

Secretariat for Analysis of  
Swedish Development Assistance

## *Några aspekter på kvaliteten hos utvärderingar av svenskt biståndssamarbete*

*Genomgång av databas upprättad vid Sekretariatet för Analys av  
Utvecklingssamarbete*

Kim Forss  
Stockholm, November 1994

**S A U**

**Sekretariatet för analys av  
utvecklingssamarbete**

**S A S D A**

**Secretariat for Analysis of Swedish  
Development Assistance**

Regeringen har tillsatt en kommitté med uppgift att analysera effekterna av och effektiviteten i det svenska biståndet. För att genomföra uppdraget inrättades den 1 mars 1993 ett sekretariat, SAU.

Sekretariatet, som har ett mandat fram till slutet av 1994, genomför policyanalyser och utvärderingar av det svenska biståndet, dess myndigheter och funktionsmekanismer. SAU:s arbete prioriterar globala studier på landnivå, sektoriella och ämnesinriktade analyser samt myndighetsspecifika utredningar som kan fungera som underlag för beslut om utformningen av det framtida svenska utvecklingssamarbetet. Bl.a. genomför SAU en utvärdering av Sveriges samarbete med Öst- och Centraleuropa.

SAU:s utgångspunkt för arbetet är att söka bättre förståelse av utvecklingsmekanismer i syfte att förstärka det svenska biståndets effekter och dess effektivitet mot bakgrund av de av riksdagen fastställda målen: resurstillväxt, ekonomisk och social utjämning, ekonomisk och politisk självständighet, demokratisk samhällsutveckling, framsynt hushållning med naturresurser och omsorg om miljön.

SAU:s analysarbete bedrivs av sekretariatets egen stab och genom studier som beställs av specialister inom de av kommittén prioriterade områdena.

SAU är underställt en kommitté, vars ordförande är generaldirektör Ingemar Mundebo.

Medarbetarna i SAU är följande:

Ambassadör Claes Sandgren, sekretariatschef  
Ekon dr Eva Julin, sekreterare  
Ekon lic Enrique Ganuza, sekreterare  
Civ ing Per Johan Svenningsson, sekreterare  
Kerstin Sandling, biträdande sekreterare  
Gunilla Jernberg, biträdande sekreterare

Postadress: Box 164 18  
103 27 Stockholm

Besöksadress: Klarabergsgatan 23, 2tr  
Stockholm

Telefon: 08 - 791 21 80  
Telefax: 08 - 791 21 88

## **Några aspekter på kvaliteten hos utvärderingar av svenskt biståndssamarbete.**

**Genomgång av databas upprättad  
vid Sekretariatet för Analys av Utvecklingssamarbete.**

**En rapport beställd av Sekretariatet för Analys av  
Utvecklingssamarbete,**

**författad av**

**Kim Forss. Andante Consultants AB. November 1994.**

ISSN 1104-697X  
Regeringskansliets Reprocentral

## Sammanfattning.

Under sitt arbete har Sekretariatet för Analys av Utvecklingssamarbete samlat in en stor mängd utvärderingar från de svenska biståndsorganen. Dessa har systematiskt gåtts igenom, klassificerats enligt en gemensam modell och lagts in i en databas. Den här rapporten innehåller en sammanställning av materialet. Databasen innehåller 177 utvärderingar, vilka beskrivits i 75 variabler, som skall spegla olika aspekter av kvalitet.

Vad säger då beskrivningen av utvärderingsarbetet? Utvärderingarna är för det mesta gjorda av oberoende experter. I en fjärdedel av rapporterna fanns personal från biståndsorganet med i utvärderingen. I knappt en tredjedel av fallen var personal från det mottagande landet med i utvärderingen. Det verkar vara sällsynt att någon person med särskild kompetens inom utvärderingsmetodik finns med i arbetet.

Utvärderingarnas formella utseende betyder mycket för deras användbarhet. De flesta utvärderingar har en sammanfattning och klara rekommendationer, antingen riktade till biståndsorgan eller mottagaren. En stor majoritet är skrivna på engelska.

Vad gäller metoder, så arbetar utvärderarna huvudsakligen med intervjuer av projektpersonal. Det är mera sällan de använder andra analysmetoder, t.ex. då direkta observationer, analys av dokumentation m.m. De brukar inte redovisa undersökningsinstrument. 16% har en ordentlig genomgång av arbetsmetoder, och det är ungefär lika många som diskuterar sin undersöknings tillförlitlighet.

Det vanligaste föremålet för utvärdering är insatsmål och nytta, vilket utreds på ett tillfredsställande sätt av cirka hälften av de 177. Det är få utvärderingar som analyserar mera långsiktiga utvecklingsmål. Ungefär en tredjedel analyserar mottagarlandets behov av insatsen, lika många analyserar effektivitet. 17% har en tillfredsställande analys av kostnadseffektivitet.

Utvärderingarna får betraktas som svaga när det gäller övergripande biståndsfrågor. Det är cirka 20% som analyserar insatserna bärkraft, dvs deras förmåga att fortleva utan externt stöd. Databasen visar att de praktiska frågor som kan underbygga slutsatser om biståndets bärkraft ofta är svagt dokumenterade i det empiriska arbetet. Cirka 20% av utvärderingarna behandlar miljöfrågor och 10% tar upp kvinnors perspektiv på utveckling. Andra s.k. tvärfackliga frågor, som t.ex. demokrati och mänskliga rättigheter får mindre utrymme.

Utvärderingen var:	Förträfflig	Tillräckligt bra	Bristfällig	Vet ej
SASDA's databas	7%	52%	21%	20%
CIDA	23%	44%	33%	-
NORAD/UD	9%	54%	29%	8%

Som en avslutande kommentar graderades utvärderingarna av oberoende forskarassistenter (långt innan den här rapporten skrevs) på skalan i tabellen ovan. Eftersom det gjorts liknande studier av andra biståndsorgan kan de svenska utvärderingarna jämföras med dessa, vilket görs i tabellen ovan. Amerikanska USAID har dessutom gjort en liknande analys, med värden snarlika kanadensiska CIDA. De 177 utvärderingar som undersökts av SASDA är på en lång rad punkter jämförbara med norska, kanadensiska och amerikanska utvärderingar, men de är också sämre på vissa områden - främst vad gäller analys av effektivitet, kostnadseffektivitet och bärkraft.

Det bör nämnas att det inte går att påvisa några kvalitetsskillnader mellan de svenska organisationerna. Däremot är utvärderingarna av östbiståndet ofta något bättre enligt dessa indikatorer. Vidare visar databasen att utvärderingarna på det stora hela taget inte förändrats. Andelen utvärderingar som bedömts vara förträffliga är lägre nu än före 1980. Å andra sidan är det färre som klassificeras som bristfälliga. Det ingick inte i mitt uppdrag att föreslå några åtgärder för utvärderingsarbetet, men det är uppenbart att åtskilligt kan göras annorlunda och förhoppningsvis bättre.

## Innehållsförteckning.

<b>Inledning</b>	1
<b>Urval av utvärderingar</b>	4
Metod	4
Regioner	4
Typ av insats	5
Sektorfördelning	5
Insatsernas storlek	6
Insatsernas längd	7
Insatsernas kanal och typ av mottagare	8
Fördelning på biståndsorgan och rapporternas ålder	8
Övrigt	9
Slutsatser	9
<b>Utvärderarna - vilka är de och var kommer de ifrån?</b>	10
Slutsatser	12
<b>Utvärderingsrapporterna</b>	12
<b>Utvärderingsmetoder</b>	14
<b>Utvärderingsobjekt</b>	17
<b>Slutomdöme om utvärderingarnas kvalitet</b>	23
<b>Sammanfattande kommentar</b>	24
<i>Noter</i>	27

Bilaga 1. Uppdragsbeskrivning

Bilaga 2. Analysschema för utvärderingarna

Bilaga 3. Utskrift av databasen

Bilaga 4. Synpunkter på utveckling av databasen.

## Inledning.

Inom utvecklingsbiståndet företas en stor mängd utvärderingar i syfte att ge såväl allmänheten som de finansierande organisationerna information om hur insatserna fungerar och om de når sina mål. Det är utomordentligt viktigt att utvärderingssystemet ger tillförlitlig information. Möjligheterna att förbättra programmen och lära av erfarenheter är helt beroende av korrekt information om vad som händer och vilka resultat som uppnåtts.

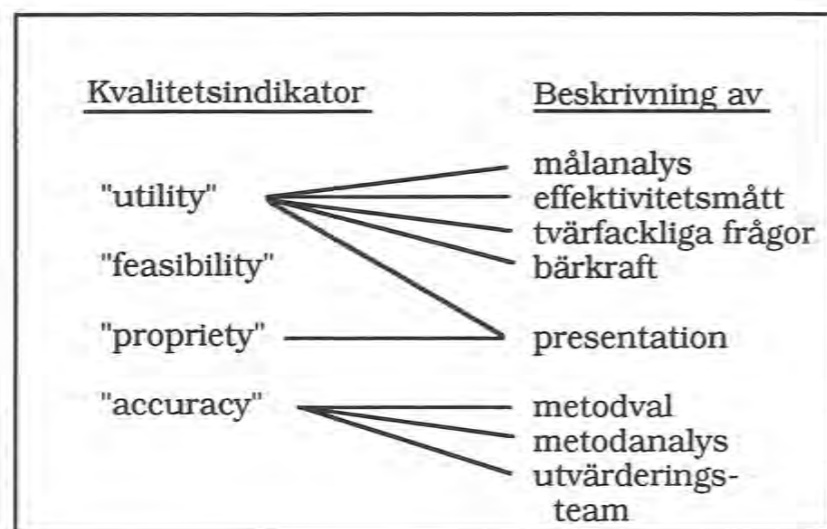
Vad vet man då om hur bra utvärderingssystemet är? Signalerna är motstridiga. Det finns studier som påvisar att man har relativt god information om resultat inom vissa delar av biståndet och det finns andra studier som påvisar stora brister inom andra delar av biståndet. Sekretariatet för Analys av Utvecklingssamarbete (SASDA) har genomfört organisationstudier av utvärderingssystem som påvisar brister, särskilt vad gäller den systematiska analysen av effekter och kopplingen mellan enskilda studier och övergripande mål för utvecklingssamarbetet (1).

Det är därför angeläget att på skilda sätt försöka "ta temperaturen" på utvärderingsarbetet. Syftet med den här rapporten är att mera formellt granska de skriftliga produkterna. Under sitt arbete har SASDA samlat utvärderingsrapporter som beskrivits enligt en standardiserad modell och lagts in i en databas. I skrivande stund innehåller databasen 177 utvärderingar som klassificerats enligt 75 variabler. Variablerna beskriver sektor, kostnad, land, typ av insats, huruvida insatsen utvärderats tidigare mm. samt information om vilka metoder som använts för utvärderingen, vilken typ av frågor som analyserats och hur utvärderarna fastställt nytta, effektivitet, mm.

Dessa data har lagts in av forskarassistenter enligt en mall utarbetad av SASDA - men vilken bygger på liknande databaser som tagits fram för analys av utvärderingar i Norge, Canada och U.S.A (2). Mallen återfinnes i bilaga 2. Det svenska biståndets utvärderingar kan således jämföras med motsvarande produkter från dessa länder för att se om de håller en godtagbar internationell kvalitet.

Frågan är då om databasen fångar in utvärderingarnas kvalitet? Det finns förvisso aspekter av kvalitetsbegreppet som inte är täckta och som knappast låter sig kvantifieras på detta sätt. Figur 1 ger en schematisk presentation av de kvalitetsindikatorer som utarbetats inom branschen (3). Det som fångas upp i den här rapporten är huvudsakligen indikatorer på tillförlitlighet, samt i viss mån användbarhet. Variablerna i mallen skall ses som indikatorer på i vilken riktning kvalitetsmått pekar. Som ett exempel; tillförlitlighet (accuracy) bedöms bl.a. genom att analysera vilka metoder utvärderingarna använder. Detta kan tas som en indikation på tillförlitlighet, men det är ju inte säkert att en utvärdering som arbetar på detta sätt når tillförlitliga resultat. Det är inte bara en fråga om vilka metoder som appliceras utan även hur det sker.

Figur 1. Kvaliteter hos utvärderingar.



Sannolikheten för att finna korrekta resultat ökar när utvärderarna arbetar metodologiskt medvetet, med redovisning av data och med en prövning av sina metoder. På samma sätt utgår vi i modellen från att en utvärdering är mera användbar som beslutsunderlag om den innehåller bedömningar av måluppfyllelse, effektivitet, bärkraft och andra frågeställningar (mer om dessa nedan).

Den formella presentationen ses delvis som en indikator på hur pass "etiskt" korrekt utvärderingen är, t.ex. om den redovisar uppdraget, om den är explicit kring sina antaganden och hypoteser, om den är skriven på ett för mottagaren begripligt språk. Vi ser också användbarheten som kopplad till frågan om det finns en sammanfattning av rapporten och om den innehåller explicita rekommendationer.

I det följande redovisas databasen och innehållet analyseras. Syftet är att beskriva hur de utvärderingar som genomförs inom det svenska biståndet ser ut, hur de kommit till, hur man arbetat och vad de ger för information. Den här rapporten är med andra orden rent deskriptiv i förhållande till utvärderingarna, den beskriver och diskuterar dem som produkter. Jag säger ingenting om själva biståndet och dess effekter.

Frågan är om man har anledning att tro någonting a priori om utvärderingarna? Eftersom det har gjorts en del studier tidigare går det att lansera ett antal hypoteser och se om det finns anledning att fortfarande tro på dem, eller om databasen ger annan information om utvärderingarna. Här följer 10 hypoteser om "hälsotillståndet" hos utvärderingarna. Jag vill betona att de är starkt subjektiva. Läsaren måste tänka efter själv och bedöma om de är rimliga; om det är för högt eller för lågt ställda krav.

*Hypotes 1. En majoritet av utvärderingarna arbetar på ett metodologiskt tillfredsställande sätt.*

*Hypotes 2. En majoritet av utvärderingarna har tillfredsställande bedömningar av måluppfyllelse.*

*Hypotes 3. En majoritet av utvärderingarna har tillfredsställande bedömningar av effekter.*

*Hypotes 4. En majoritet av utvärderingarna har tillfredsställande bedömningar av kostnadseffektivitet.*

*Hypotes 5. En majoritet av utvärderingarna har tillfredsställande bedömningar av biståndets bärkraft.*

*Hypotes 6. En majoritet av utvärderingarna har tillfredsställande bedömningar av s.k. "cross-cutting issues".*

*Hypotes 7. De svenska utvärderingarna är ungefär lika bra som de utvärderingar som görs i Norge, U.S.A och Canada.*

*Hypotes 8. Utvärderingarna har blivit bättre över tiden, dvs en allt större andel karaktäriseras som tillfredsställande enligt kvalitetsindikatorerna.*

*Hypotes 9. Det finns inte någon skillnad mellan de olika biståndsorganisationerna vad gäller kvalitetsindikatorerna.*

*Hypotes 10. Det finns inte någon skillnad mellan östeuropabiståndet och andra delar av biståndet vad gäller kvalitetsindikatorerna.*

Nu är ju inte det här en hypotesprövning i formell statistisk bemärkelse, utan hypoteserna används som ett sätt att strukturera diskussionen och de uttrycker en förväntan kring vad analysen skall visa av utvärderingarnas kvalitet. När jag skriver "majoritet" och "tillfredsställande" så menar jag att de forskarasistenter som (helt oberoende av mig) har graderat utvärderingarna på skalan "adequate", "limited", "nothing" och "irrelevant", skall ha valt svaret "adequate" minst lika många gånger som svaren "limited" och "nothing" tillsammans.

Korrelationen mellan olika variabler har gjorts med hjälp av korstabuleringar. När jag påvisar ett samband mellan variabler så har det fastställts med statistiska metoder för sambandsanalys och resultaten har befunnits signifikanta. De främsta hoten mot den här undersökningens tillförlitlighet består, för det första, av att det endast är en mindre del av biståndets utvärderingar som finns med, och vi vet inte hur resten av utvärderingarna ser ut, samt, för det andra, att de som lagt in data kan ha värderat rapporterna olika. Bilaga 3 innehåller dock en fullständig utskrift av databasen, så läsaren kan själv bilda sig en uppfattning om klassificeringarnas rättvisa genom att jämföra med respektive utvärdering.

## Urval av utvärderingar.

### Metod.

Som nämnts ovan ingår det 177 rapporter i databasen. Dessa utgör ett urval av rapporter. Den äldsta kommer från år 1969 och den senaste är från 1994. Hur många utvärderingar har gjorts inom det svenska biståndet under de åren? Tyvärr är det ingen som vet. Utvärderingsverksamheten är så pass fragmenterad och det finns ingen instans som har överblick över helheten. Det är inte ens säkert att de skilda organisationerna själva kan ge svar på hur många utvärderingar de beställt - dvs svar utan att göra kostnadskrävande undersökningar.

Om vi antar att BITS, SAREC och SwedeCorp vardera beställer ca 10 utvärderingar per år, och sedan till detta lägger ca 40 utvärderingar från SIDA, skulle vi få kring 70 utvärderingar årligen, dvs under perioden från 1980 till idag ca 1.000 utvärderingar. De utvärderingar som ligger i databasen utgör då ett urval på knappt 20%. Hur har urvalet gjorts? Pragmatiskt. SASDA har genomfört organisationsstudier, länderstudier och sektorstudier. Inom ramen för dessa har biståndsorganen översänt de utvärderingar som befunnits relevanta för respektive studie. Valet av utvärderingar i databasen är därmed till stor del påverkat av inriktningen på SASDA's verksamhet i stort.

### Regioner.

Afrika dominerar utvärderingarna i databasen. 68% av rapporterna hänför sig till den kontinenten. Av det bilaterala, länderfördelade biståndet har ungefär 52% gått till Afrika under den här tiden, vilket gör att biståndet till Afrika framstår som något överrepresenterat. Enbart 4% av utvärderingarna avser biståndsprogram i Asien, att jämföra med att ca 23% av det bilaterala biståndet gått till länder i Asien. Det beror på att SASDA inte gjort några landanalyser av asiatiska mottagarländer, och därför inte samlat in så många utvärderingar därifrån. Däremot avser 14% av utvärderingarna insatser i Latinamerika, vilket kan jämföras med att 12% av det bilaterala, regionfördelade biståndet går till den latinamerikanska kontinenten. (4)

8% av utvärderingarna avser insatser i Europa, främst då i Baltikum, Polen och Ryssland. Eftersom biståndet till Öst- och Centraleuropa fördelats utanför biståndsanslagen är det inte lika relevant att jämföra dess andel av det totala biståndet. Jag vill dock peka på att räknat i kronor är biståndet till Östeuropa ungefär lika stort som biståndet till Latinamerika, men det senare har dock ca 60% flera utvärderingar (åtminstone som det ser ut i det urval som finns i SASDA's databas).

Sedan den geografiska fördelningen är gjord återstår 8 utvärderingar som inte går att klassificera i förhållande till länder eller regioner. De kan avse globala program, bistånd till multilaterala organisationer, eller vara tematiska, t.ex. avseende miljöprogram, som finns i flera regioner.

### Typ av insats.

En utvärdering kan ha flera skilda angreppssätt beroende på vad det är som studeras. I biståndets - och utvärderingarnas - gryning var det vanligast att man talade om enstaka projekt. De utgjorde den organisatoriska ramen för insatser och det var också projekt som utvärderades. Sedan slutet av 1970-talet har det funnits en önskan att lyfta blicken från enskilda projekt för att istället sätta sektorer, program eller länder i fokus. Det har dock varit lättare sagt än gjort.

I databasen avser 50% av utvärderingarna enskilda projekt, 37% avser stöd till sektorer och insatser i programform, 7% utgör utvärderingar av länder och slutligen finns det 8 utvärderingar (5%) som antingen hänför sig till organisationer eller som inte låter sig kategoriseras på samma sätt som de övriga. Jag antar att de som avser organisationer, vilket är 6 stycken, då behandlar någon av de multilaterala organisationerna, eller möjligtvis stöd genom folkrörelseorganisationer.

I det här fallet släpar utvärderingspraxis efter utvecklingen av biståndet i stort. Det är säkerligen mindre än 10% av det totala biståndet som fördelas till projekt. Det mesta organiseras inom ramen för landprogram, sektorstöd, eller genom någon annan programmeringsform. Det kan finnas många förklaringar till att utvärderingarna fortfarande inriktas på projekt, inte minst att det är metodiskt lättare att ta tag i ett projekt.

När klassificeringarna gjordes var det sagt att frågan om typ av insats skulle beaktas ur mottagarperspektiv, det kan medföra att andelen projekt är större än vad de är ur givarperspektiv. Det SIDA kallar sektorstöd kanske kallas projekt i Tanzania - för att nämna ett exempel.

### Sektorfördelning.

Tabell 1 visar hur utvärderingarna i databasen fördelas på insatser. Det är anmärkningsvärt att nästan 40% av utvärderingarna behandlar insatser inom utbildningsområdet. Det förklaras delvis av att forskningsbiståndet klassificerats som undervisningssektor - och SAREC har gjort många utvärderingar. Totalt utgör sektorn drygt 10% av biståndet, och har följaktligen en stor andel utvärderingar. Det kan jämföras med det relativa antalet utvärderingar som behandlar infrastrukturprojekt. Sektorn utgör kring 15% av de totala biståndsvolymererna, men enbart 4% av utvärderingarna.

Både industribiståndet, hälsobiståndet och biståndet till offentlig förvaltning står för ungefär lika stor andel av utvärderingarna som respektive sektors del av biståndet. NGO-stödet är betydligt mindre utvärderat. Under åtskilliga år har närmare 1 miljard, dvs ca 7% av biståndet, fördelats genom sådana organisationer, men som vi ser har det bara gjorts en utvärdering (som finns med i databasen).

**Tabell 1. Sektorfördelning av utvärderingar i databasen.**

	<u>Antal</u>	<u>Procent</u>
Utbildning	61	39
Jordbruk	28	18
Industri	24	15
Hälsovård	13	8
Förvaltning	13	8
"Cross-cutting issues"	7	4
"Allmänt"	3	2
Tjänster	1	1
NGO-stöd	1	1
	158	100

De flesta utvärderingarna låter sig klassificeras enligt gängse sektorer. Det är bara 7 utvärderingar som avser vad som kallas "cross-cutting issues", dvs frågeställningar som skär över flera ämnesområden. Det är lågt, eftersom det i åtminstone 10 år talats om önskvärdheten av att bredda jämförelsematerialet, öka möjligheten till korsbefruktning genom att föra erfarenheter från ett område till ett annat. Dessutom är många av de svåra frågorna gemensamma över sektorer, t.ex. vad avser kunskapsöverföring, personalbistånd, hantering av konsulter mm.

#### *Insatsernas storlek.*

I databasen beskrivs den utvärderade insatsens storlek i tre kategorier; mindre än 10 miljoner, mellan 10 och 50 miljoner, samt mer än 50 miljoner. Dessutom tillkommer svarsmöjligheten att storleken inte är relevant eller att den är okänd. Eftersom de som lagt in svaren enbart haft utvärderingarna att tillgå, bygger klassificeringen på vad utvärderingen säger om programmets storlek. I en del fall redovisar inte utvärderingarna någon sådan information - faktiskt inte i 34% av utvärderingarna. Det är något förvånande, för om man inte redovisar utbetalade belopp kan man ju inte heller göra någon typ av kostnads- eller effektivitetskalkyler.

För övrigt är 33% av insatserna som utvärderats större än 50 miljoner kronor, 22% är mellan 10 och 50 miljoner kronor, samt 11% är under 10 miljoner kronor. Notera dock att dessa siffror anger hela insatsens storlek, det är inte den årliga utbetalningen. Eftersom en hel del av

utvärderingarna avser insatser som pågått länge, är det möjligt att de årliga utbetalningarna varit avsevärt mindre. Slutligen är det viktigt att komma ihåg att beloppen enbart avser den av Sverige finansierade delen av insatserna. Det är ju möjligt att aktiviteten i sig kostar mer, men då finansieras av andra organ/länder.

#### *Insatsernas längd.*

Som nämnts bör insatsernas storlek jämföras med den tid de pågår. Den största delen av de insatser som utvärderats här har pågått mer än 10 år (nästan 40%). 27% har pågått mellan 5 och 10 år; 5% mellan 2 och 5 år; samt 9% mindre än 2 år. För 5 av utvärderingarna kunde det inte sägas hur länge insatserna pågått eller så var frågan irrelevant. Tabell 2 presenterar en sammanställning av insatsernas längd och storlek. (Notera dock att nio utvärderingar inte har kategoriserats längs någon av de här två variablerna; det finns alltså ett bortfall av nio rapporter.)

Eftersom en stor del av biståndet är långsiktigt så reflekterar utvärderingarna också detta. Kanske är det ändå något förvånande att så många insatser pågått mer än tio år, men å andra sidan kan också antalet långa projekt vara överrepresenterat bland utvärderingarna. Det är troligtvis mera angeläget att beställa utvärderingar för kontroll och beslutssyfte kring insatser som pågår länge än för sådana som är korta och tidsbegränsade.

**Tabell 2. Insatsernas storlek i kronor och tid i år.**

	<u>1-10 Mkr</u>	<u>10-50 Mkr</u>	<u>&gt;50 Mkr</u>	<u>Okänt</u>	<u>Summa</u>
< 2 år	5	6	0	3	14
2-5 år	10	14	12	7	43
5-10 år	2	15	14	11	42
> 10 år	1	6	33	22	62
Okänt	1	-	1	3	5
Summa	19	41	60	46	168

Mot bakgrund av tidsaspekten kan det vara intressant att notera att minst en tredjedel av projekten utvärderats en eller flera gånger tidigare. Jag skriver "minst" eftersom det i 27% av fallen inte är känt om de utvärderats tidigare, men i 40% av fallen har det definitivt inte förekommit någon tidigare utvärdering.



#### *Insatsernas kanal och typ av mottagare.*

Det är knappast förvånande att 80% av utvärderingarna avser bilateralt bistånd. Visserligen utgör det multilaterala biståndet för det mesta ca 30% av den totala biståndsvolymen, men det multilaterala biståndet har aldrig varit utsatt för samma utvärderingsnit som övriga delar. Där är det oftast den mottagande organisationen som står för utvärderingen av insatser som den genomför.

De övriga insatskanalerna innehåller 16 utvärderingar; 2 av dessa hänförs till kommersiella kanaler, 5 till NGO-kanaler och 9 till övriga. Det visar att utvärderingen av NGO aktiviteter är något högre än vad som skulle förmodas av sektorfördelningen som redovisades ovan. Det är inte helt klart vad som menas med kommersiella kanaler. De skall troligtvis ses som en underavdelning till den bilaterala kanalen (vilket även bör vara fallet för de övriga 16 utvärderingarna).

Databasen innehåller också information om mottagaren, dvs huruvida insatsen går direkt till den privata sektorn, till den offentliga sektorn, eller till en blandning. De flesta utvärderingarna (47%) i databasen avser insatser för den offentliga sektorn - vilket är naturligt med tanke på sektorfördelningen ovan. 42% kunde inte klassificeras, eller beskrevs som blandat privat-offentligt. Endast 12% av utvärderingarna avsåg insatser där mottagaren beskrevs som tillhörande den privata sektorn. Förmodligen motsvarar detta biståndets fördelning till skilda typer av mottagare - om något är kanske andelen mottagare i den privata sektorn något överrepresenterat.

#### *Fördelning på biståndsorgan och rapporternas ålder.*

Eftersom SIDA är den äldsta och största av biståndsorganisationerna är det inte förvånande att de flesta utvärderingarna kommer därifrån. 56% av rapporterna i databasen var beställda av SIDA, 25% av SAREC och 15% av BITS. Tre utvärderingar var beställda av Utrikesdepartementet och en av SwedeCorp. Sedan fanns det några utvärderingar som inte var klassificerade efter någon av myndigheterna. Det kan bero på att de var beställda av en utländsk biståndsmyndighet (men ändå gav information om insatser finansierade av svenska biståndsmedel) eller av någon annan myndighet (t.ex. inom Östeuropa biståndet, där även andra myndigheter haft funktioner liknande de fyra biståndsorganen).

Inledningsvis nämndes att den äldsta rapporten var från 1969. I databasen har åldern indelats i fyra kategorier; sådana från före 1980, mellan 1980 och 1984, mellan 1985 och 1989, samt efter 1990. 42% av utvärderingarna kommer från tiden efter 1990, dvs de är relativt nya. 32% kommer från tiden mellan 1985 och 1989, 18% kommer från tiden mellan 1980 och 1984, samt 8% kommer från tiden före 1980.

Den absolut största delen av utvärderingarna i de två äldre kategorierna kommer från SIDA. Under perioden 1985 till 1989 finns det även ganska många utvärderingar från SAREC. De andra organisationerna har huvudsakligen med utvärderingar från 1990 och senare. Det säger sig självt att utvärderingarna av östbiståndet är från 1990-talet, och dessutom ganska sent där.

#### *Övrigt.*

De flesta utvärderingarna har klassificerats som "mid-term", dvs de har genomförts under tiden som insatsen pågick. Troligtvis har de utgjort beslutsunderlag inför nya avtal om insatserna. Endast 10% av utvärderingarna är gjorda efter det att insatsen avslutats. Det innebär att utvärderingarna sällan innehåller information om hur det ser ut på fältet när samarbetet är slut.

I en del fall innehåller samarbetsavtalen som reglerar en insats överenskommelser om att insatsen skall utvärderas efter en viss tid. 17% av utvärderingarna i databanken tillkom på grund av att man redan från början kommit överens om att göra en utvärdering.

#### *Slutsatser.*

Utvärderingarna i databasen avser många skilda typer av insatser, både inom det bilaterala och multilaterala området. Det finns dock en slagsida mot relativt små och långvariga insatser i projektform, företrädesvis sådana som ägt rum i Afrika, inom den offentliga sektorn.

Utvärderingarna i databasen behandlar därmed inte biståndet i sin helhet. Där återfinns relativt få utvärderingar avseende bistånd till Asien och Europa, samt relativt få utvärderingar av kortsiktiga insatser. Där återfinns inte många utvärderingar av det multilaterala biståndet. Påfallande få avser "nya" utvecklingsområden inom biståndet, t.ex. stöd för mänskliga rättigheter, multisektorstöd, programstöd av skilda slag.

Frågan är då om databasens utvärderingar är typiska för helheten av svensk biståndsutvärdering? Det är svårt att uttala sig med någon större säkerhet om den saken. Troligtvis finns det rapporter som täcker in en del av de nyare områdena. Vad gäller programstöd, katastrofbistånd och utvärdering av stöd för mänskliga rättigheter så finns det utvärderingar som inte är med i databasen.

Det kan vara värt att notera att det i många år påtalats att utvärderingsverksamheten borde ägnas mera åt övergripande frågeställningar, jämförelser mellan länder och sektorer, följa och bevaka nya strömningar inom "biståndsmodet" etc. (5) I den mån utsagan om utvärderingarnas huvudsakliga inriktning stämmer, så är det mot bakgrund av snart 20 års trägna försök att aktualisera, bredda, fördjupa och förnya verksamheten. Eftersom utvärderingsverksamheten i sig bör vara en källa till förnyelse och omprövning, är det förvånande

att den själv är så pass trögrörlig. Kanske är det något som förtjänar sin egen undersökning?

I de följande avsnitten skall närmare beskrivas hur utvärderingarna har genomförts, av vem och med avseende på vilka frågor. Inför den genomgången är det viktigt att läsaren håller i minnet det urval av utvärderingar som finns i databasen - och även tar ställning för sig själv till hur representativt det är för utvärderingar i allmänhet respektive för biståndet i sin helhet. När jag drar slutsatser och diskuterar svagheter respektive starka sidor hos utvärderingsverksamheten är det enbart mot bakgrund av data kring de 177 utvärderingarna i databasen.

### Utvärderarna - vilka är de och var kommer de ifrån?

Databasen innehåller information om utvärderarna; namnuppgifter och vilken institution/vilket företag de kommer ifrån. I det följande skall jag dock enbart kommentera de mera övergripande uppgifterna, som vilken typ av kunskap de representerar och vilka länder som är med i utvärderingsarbetet. Tabell 3 sammanställer materialet.

**Tabell 3. Uppgifter om utvärderingsteam.**

	Ja	Nej	Vet ej	Summa
Är biståndsorganet representerat på utvärderingsteamet?	23	67	10	100%
Är det mottagande landet representerat på utvärderingsteamet?	32	59	9	100%
Finns det någon med sektorkompetens på utvärderingsteamet?	63	4	33	100%
Finns det någon med utvärderingskompetens på utvärderingsteamet?	16	11	73	100%

En mycket omdiskuterad fråga i utvärderingssammanhang är huruvida personal från den organisation som är ansvarig för verksamheten skall vara med i utvärderingsarbetet. Är det t.ex. lämpligt att handläggare från SwedeCorp är med och tar fram de resultat om insatserna som föreläggs allmänheten, eller vem nu rapporten är avsedd för? Å ena sidan kan man hävda att slutsatsernas legitimitet komprometteras om de ansvariga handläggarna är med i utvärderingen. De skulle kunna ha ett intresse av att visa verksamheten i en bättre dager. Det är viktigt med oberoende granskning, och handläggarna är a priori inte oberoende.

Å andra sidan hävdas att utvärderingars användbarhet ökar om de direkt berörda beslutsfattarna är med och tar del av erfarenheterna. Utvärderingar är ett viktigt instrument för lärande i organisationerna, och den lär mest som är med i det empiriska och analyserande arbetet. Som siffrorna ovan visar är det dock relativt ovanligt att personal från biståndsorganisationerna är med i utvärderingsarbetet, det sker i knappt en fjärdedel av fallen.

I praktiken måste man givetvis gå en balansgång mellan kraven på oberoende och kraven på lärande/användbarhet. Bilden är mycket mera komplicerad än vad som antyds av det här resonemanget. På det stora hela taget är det knappast någon risk att det svenska biståndets utvärderingar hotas av att teamens oberoende är komprometterat. Kanske är det mera troligt att lärandepotentialen i utvärderingsarbetet inte utnyttjas i tillräckligt stor utsträckning.

Utvärderingars användbarhet - och tillförlitlighet - anses också ha ett starkt samband med en förståelse för de lokala villkoren. Dels är det viktigt att rätt tolka insatsernas miljö och de faktorer där som påverkar genomförandet, och dels är det viktigt att rätt förstå den lokala personalen, insatsens målgrupp mm. Därför bör man sträva efter att ha personal från det mottagande landet med i utvärderingsteam. Om jag gör en utvärdering i Indien skulle jag vara mån om att knyta minst en indier med god sektorkompetens till mitt utvärderingsteam - som ett exempel.

På en tredjedel av de utvärderingar som finns i databasen ingår personal från det mottagande landet. I många fall har det inte framgått eller kunnat avgöras med hjälp av namnen. Men i 60% av fallen finns det inte någon representant för det mottagande landet med (ordet representant skall inte tolkas formellt, det är inte fråga om någon officiell representant utan enbart huruvida den särskilda insikt och förståelse som infödd personal kan besitta finns representerat på teamet).

Vid en internationell jämförelse framstår detta som ganska typiskt. Utvärderingsverksamheten domineras av konsulter från hemlandet. Det kan tilläggas att branschen är helt dominerad av män. Det finns mycket få kvinnliga teamleaders och överhuvudtaget få utvärderare som är kvinnor.

Detsamma kan kanske sägas om närvaron av utvärderingsexpertis på utvärderingsteam; dvs det är sällan förekommande. Situationen är ungefär lika dålig i Norge, Canada och U.S.A. Frågan är dock vad som menas med utvärderingsexpertis? Om utvärdering definieras som en "tillämpning av samhällsvetenskapliga metoder för att analysera planering, genomförande och resultat", då skulle man troligtvis med utvärderingsexpertis mena någon person som genom sin utbildning eller erfarenhet behärskar sådana metoder.

Eftersom man knappast tillägnar sig vetenskapliga metoder på något annat sätt än genom vetenskapliga studier, är det med andra ord personer som fullgjort en forskarutbildning, eller åtminstone väsentliga

delar av en sådan, som avses. I den här studien konstateras faktum - det är sällsynt att personer med utvärderingskompetens är närvarande i utvärderingsarbetet. Det är givetvis en svaghet. Därmed är det inte sagt att biståndsorganisationerna skall lastas för det. Kanske finns sådana personer helt enkelt inte tillgängliga på arbetsmarknaden - eller åtminstone inte för den typen av korta anställningar som det här är frågan om (6). Det är därmed inte så förvånande att den kompetensen saknas.

Något mer förvånande är det att sektorkompetens inte är representerat i mer än 63% av fallen. I princip är det tre specifika kompetensområden som kan sökas; länderkunskap, sektorkunskap och metodkunskap. Säkert har det också många gånger ett värde att få en generalists synpunkter. Inom ramen för ett utvärderingsarbete skulle man dock förmoda att sektorkunskapen, vare sig det avser tropiskt jordbruk, standardiseringsrutiner, limnologisk forskning eller bankutbildning, väger tungt. Det är visserligen den typ av kompetens som oftast finns representerad på utvärderingsteamet, men den finns på knappt två tredjedelar av utvärderingarna. Jag hade a priori trott att sådan kompetens skulle finnas närvarande så gott som alltid.

#### Slutsatser

Utvärderingarna i databasen visar ett tydligt mönster. Teamen står oftast helt oberoende i förhållande till biståndsorganisationerna. Det är förhållandevis sällsynt att det mottagande landet för insatsen är representerat på utvärderingsteamet. Av de skilda kunskapsområden som i princip kan representeras på ett utvärderingsteam, var det metodologiska sämst tillgodosett. Slutligen kan vi konstatera att utvärderingsverksamheten i det svenska biståndet inte framstår som vare sig bättre eller sämre än andra länders.

#### Utvärderingsrapporterna.

Hur ser då utvärderingsrapporterna ut, den visuella produkten av arbetet? Databasen innehåller uppgifter om förekomsten av sammandrag, uppdragsbeskrivning, rekommendationer, språk, mm. Utvärderingsrapportens formella utseende är viktigt och betyder mycket för läsbarheten. Därmed är det inte sagt att alla rapporter behöver se likadana ut, och i många fall anser jag själv att standardiseringsarbetet nått en bra bit bortanför det rimligas gräns. Innehållet skall styra. Men vissa av de formella kraven är viktiga. Tabellen nedan ger ett sammandrag av hur utvärderingarna ser ut.

Svaren i tabellen talar för sig självt. Idealet vore naturligtvis att det aldrig förekom något annat än "ja" som svarsalternativ. I en del fall kanske det finns naturliga förklaringar till att man inte redovisar uppdragsbeskrivningar, redogör för arbetsplan eller skriver rekommendationer och sammanfattning. Men det är säkert inte mer än i någon enstaka procent som det vore befogat att utelämna dessa, för legitimitet och användbarhet så viktiga, aspekter.

**Tabell 4. Formalia kring utvärderingsrapporterna**

	Ja	Nej	Summa
Är uppdragsbeskrivningen med i rapporten?	62	38	100%
Finns utvärderarnas arbetsplan med i rapporten?	30	70	100%
Innehåller rapporten rekommendationer?	86	14	100%
Innehåller rapporten rekommendationer explicit riktade till biståndsorganet?	80	20	100%
Innehåller rapporten rekommendationer explicit riktade till mottagarna?	51	49	100%
Innehåller rapporten någon "executive summary"?	77	23	100%

Frågan är då om utvärderingarna har blivit bättre i dessa avseenden? Nej, i de flesta fall är det ingen skillnad mellan rapporter skrivna före 1980 och de som är skrivna senare. I två avseenden har de blivit bättre; det var bara 10% av utvärderingarna gjorda före 1980 som innehöll en sammanfattning, mot 37% av dem skrivna efter 1990. Det är också något vanligare att rapporterna innehåller uttryckliga rekommendationer efter 1990.

En stor majoritet, 88%, av rapporterna är skrivna på engelska. 8% är skrivna på svenska och 4% på det mottagande landets språk (antagligen är det då spanska eller något annat språk som även är läsbart i den svenska administrationen). Det är bra att så många är skrivna på engelska - och egentligen finns det ingen anledning att någon utvärdering alls är skriven på svenska. Eftersom biståndet definitionsmässigt är internationellt finns det alltid viktiga intressenter som inte kan ta del av en rapport på svenska. (Jag är medveten om att kasta sten i glashus.)

Det vore möjligt att se till att åtminstone en sammanfattning av en rapport på svenska finns tillgänglig på engelska - eller på det mottagande landets språk. Ingen av de här har haft en sammanfattning på annat språk än rapportens. Utvärderingarna i Norge, Canada och U.S.A är lika de svenska, dvs de innehåller både uppdragsbeskrivningar och sammanfattning. De är sammantaget mycket ovanligt att något annat språk än engelska används.

## Utvärderingsmetoder.

Låt oss då se hur utvärderarna arbetar? Det första "formella" kravet på ett metodologiskt tillfredsställande arbete är att den som skriver redogör för sitt tillvägagångssätt. Genom att presentera metoderna blir det möjligt för läsaren att bedöma tillförlitligheten i resultaten. I den mån man kan ha ett intresse av att generalisera lärdomarna för studien säger en ordentlig metodgenomgång i vilken utsträckning något sådant är möjligt.

Det vanligaste är att utvärderingarna helt enkelt inte innehåller någon metodanalys. Det är enbart 16% som innehåller en metoddiskussion som har kunnat klassificeras som "tillräcklig". Med tillräcklig menas att det finns en presentation av utvärderingsmetoderna, att datainsamlingen kopplas till de frågor man vill belysa, och att det finns en diskussion kring urvalsförfarande, bortfall, tolkning av svar och en presentation av de analysmetoder som används.

Vad innebär då en begränsad metodgenomgång? När en utvärderingsrapport säger någonting om en mindre del av metoden, t.ex. att det sänts ut en enkät till 300 hushåll, eller att man genomfört 20 intervjuer med ostrukturerade frågor, och då beskriver vilka befattningar de intervjuade haft, då ser vi det som element i en metoddiskussion, men inte tillräckligt för att kunna bedöma metodens - och därmed utvärderingens tillförlitlighet.

**Tabell 5. Utvärderingarnas genomgång av metodfrågor.**

	Tillräcklig	Begränsad	Ingenting	Irrelevant
Innehåller utvärderingen en analys av metodfrågor och presentation av hur metodiska problem har lösts under arbetet?	16%	35%	47%	2%

Finns det då något samband mellan de variabler som har med utvärderingsmetod att göra och andra? Ja, det finns några sådana samband. För det första är det vanligare att de utvärderingar som har haft personal med utvärderingskompetens (enl. definitionen ovan) har ordentliga redogörelser för metoden och diskussioner av validitet och reliabilitet. Det är mera vanligt att de arbetar med hypoteser och med vetenskapliga metoder. Men det är förstas en självklarhet.

Det finns en viss skillnad mellan organisationerna. 46% av BITS utvärderingar hade "tillräckliga" metodgenomgångar, mot bara 20% av SIDA:s utvärderingar. Däremot fanns det inget samband över tiden, dvs det finns inga belägg för att utvärderingar blivit bättre med tiden. Det fanns inte någon skillnad avseende sektorer.

Utvärderingarna av östbistånd är något bättre än övriga. Det var 20% av dem som befanns ha "tillräckligt" bra metoddiskussion. Utvärderingar av bistånd till Afrika var också metodologiskt starkare i genomsnitt; det var något fler av dem som var grundade på vetenskapliga metoder.

Men kanske utvärderarna arbetar metodologiskt kunnigt även om metoden inte redovisas? Ett sätt att arbeta metodiskt är att klargöra hypoteser om effekter och samband, och att använda sådana hypoteser för att strukturera framställningen? Det kan också vara användbart och önskvärt att klargöra eventuella förväntningar eller antaganden i förväg. I 13% av utvärderingsrapporterna fanns sådana, men övriga 87% innehöll inte någon presentation av antaganden eller hypoteser.

I vad mån använder utvärderingarna metoder som t.ex. modeller för att bedöma samhällsekonomisk effektivitet, eller miljökonsekvensanalyser, eller modeller för att bestämma kunskapsöverföring och utbildningseffekter? 15% av utvärderingarna presenterade någon form av explicit redovisad teknik för att söka svar på komplicerade frågeställningar, och följaktligen var det 85% som inte presenterade några sådana tekniker.

I tabell 6 görs en mera grundlig analys av hur utvärderarna faktiskt arbetar, dvs vilken typ av data de samlar in och hur. Det vanligaste sättet att samla in data är genom intervjuer med projektpersonal och mottagargrupper. Många har också intervjuat personal i biståndsorganisationen - åtminstone i begränsad utsträckning. Det är påfallande få som intervjuat representanter för andra biståndsorganisationer.

Klassificeringen pekar på brister i användningen av dokumentation från biståndsorganisationen. Däremot förekommer annan dokumentation ofta i tillräcklig omfattning. Slutligen är direkt observation ett ovanligt sätt att samla in data, och i den mån det utnyttjas så redovisas inte några instrument för hur man går tillväga. Så var det också i Canada och U.S.A.; mindre än hälften av utvärderingarna använde sig av observation som metod (sannolikt beroende på att utvärderingsteam har så kort tid till sitt förfogande).

Frågorna är till viss del formulerade så att de anknyter till metoder för utvärdering av projekt och program. Inledningsvis påtalade jag att en stor del av utvärderingarna avser just insatser i den organisatoriska formen. En del av rapporterna är dock av mera övergripande karaktär, och kanske är det då inte relevant att intervjua projektpersonal eller att observera projektarbete på plats. Det finns alltså en viss värdering genom frågorna till förmån för den mera traditionella insatsformen och tillhörande utvärderingsmetodologi. Det är inte helt säkert att svars-kategorin "irrelevant" fångat alla tillfällen där en viss form av datainsamling faktiskt är irrelevant.

**Tabell 6. Utvärderingarnas metoder för datainsamling.**

	Tillräcklig	Begränsad	Ingenting	Irrelevant
I vilken utsträckning bygger utvärderingarna på:				
analys av biståndsorganens dokumentation?	28%	35%	35%	2%
analys av annan dokumentation?	44%	24%	30%	2%
Intervjuer med personal från biståndsorgan?	29%	24%	43%	4%
Intervjuer med projekt personal?	56%	23%	16%	5%
Intervjuer med mottagare?	48%	19%	23%	-
Intervjuer med andra biståndsgivare?	11%	3%	81%	5%
Direkt observation av projektarbete?	27%	24%	38%	11%
Finns intervjuprotokoll eller guidelines?	13%	3%	80%	4%
Finns instrument för observationsdata?	5%	8%	72%	15%

Syftet med en noggrann metodgenomgång är att författaren själv skall kunna redogöra för hur pass tillförlitliga och relevanta slutsatserna i utvärderingen är; dvs dess validitet och reliabilitet. Genom att redovisa metoder, och föra en sådan diskussion, blir det också möjligt för läsaren att bilda sig en egen uppfattning av dessa egenskaper. Tillsammans kan detta leda till ett ökat förtroende för utvärderingen, alternativt till en mera insiktsfull debatt på en hög nivå om vad utvärderingen säger.

Både presentation av metod och analys av densamma leder till ökad transparens. Tyvärr är det 81% av utvärderingarna som inte har någon diskussion av validitet, och 75% som inte innehåller någon diskussion av reliabilitet. I båda fallen är det 7% som har en tillräcklig diskussion av dessa två begrepp. Det är också anmärkningsvärt att inte mer än 13% redovisar intervjuprotokoll eller guidelines; intervjuer är ju ändå det vanligaste sättet att samla in data.

Den norska genomgången av utvärderingar fann samma grundläggande drag. Den påvisade stora brister i metodpresentation. Det vanligaste undersökningsinstrumentet var intervjuer med projektpersonal. Det var något vanligare än i Sverige med direkt observation och samtal med målgrupper, men mindre vanligt med kontakter med andra givarorganisationer.

I jämförelse kan sägas att en tredjedel av CIDA och en fjärdedel av USAID utvärderingarna saknade metodgenomgång. De flesta använde intervjuer med projektpersonal, men för övrigt fanns det även där stora brister. 50% av de kanadensiska utvärderingarna hade inte någon kontakt med andra givare (variabeln undersöktes inte i genomgången av USAID utvärderingar). Canada var i alla fall markant bättre än Norge och Sverige i detta avseende.

#### Utvärderingsobjekt.

Låt oss då se vad utvärderingarna värderar; dvs vad de innehåller värdeutsagor om. Det är vanligt att skilja mellan utvecklingsmål och insatsmål vid planeringsarbetet, eller med en annan terminologi mellan långsiktiga och kortsiktiga mål. Den senare kategorin hänför sig till de mål som omedelbart skall uppnås genom insatsen, den förra kategorin innefattar mål som man förväntas nå på något längre sikt - troligtvis som en följd av att man når de kortsiktiga målen.

Det är knappast förvånande att tonvikten i utvärderingsarbetet ligger på de kortsiktiga målen, det som är insatsernas omedelbara effekt. Det är en logisk följd av att den större delen av utvärderingsarbetet görs under pågående insats och att det som regel är kopplat till beslut om förlängning av insatserna.

Det är också en följd av att utvärderingarna inte är särskilt omfattande, de resurser som avsätts för utvärderingar räcker bara till för att ge enkla svar på enkla frågor. Att värdera långsiktiga mål är däremot en tids- och kostnadskrävande process.

**Tabell 8. Utvärderingarnas objekt för värdeomdöme.**

	Tillräckligt	Begränsad	Saknas
Utvärderingarna analyserar:			
Utvecklingsmål	25%	29%	44%
Insatsmål	42%	42%	16%
Förutsättningar/antaganden	20%	30%	50%

Som tabell 8 visar är det ändå 54% av utvärderingarna som innehåller någon form av analys av de långsiktiga målen (tillräcklig eller bristfällig), och det kanske skall ses som glädjande. Det innebär i alla fall att kopplingen mellan kortsiktiga mål och insatsmål görs explicit på åtminstone hälften av utvärderingarna.

När man planlägger en insats är det viktigt att explicit peka på omvärldsfaktorer som kan påverka genomförandet; t.ex. vad andra organisationer bidrar med vad gäller personal och pengar, eller hur migrationsmönster kan påverka ett markvårdsprojekt mm. I utvärderingsarbetet brukar man då också påpeka att det är angeläget att se över dessa antaganden och förutsättningar.

Det är dock bara 20% av utvärderingarna som har gjort en "tillräcklig" analys av antaganden. Även om ytterligare 30% har en "begränsad" analys av sådana antaganden, är det fortfarande 50% som helt saknar det. Liksom ovan kan man då antingen glädja sig över att hälften har sådana analyser, eller beklaga att den andra hälften saknar det, alltefter vad man anser vara en rimlig ambitionsnivå.

Eftersom det är metodiskt mycket enklare att värdera antaganden och förutsättningar än att värdera utvecklingsmål kan det tyckas förvånande att det är ungefär lika bra/dåligt vad avser de två områdena. Eftersom en analys av antaganden kan vara mera operationellt användbart kanske det är mera angeläget att åstadkomma förbättringar där.

Det kan vara intressant att notera att analysen av mål inte blivit bättre, dvs det har inte blivit bättre analyser av utvecklingsmål med tiden och inte heller mera omfattande målanalyser. Den norska studien fann att analysen av utvecklingsmål var "tillräcklig" i drygt 15% av fallen, och analysen av insatsmål var "tillräckligt" i drygt 65% av fallen (dvs något sämre respektive bättre än de svenska). De kanadensiska respektive amerikanska studierna mätte enbart analys av måluppfyllelse, och särskilde inte långsiktiga och kortsiktiga mål. Analysen av måluppfyllelse befanns dock vara "tillräcklig" på 51% respektive 20% av utvärderingarna.

Hur ser det ut med andra aspekter på insatsens nytta och skilda sätt att värdera denna. Tabell 9 visar att drygt 50% av utvärderingarna innehåller en "tillräcklig" analys av insatsernas nytta. Det är väl i och för sig bra, kanske det är mera förvånande att så mycket som 10% inte innehåller någon analys alls av insatsens nytta.

En sidoaspekt på nytta är frågan om mottagarlandet har ett behov av insatsen. Det är ju inte säkert att bistånd behövs även om det gör nytta (och vice versa). Insatsen kanske hade blivit av även utan bistånd t.ex. Det är dock svårt att bedöma, och det är inte heller vanligt att utvärderingar analyserar mottagarländernas behov. I 33% av fallen gjordes en tillräcklig analys av behov, och i ytterligare 26% av fallen gjordes en begränsad behovsanalys.

En behovsanalys kan ofta kräva en jämförelse av situationer med eller utan insatsen. T.ex. har katastrofbistånd ibland byggts på jämförelser mellan byar som nåtts av hjälp och andra som blivit utan. Man har sedan kunnat värdera återhämtningsmekanismer, anpassning och beredskap att hantera nya katastrofsituationer. Det är dock sällsynt med sådana utvärderingar; enbart 6% har innehållit en tillräcklig analys av hur verkligheten hade sett ut utan insatsen, och närmre 70% innehöll inte någon sådan analys alls.

Regeringens budgetproposition har under många år påtalat vikten av att insatsernas kostnadseffektivitet blir utvärderad. I de här utvärderingarna ser vi att effektiviteten vanligtvis analyseras, men det är mera sällan man går vidare till att bedöma kostnadseffektivitet. Det i sin tur kan bero på att det inte är så vanligt att utvärderingarna analyserar projektens kostnader i tillräcklig utsträckning. Har man inte ett tillräckligt bra kostnadsunderlag kan man givetvis inte ta fram något meningsfullt mått på kostnadseffektivitet.

**Tabell 9. Utvärderingarnas innehåll.**

	Tillräcklig	Begränsad	Ingenting	Irrelevant
<u>Analytiskt innehåll i utvärderingarna:</u>				
Mottagarlandets behov av insatsen?	33%	26%	35%	6%
Jämförelser med eller utan insatsen?	6%	19%	68%	7%
Analys av insatsens kostnad?	38%	35%	23%	4%
Analys av insatsens nytta?	52%	36%	10%	2%
Bedömning av effektivitet?	30%	46%	20%	4%
Bedömning av kostnads-effektivitet?	17%	28%	52%	3%
Diskussion av bistånds-myndighetens arbete?	10%	32%	49%	9%

Databasen innehåller också information om i vad mån insatsernas "inputs" respektive "outputs" värderats på ett tillräckligt eller begränsat sätt, eller inte alls (enligt samma svarskategorier som ovan). I ungefär en tredjedel av fallen befanns måtten vara tillräckliga, men i 22% av fallen saknades mått på "inputs", och i 16% av fallen saknades mått på "output". Det kan verka förvånande att det är mera vanligt att det saknas mått på "outputs" än på "inputs", det är ju betydligt svårare att mäta det som kommer ur insatserna än det som går till dem.

Mätningar av kostnadseffektivitet var bättre förr. Av de utvärderingar som var gjorda före 1980 hade drygt 50% "tillräckliga" mått på kostnadseffektivitet. Andelen gick ned kraftigt under 1980-talet, men har sedan gått upp igen så att 30% av rapporterna från tiden efter 1990 har "tillräckliga" mått. Däremot fanns det ingen större skillnad på hur man mätt effekter, och heller inte vad gäller analys av insatsernas nytta. Det mest intressanta är väl dock att notera att det inte finns några som helst belägg för att dagens utvärderingar är bättre än de var för 20 år sedan.

De svenska utvärderingarna förefaller ligga något efter internationell praxis. 48% av CIDA's utvärderingar hade "tillräcklig" genomgång av effekter, 71% innehöll "tillräckliga" mått på kostnadseffektivitet och 50% "tillräcklig" analys av nytta. 40% av de norska utvärderingarna analyserar produktivitet och 54% innehåller "tillräcklig" analys av effekter.

Även om utvärderingarna som regel behandlar en biståndsaktivitet är det inte omöjligt att de även granskar myndighetens (dvs ofta uppdragsgivarens) arbete. Många uppdragsbeskrivningar innehåller en paragraf om att utvärderingen även skall behandla biståndsorganisationens roll, eller något liknande. Det är dock bara 10% av utvärderingarna som gör någon sådan analys. Knappt hälften behandlar inte ämnet alls.

Givet att en insats behövs, blir till nytta och är effektiv (vilket oftast inte kan besvaras med entydiga ja eller nej) blir nästa fråga i vad mån det som biståndet bidragit till att bygga upp har någon egen bärkraft? Kommer aktiviteterna/institutionerna att överleva om biståndet avslutas? Hur kan biståndet anpassas så att det leder till bärkraftiga strukturer i mottagarländerna? Dessa är några av de viktigaste - och svåraste - frågor som utvärderingar har att besvara. Tabell 10 nedan redogör för i vilken utsträckning utvärderingarna i databasen har något att säga om insatsernas bärkraft.

Det framgår att analysen av bärkraft (sustainability) är svagt utvecklad. Den mest frekventa svars-kategorin är "ingenting" vad gäller samtliga delfrågor. På det mera allmänna planet, kan vi konstatera att 21% av utvärderingarna innehåller tillräcklig analys av bärkraft, men vi måste ifrågasätta innebörden av detta. Det visar sig ju sedan att konkreta frågor om lokala parters avsikter, grad av egenfinansiering, tillgång till konvertibel valuta och alternativa finansieringskällor sällan är belysta i tillräcklig utsträckning. Detta kan jämföras med att 23% av de norska utvärderingarna innehåller tillräcklig analys av bärkraft, och 55% av de kanadensiska och 36% av USAID's utvärderingar. Knappt hälften av de norska och svenska rapporterna innehöll inte någon som helst analys av bärkraft, jämfört med 17% av CIDA's och 25% av USAID's.

Korstabuleringen visade ett starkt samband mellan variablerna ålder och analys av bärkraft. Det var vanligare att sådana analyser gjordes före 1980, och de gjordes dessutom bättre då. Det fanns också ett samband med kostnad; desto dyrare utvärderingen var, desto större

chans att den hade en bra analys av bärkraft. SIDA och SAREC befanns ha fler utvärderingar med analys av bärkraft; och det var också mera vanligt att de utvärderingar som var gjorda av "utvärderingsexperts" hade sådana analyser.

Det kan vara intressant att notera att en relativt stor andel av svaren i tabell 10 hamnar i kolumnen "irrelevant". Det kan vara förståeligt vad gäller detaljerna, t.ex. om tillgången på konvertibel valuta. Om insatsen äger rum i ett land vars valuta i praktiska avseenden är konvertibel är ju frågan irrelevant. Det är dock svårare att se hur frågeställningen som sådan skulle vara irrelevant på så mycket som 12% av utvärderingsrapporterna.

**Tabell 10. Utvärderingarnas behandling av insatsernas bärkraft.**

	Tillräcklig	Begränsad	Ingenting	Irrelevant
Behandlar rapporten:				
insatsens bärkraft när biståndet avslutats?	21%	24%	43%	12%
analys av alternativ finansiering när biståndet avslutas?	12%	14%	60%	14%
tillgång av utländsk valuta efter det att insatsen avslutas?	7%	8%	68%	17%
insatsen grad av självfinansiering?	3%	13%	55%	29%
i vilken grad lokala myndigheter/partner är beredda att prioritera insatsen?	15%	31%	42%	12%

En stor del av de faktorer som påverkar om en insats skall betraktas som bra eller dålig skär över ordinarie fackområden. Utveckling är ett tvärvetenskapligt fenomen. Regering och riksdag, liksom flera internationella organisationer, har betonat att insatser inom biståndet bör utformas så att de bidrar till utveckling i en rad skilda dimensioner. Vid utvärderingsarbete kan man som regel finna att en insats har betydelse i flera av dessa dimensioner, antingen direkt genom medveten planering, eller indirekt som en önskad eller oönskad sidoeffekt.

Tabell 11 redovisar i vilken utsträckning utvärderingarna tar upp fem sådana tvärfackliga områden, som är av betydelse för hur insatsen värderas - oavsett dess nytta, bärkraft och effektivitet.

**Tabell 11. Utvärderingarnas behandling av tvärfackliga frågor.**

	Tillräckligt	Begränsat	Ingenting
Behandlar utvärderingen:			
kvinnoperspektiv?	10%	14%	76%
miljöfrågor?	15%	16%	69%
mänskliga rättigheter?	1%	5%	94%
demokrati?	3%	10%	87%
marknadsekonomi?	7%	17%	76%

Den vanligaste av dessa "cross-cutting issues" är miljöfrågor som på antingen tillräckligt eller begränsat sätt behandlas av 31% av utvärderingarna. Därefter kommer de två frågorna om kvinnoperspektiv och marknadsekonomi. Det är visserligen vanligare att kvinnofrågorna behandlas tillräckligt, men det är vanligare att marknadsekonomi behandlas i begränsad omfattning. I vardera fallet är det ungefär tre fjärdedelar av utvärderingarna som inte har något alls att säga.

Tyvärr går det bara att jämföra behandling av miljöfrågor och kvinnoperspektiv med de andra ländernas utvärderingar. Lika sorgset är att de svenska utvärderingarna är avsevärt sämre. 31% av de kanadensiska, 33% av de amerikanska och 31% av de norska har minst "tillräcklig" analys av dessa kvinnoperspektiv, jämfört med 10% av de svenska utvärderingarna. När det gäller miljöfrågor var USAID's och CIDA's lika bra som de svenska, men de norska hade "tillräckliga" miljöanalyser i drygt 30% av fallen.

Det kan dock tilläggas att analysen av miljöfrågor blivit sämre med åren, utvärderingarna efter 1985 har här sämre värden än de som var gjorda innan dess. Med kvinnoperspektiven är det tvärtom, där är det vanligare med analyser nu. De svenska utvärderingarna har också bättre analyser av demokrati, men vad gäller de andra frågorna kan databasen inte påvisa någon skillnad.

Bedömningen av hur mänskliga rättigheter och demokratifrågor påverkas av biståndsinsatser är, som tabellen visar, ytterst styvmoderligt behandlat. Det är bara någon enstaka procent av utvärderingarna som innehåller någon information alls på dessa punkter. Frågan om vilka av de här tvärfackliga frågorna som har högst aktualitet beror förstås på vilken typ av insatser som behandlas. Som vi såg ovan är det stora flertalet här projektinsatser inom sektorerna

utbildning, jordbruk, industri, förvaltning och infrastruktur. Kanske det förklarar behandlingen av dessa områden?

#### Slutomdöme om utvärderingarnas kvalitet.

Den allra sista punkten i SASDA's analyschema består av ett sammanfattande värdeomdöme. Var det en förträfflig utvärdering, var det en tillräckligt bra utvärdering eller var det en bristfällig utvärdering? Tabell 12 visar hur de forskarasistenter som klassificerat utvärderingarna enligt ovan sedan avgivit ett slutgiltigt betyg på den tregradiga betygsskalan.

En majoritet av utvärderingarna höll måttet; de bedömdes som tillräckligt bra för sina syften. I den gruppen ingår då 7% av utvärderingarna som kunde bedömas som "förträffliga", dvs gjorda med stor metodologisk kunskap och med insamlande av empiriskt material, analys av detta samt en breddning av frågeställningarna till att omfatta bärkraft och de skilda typerna av sidoeffekter som nämndes ovan.

**Tabell 12. Övergripande omdöme om utvärderingarnas kvalitet.**

Utvärderingen var:	Förträfflig	Tillräckligt bra	Bristfällig	Vet ej
SASDA's databas	7%	52%	21%	20%
CIDA	23%	44%	33%	-
NORAD/UD	9%	54%	29%	8%

Notera att det inte ansågs vara möjligt att ge ett sammanfattande omdöme på hela 20% av utvärderingarna. Det är förstås alltid utmanande och svårt att göra en sådan generalisering, kanske är det också svårt att göra mot bakgrund av att ha fyllt i uppgifterna ovan. Det är ju inte säkert att en enskild utvärdering för att vara till nytta behöver ha "goda" värden i alla avseenden. Kanske det tvärtom vore så att en utvärdering som alltid har "tillräcklig" information och analys enligt ovanstående vore både oanvändbar och ineffektiv? Tabell 12 innehåller också en jämförelse med omdömena av de andra ländernas utvärderingar. Jag har inte tillgång till de sammanfattande omdömena i USAID's rapport, men i den kanadensiska studien sägs att de är lika CIDA's. Ungefär en tredjedel var otillräckliga, och knappt en femtedel ansågs vara "excellent".

Har utvärderingarna blivit bättre med tiden? Både ja och nej. 20% av utvärderingarna gjorda före 1980 klassificerades som "excellent", men bara 12% av dem gjorda efter 1990 befanns så bra. 1980-talet var



uppenbarligen ett mörkt decennium, det var få riktigt bra utvärderingar då. Däremot har andelen riktigt dåliga minskat hela tiden, och särskilt sedan 1990 är det få utvärderingar som fått den utmärkelsen. Det kan vara intressant att notera att det inte finns något samband mellan slutomdömet beträffande utvärderingarnas kvalitet och förekomsten av utvärderingsexpertis på teamen. Däremot är det dubbelt så vanligt att de som klassificerats som riktigt bra har deltagare från mottagarlandet med i utvärderingsteamet.

Finns det någon skillnad mellan de svenska biståndsorganisationerna? Tabell 13 visar den korstabulering som kan göras på databasen. Jag vill påpeka att det finns för få utvärderingar från några av organisationerna för att det skall vara möjligt att jämföra. Notera dessutom att de forskarasistenter som klassificerat BITS och SAREC's material inte har besvarat frågan vad gäller 50 utvärderingar.

**Tabell 13. Övergripande omdöme om utvärderingarnas kvalitet efter beställande organisation.**

Utvärderingen var:	Förträfflig	Tillräckligt bra	Bristfällig	Vet ej
BITS	1	7	-	18
SAREC	1	7	3	32
SIDA	7	61	27	2
SwedeCorp	1	-	-	-
Utrikesdepartementet	-	1	1	-

#### Sammanfattande kommentar.

Låt oss då återvända till de hypoteser om utvärderingsarbetet som lades fram på sidan 3 ovan. I vad mån kan vi nu bekräfta att de här föraningarna var riktiga eller i vad mån måste vi modifiera dem. Observera att utsagorna enbart bygger på utvärderingarna som finns i databasen och på det sätt som de av SASDA anlidade personerna har klassificerat dessa rapporter. Båda dessa faktorer ligger utanför min bedömning, men påverkar givetvis slutsatsernas tillförlitlighet.

Jag vill åter erinra om att hypoteserna utgör en form av subjektiva måttstockar. Läsaren bör själv tänka igenom i vad mån de uttrycker rimliga förväntningar på utvärderingssystemet, och anpassa sin läsning och sina slutsatser till de värden han eller hon själv vill sätta. Låt oss då se i vad mån genomgången ovan leder till att hypoteserna kan accepteras eller måste förkastas.

*Hypotes 1. En majoritet av utvärderingarna arbetar på ett metodologiskt tillfredsställande sätt.*

Nej, få av utvärderingarna har kunnat klassificeras som metodologiskt tillfredsställande. De flesta redovisar över huvudtaget inte sina metoder, och de metoder som används är ensidigt inriktade på intervjuer med projektpersonal. Intervjuprotokoll eller guidelines redovisas sällan. Få innehåller någon diskussion av tillförlitlighet.

*Hypotes 2. En majoritet av utvärderingarna har tillfredsställande bedömningar av måluppfyllelse.*

Nej, hypotesen måste förkastas. En fjärdedel av utvärderingarna innehåller en "tillräcklig" analys av utvecklingsmål, och drygt 40% innehåller en "tillräcklig" analys av kortsiktiga mål, eller insatsmål. I övriga fall är analysen bristfällig eller saknas helt.

*Hypotes 3. En majoritet av utvärderingarna har tillfredsställande bedömningar av effekter.*

Hypotesen kan i alla fall inte accepteras enkelt och självklart. 30% av utvärderingarna innehåller en tillfredsställande analys av effekter, dvs mindre än hälften. Men 52% av utvärderingarna innehåller en tillfredsställande bedömning av insatsens nytta, vilket kanske gör att hypotesen kan accepteras med tvekan. Skillnaden mellan effekter och nytta kan ha varit svår att klassificera rätt.

*Hypotes 4. En majoritet av utvärderingarna har tillfredsställande bedömningar av kostnadseffektivitet.*

Nej, den här hypotesen måste förkastas. Det är enbart 17% av utvärderingarna som innehåller en tillfredsställande analys av kostnadseffektivitet, dvs betydligt mindre än hälften.

*Hypotes 5. En majoritet av utvärderingarna har tillfredsställande bedömningar av biståndets bärkraft.*

Nej, så är det inte. Det är 21% av utvärderingarna som innehåller en tillfredsställande analys av biståndets bärkraft. Mer än hälften av utvärderingarna behandlar inte frågan alls. Resten behandlar frågan bristfälligt, eller så var det inte en relevant frågeställning för utvärderingens uppdrag.

*Hypotes 6. En majoritet av utvärderingarna har tillfredsställande bedömningar av s.k. "cross-cutting issues".*

Nej, tvärtom är det så mycket som 75% till 95% av utvärderingarna som inte alls behandlar de skilda tvärfackliga frågorna. Det varierar dock beroende på vilken fråga det rör sig om. Det vanligaste är att behandla miljöaspekter, vilket 15% av utvärderingarna gör. 10% av utvärderingarna behandlar kvinnoperspektiv på utvecklingsfrågor.

*Hypotes 7. De svenska utvärderingarna är ungefär lika bra som de utvärderingar som görs i Norge, Canada och U.S.A.*

Nu avser jämförelsen flera variabler. På många av dem finns det inte några skillnader. Men det är betydligt mera vanligt att utvärderingarna från U.S.A. och Canada innehåller tillfredsställande mått på effekter och effektivitet. De har dessutom oftare bedömning av bärkraft och mera omfattande analyser av "cross-cutting issues". Det finns ingenting som antyder att de svenska biståndsorganisationernas praxis skulle vara bättre än dessa länders, tvärtom ger denna studien anledning att tro de svenska organisationerna är något sämre.

*Hypotes 8. Utvärderingarna har blivit bättre över tiden, dvs en allt större andel karaktäriseras som tillfredsställande enligt kvalitetsindikatorerna.*

Nej, det finns ingenting i databasen som skulle styrka en sådan "utvecklingsoptimism". Korstabuleringar mellan variabeln tid och de olika kvalitetsindikatorerna tyder inte på att utvärderingarna under 90-talet är bättre än under föregående femårsperioder. Visserligen varierar det mellan de olika måtten, men som en generell slutsats kan hypotesen definitivt avfärdas. I flera aspekter har utvärderingarna tvärtom blivit sämre.

*Hypotes 9. Det finns inte någon skillnad mellan de olika biståndsorganisationerna vad gäller kvalitetsindikatorerna.*

Undersökningen har inte kunnat påvisa några statistiskt signifikanta skillnader mellan biståndsorganisationerna. Det beror delvis på urvalet rapporter, och att det finns så få från vissa myndigheter. I några fall är någon enstaka organisation bättre än de övriga; så har t.ex. SIDA:s och SAREC's rapporter oftare engelska som huvudspråk, SIDA har oftare tillfredsställande bedömning av bärkraft, BITS har oftare en tillfredsställande metoddiskussion - för att nämna några exempel.

*Hypotes 10. Det finns inte någon skillnad mellan östeuropabiståndet och andra delar av biståndet vad gäller kvalitetsindikatorerna.*

Nej, hypotesen förkastas. I många avseende finns det visserligen inga skillnader mellan östbiståndet och övriga utvärderingar. Men 14% av utvärderingarna av östbiståndet klassificerades som "excellent", vilket är dubbelt så många som genomsnittet i databasen. En majoritet av dem har en tillfredsställande analys av bärkraft, och kring 40% hade tillfredsställande analys av effektivitet respektive kostnadseffektivitet. På dessa centrala områden var rapporterna om östbiståndet alltså avsevärt bättre än utvärderingarna av bistånd till övriga regioner.

Nu skall dessa slutsatser betraktas med försiktighet och de skall snarare läsas som fingervisningar eller uppslag om särskilda områden som kan behöva undersökas närmare. Observera även att det bara är två av de fyra kvalitetsaspekter som branschorganisationerna kommit överens om som uppmärksammas här - och faktiskt också enbart till del vad gäller dessa.

## Noter.

(1) Det svenska biståndets utvärderingar behandlades i DsUD 1990:63, *Bra Beslut: om effektivitet och utvärdering i biståndet*. Betänkandet av kommittén för analys av utvecklingssamarbete, SOU 1994:102 innehåller också en genomgång av utvärderingsarbetet. Detta i sin tur bygger delvis på Bruce and Finzi; *Improving Monitoring and Evaluation in Swedish development assistance*, SASDA working paper no 4, samt Stokke; *A Perspective on SAREC's Evaluation System: Policy and Performance*, SASDA working paper no 10, samt Forss; *Evaluations in Multilateral Development Cooperation*, SASDA working paper no 11, samt Carlsson and Jörnmark; *Evaluation and Learning in BITS*, SASDA working paper no 12.

(2) Den norska analysen var en delrapport i studien *Learning from experience - a study of the feedback from the evaluation system in the Norwegian Aid Administration* (Samset, Forss and Hauglin, 1993). Den kanadensiska studien har referens Watson; *A study of the quality of CIDA's bilateral evaluations 1988-89 and 1989-90*. USAID studien gjordes internt och har titeln "Review and Quality of A.I.D. Evaluations FY 1987 and FY 1988".

(3) De standard som refereras här utsågs som frukten av ett intensivt arbete av en 17-manna-kommitté, utsedd av 12 professionella organisationer. 100-tals aktiva utvärderare deltog i diskussionerna. Dessa standards publicerades av the "Joint Committee on Standards for Educational Evaluation" (1981), men de används och sprids numera huvudsakligen genom den största branschorganisationen för "evaluation research"; The American Evaluation Association.

4. Sifferuppgifterna bygger på statistik i "Bistånd i siffror och diagram" 1992/93.

5. Se t.ex. den biståndspolitiska utredningen, SOU 1977:13; inlägg i debatten som Elzinga, A., (1981) "Evaluating the Evaluation Game"; promemorior från SIDA:s utvärderingsenhet, DAC's expertgrupp för samarbete på utvärderingsområdet, samt den ovannämnda "Bra Beslut - om effektivitet och utvärdering i biståndet", DsUD 1990:63.

6. Det finns i skrivande stund inte någon kurs i utvärdering vid något svenskt universitet eller högskola, vare sig inom ramen för forskarstudier eller grundexamen. Det närmaste man kommer är grundkurser i utredningsmetoder och sociologisk metod. De uppsattsseminarier som finns vid de flesta högskolor ligger dock kvalitetsmässigt långt från de terminskurser i "evaluation" som ges vid ett flertal universitet i Canada och U.S.A. Inom ramen för en forskarutbildning inom samhällsvetenskaperna kan man dock anta att den studerande tillägnar sig grunderna för utvärderingsmetoder.

Bilaga 2. Analysschema för utvärderingarna

### Bilaga 3. Utskrift av databasen.

Observera databasen innehåller 177 rader och 76 kolumner. Den får inte plats på en sida. Inför läsningen: varje kolumn representerar en av de variabler som utvärderingarna klassificerats efter. Kolumn 1 är namnet på den forskarassistent som klassificerat utvärderingen, kolumn 2 är tiden för inmatning i databasen. Därefter följer kolumnerna i tur och ordning respektive fråga i det analyschema som finns i bilaga 2. Kolumn 76 är en "dummy" som inte innehåller någon information. De första tretton sidorna innehåller information om respektive kolumn.

Varje utvärdering har en rad. För att följa klassificeringen av en utvärdering måste man tyvärr bläddra mellan sidorna för respektive kolumner.

Det är inte lätt att arbeta med materialet i den här formen. Den som vill tränga djupare in i det kan beställa en diskett med databasen antingen i Excel format för PC eller Statview format för Macintosh från SASDA.

Bilaga 4. Synpunkter på vidareutveckling av databasen.

Det har varit givande att utföra analysen av databasen. Det gav mer information än jag hade väntat mig. Jag anser följaktligen att arbetet bör fortsätta enligt följande;

1) Sträva efter att lägga in åtminstone 15 till 20 slumpmässigt utvalda utvärderingar varje år. Klassificera dem enligt analys-schemat med de modifieringar som föreslås nedan,

2) Uppdatera den här rapporten en gång per år och publicera resultatet som en barometer över hälsotillståndet i det svenska utvärderingssystemet. Välj ut ett litet antal variabler och följ dem noggrant över tiden. Jag föreslår att följande väljs ut i första hand:

- förekomst av metodavsnitt
- användning av skilda undersökningsinstrument
- diskussion av validitet och reliabilitet
- mätning av utvecklingsmål
- mätning av insatsmål
- mätning av nytta
- analys av mottagarlandets behov av insatsen
- mätning av effektivitet
- mätning av kostnadseffektivitet
- analys av bärkraft
- analys av cross-cutting issues
- slutomdöme om utvärderingen

Jämför kontinuerligt hur dessa värden förändras över tiden. Detta skulle var det starkaste styrinstrumentet för att påverka kvaliteten på utvärderingar inom biståndet.

3) För att genomgången skall anses tillförlitlig krävs förtroende för databearbetningen. Var noga med att de som klassificerar utvärderingarna gör det på samma sätt, enligt överenskomna kriterier. Kontrollera att det blir rätt. Utbilda dem som gör klassificeringen med ett antal seminarier och praktikfall för detta. Acceptera inte att alternativen "vet ej" och "irrelevant" används annat än undantagsvis.

4) Se över svarsalternativen under respektive variabel enligt följande:

Sid 2. Aid channel. Använd bara bilateral och multilateral.

Sid 2. Region. Middle East förekom ej som svarsalternativ. Kan tas bort. Ändra other till global/regional

Sid 2. Receiver. Tag bort semi-private. Inför "Other"

Sid 3. Type of evaluation. Tag bort research

Sid 4. Summary language. Tag bort no summary

Sid 4. Methodology issues. Tag bort not relevant

Överallt där skalan "adequate, limited, no, not relevant" används. Inför istället skalan:

not addressed  
addresses minimally  
addressed adequately  
exemplary treatment

Specificera tydligt vad vart och ett av värdena innebär. Gör klart att klassificeringen skall vara öppen, författarna av respektive rapport skall ha möjlighet att "överklaga" - men framförallt skall de inte ha någon anledning att vilja göra det.

5) Inför nya variabler eller tag bort enligt följande.

Sid 1. Kategoriindelning av utvärderingens ålder. 5-års intervall så som det använts i den här rapporten (kolumn 77)

Sid 3. Ny variabel. Sex of evaluation team: Male, Female, Mixed

Sid 3. Ny variabel. Regional/country expertise present. Yes/no/unknown.

Sid 4. Ny variabel. Summary included. yes/no.

Sid 4. Flytta fråga 14 och 15 på sid 7 till sid 4, under avsnitt 6.

Sid 5. Ny variabel. Questionnaire used. Svarsalternativ enl ovan.

Sid 5. Ny variabel. Questionnaire format included. Alternativ enl ovan.

Sid 5. Ny variabel. Use of participatory evaluation methods. Alternativ enl ovan.

Sid 5. Ny variabel. Presentation of raw data. Alternativ enl ovan.

Sid 6. Ändra variabeln "discussion of agency's monitoring" till "discussion of the aid agency's performance".

6) Hur mycket arbete krävs? Det tar uppskattningsvis 1 timme att lägga in data från en utvärdering, om det görs i samband med att man läser utvärderingen ändå. Annars bör man lägga på cirka 1-2 timmar för grundlig genomläsning. Att lägga in 20 utvärderingar bör följaktligen vara ungefär en veckas arbete, lite drygt.

Det tog två veckor att skriva den här rapporten. Det bör vara något lättare, och gå något fortare när väl den första genomgången är gjord. Med cirka 3 veckors arbete per år skulle det vara möjligt att hålla databasen uppdaterad och ge ut en rapport om tillståndet i utvärderingssystemet. Det förefaller inte kräva någon större ansträngning. Min bedömning är att det skulle vara ett kraftigt incitament för såväl handläggare som konsulter och utvärderare att göra professionellt tillfredsställande arbeten.

Kim Forss. November 1994.