1 §.
- 4 Verksförordningen (1995:1322), 7 §.
- 5 Se även ESV 2005:26: Att beställa utvärderingar - en vägledning.
- 6 Värdering av effekter behandlas till exempel i RRV: Resultatanalys.
- 7 I en del sammanhang används begreppet resultat endast i samband med effekter, ofta i utländsk litteratur.
- 8 Ibland används även begreppen intern- eller delprestation. Med detta avses en avslutad del i produktionsprocessen inom myndigheten som går in i nästa led av processen för att producera slutprestationen. Ett exempel på detta är ett underlag som tagits fram, som ska behandlas av en annan person eller grupp av personer i myndigheten.
- 9 I engelsk litteratur motsvarar effekter ofta impact eller effect medan outcome vanligen motsvarar förändring i tillståndet.
- 10 Man kan förstås även tänka sig att en effekt uppstår trots att det inte inträffat någon förändring i tillståndet. I fallet med rattonyckerhet kan det vara så att de alkoholrelaterade olyckorna ligger på samma nivå som tidigare men att utandningsproven trots detta haft effekt. Vårt kontrafaktiska resonemang innebär då att om inte utandningsproven hade funnits hade olyckorna varit betydligt fler än tidigare.
- 11 Eftersom vi i Sverige inte talar om program utan om verksamheter har ESV benämnt modellen verksamhetslogik. Tankarna bakom verksamhetslogiken presenteras närmare i ESV:s skrift Verksamhetslogik (ESV 2001:16). I den engelska utvärderingslitteraturen finns ett antal liknande uttryck, till exempel logic model, logical framework, program theory, performance framework, theory of action och chains of reasoning.
- 12 För en fördjupning i målfrågor se ESV 2006:7 Måluppfyllelseanalys.
- 13 För uppföljning använder ESV definitionen "en beskrivning av hur något har utvecklats" (ESV:s ordbok, ESV 2004:14).
- 14 Vägverket använder till exempel inom trafiksäkerhetsområdet standardmått för att värdera effekten av ett sparat liv. Exempel på en lärobok i cost-benefit-analys är Peter Bohm: Samhällsekonomisk effektivitet.
- 15 Detta diskuteras till exempel i Carol H. Weiss: Evaluation, kapitel 3.
- 16 Det är lätt att blanda ihop upplägg och analysmetoder. Exempelvis är korrelations- eller regressionsanalys inte upplägg utan statistiska metoder som ofta används för att analysera data. De omfattas därför inte av denna handledning. I vissa sammanhang kan det vara fördelaktigt att arbeta med fallstudier som berör effekterna på en enda individ eller enhet. Fallstudiemetodik behandlas dock inte heller i denna handledning. Cost-benefit-analys utförd i efterhand kan också ses som ett speciellt upplägg men behandlas inte här. Läsaren hänvisas till Peter Bohm: Samhällsekonomisk effektivitet.
- 17 Ibland används uttrycket trovärdiga resultat.
- 18 Se Rossi et al.: Evaluation - A systematic approach, kapitel 5.
- 19 Bristande reliabilitet kan orsakas av slarv fel under datainsamlingen och den efterföljande databearbetningen. Det kan till exempel handla om slarviga anteckningar, missförstånd i samband med intervjuer eller felaktigheter i samband med matematiska och statistiska beräkningar. Bristande reliabilitet kan också orsakas av ett felaktigt urval.
- 20 Indelningen bygger på Evert Vedung: Utvärdering i politik och förvaltning, kapitel 11.
- 21 I utvärderingslitteraturen talar man också om randomiserat kontrollerat experiment eller ett experiment med slumpmässig kontroll.
- 22 Även om grupperna är lika på gruppnivå före insatsen kan före-efterstudier vara av värde och ge säkrare resultat.
- 23 Den som vill läsa en mer omfattande och detaljerad beskrivning av olika upplägg hänvisas till Shadish, Cook, Campbell: Experimental and Quasi-Experimental Designs eller Isaac och Michael: Handbook in Research and Evaluation.
- 24 I bilaga 2 ges en förteckning över ett antal effektutvärderingar, sorterade efter typ av upplägg.
- 25 Se Public Policy and Administration 2002, volym 17, nummer 3.
- 26 Department of Education, US. Federal Register/Vol. 68, No. 213/Tuesday, November 4, 2003.
- 27 Före-efterstudier är även vanliga för att bedöma effekterna av trafiksäkerhetsinsatser.
- 28 Metoder för tidserieanalys behandlas i de flesta statistiska grundböcker.
- 29 BRÅ, Rapport 2003:11: Kameraövervakning i brottsförbyggande syfte.
- 30 I bilaga 1 ger vi ytterligare ett exempel på efterstudie, en utvärdering av effekterna av Skolverkets utbildningsinspektion.

- 31 I boken Svensson, Lennart et al: Projekt eller långsiktiga effekter? redovisar författarna en processorierad metod för att i dialog med de projektansvariga bedöma effekterna av ett projekt.
- 32 Institutet för personal- & företagsutveckling (IPF): Utvärdering av statens kvalitets- och kompetensråd (KKR) introduktionsprogram för nyanställda handläggare i statlig tjänst, 2002.
- 33 Rehabilitation of young immigrants in primary care - A comparison between two treatment models, Monica Löfvander et al, Scand J Prim Health Care 1997; 15:123-28.
- 34 Efficacy trial of acellular pertussis vaccines, Final Report 1991 - 2003, Swedish Institute for Infectious Disease Control (Smittskyddsinstitutet).
- 35 Det går också under benämningen "Experiment med matchad kontroll".
- 36 Institutet för social forskning, Stockholms universitet och Senter for høyere studier, Det Norske videnskapsakademi, 2001: Effekter av lördagsöppna systembolagsbutiker. Uppföljning av de första 17 månaderna.
- 37 Institutet för arbetsmarknadspolitisk utvärdering (IFAU), Rapport 2004:17.
- 38 Informationskampanjen om HIV/AIDS - samhällsinformation som styrmedel, RRV 1988:1836.
- 39 De utvärderingar som görs av Statens beredning för medicinsk utvärdering (SBU) utgörs till största delen av metaanalyser.
- 40 Nutek, R 2004: 15, Transportbidraget - en analys av motiv, effekter och alternativ.
- 41 Eurofutures/L Widell Consulting: Utvärdering av Skolverkets arbete med regeringsuppdraget Kvalitetsförbättrande insatser för gymnasieskolans yrkesutbildningar, 2003.
- 42 De sju stegen kan i praktiken komma i en annan ordning. I ett fall där budgeten är starkt begränsad kan till exempel resursöverväganden komma i det första steget.
- 43 Se Peter Esaiasson et al.: Metodpraktikan.
- 44 Det är inte självklart vilka kostnader som ska ingå i utvärderingskostnaden. I testet av vaccin är slutprestationen från testet dels en utvärdering, dels en åtgärd i form av att ett antal barn har vaccinerats. I övriga fall gäller kostnaden själva utvärderingen och inte motsvarande åtgärd.
- 45 Statens beredning för medicinsk utvärdering (SBU) arbetar med kriterierna starkt vetenskapligt underlag, måttligt starkt vetenskapligt underlag, begränsat vetenskapligt underlag samt otillräckligt vetenskapligt underlag. Se vidare Mona Britton: Så graderas en studies vetenskapliga bevisvärde och slutsatsernas styrka, Läkartidningen, volym 97, nr 40.
- 46 Syftet med utvärderingar kan vara att kontrollera och bedöma verksamheten mot till exempel mål och riktlinjer (kontrollutvärdering eller summativ utvärdering). Syftet kan också vara att främja och verka för att verksamheten förbättras (främjandeutvärdering eller formativ utvärdering).
- 47 Skolverkets utbildningsinspektion - ger den några effekter? Effekttvärdering genomförd av Ekonomistyrningsverket på uppdrag av Skolverket (2006).
- 48 Utvärderingen genomfördes i praktiken som två delutvärderingar. För att förenkla framställningen handlar denna bilaga enbart om den första delstudien.
- 49 Redovisningen bygger till stor del på erfarenheter från våra exempel i kapitel 2. Vi har också kompletterat denna bild med ett antal välkända metodproblem. För en bredare presentation av metodproblem hänvisar vi till Peter Esaiasson et al.: Metodpraktikan, Isaac, Stephen och Michael, William: Handbook in Research and Evaluation, Shadish, William R.; Cook, Thomas D. och Campbell, Donald T.: Experimental and Quasi-Experimental Designs samt till Riksrevisionsverket 1996:6: Att mäta effekter.
- 50 Se Peter Fredriksson: Om bruk och missbruk av experiment, ur Ekonomisk Debatt 7/2005.



**För mer information se Utvärderingsfrågor på
ESV:s webbplats – www.esv.se**

Handledningen kan beställas från ESV

Tel: 08-690 44 30

Fax: 08- 690 45 10

E-post: publikationsservice@esv.se

Skrifter om utvärderingar

ESV ger ut en serie skrifter om utvärdering. I denna hittar du även:

ESV 2005:26 Att beställa utvärderingar

ESV 2006:7 Måluppfyllelseanalys

Författare: Kajsa Nilsson och Lennart Widell

Redaktör: Anneli Wicksell

Kontakt: utvardering@esv.se

Layout och omslag: ETC

Foto: Peter Lilja

Tryck: Edita Sverige AB

ESV 2006:8

ISBN 91-7249-211-2

FÖR ATT UTVECKLA den statliga sektorn är det viktigt att ompröva verksamheter och bedöma effekterna av dem. Vilka effekter får till exempel en satsning på minskad arbetslöshet eller på regionalpolitiskt stöd till Norrland? Denna skrift ger ett underlag för att med hjälp av utvärdering kunna bedöma effekterna.

I handledningen presenteras ett antal olika upplägg för hur en effektutvärdering kan genomföras och vi diskuterar vilka faktorer som man bör ta hänsyn till vid valet av upplägg. Det gäller till exempel typ av åtgärd och vilka effekter åtgärden är tänkt att leda till. Ytterligare faktorer är behovet och tillgången på kompetens, tid och resurser samt kravet på säkerställda resultat.

Ekonomistyrningsverket utvecklar och implementerar en effektiv och ändamålsenlig ekonomistyrning som syftar till

- god kontroll av statens finanser.
- resursfördelning i enlighet med politiska prioriteringar.
- hög produktivitet och effektivitet.

Vi är statens experter på resultatstyrning och finansiell styrning och ansvarar för god redovisningssed. Vi har unika kunskaper om statens finanser och erbjuder kostnadseffektiva administrativa stödsystem.

Vårt arbete omfattar alla nivåer i staten och vi arbetar i nära samverkan med främst Regeringskansliet och myndigheterna. Vi bedriver även internationell verksamhet.



EKONOMISTYRNINGSVERKET

Drottninggatan 89, Box 45316, 104 30 Stockholm
Telefon 08-690 43 00 • Fax 08-690 43 50 • www.esv.se