EBA

02
2021

MÅLBILD OCH MEKANISM: VAD SÄGER UTVÄRDERINGAR OM SVENSKA BISTÅNDSINSATSERS MÅLUPPFYLLELSE?

Markus Burman

Målbild och mekanism: Vad säger utvärderingar om svenska biståndsinsatsers måluppfyllelse?

Markus Burman

Rapport 2021:02

till

Expertgruppen för biståndsanalys (EBA)

Markus Burman är utredningssekreterare på Expertgruppen för biståndsanalys och arbetade dessförinnan som konsult.

Tack

Jag vill rikta ett stort tack till studiens referensgrupp bestående av Kim Forss, Johanna Lindström (FCG Sweden), Sven Olander (Sida) och Lennart Peck (UD) för värdefulla synpunkter. Jag vill även tacka Jan Pettersson och Joakim Molander för viktiga synpunkter på rapporten.

Vänligen referera till denna rapport som: Burman, M. (2021), *Målbild och mekanism: Vad säger utvärderingar om svenska biståndsinsatsers måluppfyllelse?*, Rapport 2021:02, Expertgruppen för biståndsanalys (EBA), Sverige.

Denna rapport kan nedladdas utan kostnad på: www.eba.se

Rättigheterna för detta arbete följer Creative Commons Attribution 4.0 International License. Licensen kan läsas här: http://creativecommons.org/licenses/by/4.0/

ISBN 978-91-88143-65-5
Tryckt av Elanders Sverige AB
Stockholm 2021

Omslagets design av Julia Demchenko

Innehållsförteckning

EBAs förord **1**

Sammanfattning **3**

1. Varför denna studie? **9**

2. Vad är målbaserad utvärdering? **14**

3. Mål och uppdrag i 80 utvärderingar **25**

4. Slutsatser om måluppfyllelse **37**

5. Utvärderingarnas tillförlitlighet vid analys av måluppfyllelse **49**

6. Slutsatser i tidigare studier **63**

7. Förklaringar till svag måluppfyllelse **71**

8. Slutsatser, rekommendationer och frågor framöver **77**

9. Avslutande reflektioner om tillförlitlighet **82**

Referenser **98**

Bilaga 1 – Genomförande och metod **106**

Bilaga 2 – Analysramverket **114**

Bilaga 3 – Analyserade utvärderingar **126**

Bilaga 4 – Måluppfyllelsekriteriet **138**

Tidigare EBA-rapporter **143**

EBAs förord

EBAs uppdrag är att bidra till kunskap som grund för utveckling av biståndet. En central del inkluderar bedömning av det svenska biståndet med avseende på relevans, måluppfyllelse, effekter, hållbarhet, kostnadseffektivitet och andra kvaliteter. Då EBA i regel inte utvärderar enskilda projekt eller program görs sådana bedömningar ofta på en övergripande policynivå.

Denna studie utgår från den övergripande frågan om det svenska biståndet når sina mål. Det urval som här analyseras är Sidas decentraliserade utvärderingar under åren 2017 - 2019. En tidigare EBA-studie uppskattar att dessa utvärderingar utgör som mest 19 procent av alla utvärderingar av det svenska bilaterala biståndet, så urvalet är inte heltäckande. Det är inte heller så att alla insatser utvärderas. Därmed kan denna rapport enbart uttala sig om de slutsatser som dras i Sidas decentraliserade utvärderingar under denna period. I tillägg till detta gör utvärderingarna också andra bedömningar om insatsernas kvalitet. Denna studie kan alltså inte uttala sig om utvärderingarna som helhet utan bara de delar som berör måluppfyllelse och effekter.

Med detta sagt så finner rapportförfattaren att Sidas decentraliserade utvärderingar ofta bedömer måluppfyllelsen som god. En möjlig slutsats av detta är att det svenska biståndet i allmänhet når sina mål. En annan möjlig slutsats är att mindre lyckade projekt inte utvärderas. En ytterligare möjlighet är att utsagorna om måluppfyllelse inte är tillförlitliga. Författaren menar att de insatser som utvärderas mycket väl kan ha nått sina mål, men att detta i regel inte underbyggts i de studerade utvärderingarna.

Frågan om måluppfyllelse delas ofta upp i två delar; dels de mål som direkt skulle uppnås med insatsen, dels övergripande mål, mer långsiktiga, transformativa effekter (leder verksamheten till någon form av utveckling?). Författaren menar att det finns (åtminstone) två utmaningar i utvärdering av måluppfyllelse:

(i) Även utvärderingar av de mål som ska nås direkt måste tillförlitligt kunna visa på att observerade resultat är orsakade av insatsen, dvs. den kausala analysen är nödvändig. Författaren noterar att denna kausala

analys inte genomförs i de utvärderingar som studeras. Detta innebär att slutsatser, som självklart inte behöver vara felaktiga, brister i tillförlitlighet.

(ii) En fullständig utvärdering av insatsers mål förutsätter (kausal) analys även av långsiktiga effekter. Utan bedömning på övergripande målnivå kan man inte hävda att man utvärderat insatsernas mål. Utvärdering mot övergripande mål är nödvändig om det trovärdigt ska kunna hävdas att biståndet gör meningsfull skillnad.

En tillförlitlig resultatrapportering till riksdag, regering och allmänhet förutsätter att övergripande mål utvärderas i långt större grad än vad som idag görs idag. Författaren menar att Sidas decentraliserade utvärderingar är lämpliga för denna typ av bedömning.

Med grund i dessa slutsatser för författaren ett avslutande resonemang om möjliga orsaker till bristande tillförlitlighet och tänkbara justeringar av utvärderingsverksamheten.

EBAs förhoppning är att denna rapport ska bidra till en välbehövlig och ingående diskussion om målbaserad utvärdering inom biståndet. Rapporten vänder sig främst till personer som arbetar på övergripande eller strategisk nivå med utvärdering av bistånd, personer som genomför utvärderingar av svenskt bistånd och personer som arbetar med biståndseffektivitet eller styrning av det svenska biståndet. Vi tror också att rapporten kan finna användare hos personer som arbetar med resultatrapportering på bland annat UD och på Sida samt hos en biståndsintresserad allmänhet.

Denna rapport är författad av en person anställd på EBAs kansli. Processen är dock densamma som för övriga EBA-rapporter. Arbetet har följts av en referensgrupp under ledning av Kim Forss.

Analys, slutsatser och rekommendationer är författarens egna.

Göteborg i juli 2021,

Helena Lindholm

Sammanfattning

Syftet med denna studie är att undersöka vad nyligen genomförda utvärderingar säger och kan säga om måluppfyllelsen i Sidas biståndsinsatser och därmed svenskt bistånd. Detta utifrån den definition av måluppfyllelse (Effectiveness) som Sida använder med hänvisning till OECD DACs kriterier för utvärdering av bistånd.[1]

Det finns en omfattande litteratur om mål, måluppfyllelse och utvärdering. Delar av litteraturen är normativ och föreskriver hur mål och resultat, styrning och beslut bör beskrivas, utformas och utvärderas. En annan empirisk grundad del av litteraturen är överlag både mer skeptisk och kritisk och ger anledning att inte ha för högt ställda förväntningar på analys av målutvärdering i praktiken.

Min genomgång av *tidigare syntesutvärderingar* på biståndsområdet konstaterar att svenska decentraliserade biståndsutvärderingar beställda av Sida oftast tidigare dragit slutsatsen att måluppfyllelsen på kort eller medellång sikt varit god, trots att de i regel saknat bedömning mot övergripande effektmål.[2]

Biståndsinsatserna ansågs ofta ha otydliga målformuleringar vilket kan ha försvårat möjligheten att utvärdera måluppfyllelsen genom en effektiv operationalisering. Analyserna har också ofta vid extern

---

[1] Studien tar startpunkt i effektivitetskriteriet men en slutsats som dragits under arbetets gång är det sannolikt inte är rimligt eller eftersträvansvärt att strikt hålla isär måluppfyllelsekriteriet (*Effectiveness*) från effektkriteriet (*Impact*) avseende insatsers förväntade resultat. Utan bedömning mot effektmål har sällan någon uttömmande eller användbar bedömning av effektivitet gjorts och analyser på både nivån outcome och impact kräver hantering av kausalitet. Kriterierna är också mycket starkt överlappande, se bilaga 4.

[2] De decentraliserade utvärderingarna beställs av ambassader eller på Sidas avdelningar i Stockholm och berör oftast enskilda projekt eller program. De strategiska utvärderingarna som beslutas av myndighetens generaldirektör tar oftast ett bredare grepp, fokuserar inte på enskilda insatser och syftar till lärande på verksövergripande nivå.

kvalitetsbedömning ansetts ha metodologiska brister avseende bland annat den kausala analys som krävs för att resultat ska kunna kopplas samman med insats.

Denna nya genomgång av 80 decentraliserade och 6 strategiska utvärderingar visar på en stor pluralism och heterogenitet avseende övergripande målbild i de utvärderande biståndsinsatserna. Utvärderarna är ofta på ett eller annat sätt kritiska till hur mål/förväntade resultat formulerats avseende ambition, specificeringsgrad, mätbarhet, realism, konsistens samt acceptans bland involverade parter eller tidsperspektiv. Den otydlighet i målformulering som identifieras av utvärderarna har dock sällan föranlett rekonstruering eller förtydligande i enlighet med de metoder som finns för detta.

Oväntade (positiva eller negativa) resultat och effekter efterfrågas i cirka hälften av utvärderingarna, trots att sådana behövs för att undvika ”tunnelseende” i resultatbedömningen.

Utvärderingarna i urvalet rapporterar nästan bara om god måluppfyllelse/effektivitet i insatserna. Den vanligast förekommande slutsatsen är att mål/förväntade resultat så här långt är i huvudsak uppfyllda. Dessa bedömningar avser resultat så långt utvärderarna förmår fånga den, primärt av typen output och outcome, även om många utvärderingar, i regel med uppdragsbeskrivningen, även bedömer övergripande mål. Detta görs dock ofta partiellt och ofullständigt. Sammanställningen kan därmed inte ge någon fullständig bild av insatsernas effektivitet och måluppfyllelse.

Två faktorer som kan ha minskat reproducerbarheten i min genomgång är att utvärderingarnas bedömningar om måluppfyllelse ofta är ganska inlindande, sammansatta eller otydliga och att övergripande samlade bedömingar inte alltid görs. 10% av utvärderingarna har inte kunnat skalstegsbestämmas.

En övergripande slutsats i studien är också att bedömingarna av måluppfyllelse i Sidas decentraliserade utvärderingar förmodligen *inte* är tillförlitliga, framförallt då utvärderarna sällan lyckats visa att det är just insatsen som bidragit till eller åstadkommit observerad förändring. Utvärderarna använder sällan ny kvalitativ eller kvantitativ utvärderingsmetodik. Metodologin för besvarandet av måluppfyllelsefrågan beskrivs inte på ett transparent sätt och anpassas sällan ordentligt till den enskilda utvärderingsfrågan.

Ganska få av utvärderingarnas uppdragsbeskrivningar efterfrågar analyser av mekanismer eller bakomliggande förklaringar till måluppfyllelse/uppnådda resultat. Detta är en central fråga om syftet med utvärderingarna är lärande/användbarhet (Stern et al, 2012).

Ett litet antal utvärderingar i urvalet konstaterar svag måluppfyllelse. En separat analys av lärpotentialen i dessa ger att de bidrar med tänkbara delförklaringar till svag måluppfyllelse kopplat till insatsens styrning, utformning, organisering och uppföljning. För insatsen externa förklaringar eller mekanismer berörs dock i låg utsträckning. Utvärderarna listar möjliga faktorer, hinder eller åtgärder som man tror kan förbättra insatsen snarare än analyserar hur eller varför resultat uppnåtts/inte uppnåtts. Sammantaget begränsar det lärpotentialen.

Lärpotentialen i utvärderingarna minskar också av att de drar så likartade slutsatser att de inte ur ett samlat myndighetsövergripande perspektiv självklart kan bidra med ny kunskap om vad som fungerar var, när, hur och varför. Om nästan alla rapporter drar slutsatsen att insatserna når målen så saknas den variation i erfarenhet som är viktig för lärande.

Rapportens slutsatser och analyser summeras i form av två huvudslutsatser, fyra rekommendationer och fem frågor att analysera för framtiden:

Två övergripande slutsatser

1. Utvärderingarna i urvalet drar oftast slutsatsen att insatsen (i huvudsak) når sina mål. Denna bedömning sker främst på nivåerna output och outcome (kapitel 4).
2. Dessa slutsatser om måluppfyllelse är i regel troligtvis inte tillförlitliga (kapitel 5).

Fyra rekommendationer

1. De decentraliserade utvärderingarnas kvalitet avseende utvärderingsdesign och metodologi för analys av måluppfyllelse och förväntade resultat bör stärkas betydligt (se kapitel 5).
2. Fler svenska biståndsinsatser kan och bör utvärderas mot övergripande och långsiktiga mål. Detta tycks i hög grad också efterfrågas från Sidas personal som beställer utvärderingar (kapitel 3 och 4). Svenskt bistånd är idag i ganska hög grad fokuserat på kärnstöd, regionala och globala stödformer, insatser som är svåra att utvärdera. Sida bör därför analysera utvärderingsbarheten i dessa insatser med avseende på förutsättningarna att i framtiden utvärdera och bedöma det svenska biståndets måluppfyllelse och effektivitet (jmf. diskussionen i kapitel 5).
3. Bedömningar av förväntade resultat bör i högre utsträckning kompletteras med analyser av oförutsedda resultat och eventuella bieffekter inom biståndet (se kapitel 3).
4. Utvärderingarnas inriktning och frågeställningar bör i framtiden fokuseras kring måluppfyllelse och på att försöka förklara och förstå hur och varför insatser (inte) når sina mål (se kapitel 2, 3, 7 och bilaga 4).

Fem frågor som bör analyseras och belysas framöver

Det finns ett stort antal närliggande och besläktade frågor om svenskt bistånd, utvärdering, uppföljning, rapportering och effektivitet som kommit upp under arbetets gång men som inte fått plats i denna studie. Följande fem frågor bör belysas och analyseras särskilt framöver.

1. Hur bedöms måluppfyllelse mer generellt i svenskt bilateralt bistånd? Vilka underlag utöver utvärderingar använder sig Sida, FBA, Swedfund och andra av för bedömning och rapportering av måluppfyllelse till riksdag och regering? Hur tillförlitliga är dessa underlag och bedömningar?
2. Hur kan och bör man gå tillväga för att aggregera upp och summera Sidas resultatunderlag (utvärdering, uppföljning och rapportering) till strategi- och landnivå?
3. Hur kan Sida och andra biståndsgivare bedöma oförutsedda resultat och bieffekter på ett mer täckande, systematiskt och tillförlitligt sätt? Hur utvärderas det idag? Hur kan och bör det göras?
4. Hur tillförlitliga är slutsatserna i Sidas strategiska och partner-ledda utvärderingar?
5. Hur kan de utmanande kausalitetsbedömningarna i utvärderingar utvecklas för att uppnå nivån ”good enough” också i utvärdering av mindre eller begränsade biståndsinsatser eller där utvärderingsbudgeten är begränsad?

I rapportens slutkapitel diskuteras fyra områden som möjliga ingångsvärden vid en förklarande analys av tillförlitligheten i Sidas decentraliserade utvärderingar:

- Efterfrågan, behov och incitament.
- Systemet för upphandling, avrop, feedback och uppföljning av utvärderingstjänsterna.
- Utbudet på utvärderingsmarknaden.
- Kunskaper om utvärdering samt lärande och utveckling i verksamheten.

1. Varför denna studie?

Frågor om måluppfyllelse/effektivitet och resultat tillhör de mest och längst diskuterade i biståndet. Alla svenska statliga myndigheter har kravet att bedriva sin verksamhet på ett effektivt sätt.[3] För det internationella biståndet slår regeringens Policyramverk för svenskt utvecklingssamarbete och humanitärt bistånd (Regeringen, 2016) fast att:

> ”Information om utvecklingssamarbetets effekter är viktigt för både lärande och ansvarsutkrävande. Det förutsätter väl fungerande system för uppföljning och utvärdering i såväl Sverige som samarbetsländer och inom internationella organisationer.”

Även om forskning har visat att intresset för resultat i det svenska biståndet har gått i vågor (Vähämäki, 2017) så behöver regeringen, riksdagen, Sidas tjänstemän och allmänheten få tillförlitlig information om vad verksamheten uppnår. Förmågan att skilja framgång från misslyckande är central för lärande, anpassning och ansvarsutkrävande i demokratiska samhällsystem. Analys- och utvärderingsutredningen (SOU 2018:79) menade att regeringen i huvudsak efterfrågar kunskap om effekter, kostnadseffektivitet och just måluppfyllelse från sina analysmyndigheter.

I denna rapport studerar jag utvärdering av måluppfyllelse i ett urval av Sidas utvärderingar. Utvärderingar kan anses ha en särställning bland underlag och uppföljande processer i kraft av hög förväntad tillförlitlighet, potential för lärande, ansvarsutkrävande och oberoende bedömning. Utvärderingar är särskilt viktiga inom områden där resultatet av verksamheten inte direkt kan observeras

[3] Se Budgetlagen, myndighetsförordningen och regeringens förvaltningspolitik. ESV (2006) skriver att ”[å]tgärder som vidtas för att nå givna mål ska vara sådana att de faktiskt leder till målens uppfyllelse, inte att de bara syftar till målens uppfyllelse”.

av finansiären, som biståndsområdet.[4] Regering, riksdag och övriga aktörer är därigenom beroende av kvaliteten i bl.a. Sidas och EBAs utvärderingsarbete.

Syftet med denna studie är att undersöka *vad nyligen genomförda utvärderingar säger om måluppfyllelsen i svenska biståndsinsatser och svenskt bistånd*. Detta med utgångspunkt i den definition av *måluppfyllelse (Effectiveness)* som OECD/DAC anger för utvärdering av bistånd.

Definition av måluppfyllelse i studien

”Den utsträckning i vilken en biståndsinsats målsättningar har uppnåtts eller förväntas uppnås, med olika måls betydelse taget i beaktade.”[5]

Vår tolkning av måluppfyllelse, här kallad måluppfyllelsemodellen, är en standardtolkning som kombinerar DAC-kriteriets definition med utvärderingsfältets etablerade förståelse av måluppfyllelse enligt exempelvis Eliasson (2006), Vedung (2012) och Faugert och Sandberg (2012).

---

[4] ”En viktig skillnad mellan internationellt bistånd och annan offentligt finansierad verksamhet är att biståndet genomförs i andra länder. Det innebär att biståndets mottagare saknar direkta vägar att påverka beslut om hur pengar fördelas och används. Biståndets brukare har inte rösträtt i svenska val, de har inte rätt att överklaga svenska myndigheters beslut och de har små möjligheter att delta i svensk offentlig debatt. Svenska skattebetalare har, å andra sidan, begränsade möjligheter att följa hur biståndsmedel används och vilka resultat de leder till. Med andra ord: såväl lärande som ansvarsutkrävande försvåras av geografiska avstånd och verksamhetens organisering” (Burman och Hårsmar, 2015).

[5] I original: “The extent to which a development intervention’s objectives were achieved, or is expected to be achieved, taking into account their relative importance”. OECD/DACs utvärderingskriterier reviderades 2019. Den nya definitionen lyder: “The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups” (OECD/DAC, 2019). De utvärderingar som analyseras här beställdes när den tidigare definitionen fortfarande gällde varför denna används.

Studien tar sin startpunkt i OECD DACs effektivitetskriterium. En slutsats som dragits är att det i praktiken inte är eftersträvansvärt att strikt hålla isär måluppfyllelsekriteriet (*Effectiveness*) från effekt-kriteriet (*Impact)* med avseende på förväntade resultat. Utan bedömning mot effektmål har sällan någon uttömmande bedömning av effektivitet gjorts och analyser på både nivån outcome och impact kräver hantering av kausalitet. I bilaga 4 diskuteras definitionen av måluppfyllelse och hur kriterierna överlappar.

De två övergripande frågeställningarna i studien är:

1. *Vilka slutsatser dras om måluppfyllelse (eller uppnådda förväntade resultat) i nyligen genomförda utvärderingar av svenska bilaterala biståndsinsatser?*

Skiljer sig slutsatserna beroende på insatsens form (projekt, program, organisationsstöd), kanal (civilsamhälle, offentlig förvaltning, företag), geografi (region, låg- respektive medel-inkomstländer) eller övrig kontext (långsiktigt utvecklingssamarbete eller fragila eller konfliktfyllda stater)?

2. *Är utvärderingarnas slutsatser om måluppfyllelse tillförlitliga och, i så fall, användbara för lärande?*

Tydliggör utvärderarna hur måluppfyllelsefrågan besvarats och vilka begränsningar valda metoder har? Analyserar man hur insatsen specifikt påverkat måluppfyllelsen i förhållande till andra påverkande faktorer? Är det rimligt att utvärdera måluppfyllelse vid den tidpunkt när utvärderingar genomförs? Vilka faktorer anges ha påverkat graden av måluppfyllelse som grund för lärande?

Med ledning av ett på förhand framtaget analysramverk (bilaga 2) har frågorna besvarats genom en systematisk genomgång av 80 decentraliserade utvärderingar från åren 2017 till 2019 beställda av Sida i Stockholm eller ambassader. Utvärderingarna rör oftast

enskilda insatser eller organisationer. Dessutom ingår sex centralt genomförda ("strategiska") utvärderingar mellan 2015-2019.[6] Dessa studerar bredare teman eller återkommande modaliteter.

En viktig avgränsning är att jag inte har analyserat Sidas utvärderingsverksamhet som helhet utan studerar enskilda, främst decentraliserade, utvärderingars bedömning av måluppfyllelse. Även om nästan samtliga (93%) av den valda periodens decentraliserade utvärderingar ingår så är de inte representativa för Sidas biståndsportfölj som helhet. Sidas utvärderingar kan också förändras över tid. Urvalet ska därför inte ses som representativt för vare sig det svenska biståndet eller Sidas utvärderingsarbete bortom de år som ingår i studien. En utförlig diskussion om studiens metod, urval och representativitet finns i bilaga 1.

Jag gör inte heller anspråk på att ge en heltäckande bild av de enskilda utvärderingarna eftersom dessa också har andra utvärderingsfrågor att besvara (om relevans, kostnadseffektivitet, hållbarhet m.m.).

Studiens viktigaste begränsning är de analyserade utvärderingarnas tillförlitlighet vilket tydligt påverkar vilka slutsatser som kan dras om insatsernas måluppfyllelse. Denna omständighet görs dock till ett studieobjekt i sin egen rätt (kapitel 5).

Studien riktar sig primärt till dem som arbetar med styrning av det svenska biståndet (UD, riksdagen), Sidas ledning och personer som beställer eller genomför utvärderingar av svenskt bistånd. Andra viktiga målgrupper är personer som arbetar med resultat-rapportering, resultatbedömningar och resultatbaserad styrning i svenskt bistånd. En viktig målgrupp är också personer i den svenska allmänheten som intresserar sig för biståndets resultat.

I det följande diskuteras målbaserad utvärdering baserat på ett antal nedslag i *litteraturen* om målbaserad utvärdering (kapitel 2). Därefter (kapitel 3) görs en kartläggning av målbilden i de projekt

[6] Sedan en tid kallas dessa centrala utvärderingar på Sida. Dock ej under de år som här ligger i fokus.

utvärderingarna studerar: Vad är det för målsättningar eller förväntade resultat som utvärderas? Hur ser Sidas beställningar ut? Hur avgränsas och varieras måluppfyllelsefrågan? I kapitel 4 och 5 avhandlas *studiens två huvudfrågor* om utvärderingarnas slutsatser och tillförlitlighet och i kapitel 6 redovisas slutsatserna från *tidigare syntesstudier* om måluppfyllelse och utvärdering av måluppfyllelse i svenskt bistånd. I kapitel 7 analyseras de utvärderingar i urvalet som rapporterat svag måluppfyllelse med fokus på *förklaringar och bakomliggande faktorer* som grund för lärande. I kapitel 8 presenteras två övergripande slutsatser och fyra rekommendationer. Här föreslås också fem frågeställningar som bör studeras i *framtida studier*. I kapitel 9 görs ett antal avslutande reflektioner om tillförlitlighet. Samtliga översatta citat i rapporten är gjorda av författaren.

2. Vad är målbaserad utvärdering?

Måluppfyllelseutvärdering är en mycket vanligt förekommande form av utvärdering och ses av vissa som så grundläggande att de likställer utvärdering med bedömning mot en interventions mål (se Nachmias, citerad i Vedung, 2012).[7] Modellen har ibland ansetts fungera ”som en norm för vad utvärdering är och hur utvärdering bör genomföras” (Sandberg och Faugert, 2012).[8]

En vanlig defintion anger att *mål* är ”en beskrivning av ett framtida önskvärt tillstånd” (Krogstrup, 2017). Flera definitioner av *mål* utgår även ifrån att mål är ”beskrivningar av hur något målobjekt ska se ut eller vara vid en framtida tidpunkt” (Lindgren, 2012). Målobjektet är då ”den person, population, organisation, process eller något annat som behöver förändras och som en given intervention riktar sig till”. Målbaserade utvärderingar analyserar om en intervention nått eller bidrar till förväntade eller planerade målsättningar.

Utvärderare som studerar måluppfyllelse tar avstamp i ansvarigas nedtecknade, eller av utvärderarna rekonstruerade, målformuleringar. En vanlig uppfattning är att målen bör formuleras på ett särskilt sätt, exempelvis enligt SMART-kriterierna (Doran, 1981) som anger att målsättningar ska vara Specifika, Mätbara, Accepterade, Realistiska och Tidssatta för att kunna användas vid uppföljning och utvärdering.

Vedung (2012) understryker att vid analys av måluppfyllelse ställs i praktiken två frågor: ”1) I vilken grad har målen för utfall blivit uppfyllda (måluppfyllelsefrågan)? 2) Har insatsen bidragit till måluppfyllelsen och i så fall hur mycket (kausalitetsfrågan)?” Det går

---

[7]En sådan definition är kontroversiell. Inom biståndsutvärdering är utgångspunkten ofta OECD/DACs sex utvärderingskriterier (se OECD/DAC, 2019).

[8] Det finns flera alternativ till måluppfyllelsemodellen. Exempel är utvärdering med bedömningsmatriser (Rubrics-Enhanced Evaluation, Davidson, 2004) principbaserad utvärdering (Principles-Focused Evaluation, Patton, 2017), indikatorbaserad utvärdering samt målfri utvärdering (goal-free evaluation, Scriven, 1991).

alltså inte att utvärdera måluppfyllelsen av en insats utan att hantera eller reflektera kring kausalitetsfrågan. En vanlig missuppfattning tar avstamp i det logiska ramverket (Figur 1) och utgår ifrån att kausalitetsfrågan bara behöver hanteras vid steget Impact.

Ekonomistyrningsverket (ESV, 2006, sid. 7) skiljer på *beskrivande* och *förklarande* målanalys och slår fast behovet av att fastställa att det faktiskt är insatsen som bidragit till att ett mål uppnåtts eller påverkats. Detta gäller inte minst stegen outcome och impact i det logiska ramverket.[9]

Figur 1: Det logiska ramverket

Utvärderingar som utger sig för att bedöma måluppfyllelsen men som inte beaktar kausalitet eller externt påverkande faktorer riskerar att vara helt eller delvis missvisande. Att demonstrera insatsens egna bidrag till resultatet är avgörande för att insatsens värde och framtida inrikting ska kunna avgöras (Mayne, 2001). Utmaningen i att bedöma kausalitet och påverkan från olika faktorer i förhållande till ett specifikt mål eller utfall ökar samtidigt med tidens gång och grad av komplexitet (Forss et al. 2011). Metodutvecklingen om hur kausalitet kan hanteras med kvalitativ och kvantitativ metod i utvärderingar har dock varit snabb under senare år och antalet designalternativ ökar hela tiden (Stern et al, 2012; Sandahl och Petersson, 2016; Molander, 2017).

[9] Kausaliteten vid outputbaserade målsättningar liksom input och aktivitet kan delvis överblickas i den verksamhet som utvärderas, av målgrupper och indirekt av utvärderaren, men ett externt stöds additionalitet och påverkan på kortsiktiga aktiviteter och output är inte självklar. Det finns många exempel på offentliga interventioner med svag additionalitet på input- och outputnivå.

Den kritiska litteraturen om målbaserad utvärdering är omfattande. Tabell 1 summerar några av de styrkor och svagheter som lyfts fram i olika delar av litteraturen om måluppfyllelsemodellen.

Tabell 1: Måluppfyllelsemodellens styrkor och svagheter

	Styrkor (+)
Demokrati	Modellen bejakar uppfattningen att mål för politiken ofta spelar en central roll i demokratiska system. Skattebetalarna har demokratisk rätt att få veta om offentliga insatser lever upp till offentligt deklarerade avsikter. Modellen kan ta officiella politiska mål på allvar.
Objektivitet	Modellen kan ge en opartisk objektiv lösning på utvärdering som en normativ och värderande aktivitet.
Enkelhet och transparens	Modellen är förhållandevis lätt att tillämpa och kommunicera till beslutfattare, intressenter och allmänhet.
	Svagheter/utmaningar (-)
Luftighet, oklarhet eller förenkling	Offentliga insatsers mål är ofta så otydliga, svagt prioriterade, och många att de är oanvändbara som bedömningsinstrument.
	Verksamheter med komplexa uppdrag kan ha svårt att förenkla verksamheten genom att anta ett fåtal mål. Svårmätbara fenomen riskerar att prioriteras bort.
Inflation och dolda motiv	Mål formuleras ofta för att skapa sympati eller ge legitimitet och inte för att de ska uppnås eller ligga till grund för utvärdering.
	Modellen kan vara blind för den roll som ej redovisade motiv spelar i målformuleringar och offentliga insatser. De ”egentliga” målen formuleras eller kommuniceras inte alltid.
Oförutsedda bieffekter	Med fokus på förutsedda formella mål kan utvärderingen förbise att offentliga interventioner inte sällan har starka, oförutsedda (negativa eller positiva) bieffekter.

	Svagheter/utmaningar (-)
Toppstyrning, rigiditetet, tunnelseende och suboptiomering	Målstyrning riskerar att skapa toppstyrda, rigida suboptimerande verksamheter som inte ser behovet av lärande och anpassning och istället fokuserar på kortsiktigt enkelt mätbara eller manipulerbara resultat. Måluppfyllelseutvärdering kan då tolkas som en del i ett sådant skadligt system.

Källa: Sammanställningen baseras på Vedung (2012), Schutt (2016) och Alexius (2017).

Många utvärderare skulle nog hålla med om att luftiga, oklara och förenklande mål åtminstone utgör utmaningar även om de inte behöver vara problematiska i praktiken. Utvärderingstänkande bygger delvis på förmågan att itinitalt och löpande precisera genomarbetade beskrivningar av vad som skulle uppnås och hur något ska uppnås, med noggrann begreppsanvändning i operationaliseringen av utvärderingsfrågor (se Leeuw, 2003, om rekonstruering av mål och förändringsteorier i utvärdering). Erfarna utvärderare förväntas vara vana att hantera en oprecis målbild i dialog med beställare och insatsansvariga.

Samtidigt är det lättare att bedöma måluppfyllelse i insatser med en tydlig, precis och lagom ambitiös målbild. Operationaliseringen underlättas och förutsättningen för valida, reliabla slutsatser stärks.

Träffande är även invändningen om insatsers ”egentliga” eller outtalade mål och intentioner. Sådant kan vara svårt att hantera för utvärderaren då det kan finnas incitament hos beställare eller insatsansvariga att aktivt styra bort från frågan när en utvärdering ska genomföras.

Ekonomistyrningsverket har lyft fram att målsättningar i statlig verksamhet inte behöver leva upp till SMART-kriterierna om ansvariga har en tydlig idé om hur målet följas upp eller utvärderas (Eliasson, 2006). Otydliga mål kan preciseras genom operationalisering. Detta kräver en process för att tillsammans med intressenter och ansvariga rekonstruera vad man vill uppnå och att man operationaliserar förväntan som indikatorer, mått eller kriterier på framgång.

En potentiell komplikation berör också distinktionen mellan strävansmål och uppnåendemål.[10] Värderingen av måluppfyllelsen skiljer sig åt om målet är ett utryck för långsiktig visionär strävan eller om tanken verkligen är att målen ska uppnås. Om en projektportfölj innehåller både insatser med strävansmål och uppnåendemål så är det kanske tveksamt om måluppfyllelsen kan jämföras mellan insatser.

Easterly (2006) har beskyllt det internationella biståndet för *Utopian Planning*. Målen är, menar han, orealistiska i relation till de insatser/medel biståndet har tillhanda. De ambitiösa målen skapar ouppnåeliga förväntningar, besvikelse och försvårar utvärdering genom en svag koppling (en ”missing-middle”) mellan övergripande mål och interventionsmål. De enskilda insatserna syftar endast *indirekt* till att uppnå de mål som den politiska nivån specificerat. Regeringens mål, såsom de uttrycks i olika biståndsstrategier, skiljer sig dock åt i detta avseende.[11]

Inom diskussionen om så kallad målfri utvärdering (”goal free evaluation”, Schriven, 1991) har det framförts att målbaserad utvärdering kan ge tunnelseende (”tunnel vision”). Förespråkare för modellen, som sällan tillämpas, menar att utvärderingar bör fokusera på vad insatser faktiskt gör och uppnår. Det är inte intentionerna som räknas utan vad man åstadkommer.

---

[10] Distinktionen används exempelvis i skolans läroplan. Ett exempel på strävansmål är nollvisionen på trafikområdet.

[11] Strategin för stöd genom civila samhället för perioden 2016-2022 väntas exempelvis bidra till att uppnå ”stärkt kapacitet hos civilsamhällesaktörer i utvecklingsländer” och ett ”mer gynnsamt samhällsklimat för civilsamhällesorganisationer i utvecklingsländer.” Strategin för Sveriges globala utvecklingssamarbete inom hållbar ekonomisk utveckling 2018-2022 ska bl.a. åstadkomma ”mer effektiv inhemsk resursmobilisering”, ”ökad finansiell stabilitet” och ”motverkande av korruption”. Skillnaderna i ambition kan vara stor inom och mellan strategier.

Utvärderingsbarhet och att dra slutsatser

Karlsson Westman (2011) lyfter att valet av måluppfyllelse som utvärderingskriterium görs utifrån antagandet att ”det finns ett samband mellan mål och resultat som kan mätas”, dvs. ett antagande om utvärderingsbarhet, att måluppfyllelse kan bedömas vid den tidpunkt då utvärderingen genomförs, att relevanta data finns tillgängliga eller kan samlas in, att utvärderingsteamet har förmåga och vilja att tillförlitligt operationalisera målsättningarna och göra nödvändig kausal analys.[12]

En utmaning vid utvärdering av måluppfyllelse är ofta att komma till ett övergripande värdeomdöme om den insats som utvärderas (Forss, 2007). Dessa värdeomdömen kan vara viktiga om en givare exempelvis ska bestämma sig för om en satsning ska skalas upp eller ett stöd avslutas. Davidson (2014) menar att det inte räcker att dra slutsatser om målen uppnåtts eller inte, utvärderaren bör tydligt värdera måluppfyllelsen. Utan detta moment uppnås inte mervärdet med utvärderingen och användbarheten för beslutsfattare begränsas.

I många biståndsinsatser har utvärderaren flera målsättningar att hantera så det är inte självklart hur eller om måluppfyllelseanalysen ska summeras (exempelvis i en situation där den bedöms som god för två mål men som svag för två andra). Genomföraren och finansiären av projektet kan prioritera olika målsättningar och det kan ske målförskjutningar eller omprioriteringar över tid. En närliggande svårighet är att koppla samman måluppfyllelse på projektnivån med övergripande mål på strateginivå eller för politiken som helhet (jmfr. Christoplos et al, 2014).

[12] Holvoet et al (2018) menar i en studie av utvärderingsbarhet i belgiskt bistånd att ”majoriteten projekt har villkoren på plats för att på ett tillfredsställande sätt utvärdera inre och även yttre effektivitetet”.

Det bör poängteras att en utvärdering som enbart bedömer kortsiktig output, outcomes, del- eller aktivitetsmål snarare än övergripande förväntade slutgiltiga mål *per definition* inte gjort en fullständig bedömning av måluppfyllelse och effektivitet. En bedömning av effektivitet är en bedömning i relation till faktiska *förväntade* slutgiltiga mål/resultat för en aktuell verksamhet.

En begränsning vid måluppfyllelseutvärdering föreligger också om man inte inkluderar kostnadseffektivitet, dvs. väger in insatsens kostnader vid bedömningen av måluppfyllelsen (Sandberg och Faugert 2012). God måluppfyllelse har inte samma värde om den sker till en för hög kostnad eller om det finns mer kostnadseffektiva alternativ. Inom biståndsutvärdering försöker man som regel uppnå en mer helhetlig bedömning genom att använda samtliga eller flera av OECD/DACs utvärderingskriterier.

Kortsiktig förändring, väntade och oväntade resultat

God måluppfyllelse avseende mer kortsiktiga prestationer i ett projekt garanterar inte goda resultat på halvlång sikt och goda resultat på halvlång sikt garanterar inte heller goda långsiktiga effekter. Tvärtom säger kortsiktig förändring ofta förhållandevis lite om resultat och påverkan på lång sikt. Forss (2021) konstaterar att:

> ”det finns en risk att kortsiktiga utvärderingar missar förändring som är transformativ, som sker vid en tipping point, som är episodisk, sätter punkt, är exponentiell eller omvälvande.”

Inom biståndet placerar OECD/DACs utvärderingskrititerier oväntade, positiva och negativa, resultat och bieffekter under kriteriet Impact. Det är dock centralt att diskutera och förhålla det förväntande i relation till det oväntade också vid bedömning av måluppfyllelse (se också diskussionen i bilaga 4):

> ”Ifall utvärderingen som i måluppfyllelsemodellen koncentreras till frågan om reformens egna mål har förverkligats, kommer den per definition inte att spåra upp serendipiteter och oförutsedda bieffekter. Anledningen är att dessa inte kan vara målsatta och därmed ingå i undersökningsfältet eftersom de är oförutsedda. Detta skapar en vit fläck, som kan ge en skev bild av vad reformerna uppnått” (Vedung 2012, sid 101).

Flexibilitet och anpassning

En diskuterad svaghet i måluppfyllelsemodellen som delvis kan sättas i fråga är toppstyrning och rigiditet (se Tabell 1). Forskare har under senare år lyft behovet av flexibilitet, realism, lärande och anpassningsförmåga i biståndsinsatser (exempelvis inom ramen för Problem Driven Iterative Approach (PDIA, Andrews et al, 2015) och Doing Development Differently (DDD, Honig & Gulrajani, 2017).

En motsättning mellan målbaserad utvärdering och värden som flexibilitet, realism, lärande eller anpassning kan uppstå om utvärderaren inte väger in behov av flexibilitet och anpassning avseende ambition eller målformuleringars förhållande till varandra. Men om en insats vid en viss tidpunkt inte uppnår sina mål så uppnår den - helt enkelt - inte sina mål. Kritik mot målbaserad utvärdering ur ett ”adaptivt” perspektiv kan därmed ibland bygga på en sammanblandning av mål och medel. Det finns som Vedung (2012) slår fast inget ”i måluppfyllelsemodellen som säger att insatser ger upphov till avsedda utfall”, tvärtom är det just vad som undersöks. Givet att mål behöver justeras, ambitionsnivån sänkas eller höjas för att skapa rätt avvägning mellan utmaning och realism så kan en oflexibelt genomförd målutvärdering ändå krocka med ett adaptivt förhållningssätt. Initialt syftade också mål- och resultatstyrning tydligt till flexibilitet och delegerad implementering (Jacobsson och

Sundström, 2001). Man bör inte heller ändra strategi, arbetssätt eller process i en verksamhet vars effektivitet och resultat man inte känner.[13]

Incitament och makt

Den kanke mest berömda invändningen mot målbaserad utvärdering och styrning kallas *Goodharts lag* (Goodhart, 1975), ursprungligen formulerad för penningpolitiken som att:

> "Varje observerad statistisk regelbundenhet tenderar att kollapsa när tryck läggs på den för kontrolländamål".

Mål är kanske inte otydliga för att man inte kan formulera dem rätt utan för att man inte vill.[14] Incitamenten att ta fram tydliga mål kan vara svaga från flera intressenthåll i exempelvis biståndsinsatser. Måluppfyllelse i en verksamhet ska ofta rapporteras mot flera finansiärer varför alltför tydliga mål kan bli direkt opraktiska.[15]

---

[13] Besvärligare ur ett adaptivt perspektiv är när biståndets logiska ramverk (ex vis LFA, Logical framework approach) används som styrande projektplanerings-verktyg och då styr både mål och medel samtidigt. LFA-matriser innebär därför en starkare intervenerande styrning än traditionell mål- och resultatstyrning.

[14] Tarchys (2006) menar att: ”rapport efter rapport når därför fram till samma entoniga slutsats: Beslutsfattarna måste tydligare precisera vad de vill uppnå! Men det vill inte beslutsfattarna. Av flera olika skäl är de inte alls särskilt hågade att precisera sig. Ett kan vara att man inte vill ge slag på sig genom att ställa upp tydliga mål som sedan inte uppnås; ett annat kan vara att vaghet och flertydighet underlättar kompromissande och samstämmigt beslutsfattande; ett tredje kan vara att politikerkollektivet som sådant inte behöver utöva ett tydligt ledarskap i förhållande till förvaltningen.”

[15] Det faktum att genomgångar ständigt visat att mål i biståndsinsatser är otydliga (se nedan) talar för att förtydligande eller rekonstruktion av mål för att möjliggöra utvärdering delvis bör läggas på utvärderaren och inte bara på insatsen.

Pritchett (2002) menar att begränsade investeringar i effekt-utvärdering inom ett politikområde kan förklaras av intressegruppers incitament att inte finansiera trovärdig kunskap då sådan kan undergräva förmågan att mobilisera politiskt och budgetmässigt stöd. Det är också belagt internationellt att utvärderare ibland upplever press att leverera ”rätt” budskap till sina beställare (se exempelvis Pleger et al, 2020).

Samtidigt behöver det inte ses som illegitimt att en tjänsteman exempelvis beställer utvärderingar för att hitta belägg för en specifik insatstyps effektivitet inom ramen för en intern förhandling om resurser och inriktning för verksamheten. Detta givet att den som genomför utvärderingen upprätthåller grundläggande kvalitets-kriterier för utvärdering, som tilllförlitlighet och etiskt genom-förande.

Måluppfyllelse, lärande, användning

Ur lärandeperspektiv innebär måluppfyllelseutvärdering i praktiken ofta så kallad single-loop learning (Argyris & Schön, 1976). Mål i sig ifrågasätts inte alltid, utan endast måluppfyllelsen.[16] Det positiva i detta ligger i utvärderarens demokratiska underordning gentemot politikens mål vilket stärker värderingsmomentets legitimitet. Men det bygger då på att utvärderaren faktiskt utgår ifrån övergripande eller politiska mål i analysen, vilket inte alltid är fallet. I praktiken fokuserar utvärderingar ofta på kortsiktiga eller mellanliggande projektmål som inte formulerats i den politiska sfären.

Centralt för lärande och användbarhet är att utvärderingar bedömer *varför* och *hur* en viss måluppfyllelse uppkommit. Vilka mekanismer som förklarar graden av måluppfyllelse? Varför fungerade insatser eller varför fungerade den inte? Skulle den fungera under andra förutsättningar, i framtiden, för vem, varför, hur då, osv. (Stern et al, 2012). Har vi överhuvudtaget formulerat rätt mål? Det sista kan med

[16] I biståndet bedöms målens relevans ofta utifrån relevanskriteriet.

Argyris (1993) kallas double-loop learning. Vissa forskare anser också att det finns något man kallar Triple-loop-learning som handlar om att ”lära sig att lära” och som till exempel kan motsvara en process där man mer fundamentalt sätter ifråga en organisations förmåga till lärande genom målbaserad utvärdering (Swieringa och Wierdsma, 1992). Finns det exempelvis mer grundläggade principer, faktorer eller komponenter i utvärderingsprocesserna som gör att dessa inte stödjer organisationen i sin strävan att lära?

Summering: litteratur om måluppfyllelsemodellen

Det finns en omfattande litteratur om mål, måluppfyllelse och utvärdering. Till del är litteraturen normativ och föreskriver hur mål och resultat bör beskrivas, hur styrning och beslut ska utformas och utvärderas. En annan del av litteraturen är deskriptiv, och visar hur verkligheten har en tendens att trassla till det. Den deskriptiva litteraturen är överlag mera skeptisk och kritisk, och även empiriskt grundad. Det kan finnas alltså anledning att inte ha för högt ställda förväntningar på analys av målutvärdering i praktiken.

3. Mål och uppdrag i 80 utvärderingar

I detta första empiriska kapitel kartläggs insatsernas mål och uttryckta förväntningar på utvärderingarna i urvalet av 80 decentraliserade utvärderingar. Jag inleder dock med en kortare bakgrund till Sidas utvärderingsarbete, med fokus på decentraliserade utvärderingar.

OECD/DAC:s utvärderingskriterier har sedan de tillkom 1991 blivit en internationell standard för bedömningsgrunder inom biståndsutvärdering (OECD/DAC, 2019, se även Chianca, 2008). Som ett av sex kriterier ingår bedömning av insatsens måluppfyllelse (effectiveness).

Kriteriet hade sannolikt ingått vid utvärdering av svenska biståndsinsatser även om OECD/DACs kriterier inte hade funnits. Några exempel på processer och händelser som över tid kan ha främjat ett fokus på målbaserad utvärdering i det svenska biståndet är utvärdering som ”förvaltningspolitisk megatrend” föreskrivande utvärdering inom nästan alla områden (Vedung, 2003), cykliska resultatagendor (Vähämäki, 2017) och den internationella biståndseffektivitetetsagendan (Paris, Accra, Busan) som lyfte vikten av resultat.[17] Sida gjorde också utvärderingar av måluppfyllelse i biståndsinsatser innan OECD/DACs kriterier tillkom.

[17] Det är tveksamt att ensidigt koppla målfokuserad utvärdering till det idékomplex som kallas *New Public Management* (Hood, 1991). Vedung (2004) menar att utvärderingssverksamheten i Sverige vuxit fram genom ”sedimentering” efter 4 historiska ”böljor” från 1960 och framåt: en vetenskapsfokuserad våg (med rötter i USA och 1960-talet), en nyliberal våg, en dialogorienterad våg och en evidensvåg. På biståndsområdet studerar Vähämäki (2017) fyra ”resultatinitiativ” som lanserats på Sida under åren 1971, 1981, 1998 och 2012. I samtliga dessa initiativ tycks utvärdering och bedömningar mot mål ha varit (mer eller mindre) viktiga beståndsdelar (för ett tidigt exempel se: *Resultatutvärdering- ett programförslag*, SIDA 1971). OECD/DACs kriterier har dock sannolikt starkt bidragit till att institutionalisera måluppfyllelse som bedömningsgrund i biståndutvärdering.

Utvärdering på Sida

Sida arbetar med en användningsfokuserad utvärderingsansats (Patton, 2008). Tonvikten ligger på att ”identifiera vem de avsedda användarna av en specifik utvärdering är […] Om det är oklart vad den avsedda användningen av en utvärdering är eller vem de avsedda användarna är ska utvärderingen inte genomföras”(Sida 2018; 2020). Sida anger samtidigt på sin hemsida att utvärderingsarbetet syftar till ansvarsutkrävande (accountability) genom insyn.[18]

Myndigheten skiljer på tre former av utvärderingar: decentraliserade utvärderingar, strategiska utvärderingar och partnerledda utvärderingar. De decentraliserade utvärderingarna (som ligger i fokus för denna studie) beställs av ambassader eller på Sidas avdelningar i Stockholm och berör oftast enskilda projekt eller program. De strategiska utvärderingarna, som beslutas av myndighetens generaldirektör, tar som regel ett bredare grepp och syftar till att åstakomma lärande på verksövergripande nivå. De partnerledda utvärderingarna beställs av Sidas partners men bedömer insatser som Sida finansierat.[19]

Sidas representant i studiens referensgrupp betonar att de decentraliserade utvärderingarna ska ses som stöd i insatshanteringen, i specifika beslutssituationer eller utifrån särskilda behov. De beställs ofta inför beslut om en ny projektfas och ska enligt utvärderingsenheten ses i ett sammanhang där de ofta genomförs halvvägs i projektcykeln som en praktisk, insatsnära och behovsfokuserad verksamhet. Man betonar att även partners genomför utvärderingar och att Sida därmed är en av flera aktörer som analyserar, följer upp och utvärderar biståndet. Samtidigt

[18] https://www.sida.se/Svenska/Samarbetsparter/resurser/utvarderingar/ (Läst 2020-06-15).

[19] Inga partnerledda utvärderingar finns med i denna studie.

hänvisar Sida i hög grad till sina utvärderingar i årsredovisningen (2019) och på nya hemsidan till exempel när man kommunicerar sitt utvärderingsarbete mot mål.[20]

Ansvaret för de decentraliserade utvärderingarnas kvalitet ligger på projektledande författare, på den som beställt utvärderingen samt kvalitetsansvarig konsult. Sidas utvärderingsenhet är ansvarig för kvaliteten i de decentraliserade utvärderingarna endast i en institutionell mening genom att man upphandlar, administerar och kvalitetssäkrar det ramavtal som ligger till grund för utvärderingarna. [21] Enheten följer upp hur det fungerar med leverantörerna, man tillhandahåller metodstöd, erbjuder rådgivning, ger ut utvärderingshandboken, har dialog om uppdrags-beskrivningar, utbildar, följer upp ramavtalet och diskuterar slutsatser och kvalitet i enhetens årsrapport.

Sida utgår i sitt utvärderingsarbete från OECD/DACs utvärderingskriterier och dess definition av måluppfyllelse och preciseringar av kvalitet. Sidas utvärderingshandbok (2020) diskuterar inte effektivitetskriteriet mer än att definitionen från OECD DAC återges. Handboken är tydlig med att utvärderingar ska vara tillförlitliga och man betonar vikten av integritet från utvärderaren (Ibid sid 10).

Följande genomgång av insatsernas målbild och Sidas beställningar baseras på 80 decentraliserade utvärderingar publicerade åren 2017-2019 (för mer om urval, se bilaga 1).

---

[20] https://www.sida.se/sa-fungerar-bistandet/sa-har-mater-sida-resultat (läst 2021-03-23).

[21] Mailkorrespondens, Sven Olander, 2021-03-11. I dialog med tidigare chefsutvärderare har det framkommit att kvalitetsansvaret för de decentraliserade utvärderingarna då inte bedömts ligga på Sida UTV.

Vilka mål ska nås?

Tabell 2 listar de utvärderade insatsernas målbilder avseende övergripande insatsmål efter förekomst. Sammanställningen visar på stor pluralism avseende vilken typ av långsiktiga mål som de utvärderade insatserna har.

Ambitionsnivån skiljer sig starkt mellan insatser då några ska sprida information eller stärka kunskaper medan andra ska åstadkomma ökad ekonomisk tillväxt eller minska våld.[22] Insatserna ska ibland, men långt ifrån alltid, uppnå stora ambitösa mål som fattigdomreducering. [23] Mål som stärkta mänskliga rättigheter, minskat våld, ökad handel, nya jobb och liknande förekommer. Den största gruppen insatser har dock mål som att förändra eller förbättra offentliga institutioners arbetssätt, stärka kapacitet, stärka mobilsering, deltagande, ökad kunskapsanvändning, sprida kunskap och dyl.[24]

---

[22] Ett stort antal utvärderingar hade mål unika för den utvärderade insatsen. Tabellen innehåller endast mål som återfinns i åtminstone tre utvärderingar.
[23] Samtidigt kan fattigdomsmålet ligga implicit i insatsen.
[24]Ambitionsnivå avseende mål i biståndet bör förstås i relation till specifika sammanhang. Vad som är ett ambitiöst mål på skattemyndigheten i Afghanistan behöver inte vara det på skattemyndigheten i Tanzania.

Tabell 2: Övergripande insatsmål i utvärderingarna per förekomst

Förändra eller förbättra offentliga institutioners funktions-/arbetssätt eller kapacitet inom en sektor	12
Stärka implementering eller efterlevnad av mänskliga rättigheterna inom ett/ flera områden (tematiskt eller geografiskt)	9
Stärkt kapacitet att bedriva forskning	8
Stärkt användning av forskningsresultat	7
Skapa mobilisering, deltagande eller åstadkomma demokratiskt efterfrågetryck	5
Minskad fattigdom i region, land eller landsdel	4
Ekonomisk empowerment för särskild grupp	4
Konfliktlösning, fred, fredliga samhällen och säkerhet	3
Kunskapssprida, utbilda, kompetensutveckla, informera om område/fråga	3
Nya och bättre jobb/arbetstillfällen	3
Förbättrade levnadsförhållanden eller ökande inkomster	3
Stärkt handel	3
Minskat våld i olika former, sammanhang eller gentemot olika grupper	3

Källa: Samtliga decentraliserade utvärderingar i urvalet 2017-2019. Endast mål som förekommer tre gånger eller mer finns med i tabellen. Målen är framtagna induktivt genom att övergripande mål kodats i breda kategorier.

En komplikation studien inte fördjupat sig i men som tagits upp i referensgruppen handlar om att målen ofta formuleras som att något ska öka, stärkas, skapas, bidra eller expandera, dvs. i verbform istället för som ett tillstånd. Det är sällan som det inte händer något i ett biståndsprojekt; aktörerna jobbar på, har seminarier, utbildningar, organiserar sig, arrangerar workshops, genomför aktiviteter etc. Lite

hårdraget kunde man då dra slutsatsen att det bara är när ingenting händer eller verksamheten helt stannat av som målen i fråga inte uppnås. [25]

Av de 68 utvärderingar som enbart utvärderar en insats har 27 (40%) fler än ett övergripande och långsiktigt mål. Samtidigt ligger fokus ofta på kortsiktiga utfall och resultat. Utvärderaren har därför i praktiken betydligt fler än ett mål att förhålla sig till. Effektivitetskriteriet innehåller ibland 10 på förhand specificerande utvärderingsfrågor. Utöver detta ska utvärderaren besvara frågor om relevans, kostnadseffektivitet, hållbarhet etc.

Intrycket är att insatserna oftare har uppnåendemål än visionära mål. I utvärderingarna återkommer en förväntan om att målstrukturer och logiska ramverk ska fungera styrande för insats, uppföljning och utvärdering. Det har dock inte varit möjligt att bestämma hur stor andel av målen som är uppnåendemål respektive visionära mål, vilket kan påverka jämförbarheten mellan insatser.

Insatserna syftar ofta till att stärka olika organisationer som civilsamhällesorganisationer, universitet och myndigheter (Tabell 3). En tredjedel av de utvärderade insatserna förväntar sig enbart resultat direkt för fattiga människor, medborgare, rättighetsbärare, i natur eller ekosystem.[26]

---

[25] Om hypotesen stämmer innebär den att måluppfyllelsemodellen inte fungerar för att målen inte formuleras på optimalt sätt. Ambitiösa projekt eller välformulerade tillståndsmål riskerar att missgynnas vid jämförande analyser. SMART-kriterierna anger att mål ska vara realistiska (R), men de måste också vara utmanande.

[26] Detta ska inte ses som en värdering av insatsen.

Tabell 3: Ska övergripande långsiktiga resultat åstadkommas direkt eller i processer, organisationer och verksamheter?

I både och	29	(36%)
I processer, organisationer, verksamheter som ska stärkas, byggas el förändrasa	27	(34%)
Direkt i berörda samhällenb	23	(29%)
Ej relevant/kan ej läsas ut	1	(1%)
Totalt	80	(100%)

Noter: [a] Som därefter direkt/indirekt kan påverka människor/samhälle/natur. [b] Direkt mot medborgare, rättighetsbärare, i natur eller ekosystem. Källa: Samtliga 80 decentraliserade utvärderingar i urvalet 2017-2019.

Vad ska utvärderas?

Analys av effektivitet utgör en central del av uppdraget i alla 80 utvärderingar. Genomgången (Tabell 4) visar dock att Sida i uppdragsbeskrivningarna ofta inte explicit specificerar att det är just måluppfyllelse som avses vid bedömning av effektivitet, utan istället används andra formuleringar som förväntade/planerade resultat eller resultat-komponenter (outcome components).

Tabell 4: Efterfrågar uppdragsbeskrivningen *explicit* en bedömning av måluppfyllelse eller används andra begrepp?

Ja, måluppfylllelse efterfrågas	34	(43%)
Nej, andra begrepp används	43	(54%)
Uppdragsbeskrivning saknas	3	(4%)
Totalt	80	(100%)

Källa: Samtliga 80 decentraliserade utvärderingar i urvalet 2017-2019. Samtliga dessa utvärderingar fokuserar på effektivitetskriteriet.

Som framgår i Tabell 5 analyserar eller diskuterar utvärderarna ofta (66 fall) i någon utsträckning formuleringen av insatsens mål eller förväntade resultat. Kritiken handlar ofta om hur mål/förväntade resultat formulerats angående ambition, specificering, mätbarhet, realism, konsistens, acceptans bland involverade parter eller tidsperspektiv.

Tabell 5: Vad säger utvärderarna om formuleringen av mål/förväntade resultat?

Utvärderarna är kritiska el delvis kritiska.	54	(68%)
Utvärderarna är positiva/mycket positiva.	12	(15%)
Formuleringen diskuteras inte.	14	(18%)
Totalt	80	(100%)

Källa: Samtliga 80 decentraliserade utvärderingar i urvalet 2017-2019.

Måluppfyllelsekriteriet motsvaras i praktiken, i det som efterfrågas av Sida, av ett ganska stort antal besläktade frågor och underfrågor i utvärderingarnas uppdragsbeskrivningar (Tabell 6). Många beställare (79%) specificerar särskilda typer av resultat/mål eller nedbrytningar av resultat som utvärderaren förväntas analysera.[27]

Drygt en tredjedel av uppdragsbeskrivningarna (38%) efterfrågar analyser av mekanismer eller förklaringar till måluppfyllelsen eller resultatet. Ungefär en fjärdedel av uppdragsbeskrivningarna (23%) vill se analyser av insatsens result- och uppföljningsarbete medan en femtedel (21%) efterfrågar bedömningar av program-/förändringsteori eller logiskt ramverk. Knappt hälften av uppdragsbeskrivningarna (visas ej i tabell) efterfrågar bedömningar av oförutsedda (positiva eller negativa) resultat.

[27] Det är ibland oklart om de mål Sida vill utvärdera är de initialt förväntade eller om man fokuserar på mål eller resultat som man senare intresserat sig för.

Tabell 6: Vad efterfrågar uppdragsbeskrivningen utöver svaret på grundfrågan (måluppfyllelse) under effektivitetskriteriet?

Analys av:		
Specifika nedbrytningar av resultat	63	(79%)
Mekanismer el förklaringar (till måluppfyllelse)	30	(38%)
Insatsens resultat- och uppföljningsarbete	18	(23%)
Programteori, förändringsteori, logiskt ramverk	17	(21%)
Kan målen nås i framtiden	6	(8%)
Har aktiviteter, outputs eller resultat levererats i tid	4	(5%)

Källa: Samtliga 80 decentraliserade utvärderingar i urvalet 2017-2019. Flera val är möjliga.

Diskussionen inom biståndet under senare år har delvis handlat om behovet av adaptivitet i biståndsinsatser varför det är intressant att se om mål och förväntade resultat förändrats eller justerats under den period som utvärderas i insatserna. Av Tabell 7 framgår att målen justerats eller förändrats i knappt hälften (44%) av insatserna som utvärderas i urvalet. I tio av dessa har de förtydligats eller rekonstruerats i samband med själva utvärderingen. Otydliga målformuleringar tycks sällan ha föranlett någon rekonstruering av målen för att underlätta utvärdering.[28]

Tabell 7: Har mål/förväntade resultat i insatsen förändrats el justerats under den period som utvärderas?

Nej	34	(50%)
Ja, vid övergång mellan projektperioder eller av andra skäl	20	(29%)
Ja, de har rekonstruerats, förtydligats, strukturerats om av utvärderarna	10	(15%)
Inte möjligt att läsa ut ur rapporten/framgår inte	4	(6%)
Totalt	68	(100%)

Källa: Samtliga 68 decentraliserade utvärderingar i urvalet 2017-2019 som fokuserat på en enskild insats. 12 utvärderingar fokuserade på fler än en insats.

[28] Sida menar att de brukar lyfta behovet av målförtydligande med utvärderarna och att detta har tydliggjorts i metodstöd och vid upphandling.

I Tabell 8 anges utvärderingarnas högsta angivna resultatsteg (jmf. Figur 1). Många utvärderingar (61%) anger att de analyserar långsiktiga mål eller impact. Det är dock *inte* samma sak som att en fullödig analys av effektmål genomförs. Avsnittet om impact, som finns med är ofta kort relativt övriga delar. Utvärderingarnas huvudsakliga vikt ligger starkt på stegen output och outcome, också i utvärderingar som har impact som högsta analyssteg. En tredjedel av utvärderingarna (36%) anger outcome som högsta analyssteg och bedömer således enbart måluppfyllelse på nivån output och outcome.

Tabell 8: Högsta tidsliga resultatsteg i mål/resultatanalysen (enligt utvärderarna)

Impact	49	(61%)
Outcome	29	(36%)
Inte möjligt att läsa ut	2	(3%)
Totalt	80	(100%)

Källa: Samtliga 80 decentraliserade utvärderingar i urvalet (2017, 2018, 2019).

Det finns dock inget riktigt tydligt stöd i utvärderingarnas uppdragsbeskrivningar för att inte utvärdera övergripande målsättningar (Tabell 9). Även om beskrivingarna är bitvis otydliga så uttrycker beställarna sällan att de *inte* efterfrågar en sådan bedömning, utan att insatserna ska bedömas mot på förhand uppsatta mål i resultatramverk eller i linje med vad som överenskommits.

Tabell 9: Exkluderar uppdragsbeskrivningen uttryckligen bedömning av övergripande/långsiktiga mål?

Ja	4	(7%)
Nej	47	(80%)
Oklart/texten är motsägelsefull i detta avseende/ToR saknas	8	(14%)
Totalt:	59	(100%)

Källa: Samtliga 59 decentraliserade utvärderingar i urvalet 2017-2019 som fokuserat på en enskild insats och där slutsatsen om måluppfyllelse gått att läsa ut i rapporten. Tolv utvärderingar bedömmer fler än en insats. Nio utvärderingar har slutsatser i vilka måluppfyllelse inte går att utläsa.

Summering: mål och uppdrag i utvärderingarna

Kartläggningen visar på stor pluralism avseende övergripande målbild i utvärderade insatser. Ambitionsnivån skiljer sig i hög grad mellan projekt och utvärderarna är ofta delvis kritiska till hur mål/förväntade resultat formulerats avseende ambition, specificeringsgrad, mätbarhet, realism, konsistens mellan dokument, acceptans bland involverade parter och tidsperspektiv. [29] De övergripande målen har ibland justerats under insatsperioden vilket eventuellt kan vara ett tecken på lärande eller adaptivitet. De, enligt utvärderarna, otydliga målen tycks sällan ha föranlett något förtydligande för att möjliggöra utvärdering i enlighet med de metoder som finns för detta (se Leeuw, 2002).

[29] Långt ifrån alla uppdragsbeskrivningar och utvärderingar talar explicit och uttryckligt om mål och måluppfyllelse i utvärderingarna. Istället används begrepp så som förväntade resultat (intended outcomes) eller exempelvis resultat-komponent (outcome component). Vad som skiljer mål från förväntade resultat här är oklart. En skillnad ligger möjligen i måluppfyllelsemodellens mer regelfästa idéer och föreställningar om hur mål formuleras och ordnas för användbarhet vid styrning, uppföljning och utvärdering. Referensgruppen till denna studie menade att det är en fråga om språkbruk och att skillnaden i praktiken är liten.

Uppdragsbeskrivningarna utrycker sällan att man inte efterfrågar analys av övergripande måluppfyllelse. Insatserna ska bedömas mot på förhand uppsatta mål i resultatramverk eller i linje med vad som överenskommits. Fokus på oväntade resultat och effekter efterfrågas i cirka hälften av uppdragsbeskrivningarna. Drygt en tredjedel efterfrågar analyser av bakomliggande förklaringar till måluppfyllelse/uppnådda förväntade resultat vilket är en viktig fråga om syftet med utvärderingarna är lärande.

4. Slutsatser om måluppfyllelse

Vilka slutsatser drar då utvärderarna om insatsernas måluppfyllelse? I detta kapitel är ambitionen att besvara studiens fråga 1. I urvalet ingår både decentraliserade och strategiska utvärderingar.

Två viktiga begränsningar

Två begränsningar i utvärderingarna bör nämnas innan slutsatserna om måluppfyllelse presenteras. Dessa försämrar delvis förutsättningarna för en reproducerbar syntes:

1. Bedömningarna om måluppfyllelse är ofta otydliga, inlindade, motsägelsefulla eller sammansatta. Nio utvärderingar var så otydliga att deras slutsatser inte kunnat skalstegsbedömas.
2. Övergripande samlade bedömningar görs inte alltid.

Läsarens förståelse utmanas när frågeställningar och slutsatser presenteras utan att insatsens initialt formulerade mål redovisas. Kopplingen mellan mål och slutsatser är i flera fall oklar.

Dessa faktorer ger genomgången nedan ett inslag av helhetsbedömningar som till viss del kan försvåra reproducerbarhet av syntesen kring måluppfyllelsefrågan (fråga 1).

Till detta ska läggas att bedömningarna sällan görs i förhållande till de övergripande målen utan fokuserar på vad som uppnåtts *så här långt* i varje aktuell process på främst nivån output och outcome.

Resonemang och slutsatser om måluppfyllelse i utvärderingarna exemplifieras i Tabell 10. Därefter presenteras resultaten från genomgången av de decentraliserade utvärderingarna, följt av de strategiska utvärderingarna.

Tabell 10: Slutsatser om måluppfyllelse i fyra utvärderingar

Fullt uppfyllda mål/resultat	Verksamheten ”har genomfört ett program som kännetecknas av mycket hög effektivitetsnivå. […] Programmet lyckades uppnå intressanta och ibland imponerande resultat. […] De mycket relevanta aktiviteterna gav direkta resultat av hög kvalitet som påvisar förändring på resultat och effektnivå” (Gouzou et al, 2017).
I huvudsak uppfyllda mål/resultat	”Även med många förseningar så slutfördes huvuddelen av planerade insatserna, nästan alla kortsiktiga resultat har levererats och programmet var i huvudsak effektivt. Programmet var dock mest effektivt inom […] som har internaliserats och institutionaliserats. […] Samma sak kan inte sägas om aktiviteterna inom […]” (Moran et al, 2019).
I huvudsak ej uppfyllda mål/resultat	”Utvärderarna har försökt att utvärdera resultaten i förhållande till de fem strategiska målen. […] [Verksamheten] har uppnått mål nummer 2 och delvis mål nummer 4. De specifika målen nummer 1 och 3 har dock inte uppnåtts. […] Avseende mål nummer 5, med fokus på att konsolidera den finansiella och operativa hållbarheten så har förbättringar gjorts […] [men] målet kan inte sägas ha uppnåtts” (Nyberg och Gunnarsson, 2019).
Inte uppfyllda mål/resultat	”Det har gjorts vissa framsteg när det gäller att uppnå kortsiktiga utfall, men mycket mindre så när det gäller att uppnå faktiska resultat. Även om den svåra kontexten delvis hjälper till att förklara de begränsade framgångarna i programmet. […] Det har varit svårt att hitta bevis på direkta bidrag […] till exempelvis fattigdomsminskning och resultat på jämställdhetsområdet” (Watkins och Azarbaijani-Moghaddam, 2019).

Källa: Fyra utvärderingar i urvalet av 80 decentraliserade utvärderingar 2017-2019.

Decentraliserade utvärderingar

Sammanställningen visar att de decentraliserade utvärderingarna i urvalet i regel har rapporterat god måluppfyllelse och effektivitet.

Av de 59 utvärderingar som kunnat skalstegsbestämmas (Tabell 11) så har hela 50 stycken angett att förväntade mål/resultat uppnåtts, fullt eller i huvudsak, dvs. skalsteg 1 eller 2, till och med den tidpunkt då utvärderingen genomförs. Den vanligast förekommande slutsatsen (37 utvärderingar) är att målen, så här långt, är i huvudsak uppfyllda (skalsteg 2). Bara nio insatser är inte eller i huvudsak inte effektiva enligt sammanställningen (skalsteg 3 eller 4).

Tabell 11: I vilken utsträckning har mål/ förväntade resultat uppnåtts så här långt i insatsen, enligt utvärderaren?

Skalsteg	Specifikation	Antal	Andel
1. Fullt uppfyllda mål/resultat	Med undantag för något eller några enskilda, marginella, mål/resultat är bedömningen att samtliga mål/ resultat är uppfyllda eller på väg att nås/ kommer att nås, alt: insatsen beskrivs nästan utan undantag som effektiv.	13	(16%)
2. I huvudsak uppfyllda mål/resultat	Utvärderaren bedömer att majoriteten mål/förväntade resultat (eller de mål som ut-värderaren angett som viktigast) är uppfyllda eller på väg att nås/ kommer att nås, alt insatsen beskrivs huvudsakligen eller främst som effektiv.	37	(46%)

Skalsteg	Specifikation	Antal	Andel
3. I huvudsak ej uppfyllda mål/resultat	Utvärderaren bedömer att majoriteten mål/förväntade resultat, alternativ de mål som utvärderaren angett som viktigast, inte är uppfyllda eller är på väg att nås/ kommer att nås, och/eller: Insatsen beskrivs i huvudsak (men inte bara) som ineffektiv.	8	(10%)
4. Inte uppfyllda mål/resultat	Inget eller endast något enstaka, eller för insatsen marginella, mål/resultat, bedöms ha uppnåtts/vara på väg att nås/kommer att nås, och/eller: Insatsen beskrivs nästan utan undantag som ineffektiv.	1	(1 %)
Går ej att läsa ut i rapporten/ slutsatser för otydliga för skalstegbestämnig	Önskad information kan ej läsas ut i rapporten. Alt. Slutsatserna är mycket otydliga. Alt utvärderingen har inte gjort någon bedömning.	9	(11%)
Fler än ett utvärderingsobjekt	Utvärderar flera olika insatser och bedöms därför inte.	12	(15%)
Totalt		80	(100%)

Källa: Samtliga 80 decentraliserade utvärderingar i urvalet (2017, 2018, 2019).

Tabell 12 visar att utvärderarna rapporterat full måluppfyllelse i en något större andel projekt- och organisationsutvärderingar än programutvärderingar.[30] Observationerna är samtidigt för få för att några slutsatser ska kunna dras om skillnader mellan olika modaliteter avseende måluppfyllelse.

Tabell 12: Bedömning av måluppfyllelse per modalitet

	Uppfyllda?				
Modalitet	Fullt	I huvud-sak	I huvud-sak inte	Inte	Antal
Program	1 (4%)	20 (80%)	4 (16%)	0 (0%)	25 (42%)
Projekt	5 (33%)	10 (67%)	0 (0%)	0 (0%)	15 (25%)
Organisation	6 (50%)	5 (42%)	1 (8%)	0 (0%)	12 (20%)
Strategi/plan	1 (33%)	2 (67%)	0 (0%)	0 (0%)	3 (5%)
Garanti	0 (0%)	0 (0%)	2 (100%)	0 (0%)	2 (3%)
Oklart	0 (0%)	1 (50%)	0 (0%)	1 (50%)	2 (3%)
Totalt					59 (100%)

Källa: 59 decentraliserade utvärderingar i urvalet. Utvärderingar av fler än en insats samt med otydliga slutsatser har sorterats bort.

Jag finner inte heller skillnad i rapporterad måluppfyllelse mellan de regioner där det svenska biståndet är aktivt (Tabell 13). Andelen insatser med uppfyllda mål är dock marginellt högre i Öst- och centraleuropa än i Afrika SoS.

[30] De två garantiutvärderingar som ingår uppfattas ha visat något mer klen måluppfyllelse/effektivitet. Program har som regel också en bredare målbild än projekt.

Tabell 13: Bedömning av måluppfyllelse per region

	Uppfyllda?				
Region	Fullt	I huvud-sak	I huvud-sak inte	Inte	Antal
Afrika SoS	4 (14%)	21 (72%)	4 (14%)	0 (0%)	29 (49 %)
Öst- och Central-europa	3 (30%)	7 (70%)	0 (0%)	0 (0%)	10 (17 %)
Globalt	3 (33%)	4 (44%)	2 (22%)	0 (0%)	9 (15 %)
Öst-/Syd- och Centralasien	1 (17%)	4 (67%)	0 (0%)	1 (17%)	6 (10 %)
MENA	1 (33%)	0 (0%)	2 (67%)	0 (0%)	3 (5 %)
Latinamerika	1 (50%)	1 (50%)	0 (0%)	0 (0%)	2 (3 %)
Totalt					59 (100%)

Källa: 59 decentraliserade utvärderingar i urvalet. Utvärderingar av fler än en insats samt med oklara slutsatser har sorterats bort. Globalt avser här mer än en region.

Sidas utvärderingar görs ibland i form av halvtidsutvärderingar (mid-term reviews) med syftet ”att bedöma framstegen i ett pågående projekt för att underbygga beslut om hur projektimplementeringen kan justeras och förbättras” (Sidas utvärderingshandbok). Om projekt är halvvägs pågående så är det kanske inte mycket som kan sägas om övergripande måluppfyllelse?

I urvalet är 19 utvärderingar (24%) halvtidsutvärderingar.[31] Dessa hade en likartad genomsnittlig budget som övriga utvärderingar varför de inte kan ses som en enklare utvärderingsform.[32]

De insatser som utvärderats har i medianfallet pågått i fyra år. Uppdelat efter typ av utvärdering är medianen för de 19 halvtids-utvärderingarna 3 år och fem år för övriga utvärderingar i urvalet.

[31] Data inhämtad från Sida 2020-11-17.

[32] Genomsnittsbudgeten för de 19 halvtidsutvärderingarna i urvalet var 788 tkr, jämfört med c:a 700 tkr för hela urvalet.

Utvärderarna verkar inte bedöma måluppfyllelsen som högre eller lägre när insatser pågått längre tid (Tabell 14). Graden av måluppfyllelse verkar därför inte självklart kunna kopplas samman med att det rör sig om en halvtidsutvärdering eller att insatsen pågått länge.

Tabell 14: Måluppfyllelse och antal år som utvärderas

	Uppfyllda?					
Antal utv. år	Fullt	I huvud-sak	I huvud-sak inte	Inte	Antal	MTR
1- 3 år	8 (40%)	9 (45%)	3 (15%)	0 (0%)	20	10
4-6 år	4 (13%)	22 (73%)	3 (10%)	1 (3%)	30	3
7-10 år	1 (11%)	6 (67%)	2 (22%)	0 (0%)	9	0
Totalt	13	37	8	1	59	13

Källa: 59 decentraliserade utvärderingar i urvalet. Utvärderingar av fler än en insats samt utvärderingar med oklara slutsatser har sorterats bort. MTR = Mid-term review, dvs. halvtidsutvärderingar. Antal utvärderade år motsvarar här analysramverkets fråga 6.

En förklaring till att måluppfyllelsen inte är högre i insatser som pågått längre skulle bl.a. kunna vara att utvärderingarna av insatser som pågått kortare tid bedömt mer kortsiktiga resultat och att måluppfyllelsen då eventuellt ändå kan ses som över tid ökande. Detta stöds av Tabell 15 som visar att man vid utvärdering av insatser som pågått längre i något högre grad utvärderar impact. Tonvikten även i dessa utvärderingar ligger dock på output och outcome (se diskussionen i kapitel 3 vid Tabell 8).

Tabell 15: Antal utvärderade år och analyssteg

Antal utv. år	Output/Outcome	Output/Outcome/ Impact	Antal
1- 3 år	12 (60%)	8 (40%)	20
4-6 år	8 (27%)	22 (73%)	30
7-10 år	2 (22%)	7 (78%)	9
Totalt	22	37	59 (100%)

Källa: 59 decentraliserade utvärderingar i urvalet. Utvärderingar av fler än en insats samt med oklara slutsatser har sorterats bort.

Förutsättningarna för att dra slutsatser genom en jämförelse mellan mer och mindre effektiva insatstyper baserat på detta material tycks vara *begränsad*. Som framgår i Tabell 11 är det endast nio insatser som bedöms mindre effektiva. Tabell 16 listar dessa insatser.

Tabell 16: Nio mindre effektiva insatser

Geografi	Modalitet	Sektor	Genom-förande organisa-tion	Antal utv. år
Zambia	Garanti	Marknad, företagande, handel, innovation	Privat	5
Uganda	Garanti	Flera områden	Privat	7
Afghanistan	Oklart/ övriga	Jämställdhet	Multilateral /bank, mellan-statlig	4
Globalt	Organisation	Konflikt, fred, säkerhet	Civil-samhälle	5
Kenya	Program	Hållbar samhällsbyggnad, urban utveckling	Gränsöver-skridande samarbete	3

Geografi	Modalitet	Sektor	Genom-förande organisa-tion	Antal utv. år
Globalt	Program	Marknad, företagande, handel, innovation	Privat	5
Regionalt Afrika SoS	Program	Vatten och sanitet	Offentlig	9
Regionalt MENA	Program	Marknad, företagande, handel, innovation	Multilateral /bank, mellan-statlig	3
Regionalt MENA	Projekt	Marknad, företagande, handel, innovation	Multilateral /bank, mellan-statlig	2

Källa: Nio decentraliserade utvärderingar i urvalet där mål/förväntade resultat i huvudsak inte eller inte alls har uppnåtts.

Geografiskt visar Tabell 16 att de mindre effektiva insatserna funnits på flera kontinenter, i stort oberoende av hur länge insatsen hunnit verka, i olika sektorer och bland bilaterala, regionala och globala insatser. Det begränsade antalet observationer gör att eventuella skillnader kan ha uppkommit av slumpen varför inga systematiska skillnader kan spåras. Det tydligaste mönstret i urvalet av 80 decentraliserade utvärderingar av insatser i olika sammanhang och genom olika modaliteter är att insatserna enligt utvärderarna nästan alltid bedöms effektiva och *i huvudsak* uppnår sin mål.

Strategiska utvärderingar

Alla Sidas *strategiska utvärderingar* åren 2015-2019 som bedömt effektivitet (6 st.) ingår också i sammanställningen. Eftersom dessa ofta haft ett brett fokus på ett flertal insatser i olika länder och presenterar förhållandevis sammansatta slutsatser om effektivitet så har vi utgått från grövre skalsteg vid sammanställningen av de strategiska utvärderingarna.

Sammanställningen av strategiska utvärderingar visar på en ungefär lika stark rapporterad måluppfyllelse som i de decentraliserade utvärderingarna.

Tabell 17: Strategiska utvärderingars måluppfyllelse (2015-19)

Utvärdering	Övergripande bedömning	Narrativ slutsats:
Policydialog (2015)	Fullt eller i huvudsak uppfyllda	”politisk dialog har varit ett effektivt verktyg […] svenska ambassader har kunnat använda det i olika sammanhang för att nå Sveriges mål i utvecklings-samarbetet. […] Det finns även tydliga bevis för att det bidragit till ökad jämställdhet.”
Kapacitets-utveckling (2015)	Fullt eller i huvudsak uppfyllda	”det svenska stödet gav ett mycket viktigt bidrag till utvecklingen av de studerade partnerorganisationernas kapacitet. […] Sammantaget bedömdes insatserna ha medel-hög till hög total effektivitet.”

Utvärdering	Övergripande bedömning	Narrativ slutsats:
Garantier (2016)	Fullt eller i huvudsak uppfyllda (?)	"Sidas garantier är användbara instrument som positivt bidrar till utveckling av den privata sektorn. […] När det gäller effektivitet, effekter och hållbarhet visade interventionerna på blandade resultat."
ITP-program (2017)	I huvudsak till i huvdsak ej uppfyllda	"ITP har varit effektiv för kapacitetsutveckling av individer men mindre effektiv för kapacitetsutveckling av organisationer"
Challenge Funds (2018)	Fullt eller i huvudsak uppfyllda	"de avsedda resultaten för majoriteten av fonderna har i stort sett uppnåtts."
Fredsbyggande (2019)	Går ej att läsa ut i rapporten/slutsatser för otydliga för att skalstegbestämning	Slutsatserna ej möjliga att kortfattat summera i termer av huvudsaklig slutsats.

Källa: Urvalet av sex strategiska utvärderingar 2015 till 2019.

Det faktum att de strategiska utvärderingarna analyserar flera insatser samtidigt kan ha gjort det svårare att dra samlade och övergripande slutsatser om måluppfyllelse.

Summering: slutsatser om måluppfyllelse

Utvärderingar av Sidas insatser (decentraliserade och strategiska) rapporterade nästan alltid god måluppfyllelse och effektivitet i utvärderade insatser de studerade åren. Den vanligast förekommande slutsatsen i de decentraliserade utvärderingarna är att mål/förväntade resultat så här långt är i huvudsak uppfyllda. Man bör dock komma ihåg att de mål som bedöms uppnådda oftast är

mål på output- och outcomenivå. Det är mer sällan utvärderingarna bedömer om mål på impactnivå är uppnådda, även om detta verkar efterfrågas i beställningen av utvärderingar (se kapitel 3).

De insatser som beskrivs som mindre effektiva är så få att det blir metodologiskt tveksamt att dra slutsatser om exempelvis mer eller mindre effektiva projekttyper.

5. Utvärderingarnas tillförlitlighet vid analys av måluppfyllelse

Detta kapitel avser besvara studiens andra fråga, om utvärderingarnas tillförlitlighet. Mitt resonemang utgår från fyra kriterier för tillförlitlig bedömning av måluppfyllelse:

1. Tydliga avgränsade mål/förväntade resultat för insatsen (initialt eller rekonstruerade av utvärderaren).
2. En någorlunda avgränsad och väldefinierad intervention, som pågått så lång tid att resultat kan fångas.
3. En testbar teori om hur insatsen ska påverka mål/resultatområde (initialt eller rekonstruerad av utvärderaren), möjliga externt påverkande faktorer i kontexten tagna i beaktande.
4. En genomtänkt utvärderingsdesign för att bedöma om det är insatsen som orsakat eller påverkat observerade resultat.

Utgångspunkten i analysen är att en tillförlitlig bedömning av måluppfyllelse, med primärt kvalitativ utvärderingsdesign, försvåras påtagligt om någon av förutsättningarna ovan inte finns på plats.[33]

Under respektive kriterium utvecklas varför just detta bör vara uppfyllt för att en tillförlitlig måluppfyllelsebedömning ska kunna göras. Utöver dessa kriterier bör utvärderingen vara transparent om hur måluppfyllelsefrågan besvarats.

[33] Vi har utgått ifrån kvalitativ utvärderingsmetod då samtliga utvärderingar i urvalet genomför sina analyser av måluppfyllelse med primärt kvalitativ metod. Inga eller nästan inga regressionsanalyser, ekonometriska metoder, experimentella eller kvasiexperimentella studier förekommer, men i viss mån beskrivande statistik, indikatoruppföljning samt sammanställningar av enkätdata.

Kriterium 1: Tydliga, väldefinierade mål

Mål behöver vara specifika och tydliga i utvärderingar för att underlätta operationaliseringen och för att effektivt kunna kopplas samman med specifik, observerbar, empirisk evidens. Tydliga mål gör det lättare att formulera bra frågor, att hitta en lämplig utvärderingsdesign och användbara mått eller indikatorer, att framställa testbara teorier eller hypoteser samt att ställa rätt frågor under intervjuer eller fokusgrupper. Förutsättningarna för god validitet (att man mäter det man avsett) ökar då betydligt. Tydliga mål kan även ge tydligare avgränsningar i tid och rum och ökad förståelse mellan aktörer om vad som diskuteras (och varför) i utvärderingsprocessen.

Samtidigt är det sedan länge känt att mål i offentliga interventioner och inom biståndet ofta är otydliga (och formulerade på ett för utvärdering ej optimalt sätt) varför de behöver rekonstrueras eller förtydligas som ett tidigt steg i utvärderingsarbetet och av utvärderare, helst i nära dialog med den insats eller organisation som utvärderas.

Utvärderarna i materialet är som framgått ofta kritiska till hur mål/förväntade resultat formulerats avseende sådant som ambition, specificeringsgrad, mätbarhet, realism, konsistens och tidsperspektiv. De otydliga målen tycks dock sällan ha föranlett något förtydligandearbete för att stärka utvärderbarheten i enlighet med de metoder som finns för detta.

Slutsatsen är att denna grundförutsättning (kriterium 1) ofta, men inte alltid, saknats vid utvärderingstillfället.

Kriterium 2: En tydlig avgränsad insats som pågått lämplig tid

Tillförlitliga måluppfyllelsebedömningar är lättare att åstadkomma i enkla, sammanhållna, geografiskt och tidsligt avgränsade insatser. Verksamheten bör också ha pågått under så lång tid att resultat hunnit uppkomma.[34]

I urvalet har 16 utvärderingar (20%) ett globalt fokus (definierat som att minst två kontinenter ingår) och 18 utvärderingar (23%) ett regionalt fokus (definierat som att minst två länder på en kontinent ingår i projektet). Många av de regionala organisationer och program som utvärderas arbetar i ett mycket stort antal länder. Drygt hälften av utvärderingarna (46 stycken) har ett nationellt eller lokalt fokus i de insatser som studeras. Ett intryck är också att många insatser arbetar förhållandevis brett vad gäller antal områden och/eller antal mål på kort och halvlång sikt (jmf kap. 3). I nästan 80% (63) av utvärderingarna efterfrågar beställaren specifika typer av resultat eller nedbrytningar av resultat/måluppfyllelse. I 12 insatser (15%) ska utvärderaren bedöma måluppfyllelsen i fler än en insats när utvärderingen görs.

Slutsatsen är att analysen av måluppfyllelse kompliceras av den förhållandevis stora områdesmässiga och geografiska spridningen inom ramen för insatserna och för varje enskild utvärdering. Här finns ett behov av att fokusera måluppfyllelseanalysen kring viktiga och prioriterade resultat om tillförlitliga bedömningar ska kunna åstadkommas.

---

[34] Det behöver inte ta lång tid för biståndsinsatser och andra offentliga program att generera resultat. Pawson (2013) menar att tidsdimensionen är en av totalt 6 komplexitetsskapande dimensioner för utvärderare. Tidsdimensionen hade inte varit ett exempel på komplexitet om det alltid var just långsiktighet som krävdes för resultatbedömning. Om effektmål kan väntas uppnås efter 1 år eller 7 år beror på insatsen och kontexten.

Är det då för tidigt att bedöma övergripande måluppfyllelse vid den tidpunkt när utvärderingarna genomförs? 56 insatser (70%) hade totalt *fem år eller mer* med finansiering från Sida vid utvärderingstillfället. Det innebär inte att den tidsperiod som utvärderades var så lång. Det genomsnittliga antalet år som utvärderades var 4,5 år.[35] Nio insatser (11%) hade pågått i 7 år eller mer medan åtta insatser (10%) hade pågått i två år (jmf diskussionen vid Tabell 14).

Svaret på frågan om bedömningen efterfrågas för tidigt är att det skiljer sig åt mellan insatser. Insatserna har olika ambitionsnivå. Det finns dock inga starka skäl att låta bli att utvärdera mot den övergripande målnivån i projekt som pågått över fem år och som till exempel ska stärka institutioners arbetssätt, kapaciteten i civilsamhället eller i bedriven forskning (vanliga mål i urvalet). En förhållandevis stor andel av projekten (framför allt de utvärderingar som inte är halvtidsutvärderingar, med en mediantid på fem år) skulle alltså kunnat bedömas mot den övergripande målnivån vid den tidpunkt när utvärderingar genomfördes. Det finns också ett antal projekt där det sannolikt hade varit för tidigt (som halvtidsutvärderingarna med en mediantid på tre år).

Baserat på insatslängd och ambitionsnivå i urvalets insatser är en grov skattning är att en tredjedel av insatserna borde ha kunnat utvärderats utförligt mot den övergripande målnivån vid den tidpunkt när de genomfördes. Det kan ibland vara lämpligt att fokusera på resultat på medellång eller kort sikt och vissa fall kan nivån outcome vara den högsta. Men det finns inga egentliga skäl att aldrig bedöma eller systematiskt avstå från att bedöma biståndsinsatser i förhållande till de initialt formulerade övergripande målsättningar som var utgångspunkt för finansiering. Detta innebär inte heller att man måste fokusera på samtliga övergripande mål för insatsen. Tvärtom kräver ibland systematik i utvärdering att man fokuserar på något centralt mål.

---

[35] De 19 halvtidsutvärderade insatserna hade pågått i genomsnitt 3,5 år (medianvärde 3 år). Sex av dessa (32%) har på något sätt bedömt resultat på nivån impact.

Slutsatsen för kriterium 2 är att bedömningen av måluppfyllelse försvårats av brist på geografisk och områdesmässig fokusering inom insatserna men att en förhållandevis stor andel projekt (minst 30%) borde kunnat utvärderats mot den övergripande målnivån vid den tidpunkt när uvärderingarna gjordes. Större fokusering, avgränsning och anpassning av måluppfyllelseuppdragen avseende bredd samt mål och målnivåer i fokus för analysen förefaller nödvändig. Kriterium 2 har endast delvis varit uppfyllt när utvärderingarna genomfördes.

Kriterium 3: En teori om förändring

Kriterium 3 berör relationen mellan insatsen och det förväntande resultatet och ska lägga en grund för att undersöka om och i så fall hur insatsen kan leda till, leder till eller påverkar det resultat man tänkt sig. Med Pawson och Tilleys (2001) ord, *"interventioner är teorier, och utvärderingar är dessas test"*.

En absolut majoritet etablerade kvalitativa designalternativ[36] för utvärdering av måluppfyllelse, resultat och effekter idag förordar att arbetet konstrueras runt en *förändringsteori.*[37] Syftet är att koncentrera materialinsamlingen kring förändringsantaganden som är möjliga att testa empiriskt, successivt utveckla (ofta i flera steg) och som kan möjliggöra svar på de "hur och varför"-frågor som är viktiga för

[36] Exempelvis: Contribution Analysis (Mayne, 2012), Realist Evaluation (Pawson och Tilley, 1997), Process tracing (Beach och Pedersen, 2013), Systems based evaluation (Hummelbrunner, 2015), QUIP (Copestake et al, 2019). Men även Mixed-methodsutvärdering där den kvantitativa komponenten är experimentiell eller kvasiexpertimentell. En bra kortfattad introduktion till användningen av förändringsteori i utvärderingar finns i Vaessen et al (2020), sid 157.

[37] Jag använder här begreppen förändringsteori, teori om förändring och programteori synonymt.

lärande. [38] En kvalitativ måluppfyllelseanalys utan förändrings-teoretisk grund riskerar att bli ofokuserad, att ha svagare validitet, att få svårt att länka samman insats och resultat och att inte bidra till lärande.[39] Förändringsteorin måste även väga in externt påverkande eller drivande faktorer i kontexten.[40]

Man bör dock skilja på insatsens egen ”officiella” förändringsteori och den teori utvärderaren arbetar med, utvecklar och testar. Utvärderingsforskaren Carol Weiss beskrev redan 1972, när hon introducerade diskussion om offentliga programs underliggande antaganden, att dessa ofta baseras på otydliga och oprecisa idéer om hur verkligheten ska förändras. Det är en slutsats som sedan verifierats många gånger. Grundantagandet är begränsad rationalitet (Lindblom, 1959, Stame, 2004, Weiss 2007). Offentliga interventioner formuleras inte i alltigenom rationella processer av utvärderare och styrningsexperter utan speglar de politiska och sammansatta omgivningar där de arbetats fram. Det är alltså inte nödvändigt att en förändringsteori för insatsen finns på plats när utvärderaren träder in utan utvärderaren bör, precis som med målen, ha beredskap att rekonstruera eller förtydliga den underliggande

---

[38] Utvärderaren frågar sig exempelvis om teorin är rimlig, ligger i linje med övrig kunskap på området, om insatsen har genomförts i linje med teorin, kan verifieras empiriskt och vilken roll externa påverkande faktorer spelar. Se rad 2-12 i Tabell 19.

[39] Förändringsteorier är viktiga även då de möjliggör mer gemensam förståelse mellan utvärderare och insatsens personal avseende hur eller på vilket sätt insatsen är tänkt att leda till förändring. Det är viktigt för att underlätta interaktion, lägga en grund för kommunikation om slutsatser och minska spänningar i arbetet (Vaessen, 2020, sid 157).

[40] Förekomsten av en förändringsteori kan inte tas som intäkt för att utvärderingen hanterar kausalitet. Ett återkommande problem är att ”förändringsteorier” helt eller delvis saknar realistiska och genomtänkta kopplingar till kontext och externt påverkande faktorer. Det mest centrala problemet i förändringsteoretiska utvärderingar är sannolikt att de i praktiken ofta helt misslyckas med att presentera en teori om förändring, bortom några pilar och boxar (Ibid).

teorin, vid behov i flera steg, för att därigenom möjliggöra en valid bedömning av måluppfyllelse eller resultat (se Leeuw, 2003, Weiss, 2007, Vedung, 2012).[41]

Det är därför inte heller en hjälpsam slutsats i en utvärdering att förändringsteorin för insatsen är otydlig om den inte efterföljs av försök att förtydliga (rekonstruera) och empiriskt pröva förändringsteorin.[42]

Tabell 18 visar att det i 16 av urvalets utvärderingar sannolikt fanns någon typ av förändringsteori på plats som också prövats empiriskt av utvärderaren. Att dessa varit tämligen oprecisa kan vi anta baserat på bland annat Christopolos et al (2014) och tidigare litteratur presenterad ovan.

Slutsatsen för kriterium 3 är att det ofta funnits något form av oprecis, ibland rekonstruerad förändringsteori, som diskuterats och bedömts av utvärderarna men som sällan prövas ordentligt mot empiriska data. Förändringsteorin integreras därför inte i utvärderarens analys vilket troligen gjort analysen mindre fokuserad, sänkt validiteten och förvårat sammanlänkning av insats och resultat.

Tabell 18: Har insatsen en förändringsteori som prövas mot empiriska data i måluppfyllelseanalysen?

Ja (sannolikt)	16 (27%)
Nej (sannolikt inte)	42 (71%)
Oklart/kan ej läsas ut	1 (2%)
Totalt	59

Källa: 59 decentraliserade utvärderingar i urvalet. Utvärderingar av fler än en insats och med otydliga slutsatser ingår inte. Sannolikhet anges då frågan varit svår att läsa ut ur rapporterna.

[41] Det finns dock en gräns där det inte längre går att rekonsturera en förändringsteori som är allt för ogenomtänkt, absurd eller orimlig. Men sådana är troligen ganska ovanliga.

[42] Trots det dras slutsatsen ofta i svenska biståndsutvärderingar. Det som bör vara utgångspunkt i arbetets inledningsfas presenteras då som en konklusion.

Kriterium 4: En ändamålsenlig utvärderingsdesign

Utvärderaren behöver på en överordnad nivå också en *utvärderingsdesign* för sin måluppfyllelseanalys. En sådan för samman utvärderingsfråga, förändringsteori, data och dataanvändning till en helhet enligt en övergripande logik.[43] Utvärderaren försöker besvara måluppfyllelsefrågan genom att testa en förändringsteoris bärkraft mot data (dokument, intervjuutskrifter, enkätsammanställningar el dyl) enligt designalternativets preciserade utgångspunkter och processer med erfoderlig situations- och kontextanpassning.

För att måluppfyllelsefrågan om insatsens bidrag ska kunna besvaras på ett tillförlitligt sätt krävs en design som inte bara tar hänsyn till insatsen utan också övriga potentiellt påverkande faktorer. Biståndsinsatser fungerar som regel tillsammans med andra påverkande eller stödjande faktorer i ett sammanhang.[44]

När jag undersökt vilka designalternativ utvärderarna hänvisar till för bedömningen av måluppfyllelse framgår att hälften av utvärderingarna (41 stycken, 51%) saknar hänvisning till något av de etablerade designalternativ som finns tillhanda för analys av måluppfyllelse/förväntade resultat (Tabell 19).

---

[43] En utvärderingsdesign skiljer sig i grundkomponenterna inte från en forskningsdesign i samhällsvetenskaplig forskning som enligt King, Keohane och Verba (1994) består av forskningsfråga, teori, data och datanvändning.

[44] Insatsen är sällan det enda som påverkar eller driver ett resultat. "De flesta utvecklingsinsatser är 'bidragande'. De fungerar som en del av kausala paket i kombination med andra 'stödjande faktorer' såsom intressenters beteende, relaterade eller närliggande program och policyer, institutionell kapacitet, kulturella faktorer och socioekonomiska trender. Design och metoder för utvärdering måste därför kunna plocka isär sådana orsakspaket" (Stern et al 2012, sid 7).

Tabell 19: Vilken design hänvisar utvärderaren till i måluppfyllelseanalysen?

1.	Övriga/Oklart/inget	41 (51%)
2.	Contribution analysis	14 (18%)
3.	Kombination av två alternativ (2-8 i tabellen)	11 (14%)
4.	Theory based evaluation (generell)	7 (9%)
5.	Outcome harvesting	4 (5%)
6.	Outcome mapping	2 (3%)
7.	Most significant change	1 (1%)
8.	Process tracing	0 (0%)
9.	Realist evaluation	0 (0%)
10.	Systems based Evaluation	0 (0%)
11.	QCA (Qualitative Comparitative Analysis)	0 (0%)
12.	QUIP (The Qualitative Impact Assessment Protocol)	0 (0%)
13.	Regressionsbaserad analys (någon form)	0 (0%)
14.	Kvasiexperiment (någon form)	0 (0%)
15.	Experimentell design	0 (0%)
Totalt		80

Källa: Samtliga 80 decentraliserade utvärderingar i urvalet. Exempel på översikter avseende designalternativ, metoder och ansater för att utvärdera bistånd med kvalitativ och kvantitativ metod är Stern et al (2012) och Vaessen et al (2020).

Flertalet av de utvärderare som hänvisar till ett designalternativ gör en delvis oortodox tolkning av det man hänvisar till. John Maynes *Contribution Analysis* (Mayne, 2001) är det vanligast förekommande designalternativet.[45]

Av de 14 utvärderingarna som uppgett Contribution Analys så är det endast fyra som utryckligt refererar till John Maynes artiklar, sekundärlitteratur eller någon precisering om hur metoden specifikt används och endast tre nämner Maynes grundläggande sexstegs-

[45] Att alternativ 13-15 inte refereras är väntat då alla 80 utvärderingarna är primärt kvalitativa. Utvärderarna uppger ofta att de använder mixed methods. I praktiken används nästan aldrig statistiska eller ekonometriska analysmetoder en kvantitativa beskrivande data eller enkätsammanställningar förekommer förhållandevis ofta. Alternativ 11 (QCA) används jämförande varför det endast är relevant när flera insatser utvärderas samtidigt.

process.[46] Två utvärderingar (14%) ger en kortare motivering varför man inte följer processen för den design man hänvisat till. Preciserar utvärderarna då på annat sätt hur de besvarat måluppfyllelsefrågan? Tabell 20 presenterar data för de 41 utvärderingar som *inte* hänvisat till något designalternativ.

Tabell 20: Beskriver utvärderarna specifikt och utförligt hur måluppfyllelsefrågan besvarats?

Ja	8 (20%)
Nej	33 (77%)
Totalt:	41

Källa: Samtliga decentraliserade utvärderingar i urvalet som inte hänvisar till någon design i Tabell 19.

20% av utvärderingarna har beskrivit specifikt och utförligt hur måluppfyllelsefrågan besvarats. Samtliga har i enlighet med Sidas anvisningar ett metodkapitel och ofta även en Inception report som ska precisera konsultens metod och genomförande. Ganska många utvärderare har bifogat en utvärderingsmatris.[47] I flera fall framgår det att det finns en sådan men att den inte bifogats rapporten vid publicering. Det som beskrivs där är ofta kortfattad mellan utvärderingsfrågor likartad information om hur data samlats med skralt motiverade källor och indikatorer kopplat till varje fråga.

Den övergripande logiken för materialsamlingen är ofta oklar. Beskrivningar eller förklaringar av varför specifika metoder och källor används, hur data bidrar till att besvara enskilda frågor eller till att pröva insatsens förändringsteori och hur externt påverkande faktorer hanteras är ofullständiga. Ofta beskrivs hur data samlats in

[46] Contribution Analysis är förhållandevis specifikt definerat i termer av sex grundläggande analyssteg (se exempelvis https://www.betterevaluation.org/en/plan/approach/contribution_analysis)

[47] Dessa är ett verktyg för att planera och organisera en utvärdering. I praktiken ofta en tabell med en rad för varje utvärderingsfråga och kolumner som behandlar metod för datainsamling, källa, valda indikatorer etc. Tanken är att matrisen ska beskriva varje utvärderingsfrågas operationalisering.

men inte utvärderingsdesignen. Materialinsamlingen verkar ofta anpassad till utvärderingsuppdraget som helhet snarare än till de enskilda frågorna.

70% (56) av utvärderingarna tar upp eller nämner begränsningar i metoden som direkt berör analysen av måluppfyllelse.[48] Utöver kausalitetsfrågan (se nedan) är de vanligaste diskuterade begränsningarna brister i insatsernas uppföljningsdata, rapportering eller dokumentation (30 utvärderingar), resultatramverk eller logiska ramverk (17) eller att utvärderarna anser att utvärderingen görs för tidigt för att måluppfyllelseanalys ska vara möjlig (9). Det som lyfts fram är alltså externa faktorer man inte kunnat påverka själv.

Slutsatser om brister i uppföljning, rapportering och dokumentation som ett hinder i utvärderingsarbetet tycks ha förekommit länge inom svensk biståndsutvärdering (se Peck, 2007, Forss et al 2008, Christoplos et al 2014). Man kan därför tänka sig att erfarna Sida-utvärderare ser dessa som återkommande ingångsvärden som de egna designförslagen bygger på att lösa eller hantera snarare än som oväntade hinder. Det kan vara ett problem om insatserna, som utvärderarna tycks mena, saknar effektivt uppföljningsarbete eller baslinjemätningar, men en utvärderare bör kunna anpassa sin design utifrån förutsättningarna i biståndets verklighet och efter lärande från tidigare uppdrag. Utvärdering av måluppfyllelse behöver inte göra så beroende av insatsens uppföljning och rapportering.

Hur hanterar utvärderarna då den svåra men viktiga kausalitetsfrågan (jmf Vedung, 2012)? En begränsad men delvis belysande bild kan ges genom en kartläggning av begreppsanvändningen i rapporterna. Tre viktiga begrepp i utvärderingsvokabulären vid bedömningar av resultat på nivån outcome och impact är kausalitet, kontrafakticitet och additionalitet (causality, counterfactual och additionality, se

[48] Utvärderarna tar upp flera andra begränsningar, med mer oklar eller outtalad koppling till effektivitetsanalysen.

Sandahl och Petersson, 2016, eller Sandberg och Faugert, 2012).[49] Begreppen kan användas för verifiera att man arbetar på ett visst sätt, för att kritisera eller diskutera alternativa metodologiska synsätt eller beskriva begränsningar i den egna rapporten. I urvalet förekommer dessa begrepp mycket sparsamt. I 70 utvärderingar saknas kontrafakticitet och additionalitet helt och i 55 rapporter förekommer inte begreppen kausal eller kausalitet. I 26 av de 49 utvärderingar som menar att de analyserat långsiktiga effekter förekommer aldrig begreppen kasual eller kausalitet. I några få fall talar man om exogen påverkan eller exogena faktorer.[50]

Endast en tredjedel (25) av utvärderingarna nämner i begränsningsavsnittet någon form av metodbegränsningar som hänger samman med kausalitet och analysen av måluppfyllelse.[51] Bland de rapporter som nämner kausalitetsfrågan i begränsningskapitlet så är det vanligaste konstaterandet *att* kausaliteten varit svår att hantera (i några fall hävdar man att det är omöjligt) utan att sedan tydliggöra varför det varit svårt eller hur man sökt hantera saken och utan att diskutera hur detta påverkar slutsatsernas tillförlitlighet. Ingen decentraliserad utvärdering diskuterar utförligt om det är möjligt att bedöma insatsens måluppfyllelse på nivån outcome och impact utan att hantera kausalitet och externt påverkande faktorer.

---

[49] Även vid kvalitativ utvärdering kan kontrafakticitetsbegreppet ha viss relevans genom att öppna för en diskussion kring alternativtscenariot utan finansiering, en hypotestisk övning som ibland är fruktbar.

[50] Två begrepp som förekommer något oftare är contribution och attribution. Begreppen har endast kvantifierats för 2017 då särskilt begreppet contribution visat sig inte spegla kausalitetsdiskussion i rapporterna. Begreppet förekommer ofta i sammanhang som inte hör samman med frågan om kausalitet eller orsakssamband. Begreppet attribution förekom 2017 i 20 av 25 rapporterna, inte sällan då i form av ett kort nämnande eller en friskrivning i stil med ”vi fokuserar på contribution, inte attribution”.

[51] Nästan alla rapporter har en underrubrik i metodkapitlet där utvärderaren väntas ta upp metodologiska begränsningar.

Av de 51 rapporter som helt saknade diskussioner om metodbegränsningar rörande kausalitet så bedömde 27 stycken Impact. Dessa lyfte därmed inga metodologiska begränsningar avseende kausalitet trots att man delvis fokuserat på långsiktiga resultat som med stor sannolikhet påverkas av fler faktorer än den insats som utvärderats.

Den samlade genomgången visar att 54 av 80 utvärderingar inte beskriver hur de bedömt eller försökt hantera kausalitet och externa potentiellt påverkande faktorer vid analysen av måluppfyllelse.

Slutsatsen är att utvärderarna i regel inte lyckats övertyga läsaren om att de utvärderat *insatsens* måluppfyllelse eller bidrag till måluppfyllelsen. I urvalet har få utvärderingar en utvärderingsdesign som effektivt kan operationalisera måluppfyllelsefrågan, testa teorin om förändring och hantera kausalitet, i linje med kriterium 4.

Summering: utvärderingarnas tillförlitlighet

Slutsatsen från genomgången av utvärderingarna utifrån kriterierna är att bedömningarna av måluppfyllelse förmodligen inte är tillförlitliga.

Insatserna har ofta inte haft tydliga och avgränsade mål, och utvärderingarna tycks sällan ha rekonstruerat sådana. Måluppfyllelseanalysen försvåras av en brist på geografisk och områdesmässig fokusering i utvärderingsuppdrag och insatser. En majoritet utvärderingar hänvisar inte till någon etablerad utvärderingsdesign och i övriga fall beskrivs designalval och metod bristfälligt. Intrycket är att metodologin är anpassad till uppdraget som helhet snarare än enskilda utvärderingsfrågor. Förändringsteorin integreras sällan i analysen vilket minskar validiteten och försvårar sammanlänkning av insats och resultat.

Få utvärderingar visar hur kausalitet och externt påverkande faktorer hanteras. Utvärderarna lyckas därmed inte göra sannolikt att det är den utvärderade insatsen som gjort skillnad för, påverkat eller åstadkommit resultatet. Det finns därför en risk att rapporterna beskriver förändringar i de sammanhang där insatserna verkat men som kan vara helt eller delvis oberoende av den utvärderade insatsen.

6. Slutsatser i tidigare studier

Vi har kunnat konstatera att utvärderingarna i urvalet i huvudsak rapporterat om god måluppfyllelse, men att slutsatserna (vad gäller Sidas decentraliserade utvärderingar) enligt de fyra kriterier för en tillförlitlig utvärdering som behandlades i förra kapitlet, förmodligen inte är tillförlitliga. Hur pass förvånande är det? Motsäger det tidigare studier eller är det, mot bakgrund av tidigare genomgångar av liknande frågor, vad som kunde förväntas?

Nedan presenteras en genomgång av ett femtontal sammanställningar, metautvärderingar och analyser av svenska biståndsutvärderingar från den senaste 25 åren. Sammanställningen fokuserar på sk. grå litteratur och gör inte anspråk på att vara uttömmande. Den baseras på en enkel men extensiv webb- och databasökning för att hitta så många objekt som möjligt. Samtliga funna sammanställningar har inkluderats. De databaser som använts är Openaid, Sidas publikationsdatabas och Föreningen för utvecklingsfrågors (FuF) webbarkiv. Genom funna publikationer har nya kunnat lokaliseras utifrån dessas referenser till tidigare litteratur.

Tidigare slutsatser om måluppfyllelse

Utvärderingen Promoting Development by Proxy (Riddell et al, 1995) analyserade 37 projekt inom Sidas civilsamhällesstöd. En övergripande slutsats var att:

> "den överväldigande majoriteten har antingen uppnått, eller är på god väg att uppnå de uttalade och mellanliggande mål till vilka Sidas NGO-avdelning tillhandahöll medel". Författarna fortsätter dock: "uppnådda projektmål ger förhållandevis lite vägledning med avseende på den totala utvecklingseffekten av projekten ur ett bredare perspektiv. När framstegen bedömdes mot ett större antal om nio kriterier, så sjönk den sammanlagda prestationsskattningen markant".

Nilén, Nilsson och Nybergs (2002) genomgång av 171 svenska CSO-utvärderingar under åren 1999-2001 fann att: *"[ä]ven om dessa mål ofta endast är allmänt formulerade kommer utvärderarna i allmänhet fram till att projekten nått sin målsättning"*.

En metautvärdering av nio utvärderingar av Sidafinansierade forskarnätverk i Afrika drog slutsatsen att:

> "Det dessa nio utvärderingar har berättat för oss är att projekten är relevanta och att de har uppnått omedelbara mål. Men de ger ingen väsentlig information om kostnadseffektivitet, långsiktiga resultat eller hållbarhet ... Vår slutsats är därför att Sidas mål att stödja regionala nätverk för att förbättra nationell forskningskapacitet inte har behandlats och analyserats tillräckligt i utvärderingarna" (Carlsson och Wohlgemuth, 1996).

Tre utvärderingssammanställningar genomförda av Sida UTV (Peck, Jassey och Johansson, 2005; Johansson och Lindgren, 2006; Peck och Knippel, 2007) finner en hög måluppfyllelse i majoriteten av utvärderingarna.[52]

På landnivå gjorde Pain (2015) en genomgång av sju utvärderingar med fokus på biståndet till Sveriges största biståndsland Afghanistan. De insatser som utvärderades bedömdes fånga en stor andel av Sveriges portfölj till landet vid denna tidpunkt. Studien fokuserade inte uttryckligt på måluppfyllelse men drog slutsatsen att:

---

[52] "In about half of the evaluations the level of effectiveness of the interventions was said to have been high." (34 utvärderingar, 2005), "For a majority of the evaluations, the level of effectiveness of the interventions…was said to have been high. In one-third of the cases the results were more mixed." (58 utvärderingar 2006) , "Regarding the achievement of short- and medium-term objectives at the output and outcome level, effectiveness was found to be high for a majority of the interventions. However, the achievements of overarching objectives were more uncertain."(48 utvärderingar, 2007).

> ”interventioner med ett instrumentellt fokus (dvs. direkt påverkande förändring) har varit mer effektiva för att uppnå resultat. De som istället försöker skapa förståelse, påverka debatter, bygga kapacitet och förändra beteenden har uppnått mindre. […] Medan ett antal interventioner har förbättrat tillgången till offentliga tjänster såsom utbildning, hälsa och förbättrad fysisk säkerhet, har det varit få effekter på livsmedelssäkerhet och inkomst, detta begränsar de totala effekterna på fattigdomsminskning”.

Tidigare slutsatser om fokus på övergripande målsättningar

En vanlig slutsats i sammanställningar av svenska bistånds-utvärderingar är att studierna har ett svagt fokus på övergripande och centrala mål i biståndet. Bedömarna verkar ha saknat en ordentlig diskussion i utvärderingar om insatsernas bredare bidrag till övergipande mål för det svenska biståndet.[53]

I en genomgång av 80 Sida-utvärderingar konstaterades att 30, på något sätt, försökte vissa på effekter på fattigdom, utvärderarna *”uppmanades aldrig att ta reda på om det främsta målet i det svenska biståndet uppfylldes”* (Eduards, 1995, citerad i Carlsson, 1998).

Ett annat exempel är Tobisson och De Vylder (1997) som menade att: *”lite uppmärksamhet ägnas åt fattigdom och genusfrågor i dessa analysdokument och utvärderingar. Detta är överraskande med tanke på den långa traditionen av tonvikt på fattigdom i det svenska utvecklingssamarbetet”.*

[53] Med det avses både kopplingen till nationellt framtagen strategi och regeringens och riksdagens mål för biståndet och fokus på den övergripande målbilden i själva insatserna.

En studie om Sidas stöd till mänskliga rättigheter och demokrati, baserad på interna system, dokument och utvärderingar, pekade på:

> ”Otillräcklig uppmärksamhet på övergripande mål och hur program ska tänkas nå sina effekter. […] De flesta projekt verkar vara väl utvecklade och övervägda med avseende på deras direkta genomförande, men mindre så när det gäller hur de ska påverka demokrati och gott styre” (Uggla och Wulfing, 2007).

I en sammanställning av 84 Sida-utvärderingar drogs slutsatsen att:

> “Mindre än hälften av utvärderingsrapporterna tittade på insatsernas bidrag till fattigdomsminskning, och de flesta av dessa analyserade endast fattigdomsminskning i begränsad omfattning” samt att “[e]n majoritet av interventionerna verkar sakna en övergripande logik i utformningen och implementeringen. Detta berodde ofta på ett fokus på aktiviteter. Många insatser (och utvärderingar) spekulerar endast om insatsens bidrag till resultatet” (Christoplos, Liljelund Hedqvist och Rothman, 2014).

Återkommande är även slutsatsen att Sidas insatser styrs av otydliga eller oprecisa målformuleringar (se Bandstein, 2006, samt Peck, 2007). Bandstein (2006) menade samtidigt att bedömningar av måluppfyllelse är något av det som Sidas personal efterfrågar mest av utvärderingsverksamheten.

Tidigare slutsatser om tillförlitlighet

Forss (1994) kunde utifrån 177 utvärderingar ”förkasta hypotesen” att (majoriteten) svenska biståndsutvärderingar gör tillfredställande bedömningar av måluppfyllelse.

I en annan tidig studie beställd av Sida om de egna utvärderingarnas kvalitet konstaterades att: ”*en systematisk analys av utvärderingskvalitet gav en annorlunda bild än om man direkt frågade programansvariga om utvärderingars tillförlitlighet. Programansvariga bedömde att Sidas utvärderingar var tillförlitliga medan den systematiska analysen visade att rapporterna var metodologiskt väldigt svaga”* (Carlsson et al. 1997).

Samma år analyserade Carlsson och Forss (1997) hela 277 biståndsutvärderingar från Sida, UD och andra svenska biståndsmyndigheter och konstaterade att resulaten var ”*nedslående. Utvärderingssystemet har genom åren inte levererat valida och reliabla slutsatser om utvecklingssamarbetets resultat*”.

En genomgång av 34 av Sidas utvärderingrapporter tio år senare konstaterade att ”*[f]lertalet av utvärderingarna täcker måluppfyllelse på ett adekvat sätt (62%), även om detta ofta sker i meningen måluppfyllelse på nivån output eller nära outcome*”. Författarna menade dock att utvärderingarna inte ger *”frågan om attriburering tillräckling uppmärksamhet, dvs. de visar inte på någon empirisk evidens för att det faktiskt är biståndinsatsen som haft en inverkan”*. Den övergripande bedömningen av utvärderingarna var att *“utvärderingarnas kvalitetssäkring bör förbättras hos Sida. Det finns ett behov av mer och bättre empirisk bevisföring och systematisk användning av sådan information i en majoritet av de granskade rapporterna. Det är särskilt oroande att så få av utvärderingarna innehöll tillräckligt med information om de använda metoderna. Detta gjorde det svårt att bedöma om slutsatserna var tillförlitliga och tydligt härledda från dataunderlaget”* (Forss et al, 2008).

Vid en uppföljande analys av 114 decentraliserade utvärderingars kvalitet fem år senare konstaterades *“betydande brister inom viktiga områden för utvärdering, särskilt metodik, insamling och analys av data samt evidensbaserade slutsatser och rekommendationer”* (La Guardia, 2013).

I en syntes av 84 decentraliserade Sida-utvärderingar konstaterade Christopolos et al (2014) att *"vissa program är för ambitiösa avseende geografiskt spridning, komplexitet och antal förväntade resultat för att konstituera en realistisk eller verifierbar förändringsteori"*. Angående insatsernas befintliga förändringsteorier har en stor del av insatserna *"en svag koppling mellan aktiviteter, utfall, resultat och effekt (resultatkedjan) och saknar en övergripande logik i design och implementering. […] Man misslyckas ofta med att analysera antagandena som underligger program och projekt. […] Det är ofta oklart vem insatsen är riktad till i form av direkta och indirekta stödmottagare och partner och hur dessa förväntas dra nytta av interventionen"*. Utvärderingarna verkar dock, ibland, "användas som en möjlighet att konstruera, rekonstruera eller på annat sätt uppmärksamma explicita eller implicita förändringsteorier, också när det aktuella programmet inte tidigare reflekterat över dessa".[54]

Inga systematiska bedömningar av kvaliteten i Sidas utvärderingar har gjorts efter 2014. Ett undantag är Burman (2017) som studerade bedömingar av hållbarhetskriteriet och konstaterade att tillförlitligheten vid bedöming av insatsernas hållbarhet var låg.

Summering: slutsatser i tidigare studier

Sammanfattningsvis kan det konstateras att svenska biståndsutvärderingar beställda av Sida oftast över tid dragit slutsatsen att måluppfyllelsen på kort eller medellång sikt varit god. Flera tidigare analyser visar att utvärderingarna inte fokuserar på övergripande långsiktiga målsättningar för insatserna.

Det faktum att insatserna också ofta anses ha otydliga målformuleringar kan ha försvårat möjligheten att utvärdera måluppfyllelse genom en tydlig operationalisering. Externa

[54] Jag tolkar det som att rekonstruerade förändringsteorier var förhållandevis ovanliga när den studien gjordes. En delförklaring kan vara att Sida under lång tid arbetat med Logical Framework Approach (LFA) som ett slags styrinstrument.

kvalitetsbedömningar identifierar ofta och återkommande tydliga metodologiska brister, bland annat rörande frågor om kausalitet.

Slutsatserna ovan är inte förbehållet svenskt, eller Sida-finansierat, bistånd. Angående enskilda biståndsinsatsers måluppfyllelse konstaterade Riddell (2007) att:

> ”Nu tillgänglig evidens tyder på att biståndsprojekt uppnår sina omedelbara mål. Bevisen är emellertid långt ifrån entydiga och troligen också något vridna till förmån för en mer smickrande bild av projektbiståndets resultat än de förtjänar. Dessutom har ett betydande antal insatser problem med att upprätthålla de kortfristiga resultaten under en längre tidsperiod”.

Det finns enligt Riddell (2007, sid 186) belägg för att givarnas egna interna bedömningar av projektens måluppfyllelse överskattar andelen framgångsrika insatser.[55]

Colin, Sandström och Wennberg (2021) ger oss en jämförelse med ett annat politikområde när de studerat 110 utvärderingar av svenska näringspolitiska insatser 2009-2019 och finner att privata konsulter är ”*övervägande mer positiva till de utvärderande insatserna*.”[56] Att ”*de studier som granskats inte är metodologiskt tillförlitliga bör sättas i relation till att utvärderarna likväl i så betydande utsträckning verkar landa i positiva slutsatser*”. ”*Det som påverkat resultaten är just att det är konsulter som genomför utvärderingarna*”, hävdade författarna.

I en studie om decentraliserade utvärderingar i norskt bistånd drogs slutsatsen att: ”*beslut om norska biståndsprojekt fattas baserat på utvärderingsresultat och rekommendationer som inte grundats i god bevisföring*” (Chapman et al, 2017). Utvärderingarna användes av beslutfattare,

---

[55] Givarinterna bedömingar baserade på intern uppföljning bör hållas isär från sammanställningar av utvärderingar, som denna. Det framgår inte tydligt i Riddells sammanställning hur han viktat givares rapportering relativt syntesutvärderingar.
[56] Jämfört med utvärderingar genomförda av analysmyndigheten Tillväxtanalys.

menade författarna, men de var inte tillförlitliga. Tidigare analyser av olika givare och större internationella organisationer över tid har också visat på generella och utbredda kvalitetsbrister i utvärderingar på biståndsområdet (Forss och Pollard, 2020).

Sammanfattningsvis kan det konstateras att slutsatserna i denna rapport ligger tydligt i linje med tidigare studier.

7. Förklaringar till svag måluppfyllelse

Kan vi då lära oss något av de nio insatser där utvärderarna rapporterat lägre effektivitet? Lär utvärderingarna oss något om hur projekten bör genomföras som biståndets aktörer, till exempel en handläggare på en ambassad, kan ha nytta av vid fortsatt beredning eller i andra framtida insatser? Kan utvärderingarna förklara eller begripliggöra varför insatserna haft svag effektivitet?

Fokus här ligger på olika *hindrande faktorer* utvärderarna lyft fram. Dessa summeras i stiliserad form i Tabell 21.

Tabell 21: Hinder vid svag måluppfyllelse i – nio insatser

Geografi och sektor	Hindrade faktorer
Zambia, Marknad, företagande, handel, innovation	Otydlig styrning, givare ej överens om vad insatsen skulle uppnå, svårigheter att ändra beteende i genomförandeled, brist på resurser och kunskap. Viss ekonomisk riskaversion (på målgruppssidan) och låg konkurrens. Svag uppföljning.
Afghanistan, Jämställdhet	Utmananade kontext, svag ledning i insats, för ambitiösa mål samt bristande fokus, resursbrist, svagt ägarskap, för lite personal samt hög personalomsättning, bristfällig riskanalys, uppföljning och kommunikation.
Globalt, Konflikt, fred, säkerhet	För ambitiösa mål, få aktiviteter, brist på finansiering, kapacitet och fokus. Svag modell för uppföljning av resultat.
Kenya, Hållbar samhällsbyggnad, urban utveckling	För ambitiösa mål på flera delområden, förseningar, finansieringsproblem, inaktivitet på partnersidan, svaga kacacitetsinsatser, förändringar i fokus och inriktning.

Geografi och sektor	Hindrade faktorer
Globalt, Marknad, företagande, handel, innovation	Delvis oklar inriktning och bristande fokus i program, olikartade intressen och ägarskap hos målgruppen, svag koppling till parallella initiativ och institutioner.
Uganda, Flera områden i samma	Bristande analys av grundproblem, sektor och efterfrågan. Bitvis felaktig ekonomisk analys. Avsaknad av kompletterande tekniskt stöd samt ägarskap i genomförandeledet.
Regionalt Afrika, Vatten och sanitet	Bristande fokusering, svaga analyser och förstudier i flera delinsatser, svag koordinering, fel aktörer involverade i delprocess, enskilda insatser ej finansiellt hållbara, begränsad motivation hos personal i någon insats, generella problem kring finansiell och ibland materiell hållbarhet.
Regionalt MENA, Marknad, företagande, handel, innovation	Bristfällig grundläggande problemanalys, fel aktörsmedverkan, svag organisatorisk kapacitet, bristfälligt kommunikationsarbete, fel aktiviteter, svagt ägarskap och bristande inre effektivitet.
Regionalt MENA, Marknad, företagande, handel, innovation	Låg aktivitetsgrad och förseningar, bristande projektdesign i form av svag problemanalys vilket gav fel aktiviteter, bristande fokusering, för lite och fel personal, svag uppföljning.

Källa: De nio decentraliserade utvärderingar i urvalet där mål/förväntade resultat i huvudsak/inte alls uppnåtts. Tabellen beskriver hindrande faktorer som på olika sätt tas upp i utvärderingarna. Den listar hinder utan att värdera dessas betydelse i förhållande till varandra. Hindren avser inte alltid insatsen som helhet utan ibland enskilda insatser inom projektet.

När faktorerna ovan kategoriseras och grupperas finner vi ett antal återkommande tänkbara delförklaringar till bristande effektivitet i insatserna:

1. Bristande fokus, styrning el ledarskap (6 insatser)
2. Organisatorisk svaghet el brist på kapacitet (6 insatser)
3. Svagt ägarskap eller efterfrågan (5 insatser)
4. Svag problem- eller grundanalys i insats (4 insatser)
5. Svag inre effektivitet, låg aktivitetsgrad, förseningar (4 insatser)
6. Svag uppföljning eller monitorering 4 insatser)
7. Svag koordinering, samverkan, fel aktörsmedverkan (3 insatser)
8. Svag kommunikation (2 insatser)
9. Fel aktiviteter (2 insatser)

Utvärderarna lägger alltså vikt vid ett antal olika i huvudsak *insatsnära faktorer* i kombination kopplade till främst projektens grundanalys, styrning, organisering, genomförande och uppföljning. Tidsligt återfinns hindren i nästan hela den så kallade *policycykeln* (Vedung, 2016) från problemanalys, utformning och beslut om intervention, intervention och implementering, eftergranskning, uppföljning, utvärdering, spridning av eftergranskning etc.

Avseende förklarande utvärdering brukar man ibland skilja på *teorifel* och *implementeringsfel* (Sandberg och Faugert, 2012). De nio insatserna tycks då kombinera teorifel i form av svaga problem- eller grundanalyser samt svagt ägarskap eller efterfrågan med implementeringsfel som brister i styrning, fokus och ledarskap, organisatorisk svaghet/brist på kapacitet och låg aktivitetssgrad alt. svag inre effektivitetet. Insatserna tycks därmed, i några fall, ha varit fel planerade och sedan inte kunnat genomföras enligt plan på grund av olika kombinationer av processnära hinder. Tyngdpunkten i angivna hindrade faktorer ligger i implementeringsfasen.

Lär då dessa utvärderingar målgruppen, exempelvis en handläggare på ambassad, någonting användbart om vad som påverkat effektivitet som därmed kan förklara bristande framgång?

Ja delvis, utvärderingarna bidrar med flera tänkbara *delförklaringar* till den svaga måluppfyllelsen men väldigt lite fokus ligger på externa förklaringar eller mekanismer *i och i förhållande till* insatsens kontext och sammanhang. Detta är värt att begrunda då man i diskussionen om bistånd ständigt poängterar vikten av kontext och kontextspecifika mekanismer för att förstå resultat och för att lära (två EBA-rapporter som diskuterar sådana är Samoff, 2016 och Hydén, 2019).

Pawson och Tilley (1997) betonar hur kontext, mekanism och resultat bör ses som en tidslig och rumslig helhet vid kvalitativ teoribaserad utvärdering.[57] Det finns sällan eller aldrig uttömmande kontextlösa förklaringar till resultat av biståndsinsatser. För att kunna förklara eller förstå måluppfyllelse och för att åstadkomma lärande behöver externa tänkbart påverkande faktorer i kontexten beaktas och prövas på ett tillräckligt djupgående sätt tillsammans med och i relation till interna processnära faktorer. För att finna och ”plocka isär” kombinationsförklaringar (*contributory causes* eller *causal packages*) behöver man någon form av teori (se Stern et al, 2012) som väger in och prövar alternativa förklaringar och externa faktorer (jmf kapitel 6) samt troligen domänspecifik kunskap. Förändringsteorin i sig bör utvärderaren, som tidigare konstaterats, ofta rekonstruera då dessa sällan kan hämtas direkt från exempelvis LFA-matriser.

Slutsatserna i en utvärdering är delvis ett resultat av det analysramverk man tar med sig in i arbetet. Utvärderarna i materialet tycks ofta utgå från olika rationella standardmodeller för hur insatser bör utformas (Policycykeln, SMART-kriterier, olika logiska ramverk) som läggs som ett raster i analysen och binder ihop summativ och formativ analys. Risken är att utvärderarna genererar likartade slutsatser över tid med ständigt otydliga problemanalyser, oSMARTa mål och oklara förändringsteorier (insatsen lever inte upp till

[57] Man talar om CMO-configurations dvs. Context- Mechanism- Outcome. Utgångspunkten är (starkt förenklat) en hypotes om att program uppnår resultat (Outcome) (O) på grund av någon underliggande mekanism (M) som aktivierats i vissa, inte alla kontexter (C).

standardmodellen), men utan att generera nytt domän- eller kontextspecifikt lärande om hur enskilda insatser verkar eller inte verkar i sina sammanhang.

Dessa slutsatser är inte nya. Rikrevisionsverket (1988) konstaterade att det då fanns viss kritik mot utvärderingskonsulter inom SIDA och citerar en handläggare: ”*utvärderingarna presenterar ofta samma recept klätt i utvärderingsspråkets termer. Få utvärderingar tränger under ytan. Det finns dock undantag*” (sid 101).

En studie av Sidas utvärderingar 10 år senare (Carlsson et al. 1999) hittade: ”*få exempel där utvärderingarna faktiskt bidrog med något nytt i termer av kunskap..*”

Utifrån en syntes av 84 decentraliserade Sidautvärderingar konstaterade Christopolos et al (2014) att *”kontext och orsaker bakom fattigdom inte analyseras särskilt väl”*.

En angränsande slutsats är att de hindrade faktorerna ovan (Tabell 21) ganska sällan berör sektors- eller domänspecifika faktorer och mekanismer. De nio utväderingarnas hindrande faktorer rör i låg grad sakområdets substans utan tycks i ganska hög grad bygga på antaganden om överförbara styr- och utvärderingsmodeller.

Utvärderarna tar upp ett förhållandevis stort antal faktorer men ofta utan att specifikt konkludera vilka faktorer som specifikt förklarar eller begripliggör den svaga måluppfyllelsen. I den meningen gör utvärderarna sällan förklarande analyser utan de lyfter hindrande (och i mindre grad främjande) faktorer baserat på egen erfarenhet, utsagor i intervjumaterial eller standardmodellen. Bidraget till lärande blir generiskt men ofullständigt inriktat på generell projektmetodik som i praktiken aldrig på ett uttömmande sätt kan begripliggöra eller förklara framgång eller misslyckande.

Summering: förklaringar till svag måluppfyllelse

Utvärderingarna bidrar med möjliga delförklaringar till den begränsade måluppfyllelsen och som kan medverka till visst lärande kring insatsernas konstruktion, styrning, organisering, implementering och uppföljning. Låg vikt läggs dock på externa förklaringar eller mekanismer i eller i gränslandet till insatsens vidare kontext och sammanhang. Domänspecifika faktorer, kunskaper och mekanismer tycks delvis också saknas i analyserna. Utvärderarna listar ett antal möjliga faktorer eller hinder som man menar kan förbättra verksamheten snarare än att analysera hur eller varför resultat inte uppnåtts. Detta begränsar delvis utvärderingarnas lärpotential och användbarhet.

8. Slutsatser, rekommendationer och frågor framöver

Det huvudsakliga syftet med denna studie är att besvara två övergripande frågor:

1. *Vilka slutsatser dras om måluppfyllelse (eller uppnådda förväntade resultat) i nyligen genomförda utvärderingar av svenska bilaterala biståndsinsatser?*
2. *Är utvärderingarnas slutsatser om måluppfyllelse tillförlitliga och, i så fall, användbara för lärande?*

Två slutsatser

I kapitel 4 besvaras den första frågan som att:

1. Urvalet av Sidas utvärderingar (decentraliserade och strategiska) drar oftast slutsatsen att insatserna (i huvudsak) når sina mål, så långt det gjorts en bedömning.

Denna slutsats ligger i linje med ett stort antal tidigare utvärderingssammanställningar i svenskt och internationellt bistånd (se kapitel 6).

Slutsatserna i de utvärderingar jag gått igenom begränsas av att bedömningarna sällan görs mot övergripande mål. De insatser som beskrivs som mindre effektiva är få och variationen så liten att det blir metodologiskt svårt att dra slutsatser om mer eller mindre effektiva projekt- eller organisationstyper, geografier eller arbetssätt.

Lärpotentialen minskar genom att utvärderingarna drar så likartade slutsatser att de inte kan bidra med ny kunskap om vad som fungerar var, när, hur och varför. Om nästan alla rapporter drar slutsatsen att

insatserna når målen så saknas den variation i erfarenhet som är central för lärande och som motiverar återkommande studier kopplat till ett specifikt kriterium?[58]

Den andra frågan besvaras i kapitel 5 genom att relatera urvalet till fyra kriterier för tillförlitlig utvärdering. Den sammanfattande bedömningen är att:

2. Slutsatserna om måluppfyllelse i urvalet av Sidas decentraliserade utvärderingar troligtvis inte är tillförlitliga.

Även denna slutsats ligger i linje med tidigare studier (kapitel 6) som nästan alltid ifrågasatt tillförlitligheten i svenska och internationella utvärderingar av måluppfyllelse inom biståndet. Frågan är dock om denna slutsats går att generalisera till Sidas utvärderingar i allmänhet och i så fall med vilka begränsningar. Det är här viktigt att beakta att denna studie inte behandlar andra typer av utvärderingsverksamhet eller andra möjliga källor till information om måluppfyllelse som inte nödvändigtvis är utvärderingar. Det är också så att utvärderingsfunktionen (inklusive kvalitetssäkring) utvecklas kontinuerligt vilket gör att tidigare urval inte nödvändigtvis speglar nuvarande kvalitet. Ytterligare reflektioner om dessa frågor ges i kapitel 9.

Fyra rekommendationer

Analysen och slutsatserna i tidigare kapitel leder till fyra rekommendationer. De riktar sig i första hand till Sidas utvärderingsenhet, men även till Sidas ledning och UD för intern och extern efterfrågan och styrning för mer tillförlitliga utvärderingar mot övergripande förväntade mål.

---

[58] En parallel är utvärderingskriterieriet *relevans* (relevance) där årsrapporten för Sidas utvärderingsenhet 2019 drog slutsatsen att av 28 utvärderingar som studerat kriteriet har samtliga dragit slutsatsen att insatsen varit relevant eller starkt relevant. Denna slutsats är likartad åren dessförinnan (2017 och 2018). Man kan ifrågasätta vad en beställande handläggare lär sig om relevansen från en projektutvärdering, om hen inte känner till att nästan alla Sidas insatser bedöms relevanta. Vad motiverar att kriteriet används? Se även Samoff (2021).

Rekommendation 1: De decentraliserade utvärderingarnas metodologiska kvalitet bör stärkas betydligt (se kapitel 5).

En förbättring av den metodologiska kvaliteten framöver är viktig för att man på alla nivåer inom biståndet ska kunna lita på slutsatserna i Sidas utvärderingar. Utan tillförlitliga decentraliserade utvärderingar saknas oberoende kunskap om effektiviteten i svenskt bistånd på insatsnivå. Det är inte troligt att partnerorganisationers utvärderingar, EBAs utvärderingar, forskning eller internt uppföljningsarbete kan kompensera för detta.

Rekommendation 2: Fler insatser bör utvärderas mot övergripande och långsiktiga mål (se kapitel 3 och 4).

Utan utvärderingar mot övergripande förväntade mål så görs i praktiken ingen fullständig bedömning av effektiviteten. Det är för framtiden angeläget att tydliggöra varför Sida i låg grad utvärderar sina insatser mot övergripande mål. Jag har i studien visat att detta troligen inte handlar om brist på efterfrågan från Sidas handläggare eller övriga beställare.

Ett alternativ, komplement eller ett tillägg till analyser av mål på den övergripande nivå kan vara att analysera insatsen inom ramen för en så kallad effektkedjeanalys (Sandahl, 2019). På så sätt behöver man inte alltid invänta eller tidsmässigt korrekt avpassa utvärderingen. Det kan ibland vara lämpligt att fokusera på resultat på medellång eller kort sikt och vissa fall kan nivån outcome vara den högsta. *Men* det är problematiskt om man aldrig bedömer eller systematiskt avstår från att bedöma biståndsinsatser i förhållande till initialt formulerade övergripande målsättningar. Min slutsats är att ungefär en tredjedel av insatserna bör kunna utvärderas utförligt mot den övergripande målnivån vid den tidpunkt när utvärderingarna görs.

Svenskt bistånd är idag i ganska hög grad fokuserat på kärnstöd, regionala och globala stödformer, insatser som är svåra att utvärdera. Sida bör därför analysera utvärderingsbarheten i dessa insatser med avseende på förutsättningarna att i framtiden utvärdera och bedöma det svenska biståndets måluppfyllelse och effektivitet (jmf diskussionen i kapitel 5).

Det faktum att utvärderingarna fungerar som stöd i insatshanteringens beslutssituationer, utifrån särskilda behov, ska fungera deltagande och ibland genomförs halvvägs genom projektcykeln utgör inte starka argument mot att utvärdera mot övergripande mål. Det finns ingen nödvändig motsättning mellan Sidas verksamhetsnära utvärderingsbehov och användarfokusering och möjligheten att i högre grad ta fram fullständiga bedömningar av måluppfyllelsen i insatserna.

Rekommendation 3: Bedömningar av förväntade resultat bör i högre utsträckning kompletteras med analyser av oförutsedda resultat och eventuella bieffekter inom biståndet (se kapitel 3).

Det är sedan länge belagt att offentliga interventioner och projekt potentiellt kan ha negativa oförutsedda bieffekter och detta gäller också för biståndet. Det finns också starka etiska skäl till att bedöma oförutsedda effekter av bistånd. Sådana analyser kan delvis genomföras separat i särskilda studier för att inte göra Sidas decentraliserade utvärderingar spretiga.

Rekommendation 4: Utvärderingarnas inriktning och frågeställningar bör i framtiden fokuseras kring måluppfyllelse och på att försöka förklara och förstå hur och varför insatser (inte) når sina mål (se kapitel 2, 3, 7 och bilaga 4).

Lärpotentialen i utvärderingarna kan stärkas genom att utvärderingarna i högre grad fokuserar på just måluppfyllelse/förväntade resultat samt frågor om varför och hur resultat nås eller inte nås i insatsen, vilka faktorer eller mekanismer som förklarar framgång och misslyckade och frågor om vem som drog nytta av insatsen, var, när, hur och varför.

Utvärderingarna bör, lämpligen utifrån teoribaserade utgångspunkter (Pawson och Tilley, 1997, Stame, 2004, Mayne, 2012, Stern, 2012, Beach och Brun Pedersen, 2013) fokusera på lärande genom att så långt det är möjligt söka öppna den ”svarta låda” av mekanismer, kontext och processer, som kan bidra med förklaringar till biståndets troligen olikartade resultat i olika specifika sammanhang, vid olika tidpunkter och för olika målgrupper.

Fem frågor för fortsatta studier

Det finns ett stort antal närliggande och besläktade frågor om svenskt bistånd, utvärdering, uppföljning, rapportering och effektivitet som kommit upp under arbetets gång men som inte fått plats i denna studie. Följande fem frågor bör belysas och analyseras särskilt framöver.

1. Hur bedöms måluppfyllelse mer generellt i svenskt bilateralt bistånd? Vilka underlag utöver utvärderingar använder sig Sida, FBA, Swedfund och andra av för bedömning och rapportering av måluppfyllelse till riksdag och regering? Hur tillförlitliga är dessa underlag och bedömningar?
2. Hur kan och bör man gå tillväga för att aggregera upp och summera Sidas resultatunderlag (utvärdering, uppföljning och rapportering) till strategi- och landnivå?
3. Hur kan Sida och andra biståndsgivare bedöma oförutsedda resultat och bieffekter på ett mer täckande, systematiskt och tillförlitligt sätt? Hur utvärderas det idag? Hur kan och bör det göras?
4. Hur tillförlitliga är slutsatserna i Sidas strategiska och partnerledda utvärderingar?
5. Hur kan de utmanande kausalitetsbedömningarna i utvärderingar utvecklas för att uppnå nivån ”good enough” också i utvärdering av mindre eller begränsade biståndsinsatser eller där utvärderingsbudgeten är begränsad?

9. Avslutande reflektioner om tillförlitlighet

Ett viktigt budskap i denna rapport är att Sida bör öka sina ambitioner att genomföra tillförlitliga utvärderingar av insatsernas måluppfyllelse, och då med mer vikt på övergripande mål. Är detta en rimlig och realistisk förväntan? I detta avslutande kapitel görs ett antal reflektioner om möjliga orsaker till bristande tillförlitlighet inom fyra områden. Orsakerna diskuteras förutsättningslöst och utan anspråk på slutgiltiga eller definitiva slutsatser. De ska ses som tentativa utgångspunkter eller ingångsvärden för en fortsatt diskussion och analys.

Det finns idag en samhällellig och politisk förväntan om att Sida ska utvärdera sina insatser. Att regering eller riksdag skulle tona ner sina ambitioner på utvärderingsområdet för biståndet är troligen orealistiskt.

En av de mest använda och inarbetade definitionerna av utvärdering, som Sida också använder, är: *"a systematic and objective approach of determining the merit, worth or value of an object"*. Idéen om utvärdering innehåller alltså i sig en förväntan om systematik, att det görs djupgående, oberoende och (någorlunda) uttömmande så att insatsens värde så långt möjligt kan klarläggas.

Sida beskriver på sin hemsida det egna utvärderingsarbetet som ett "viktigt verktyg": *"Utvärdering skiljer sig från uppföljning genom att den genomförs vid valda tillfällen, har ett fokus på de mer långsiktiga resultaten och kan både förklara, ifrågasätta och värdera dem på ett systematiskt sätt."*[59]

Denna studie har inte undersökt *varför* utvärderingarnas måluppfyllelsebedömningar inte är tillförlitliga eller om det finns någon koppling mellan slutsatser och metodologi. Jag har inte heller

[59] https://www.sida.se/partner-till-sida/samarbeta-med-sida/utvardering-av-svenskt-utvecklingssamarbete, läst 2021-04-21.

studerat tillförlitligheten i de strategiska eller partnerledda utvärderingar som Sida ser som viktiga i arbetet med att utvärdera det svenska biståndet. Ett utvecklat utvärderingsarbete förutsätter dock ytterligare analys av vad den bristande tillförlitligheten i de decentraliserade utvärderingarna beror på. Kontinuiteten i problemen understrycker betydelsen av att ansvariga tar tag i frågan på ett genomgripande sätt.

Begränsningarna finns potentiellt på både på efterfrågesidan (inom Sida) och på utbudsidan (hos utvärderingskonsulterna samt på marknaden för utvärderingstjänster) varför svaret bör sökas i båda dessa sfärer och i deras samspel. Varken beställarna eller utvärderarna antas ha fullständig kontroll över slutprodukten, även om den *ansvariga* för kvaliteten i en enskild utvärdering alltid kan anses vara rapportens huvudförfattare.

Fyra områden som kan utgöra möjliga ingångsvärden i en förklarande analys av tillförlitligheten i Sidas decentraliserade utvärderingar är frågor om:[60]

1. Efterfrågan, behov och incitament.
2. Systemet för upphandling, avrop, feedback på samt uppföljning av utvärderingar.
3. Utbudet på utvärderingsmarknaden.
4. Kunskaper om utvärdering samt lärande och utveckling i verksamheten.

[60] Varje utvärdering har grund i en efterfrågan som tar form inför och i upphandlingsprocessen eller ett avrop utifrån institutionaliserade arbetsssätt, kunskaper och erfarenheter inom ramen för befintliga system och strukturer. Avropet genererar ett ”svar” från utbudssidan i form av en rapport som tas emot, följs upp, får ett svar, publiceras, ligger till grund för lärande i form av nya studier, framtida ramavtal för utvärderingar osv. Systemet kan i teorin ha inslag av både rationellt effektivitetsmaximerande, institutionell tröghet, spårbundhet, ”muddling trough” och ”lämplighetslogiskt” beslutsfattande (Sundström, 2001).

Områdena fokuserar på beställarnivån och relationen med utvärderare på marknaden för utvärderingstjänster. Utgångspunkten är att tillförlitlighet i Sidas utvärderingar mest troligt påverkas av ett flertal samverkande faktorer men att visa aspekter är viktigare än andra. Detta gör att problemet inte löses av enskilda åtgärder utan man kommer efter en gedigen analys behöva överväga flera samverkande åtgärder i kombination. Svensk biståndsutvärdering är också del i en internationellt standardiserande verksamhet där OECD/DAC är en central aktör och marknaden för utvärderingstjänster är internationell, varför den internationella nivån också bör beaktas.

1. Efterfrågan, behov, incitament

Sida arbetar med en användningsfokuserad utvärderingsansats. Fokus ligger på att identifiera vem de avsedda användarna av en specifik utvärdering, att det finns en tydlig efterfrågan och ett specifikt behov. På en övergripande nivå kräver tillförlitlighet att beställaren har tydliga förväntningar på tillförlitlighet och får genomslag i avtal och dialog för denna förväntan**.** Slutproduktens kvalitet är i hög grad ett utfall av vilken kvalitet man som kund efterfrågat och beställt. Detta påverkas bland annat av beställarens förkunskaper och beställningens tydlighet och detaljeringsgrad. Det kräver dock också en process som löpande arbetar för tillförlitlighet och som svarar på och ingriper mot brister.[61]

I utvärderingssammanhang kan man anta att beställningarnas fokus avseende utvärderingskriterier, syfte och frågeställningar kan påverka kvalitet och tillförlitlighet. Kvaliteten är beroende av beställarens förmåga att balansera en ofta pluralistisk förväntansbild mot det

[61] Inköp av utvärderingar kan analyseras på samma sätt som andra marknader (Nielsen et al, 2019) och utgå från köparens inflytande och position på marknaden för utvärderingstjänster. Den påverkas då av antalet andra köpare, den finansiella storleken på inköp, skillnader i utbud mellan konkurrenter, priskänslighet, förutsättningarna att byta leverantör eller att ersätta en tjänst med en annan.

behov av fokus och tydlighet som möjliggör metodologisk stringens.[62] OECD/DAC har en påverkan på efterfrågessidan genom de utvärderingskriterier Sida och andra givare arbetar med. För trots att både OECD och Sida anger att kriterierna ska användas med urskiljning så är det troligt att kriteriernas tillämpning drivit eller i alla fall upprätthållit en brist på fokusering i Sidas decentraliserade utvärderingar. Detta kan i sin tur ha begränsat förutsättningarna för djupgående analyser av viktiga utvärderingsfrågor som mål-uppfyllelse (se Burman och Hårsmar, 2015).[63]

Det faktum att utvärderingarna ofta beställs och genomförs under insatstiden eller i anslutning till att insatsen avslutats (jmfr kap 3) förvårar ofta bedömningen mot övergripande mål. Förhållandevis många av Sidas uppdragsbeskrivningar anger specifikt hur utvärderingarna ska genomföras metodologiskt, vem som ska samla vilken data när och liknande. Det kan komplicera för utvärderarna att göra ett bra jobb.[64]

Finns det svagheter i Sidas riktlinjer eller policy som kan förklara brister? Sidas utvärderingshandbok (2020), som har utvärderare och beställare som målgrupp, klargör tydligt betydelsen av tillförlitlighet och integritet hos utvärderaren och hänvisar till OECD/DACs kvalitetsdokument som betonar att frågor ska besvaras utifrån tillförlitlig evidens. Även om handboken knappt diskuterar måluppfyllelsekriteriet så poängteras de metodologiska frågorna och

---

[62] Samtidigt kan man hävda att det är en viktig kompetens hos en erfaren utvärderare att respondera på och kommunicera svårigheten med exempelvis för många utvärderingsfrågor givet en viss budget och tidsram.

[63] Ansvaret kan inte läggas på OECD då en statlig myndighet vid tillämpningen av internationellt överenskomna , ej bindande, kriterier bör säkerställa att dessa implementeras som en övervägd anpasssad ”översättning” till svenska förhållanden. Beställare i organisationen behöver förstå hur kriterierna ska användas och ha kunskap om hur man bäst prioriterar mellan kriterier etc.

[64] Exempelvis slår flera uppdragsbeskrivningar i materialet fast att utvärderarna ska genomföra en enkätundersökning, samtidigt tydliggör beställaren inte varför enkäten behövs i det specifika fallet.

tillförlitligheten. Troligen visar detta att förbättrade policy- och styrdokument har en ganska begränsad roll i fortsatt utvecklingsarbete.

Bedömningar av måluppfyllelse i specifika organisatoriska sammanhang kan samtidigt ses i relation till andra organisatoriska mål, interna maktrelationer och incitament. Det är inte alltid självklart att tillförlighet trumfar ”rätt” slutsats precis som att djup inte alltid trumfar bredd. Nobelpristagaren Elinor Ostrom menade i en studie om incitamentsbilden på Sida att *”ingen effektiv efterfrågan på meningsfull utvärdering existerar inom biståndet”*. (Ostrom et al, 2002, sid 247).[65] Ostrom hävdande att de sammantagna incitamenten bland tongivande aktörer leder till ”urvattande” utvärderingsprocesser där de aktörer som i kommer till tals (givare, utvärderingskonsulter och personal i insatserna) kan sakna incitament att beskriva insatsen på ett kritiskt och tillförlitligt sätt, att avsluta insatser etc.

Flera forskare menar att utvärdering har sin egen politiska ekonomi:

> ”Utvärdering som aktivitet äger rum i ett politiskt sammanhang. Policys och program som utvärderas föreslås, definieras, diskuteras, antas och finansieras genom politiska processer. Utvärderingen i sig får en politisk udd eftersom den gör implicit politiska yttranden om frågor såsom legitimiteten för programmål och användbarheterna i olika strategier. […] En utvärdering av effektivitet bör möjligen även poängsätta effektivitet med avseende på politisk-organisatorisk överlevnad” (Carlsson et al, 1994, sid. 176).

[65] Det är oklart om Ostrom drog sin slutsatser baserat bara på studier av Sida.

Beställarna lägger uppdrag på utvärderare som agerar inom en primärt affärsmässig snarare än vetenskaplig logik och kvalitetssäkring där incitamenten att framföra kritiska synpunkter till en trogen kund säkert kan diskuteras.[66]

På beställarsidan kan man i vissa fall uppfatta att ungefärliga bedömningar på kort eller medellång sikt av resultat är tillräckliga för de behov man har, handläggare kanske inte alltid intresserar sig för metodfrågor etc.[67]

Om man landar i det är faktorer på efterfrågesidan, incitament eller politisk ekonomi som påverkar tillförlitligheten negativt så kan lösningarna handla om sådant som:

- Insatser för att stärka Sida som kund på marknaden för utvärderingar (Nielsen et al 2019).
- Överenskommelser mellan UD och Sida och Sida och partners om att beslutsfattande i den omfattning de baseras på utvärderingar, ska vara grundade i god evidens.
- Insatser för att övertyga chefer och handläggare om vikten av tillförlitlig metodologi i utvärderingar.
- Fortatt arbete med fokusering och avgränsning av utvärderingskuppdrag och antalet frågeställningar i uppdragsbeskrivningar kopplat till budget, genomförandetid och frågeställningarnas komplexitet.

[66] Sida anför att det är partners och inte Sida som utvärderas och att det försvagar den typen av incitament. Samtidigt bör Sida ses som intressent också när man utvärderar partners.

[67] Slutsatsen ligger i linje med Carlsson et al. 1997 som rapporterade om tydliga skillnader mellan en extern kvalitetsbedömning av Sidas utvärderingar och den bedömning handläggarna själva gjorde.

- Utveckling av processen för att prioritera mellan OECDs utvärderingskriterier och frågeställningar i enskilda utvärderingar (Burman och Hårsmar, 2015).
- Insatser för att begränsa ev. internpolitisk påverkan på design eller slutsatser i utvärderingarna.
- Ytterligare kodifiering av krav på integritet, objektivitet, oberoende och självständighet.
- Förstärkta processer för att förtydliga och avgränsa insatsers målbild innan utvärderingsarbetet inleds.

2. Systemet för upphandling, feedback och uppföljning

Svårigheter att säkerställa tillförlitlighet kan hänga samman med system och process för upphandling/avrop (ramavtal), utvärderingarnas budget, utvärderarnas ägarskap, återkoppling till leverantör och resurssättning.

Viktigt är troligen hur Sidas *ramavtal* för uvärderingskonsulter är designat angående incitament att leverera tillförlitliga utvärderingar och att Sida via ramavtalet ges tillgång till bästa möjliga utvärderare internationellt. En grundfråga är i vilken i utsträckning det får återverkning att inte leverera av god kvalitet, givet att tillförlitlighets-problemen varit återkommande över tid. Ramavtalet för utvärderingstjänster ska också balansera behovet av konkurrens-utsättning mot betydelsen av att det är enkelt för personal att beställa utvärderingar.[68]

[68] Upprepad konkurrensutsättning kan ge mer arbete för beställaren vilket kan leda till att färre utvärderingar genomförs samtidigt som brist på konkurrens under ett ramavtal kan sänka kvaliteten.

När Sida lät utvärdera sitt ramavtal för utvärderingar 2014 var slutsatsen att ramavtalet "*inte varit en tillräcklig mekanism i sig för att förbätta utvärderingarna kvalitet som avsetts*" (LaGuardia et al, 2014, sid. 9). Mer specifikt menades att rankingsystemet för leverantörer inte utgjort en "*adekvat upphandlingsmodell för att främja kvalitet i utvärderingstjänsterna*" (sid. 10). Ramavtalet ser annorlunda ut idag och vi saknar underlag för att bedöma om det främjar eller motverkar metodologisk kvalitet. Otvetydigt har ramavtalsupphandlingen och avtalets design dock stor potential att påverka, till det bättre eller sämre. Ramavtalet och dess uppföljning är också ett verktyg där Sidas utvärderingsenhet har stor möjlighet att påverka kvaliteten i decentraliserade utvärderingar.

En viktig fråga är om *utvärderingarnas budget* är tillräckligt väl tilltagna (särskilt per utvärderingsfråga). Genomsnittpriset för Sidas decentraliserade utvärderingar har ökat från cirka 500 tkr åren 2011-2014 (La Gaurdia et al, 2014) till cirka 700 tkr åren 2017-2019 (median 680 tkr, min 230 tkr, max 1,5 mkr). Uppskattningsvis har utvärderingsteamen därmed ca. 15 arbetsveckor, eller knappt 4 månader, på sig att genomföra en genomsnittlig utvärdering utan att behöva sänka sin timkostnad.[69] En tillförlitlig utvärdering behöver ha tillräckligt med tid för de olika faserna i arbetet: planera, samla data, analys, rapportering och kvalitetssäkring.

Forss et al (2006) studerade Sidas utvärderingar men fann inget tydligt samband mellan kostnad och kvalitet. De bästa utvärderingarna var inte alltid de dyraste och kvalitetsproblemen ansågs inte kunna åtgärdas med större budget. Om Sida inte har möjlighet att minska antalet frågor i de decentraliserade utvärderingarna så bör man kanske analysera om budgeten per fråga eller utvärderingskriterium behöver öka ytterligare.

[69] Vi har räknat på ett timpris om 1200 kr/timme för konsultarvoden men inte räknat in kostnader i projekten utöver personal. Med tre personer i teamet blir det ca 5 arbetsveckor per person.

Det krävs inte bara en rätt avvägd budget utan också kalendertid och kontroll, ägarskap, över egen tidsanvändning om en utvärdering av god kvalitet ska kunna åstadkommas. Det är inte självklart att dessa går hand i hand. Om ambassaden på förhand specificerar en veckas arbete i fält för utvärderingsteamet utan att det är förankrat i den design och tidsanvändning utvärderaren behöver så minskar möjligheterna att göra ett bra arbete trots att budgeten som helhet är rätt avvägd. Om det är ett problem kan Sida bl.a. överväga att öka den genomsnittliga kalendertid som ska finnas tillhanda för decentraliserade utvärderingar alternativt att utvärderingsenheten i sin rådgivning trycker på utvärderarens ägarskap över egen tidsanvändning i termer av bl.a. datainsamling i fält.

Även om det finns ett centralt stöd via Sida UTV i form av bl.a. handboken, guider, mallar samt möjlighet till rådgivning så påverkas förutsättningarna för tillförlitlighet av att beställarna på ambassader eller bland handläggare ofta är ovana beställare av utvärderingar (jmf. Burman och Hårsmar, 2015 samt Burman, 2017).

Efter att en Inception report och rapportutkast levererats är det viktigt att utvärderaren får en genomarbetad *respons* och kvalificerade synpunkter. Hela situationen innehåller dock en obalans då utvärderaren som regel har lång erfarenhet av utvärdering och situationen som helhet, medan ansvarig handläggare sällan har detta.

Jag har uppfattat att Sida inte anser sig inte ha tillräckligt med resurser för att utförligt stötta eller kvalitetsgranska alla decentraliserade utvärderingar (ca 30 per år). Tidigare analyser av Sidas utvärderingar (Forss et al, 2007 och La Guardia et al, 2014) rapporterade om begränsad kapacitet bland Sidas tjänstemän att tillhandahålla teknisk support till utvärderingsprocessen. Beställaren av en utvärdering bör samtidigt ha en förmåga och kapacitet att bedöma vad som är en bra eller dålig levererad tjänst. Om utvärderaren sällan eller aldrig får feedback på leveransen eller om det saknas koppling mellan kvalitet och återkopplingens innehåll så är det troligt att detta på sikt undergräver tillförlitligheten. Värdet av

att efterfråga tillförlitlighet minskar om man som köpare i praktiken inte svarar på kvalitetsbrister. Men jag har inte heller här någon uppdaterad kunskap om hur situationen ser ut idag i detta avseende.

Man bör dock analysera om de decentraliserade utvärderingarna inte kräver ett starkt inslag av centralt ansvarstagande om man vill att tillförlitligheten i de decentraliserade utvärderingarna ska stärkas. Man behöver då analysera om det finnas tillräckliga *finansiella och personella resurser* på Sida UTV för att säkerställa kvaliteten i de decentraliserade utvärderingarna. Situationen har förändrats över tid, men när OECD/DAC år 2016 gjorde en kartläggning över utvärderingsfunktionerna hos ett stort antal bilaterala och multilaterala givare så rapporterade Sida UTV lägst budget av OECD/DACs samtliga givare som andel av total biståndsbudget (OECD/DAC, 2016 sid 230).[70]

Om man landar i att det är ramavtal, upphandlingsprocess, feedback, resurser och uppföljning av utvärderingarna som håller tillbaka tillförlitligheten så kan åtgärderna bland annat handla om:

- Översyn av ramavtalet med fokus på att försöka stärka incitamenten att leverera tillförlitliga utvärderingsrapporter.
- En stärkt process för utvärderingskunnig återkoppling och feedback på levererade rapporter. Exempelvis genom en förstärkningar av Sida UTVs kapacitetet och förutsättningar att stötta beställare.
- Ökade finansiella och personella resurser till Sida UTV (för att stödja arbetet med att säkerställa kvaliteten i decentraliserade utvärderingar).

[70] Ett skäl till detta kan ha varit att Sidas utvärderingsbudgetar tas på sakanslaget. Sida UTV har efter det organiserats om varför resurssituationen kan ha förändrats.

- Ytterligare stärka förutsättningarna för utvärderare att, baserat på sin sakområdeskompetens, planera sin egen tid, designa sitt upplägg och genomförande inom ramen för utvärderings-uppdraget inklusive materialinsamling i fält.
- Ökad kalendertid som norm för varje utvärdering.
- Ökad utvärderingsbudget per fråga eller utvärderingskriterium.

3. Utbudet på utvärderingstjänster

Framgångsrik upphandling av en utvärderingstjänster påverkas även av säljarens position på marknaden relativt köparens och av den konkurrens som råder (Nielsen et al 2019).[71] Mer specifikt styrs utbudet av utvärderare i Sidas decentraliserade utvärderingar av vilka som efter ramavtalupphandling tilldelas ramvavtal med Sida samt de enskilda utvärderare som tillhandahålls under ramavtalet. Kvalitetsgranskningen av Sidas utvärderingar 2014 (La Guardia et al) rapporterade om en över tidperioden 2008-2013 oförändrad situation med:

> ”Ett begränsat antal kvalificerade konsulter, och brist på färdigheter och kapacitet bland utvärderarna med vilka Sida arbetar, vilket eventuellt återspeglar en brist på konkurrens och en liten variation bland konsulterna”(Ibid, sid. 96).

Jag har inga data om hur situationen ser ut idag avseende till exempel omsättning på kontrakterade utvärderare över tid eller timprisernas utveckling för utvärderingskonsulttjänster.[72]

---

[71] Antalet säljare, dessas relativa storlek och kapacitet, skillnader i unicitet och kvalitet mellan leverantörer, hot från nya aktörer, trösklar in på marknaden etc.
[72] Timpriserna kunde ha gett en indikation på konkurrensnivå och marknadens mognadsgrad.

Något som talar för att konkurrensen *potentiellt* kan vara stor är att utvärdering av bistånd är en internationell verksamhet som inte behöver genomföras av svenskar eller via svenska företag. Tjänsten är inte bunden till en specifik geografi annat än vid datainsamlingen. Här bör man kanske analysera om det finns ”trösklar” i ramavtalssystemet som gör att Sida inte fullt ut drar nytta av internationell konkurrens och kapacitet på utvärderingsområdet. Man kan analysera om det från konsultföretagens sida finns sätt att rekrytera till ramavtal som minskar inflödet av internationell metod- och utvärderingsexpertis eller som begränsar omsättning eller inflöde av utvärderingskompetens. I sammanhanget bör man bedöma hur stor vikt som ska läggas på Sidaspecifik utvärderingserfarenhet då sådana krav kan belöna leveratörer som tidigare inte levererat god kvalitet.

Om man landar i att brister i utbudet på utvärderare eller utvärderingsfirmor kan ha påverkat tillförlitligheten så kan det föranledda åtgärder som:

- En förändrat process för framtida ramavtalsupphandling med fokus på att bredda utbudet av utvärderare och öka konkurrensen med fokus på en stärkt utvärderings- och metodkunskap.

- Förstärkta insatser för att nå ut till nya grupper av leverantörer av utvärdering internationellt i samband med ramavtals-upphandling.

- Fortsatt och utökad dialog med befintliga ramvatalsaktörer om att bredda utbudet av utvärderare med fokus på att stärka och bredda metodkompetensen hos både utvärderare och hos kvalitetsgranskare.

4. Kunskap, lärande och utveckling i verksamheten

De tre områden som hittills diskuterats berör i hög grad *formerna* för Sidas inköp av utvärderingar. Innehållet i arbetet påverkas dock i mycket hög grad av kunskapen i systemet och av det lärande som sker undervägs i verksamheten.

Att bedöma måluppfyllelse i en offentlig intervention är en kvalificerad uppgift som kräver kunskaper om avancerad utvärderingsmetodologi, sakområdesfrågor och kontext. Samtidigt saknas det idag accrediteringssystem för utvärderare, det finns inga krav på formell utbildning för den som vill kalla sig utvärderare och de kvalitetssäkringssystem som finns (inom exempelvis OECD DAC) kan och bör inte exakt precisera hur en enskild utvärdering bör genomföras.[73]

En viktig kvalitetsfunktion i de decentraliserade utvärderingarna är ramavtalsleverantörernas kvalitetssäkringsfunktioner. Samtidigt finns det flera potentiella fallgropar om kvalitetsansvarig kommer in sent i arbetet, om kvalitetsgranskningen saknar systematik, om kvalitetsansvarig inte använder vedertagna kriterier eller anpassar sig till en informellt förväntad (lägre) kvalitetsnivå utifrån hur hen är van att arbeta. Utvärderarna kan sakna incitament att lyssna på kvalitetsansvarig och kvalitetsgenomgången kan bli formalistisk eller sakna ett substans- eller sektorsspecifikt perspektiv (Forss och Pollard, 2020). Alternativt att kvalitetssäkringen blir ”window dressing” med begränsad påverkan på faktiskt utvärderingsarbete.

[73] Diskussionen om utbildning, professionalisering och accreditering av utvärdering och utvärderare verkar ha pågått nästan lika länge som diskussionen om utvärdering (se Schwandt, 2019). En utvärderingsdesign måste anpassas till den enskilda insatsen och uppdraget vilket begränsar användbarheten av generiska checklistor och centralt definierade processer.

Ett övergripande intryck från denna rapports genomgång av 80 rapporter är att samtida tänkande om utvärdering av biståndsinsatser med avseende på ansatser, designval och metodologi inte riktigt får genomslag i Sidas decentraliserade utvärderingar. Några aspekter på detta är att man i rapporterna sällan föreslår någon egentlig utvärderingsdesign, bristen på referenser till modern utvärderingsmetodik och litteratur (som ofta tagits fram för kvalitativa analyser på biståndsområdet, se exempelvis Vaessen et al, 2020 eller Stern et al, 2012). Men också bristen på transparens om hur enskilda utvärderingsfrågor besvarats, diskussioner om orsakssamband som när de förekommer är knapphändiga och bristfälliga. Många av de designalternativ som idag finns tillhanda för kvalitativ och kvantitativ utvärdering av biståndsinsatser används inte under de år vi studerat. I de fall man i rapporten anger användning av en etablerad metodologisk ansats så slutar det ofta med att den i praktiken inte tillämpas.

Angående Sidas lärade om utvärderingsverksamheten kan det konstateras att denna studie är en i rad av studier som dragit likartade slutsatser om kvaliteten i Sidas utvärderingar de senaste 25 åren (Burman, 2017, Christoplos, 2014, LaGuardia et al, 2014, Forss et al., 2008, Carlsson et al, 1997, Forss, 1994). Intrycket är att dessa studier inte haft någon större effekt på utvärderingarnas kvalitet vid besvarande av måluppfyllelsefrågan. Detta trots att majoriteten studier beställts av och därför på något sätt efterfrågats av Sida själva. Det betyder inte att Sida inte lärt av studierna men att ett sådant lärande inte resulterat i tydligt mer tillförlitliga utvärderingar.

Samtidigt är det troligen utmanande att åstadkomma lärande om utvärdering ute i Sidas organisation varför det behövs ett centralt ansvarstagande för decentraliserade utvärderingar. En enskild

handläggare på Sida beställer utvärderingar ganska sällan och de lärdomar som dras kommer kanske till användning några få gånger gånger under en karriär på myndigheten.[74]

Om utvärderarnas lärande kan det konstateras det över tid funnits en tendens att, som det framstår, lyfta fram likartade hinder för genomförandet av utvärderingen med fokus på otydliga mål, bristfälliga resultatramverk, avsaknad av baslinjer, otydliga förändringsteorier och svag projektintern uppföljning. När det delvis är samma utvärderare som över tid genomfört dessa utvärderingar och drar samma slutsatser om utvärderingsbarhet år efter år så framstår det i vissa fall som en brist på lärande. Biståndsutvärderarnas egen planering och analys av återkommande likartade frågeställningar och projekt borde över åren ha förbättrats mer om man tagit vara på tidigare slutsatser, förutsättningar och erfarenheter, försökt lära och utveckla verksamheten. Många av slutsatserna borde bli ingångsvärden till, snarare än konklusioner i, ytterligare ett metodkapitel.

Om man landar i att det är brister avseende kunskaper och lärande som över tid påverkat tillförlitligheten i Sidas utvärderingar så skulle det exempelvis kunna föranledda åtgärder som:

- Extern kvalitetsbedömning av ett slumpmässigt urval utvärderingar varje år genomförd av internationell oberoende utvärderingsexpertis.
- Utökad användning av forskningsmiljöer, forskare och forskarutbildade i utvärderingsarbetet.

---

[74] Cirka 4% av Sidas anställda (25 av 643 årsarbetskrafter 2020) kan väntas ansvara för att beställa en decentraliserad utvärdering ett enskilt år. Och om man antar att behovet är spritt och skiftande över organisationen och år och att få medarbetare arbetar längre än tio år på myndigheten så är det troligen få enskilda anställda (utanför UTV) som hinner med att beställa mer än två-tre decentraliserade utvärderingar under sin karriär på myndigheten och rimligen sker detta med ganska långa tidsliga mellanrum.

- Fortsätt med konsultgemensamma/ramavtalsöverskridande lär-aktiviteter med fokus på metodologi och att lära av erfarenheten från utvärdering av biståndsinsatser.

- Kunskapsprov vid ramavtalsupphandlingar på utvärderings-området framtaget av internationella oberoende utvärderings-expertis.

- Framtagande av en specifik kvalitetspolicy för Sidas utvärderingar med fokus på tillförlitlighet och metodologi. Alternativt en utökning av nuvarande utvärderingshandbok.

Referenser

Alexius, S (2017), Att styra komplexa uppdrag mot mål och resultat - en forskningsöversikt, Trafikanalys.

Andrews, M., Pritchett, L., Woolcock, M (2012), *Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA*), HKS Faculty Research Working Paper.

Argyris, C. och Schön, D (1974) Theory in practice: Increasing professional effectiveness, San Francisco: Jossey-Bass.

Bandstein, S (2006), *Views on Evaluation*, Sida.

Beach, D och Brun Pedersen, R (2013), *Process-Tracing Methods. Foundations and Guidelines*. The University of Michigan Press.

Burman, M och Hårsmar, M (2015), *Utvärdering av svenskt bistånd - en kartläggning*, Rapport 2015:02 från Expertgruppen för biståndsanalys.

Burman, M (2017), *Livslängd och livskraft: Vad säger utvärderingar om svenska biståndsinsatsers hållbarhet?,* Expertgruppen för biståndsanalys.

Befani, B (2021), *Credible explanations of development outcomes: improving quality and rigour of Theory-Based Evaluation with diagnostic approaches and Bayesian Updating*, Forthcoming, Expertgruppen för biståndsanalys.

Befani, B (2016), *Pathways to Change: Evaluating development interventions with Qualitative Comparative Analysis (QCA),* Expertgruppen för biståndsanalys.

Carlsson, J., Eriksson-Baaz, M., Fallenius, A-M., Lövgren, E (1999), *Are Evaluations Useful? Cases from Swedish development cooperation*, Sida.

Carlsson, J (1998), *Swedish Aid for Poverty Reduction: A History of Policy and Practice*, ODI.

Carlsson, J., Forss, K., Metell, K., Segnestam, L., Strömberg, T (1997), *Using the Evaluation Tool. A survey of conventional wisdom and common practice at Sida*. Sida.

Carlsson, J och Wohlgemuth, L (1996), *Capacity Building and Networking. A meta-evaluation of African regional research networks*. Sida.

Carlsson, J., Köhlin, G., Ekbom, A (1994), *The Political Economy of Evaluation. International Aid Agencies and the Effectiveness of Aid*, Palgrave Macmillan UK.

Carlsson, J and Forss, K (1997), *The Quest for Quality- Or Can Evaluation Findngs be Trusted?* Evaluation.

Chapman, N., Lloyd, R., Villanger, E., Gleed, G (2017), *The Quality of Reviews and Decentralised Evaluations in Norwegian Development Cooperation*, Norad.

Chianca, T (2008)*, The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement*, Journal of MultiDisciplinary Evaluation.

Christoplos, I, Liljelund-Hedqvist, A och Rothman, J (2014), *Lessons and Reflections from 84 Sida Decentralised Evaluations 2013 – a Synthesis Review*, Sida Studies in Evaluation.

Colin, E., Sandström, C., och Wennberg, K (2021), *Utvärderingar av näringspolitik - en intressekonflikt mellan myndigheter, konsultföretag, politik och skattebetalare?* Ekonomisk debatt, 4/2021.

Copestake, J., Morsink, M., Remnant, F (2019), *Attributing Development Impact. The qualitative impact protocol case book,* Practical Action Publishing.

Davidson, E.J. (2004) Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation. Beverly HIlls, CA: Sage Publications.

Davies, R. and Dart, J. (2005) *The 'Most Significant Change' Technique - A Guide to Its Use.*

Dijkstra, G (2018), *Aid and good governance: Examining aggregate unintended effects of aid*, Evaluation and Program planning.

Doran, G. T. (1981), *There's a SMART way to write management's goals and objectives*. Management Review.

Easterly, W (2006), *Planners versus Searchers in Foreign Aid.* Asian Development Review.

Easterly, W (2007), *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*, Penguin Group.

Eliasson, K (2006) *Måluppfyllelseanalys. Hur måluppfyllelse, effekter och effektivitet kan undersökas och rapporteras*, Ekonomistyrningsverket.

Ellström, P-E, (2011), *Informal Learning at Work: Conditions, Processes and Logics*, The SAGE Handbook of Workplace Learning, 2011.

Forss, K (1994), *Några aspekter på kvaliteten hos utvärderingar av svenskt biståndssamarbete.* Rapport till Sekretariatet för Analys av Utvecklingsamarbete (SASDA).

Forss, K (2007), Utvärdering som hantverk. Bortom mallar och manualer. Studentlitteratur.

Forss, K., Vedung, E., Kruse, S-E., Mwaiselage, A., Nilsdotter, A (2008), Are Sida Evaluations good enough? An Assessment of 34 evaluation reports, Sida Studies in Evaluation, 2008:1.

Forss, K., Marra, M och Schwartz, R (2011), Evaluating the Complex: Attribution, Contribution, and Beyond, 1st edition, Routledge.

Forss, K och Pollard, A (2020), Responses to Decline. The Persistence of Quality Problems, kapitel 10 i Perrin, B. och Tyrrell, T (red.) Changing Bureaucracies, Adapting to Uncertainty, and How Evaluation Can Help, 1st edition, Routledge.

Forss, K (2021), From Measuring impact to Understanding Change in Long Term Perspectives in Evaluation. Increasing Relevance and Utility, Kim Forss, Ida Lindkvist and Mark McGillivray, Routledge.

Goodhart, C (1975). Problems of Monetary Management: The U.K. Experience. Reserve Bank of Australia.

Holvoet, N., Van Esbroeck, D., Inberg, L., Popelier, L., Peeters, B.,Verhofstadt, E (2018). *To evaluate or not: Evaluability study of 40 interventions of Belgian development cooperation*, Evaluation and Program Planning.

Honig, D and Gulrajani, N (2017), Making Good on Donors' Desire to Do Development Differently. Third World Quarterly.

Hood, C (1991), A public management for all seasons? Public Policy and Administration.

Hydén, G (2019), Democracy in African Governance: Seeing and Doing It Differently, Expertgruppen för biståndsanalys, 2019.

Isaksson, A-S (2017), Geospatial analysis of aid: A new approach to aid evaluation, Expertgruppen för biståndsanalys.

Jacobsson, B och Sundström, G (2001) Resultat utan lärande: erfarenheter från tre decennier av resultatstyrning, Stockholms centrum för forskning om offentlig sektor (SCORE).

Johansson, L och Lindgren, M (2007) Sida Evaluations and Audits 2006, A synthesis of findings and conclusions, Sida.

John W. Evers (1980), A Field Study of Goal-Based and Goal-Free Evaluation Techniques. Western Michigan University.

Karlsson Vestman, O (2011), Utvärderandets konst., Studentlitteratur.

Karlström, U., Eriksson, C., Wockelberg-Hedlund, J (2018), Analyser och utvärderingar för effektiv styrning. Betänkande av utredningen om Regeringens Analys- och utvärderingsresurser, SOU 2018:79.

King, G, Keohane, R and Verba, S (1994), Designing Social Inquiry. Scientific Inference in Qualitative Research, Princeton Univrsity Press.

Krogstrup, H-K (2017), Utvärderingsmodeller, Studentlitteratur.

LaGuardia, D., Becker-Thierry, S., Bonoll, F (2014), Mid Term Review of the Framework Agreement for Sida Reviews, Evaluations and Advisory Services on Results Frameworks Final Report, Sida Decentralised Evaluation, 2014:13.

Leeuw, F (2003), Reconstructing Program Theories: Methods Available and Problems to be Solved, American Journal of Evaluation.

Lindblom, C (1959), The Science of `muddling through.` Public Administration Review.

Lindgren, L (2012), Terminologihandbok för utvärdering, Studentlitteratur.

Mayne, J (2012) Contribution analysis: Coming of age?, Evaluation.

Molander, P (2017), Dags för omprövning - en ESO-rapport om styrning av offentlig verksamhet, ESO-rapport 2017:1.

Mosley, P (1986). Aid-effectiveness: The micro-macro paradox. Institute of Development Studies Bulletin.

Nielsen, S-E., Lemire, S., Christie, C-A (2019) The Commercial Side of Evaluation. Evaluation as an industry and as a Professional Service. I Furubo, J-E. och Stame, N, The Evaluation Enterprise. A Critical View. Routledge.

Nilén, H., Nilsson, P-U., Nyberg, J (2002), Syntes av svenska enskilda organisationers utvärderingar, Sida.

Nilsson, K och Widell, L (2006), Effektutvärdering. Att välja upplägg, Ekonomistyrningsverket.

OECD/DAC (2019), Better Criteria for Better Evaluation. Revised Evaluation Criteria Definitions and Principles for Use, OECD/DAC Network on Development Evaluation, 10 December, 2019.

OECD/DAC (2016), Evaluation Systems in Development Co-operation. 2016 Review, OECD/DAC.

Ostrom, E., Gibson, C., Shivakumar, S., Andersson, K (2001), Aid, Incentives, and Sustainability. An Institutional Analysis of Development Cooperation, Sida Studies in Evaluation, 2002:01.

Pain, A (2015), Review of Sida's Support to Afghanistan - Lessons and Conclusions from 7 Evaluations, Final Report, Sida.

Patton, M. Q. (2008), Utilization-Focused Evaluation, Fourth Edition, Sage Publication.

Patton, M. Q. (2018). Principles-Focused Evaluation - The GUIDE. New York: Guilford Press.

Pawson, R. och Tilley, N (1997), Realistic Evaluation. London: Sage.

Pawson, R (2013), The Science of Evaluation. A Realist Manifesto. Sage Publication Ltd.

Peck, L., Jassey, K., Johansson, J (2006), Sida Evaluations and Audits 2005, Sida.

Peck, L och Knippel, V (2008), Sida Evaluations and Audits, 2007. A Synthesis of Findings and Conclusions. UTV Working Paper, Sida.

Peck, L (2007), Erfarenheter av resultatstyrning. En genomgång av utvärderingar och studier, Sida.

Petersson och Sandahl (2016), Kausalitet i filosofi, politik och utvärdering, Studentlitteratur.

Pleger, L., Sager, F., Morris, M., Meyer, W., Stockmann, R (2016), *Are Some Countries More Prone to Pressure Evaluators Than Others? Comparing Findings from the United States, United Kingdom, Germany, and Switzerland.* American Journal of Evaluation.

Pritchett, L (2002), It pays to be ignorant: A simple political economy of rigorous program evaluation, The Journal of Policy Reform.

Riddell, R., Bebbington, A., Peck, L (1995), Promoting Development by Proxy. An Evaluation of the Development Impact of Government Support to Swedish NGOs, Sida Evaluation Report, 1995:02.

Riddell, R (2007), Does Foreign Aid Really Work?, Oxford University Press.

Riksrevisionsverket (1998), Lär sig SIDA? En granskning av SIDAs förmåga att lära sig av erfarenheterna: revisionsrapport, Riksrevisionsverket.

Rogers, P och Weiss, C. (2007), Theory-Based Evaluation: Reflections Ten Years On. Theory-Based Evaluation: Past, Present, and Future. New Directions for Evaluation.

Samoff, J.(2021), Relevant? Almost always. Relevance i Development Cooperation. Working Paper. Expertgruppen för biståndsanalys.

Samoff, J., Leer, J., Reddy, M (2016), Capturing Complexity and Context: Evaluating Aid to Education, Expertgruppen för biståndsanalys.

Sandahl, R (2019), Utvärdering – en balansakt, PM publicerat av Tillväxtverket.

Sandberg, B och Faugert, S (2012), Perspektiv på utvärdering, Upplaga 2, Studentlitteratur.

Scriven, M (1991), Pros and Cons about Goal-Free Evaluation, Evaluation Practice, 1991.

Schwadt, T.A (2019) The concerted effort to Professionalize Evaluation Practise. Whiter Are We Bound? The Evaluation Enterprise. A Critical View. Edited by Jan- Eric Furubo and Nicoletta Stame. Routledge.

Shutt, C (2016), Towards an Alternative Development Management Paradigm?, Expertgruppen För biståndsanalys.

Stame, N (2004), Theory based Evaluation and Types of Complexity, Evaluation.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B (2012) Broadening the Range of Designs and Methods for Impact Evaluations, Working paper 38, Department for International Development UK.

Sundström, G (2001) ”Ett relativt blygsamt förslag” - Resultatstyrningens framväxt ur tre perspektiv. Score Rapportserie 2001:2.

Tarschys, D (2006), Mål utan mening? Om ordstyrning i landet där allt är prioriterat, i Förvaltningens byggstenar – rapport från ett seminarium, Statskontorets småskrifter, Statskontoret, 2006.

Tobisson, E och De Vylder, S (1997), Poverty Reduction and Gender Equality. An Assessment of Sida´s Contry Reports and Evaluations in 1995-96. Sida.

Uggla, F och Wulfing, M (2007), Assessment of Sida’s Support for Human Rights and Democracy. Sida.

Vaessen, J, Lemire, S och Befani, B (2020), Evaluation of International Development Interventions. An Overview of Approches and Methods. Independent Evaluation Group.

Vedung, E (2004), Utvärderingsböljans former och drivkrafter, FinSoc Working Paper.

Vedung, E (2012), Utvärdering i politik och förvaltning, tredje upplagan, Studentlitteratur.

Vedung, E (2016), Implementering i politik och förvaltning, Studentlitteratur.

Vähämäki, J (2017), Matrixing Aid: The Rise and Fall of 'Results Initiatives’ in Swedish Development Aid, Stockholms universitet.

Wilson-Grau, R (2018), Outcome Harvesting: Principles, Steps, and Evaluation Applications., Information Age Publishing.

Offentliga tryck och webbsidor

Blogg om begreppen contribution och attribution i biståndsutvärdering: https://aidleap.org/2015/11/23/contribution-vs-attribution-a-pointless-debate/ (läst 2020-12-18).

Cochrane Handbook for Systematic Reviews of Interventions, Version 6.2, 2021 Cochrane Training, https://training.cochrane.org/handbook/current

Om Contribution Analysis: https://www.betterevaluation.org/en/plan/approach/contribution_analysis

Ray Pawson om komplexitet och tid i utvärdering: https://www.youtube.com/watch?v=c8TReTAbEQw&feature=youtu.be

Regeringens Policyramverk för svenskt utvecklingssamarbete och humanitärt bistånd, Regeringens skrivelse 2016/17:60.

Resultatutvärdering- ett programförslag, SIDA 1971.

Sida's Evaluation Handbook. Guidelines and Manual for Conducting Evaluations at Sida. 2018, Sida.

Strategi för stöd genom svenska organisationer i det civila samhället, 2016-2022, Utrikesdepartementet, 2016.

Bilaga 1 – Genomförande och metod

Studien har genomförts i fem steg:

1. Genomgång av tidigare studier med fokus på måluppfyllelse och måluppfyllelseanalys i och utanför svenskt bistånd.
2. Nerladdning och urval av utvärderingar (decentraliserade och strategiska) från Sidas publikationsdatabas.
3. Framtagande och utprövning av analysramverket.
4. Genomgång av utvärderingarna utifrån ramverket.
5. Sammanställning och analys av data samt författande av rapport.

Nedan beskrivs och diskuteras genomförandeprocessen i detalj (steg för steg enligt ovan).

Metoden i rapporten är en utveckling av den som användes i *Livslängd och livskraft: Vad säger utvärderingar om svenska biståndsinsatsers hållbarhet?* (EBA 2017:12), med fokus på OECD/DACs utvärderingskriterium Sustainability.

Arbetet som helhet genomfördes under perioden maj 2019 till april 2021.

1. Genomgång av tidigare studier

Inledningsvis i arbetet gjordes en genomgång av:

1. Litteratur om målbaserad utvärdering.
2. Studier av måluppfyllelse i svenskt bistånd (primärt sådana som fokuserat på flera insatser samtidigt eller som syntetiserat utvärderingar) samt,
3. Tidigare studier om Sidas utvärderingsarbete och kvaliteten i Sidas utvärderingar.

Punkt 2 och 3 ovan avser nästan enbart så kallad grå litteratur och som ej publicerats i vetenskapliga tidskrifter. Många av dessa studier är beställda av Sida själva. Genomgången avrapporteras i kapitel 2 och 3.

Litteraturgenomgången (punkt 1 ovan) gör inte anspråk på att fånga den samlade forskningslitteraturen om utvärdering av mål-uppfyllelse, även om större delen av det som finns tillhanda i bokform på svenska gåtts igenom (se litteraturförteckning).

Syftet med genomgången av tidigare studier har varit att skapa en enkel teoretisk grund för det analytiska ramverket samt att sätta in studiens slutsatser i ett längre och bredare perspektiv. Det sist nämnda syftet gör att vi inte avser att strikt tillämpa all kunskap från teoridelen på det empiriska materialet.

Litteraturgenomgången har varit viktigt då studien bör ses i ett vidare sammanhang där givare över tid återkommande beställt studier som summerat slutsatser från utvärderingar, analyserat utvärderingarnas kvalitet och då ofta (som det visat sig) dragit likartade slutsatser.

Litteratursökningar har genomförts med med hjälp av google, google scholar, i Sidas publikationsdatabas, på Openaid.se, på biståndsdebatten.se/fuf.se samt via referenslistor i av författaren redan känd litteratur.

2. Urval och nerladdning av utvärderingar

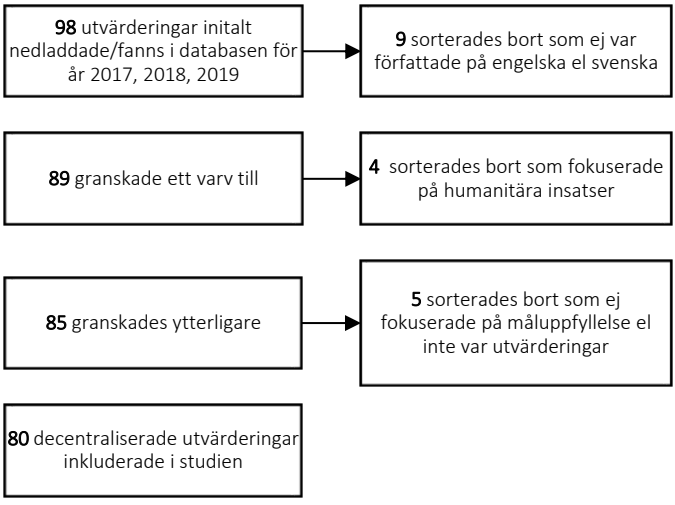
De decentraliserade utvärderingarna i urvalet laddades ner den 24/1-2020 från sidas publikationsdatabas och de strategiska utvärderingarna den 7/2-2020. Samtliga decentraliserade och strategiska utvärderingsrapporter som vid denna tidpunkt fanns i Sidas publikationsdatabas för aktuella år laddades ner. Totalt 98 decentraliserade utvärderingar för åren 2017 (35), 2018 (30) och 2019 (33) samt 8 strategiska utvärderingar för åren 2015 (2), 2016 (2), 2017 (1), 2018 (2) och 2019 (1).

Därefter sorterades rapporter bort som 1) ej var författade på svenska el engelska, 2) ej fokuserade på måluppfyllelse, som 3) fokuserade på humanitärt bistånd och 4) som inte visade sig vara utvärderingar.

Totalt återstod därefter 80 decentraliserade utvärderingar och 6 strategiska utvärderingar som grund för studien.

Figur B1:1 beskriver hur urvalet om 80 studier togs fram.

Figur B1:1: Urvalet decentraliserade utvärderingar, från 98 till 80



Källa: Författaren.

Av de 80 decentraliserade utvärderingarna hade 25 rapporter publicerats 2017, 29 2018 och 26 2019.

Bland de initalt 9 *strategiska utvärderingarna* valdes tre rapporter bort. Sida 2016:2 valdes bort för att den fokuserade på Sidas insatshantering snarare än insatserna i sig. Utvärderingen av Market Systems development approach (MSD) analyserade inte måluppfyllelse och har tagits bort. Årsredovisningen för 2017

nämner en utvärdering av samspelet mellan humanitärt och långsiktigt utvecklingssamarbete men som inte verkar ha publicerats och som inte heller ingår.[75]

Urval och representativitet

Det svenska biståndet består lite förenklat av fyra större områden: 1) det bilateral långsiktiga utvecklingssamarbetet (inklusive sk. Mult-bi), (2) det humanitära biståndet, 3) det multilaterala biståndet (inkl ex.vis EU-biståndet), samt 4) olika kostnader inom andra utgiftsområden men som ingår i den så kallade biståndsramen. De studier som ingår i vårt urval fokuserar endast på område 1 ovan och det långsiktiga utvecklingssamarbetet som hanteras av Sida. Övriga delar (2-4) fångas inte av de utvärderingar vi gått igenom.

Urvalet kan inte heller ses som representativt för Sidas långsiktiga utvecklingssamarbete som helhet under de år som studeras i rapporten. 2015- 2019 för strategiska utvärderingar och 2017- 2019 för decentraliserade utvärderingar. Sidas utvärderingsmanual fastslår att utvärderingar ska görs när det finns en tydlig användning och det kan var lämpligt att göra utvärderingar, exempelvis:

- ”För projekt eller program med okända eller omtvistade resultat.
- Vid stora och dyra insatser.
- I pilotinitiativ, innovativa projekt eller program.
- När Sida har ett strategiskt intresse.
- Där intressenterna är ängelägna om en utvärdering”

Tidigare studier (EBA 2015:02 och EBA 2017:12) har också visat att hälsoinsatser förhållandevis sällan utvärderas medan forskningsinsatser ganska ofta utvärderas liksom insatser i Öst- och centraleuropa. Tre stora områden i vårt urval är Nationell, regional eller lokal förvaltning (14/80 utvärderingar), Marknad, företagande,

[75] Rapporten ses enligt Sida som en intern produkt, tillgänglig men ej publicerad.

handel, innovation (12/80 utvärderingar), Forskning och högre utbildning (10/80 utvärderingar) samt Mänskliga rättigheter (8/80 utvärderingar). 2 utvärderingar fokuserar på jämställdhet och två utvärderingar på SRHR.

Geografiskt är Afrika (söder om Sahara) största region med 37/80 utvärderingar (46%) och näst största region är central- och östeuropa med 15/80 utvärderingar (19%).[76]

Sidas utvärderingsenhet menar att alla rapporter läggs upp på hemsidan. Det går dock inte att utesluta att det finns enskilda utvärderingar som av något skäl inte publicerats och som därmed missats i studien. Sida i Stockholm har inte alltid kännedom om ambassadernas upphandlingar och studier kan saknas av sekretessskäl.

I det decentraliserade utvärderingsurvalet har 9 utvärderingar sorterats bort för att de inte varit författade på engelska eller svenska. Det finns inget egentligt skäl att tro att dessa rapporter skulle avvika starkt med avseeende på tillförlitlighet eller måluppfyllelse-bedömning. Utvärderarna hämtas från samma konsultföretag, samma ramavtalsupphandling och flera av författarna återfinns även på de engelskspråkiga rapporterna i urvalet. Genomgången visar på små skillnader i slutsatser om måluppfyllelse mellan regioner.

Den sammantagna slutsatsen om representativitet är att även om utvärderingsurvalet inte ska ses som representativa för Sidas insatsportfölj som helhet så speglar urvalet på ett täckande sätt Sidas utvärderingsarbete (men ej humanitära insatser) för åren 2017 till 2019.

[76] Att Latinamerika endast har med 2 utvärderingar beror bara delvis på att spanska eller portugiska studier sorterats bort då endast 3 sådana publicerats under perioden som ingår här.

3. Framtagande och utprövning av analysramverket

Baserat på en inledande inläsning av litteratur (avrapporterad i kapitel 2 och 6 i huvudrapporten) togs ett analysramverk fram som styrt och gett systematik till genomgången (bilaga 3). Delar av ramverket (främst bakgrundvariablerna) utgör en utvecklad och förbättrad version av det ramverk som användes i EBA 2017:12.

Utgångspunkten för rapporten har varit OECD DACs tidigare definition av måluppfyllelse för biståndsutvärderingar (Effectiveness) (se kapitel 1 samt bilaga 4). OECD DACs nya definition från 2020 har inte använts då utvärderingarna i urvalet är beställda och genomförda innan förändringen. Den nya definitionen av effektivitetskriteriet diskuteras delvis i bilaga 4.

Ramverket testades och utvecklades först på ett slummässigt urval om 8 utvärderingar ur bruttourvalet och har sedan utvecklats successivt under arbetets gång.[77] Utgångspunkten var att åstadkomma enkla, tydliga frågor där det bedömts relevant och möjligt att söka svaren i rapporterna. En precisering av varje skalsteg har gjorts för måluppfyllelsefrågan.

4. Genomgång av utvärderingarna utifrån ramverket

Därefter har utvärderingarna gåtts igenom systematiskt utifrån analysramverket. Grunddata från genomgången har sammanställts och dokumenterats i en Exceldatabas som finns tillgänglig från EBA. Bedömningen replikerades för 80 slumpmässigt valda observationer brett över datasetet för kontroll av reproducerbarhet. I kapitel fyra diskuteras ett par faktorer som delvis kan ha sänkt reproducerbarheten vid syntes av måluppfyllelsefrågan: att konsulternas

[77] All randomisering baseras på siffror från www.Random.org

bedömningar ofta är otydliga, inlidande och sammansatta och att övergripande samlade bedömingar inte alltid görs. Ca. 10% (nio decentraliserade och en strategisk) av utvärderingarna har haft så otydliga slutsatser att bedömningarna inte kunnat skalstegs-bestämmas.

En slutsats efter EBA 2017:12 (se EBA 2021:1) var att syntesen inte blir fullt reproducerbar om man inte till frågor och svarsalternativ anger preciseringar av hur genomgången gjorts för varje enskild fråga och hur frågan specifikt uttolkats av den som gör genomgången. Ett flertal frågor har behövt preciseras ytterligare när olika tolkningar av frågan visat sig möjlig i mötet med texten. Graden av reproducerbarhet skiljer sig därmed mellan frågor i ramverket. Utvärderingarna är kvalitativt orienterade, inte sällan ganska otydliga texter. Studien kan därför inte jämföras med en systematisk översikt[78] med avseende på systematik och process för säkerställande av reproducerbarhet och validitet. Den viktigaste svagheten är att genomgången gjorts av en person och inte flera som är regel vid exempelvis systematisk översikt. Det är också centralt att svar på fråga 1 i genomgången (måluppfyllelsen) ses i ljuset av fråga 2 (tillförlitligheten).

Studien har bara analyserat publicerade utvärderingsrapporter. Om utvärderarna beskrivit sitt arbete i en ej bifogad sk. Inception Report så fångas inte det av analysen. Flera rapporter har saknat IR. Det bedöms spela mindre roll för analysens slutsatser om tillförlitlighet eftersom det är grundläggande kvalitetspraxis i utvärderingar att metodbeskrivningen finns att läsa i rapporten eller som bilaga.

[78] Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training.

5. Sammanställning och analys av data samt författande av rapport

Sista arbetssteget bestod i att med hjälp av Excel analysera och sammanställa data, samt att med genomgången av tidigare studier och utvärderingar som bakgrund författa rapporten.

Arbetet har följts av en referensgrupp som kommenterat på intitalt projektförslaget, på analysramverk samt rapportutkast. De faktum att en medlem i referensgruppen är anställd på Sidas utvärderingsenhet och att en referensgruppsmedlem arbetar som utvärderingskonsult åt Sida kan ha påverkat diskussionen i referensgruppen.

Det bör påpekas att de fem stegen (ovan) i praktiken gått in i varandra. Exempelvis bearbetades kapitel 1, 2 och 6 som utkast när genomgången av utvärderingarna pågick. Studien har på så sätt vuxit fram i samspel mellan läsning av tidigare studier på området och genomgången av utvärderingarna.

Bilaga 2 – Analysramverket

Nedan listas analysramverkets fråga med svarsalternativ, precisering, begränsningar och en grov skattning av konfidensen per fråga (konfidensen är: mycket god, god, måttlig, låg).[79]

1. Utvärderingens titel

Precisering och begränsningar: öppet svar, konfidens: mycket god

2. Utvärderingens nummer

Precisering och begränsningar: 1- 80, precisering: slumpmässig tilldelning baserad på slumpgenerator, konfidens: mycket god

3. Land:

Precisering och begränsningar: öppet svar, samtliga länder som studerats i utvärderingen anges, konfidens: mycket god

4. Region

Afrika SOS
Asien inkl. Centralasien
Central och Östeuropa
Latinamerika
MENA
Global

Precisering och begränsningar: baserat på fråga 3 ovan, global motsvarar mer än en region, konfidens: mycket god

[79] Graden av reproducerbarhet minskar troligt med konfidensnivån.

5. Utvärderingsobjektets geografiska fokus

Land alt subnationell nivå

Region (fler än ett land inom en region)
Global (flera regioner)

Precisering och begränsningar Baseras på fråga 3 ovan, global motsvarar mer än en region, region motsvarar fler än ett land inom en region, konfidens: mycket god

6. Tidsperiod som utvärderas

Precisering och begränsningar: öppet svar, antal hela år, ibland svårt att läsa ut vid flera insatser och gränsfall, avrundas till helår, konfidens: god

7.Analyseras flera insatser samtidigt?

En insats
Flera samtidigt

Precisering och begränsningar: antalet utvärderingsobjekt i utvärderingen (projekt, program etc.), oproblematiskt att läsa ut, konfidens: mycket god

8. Utvärderingsobjektets totala antal år med finansiering från Sida vid tidpunkt när utvärderingen genomförs

1 - 4 år
5 år eller mer
Det framgår inte i rapporten /går ej att läsa ut/ej relevant

Precisering och begränsningar: avser år med finansiering för verksamheten vid utvärderingstillfället oavsett projekt eller program, inte exempelvis specifika mål som kan ha förändrats över tid, se även fråga 6. Konfidens: god

9. Objekt för utvärderingen

Program
Projekt
Organisation
Strategi el plan
Garanti
Fond
Oklart och övriga

Precisering och begränsningar: följer helt den kategorisering/benämning av utvärderingsobjektet som utvärderaren använder i rapporten, konfidens: mycket god

10. Implementerande partner

Offentlig (exklusive multi)
Civilsamhälle
Privat sektor
Multilateral organisation, utvecklingsbank, mellanstatlig organisation
Gränsöverskridande samarbete (fler av ovan)

Precisering och begränsningar: organisationsform/-typ som driver eller genomför insatsen, organisationer som inte är offentliga, privata företag, utvecklingsbanker eller multilaterala organisationer har klassats som civilsamhälle, konfidens: mycket god

11. Vilka frågor ska besvaras under effectiveness-kriteriet enligt utvärderingens uppdragsbeskrivning?

Precisering och begränsningar: öppet svar klippt från rapport, samtliga utvärderingsfrågor i uppdragsbeskrivningen som berör kategorierna i fråga 13 plus resultat, måluppfyllelse och effekter, inslag av personlig bedömning avseende vad som hänger ihop med effektivitetskriteriet, konfidens: god

12. Efterfrågar uppdragsbeskrivningen explicit en måluppfyllelsebedömning som inte fokuserar på övergripande eller långsiktiga målsättningar?

Ja

Nej

Oklart/texten är motsägelsefull i detta avseende

Precisering och begränsningar: frågan avser endast utvärderingar som studerat en enskild insats. Avser endast det som explicit/uttryckligt sägs i uppdragsbeskrivningen. En begräsning kan ligga i den tvetydighet som begreppen outcome, Impact och effectiveness innebär. Mer specifikt formulerade uppdragsbeskrivningar kan också beskriva exakt vilka resultat som ska följas upp och har inte något skäl att explit ange målnivån. Språkbruket har ofta varit otydligt. Konfidens: god

13. Efterfrågar uppdragsbeskrivningen explicit en bedömning av insatsens måluppfyllelse, eller används andra begrepp vid formulering av utvärderingsfrågorna?

Ja (en bedömning av måluppfyllelse efterfrågas)

Nej (en bedömning av måluppfyllelse efterfrågas inte explicit, man trycker istället på förväntade resultat eller anger specifika resultat som ska ligga i fokus för bedömningen)

Precisering och begränsningar: avser vad som explicit/uttryckligt sägs i uppdragsbeskrivningen, gränsfall kan dock förekomma, konfidens: god

14. Vad efterfrågas i uppdragsbeskrivningen utöver svaret på grundfrågan (måluppfyllelse/uppnådda resultat)? (Flera svar är möjliga)

Analys avseende:

Specifika typer el nedbrytningar av resultat/måluppfyllelse
Mekanismer/förklaringar (till måluppfyllelse/resultat)
Programteori/ förändringsteori/ LFA
Insatsens resultat- och uppföljningsarbete eller indikatorer

Bieffekter, negativa eller oförutsedda effekter
Kan målen nås i nästa fas, framöver/i framtiden
Vad kan göras för att mål ska nås
Om målen eller förväntade resultat är tydligt formulerade
Har olika outputs eller resultat levererats i tid

Inget annat
Annat, vad:

Precisering och begränsningar: fokus ligger helt på de utvärderingfrågor som räknas upp i uppdragsbeskrivningen. Inslag av personlig bedömning avseende vilket svarsalternativ som ska användas Kategoriseringen borde dock ha innehållit fler svarsalternativ och det finns en liten risk att frågeställningar om måluppfyllelse missats för att de av beställarna kategoriserats under andra utvärderingskriterier, konfidens: god

15. Har insatsens mål/förväntade resultat förändrats eller justerats under den period som ligger i fokus enligt utvärderingen?

Ja, de har rekonstruerats, förtydligats eller strukturerats om av utvärderarna

Ja, men primärt av andra skäl (ex. vis vid övergång mellan projektperioder)

Nej, sannolikt inte

Utvärderingen fokuserar på fler än en insats (analyseras därför inte)

Inte möjligt att läsa ut ur rapporten/framgår inte

Precisering och begränsningar: delvis svårt att läsa ut ur rapporterna, delvis vagt språkbruk i rapporterna. Det är möjligt att målen preciserats eller förändrats utan att författarna nämner det. Konfidens: måttlig

16. Vilket tidsligt utvecklingssteg ligger som högst i fokus för måluppfyllelse/resultatanalysen*?*

Output
Outcome
Impact
Inte möjligt att läsa ut

Precisering och begränsningar: avser det utvärderarna själva anser/anger ligger i fokus på för analysen i utvärderingstexten. Inte nödvändigtvis vad som faktiskt analyserats. En begränsning här ligger i viss begreppsförvirring på området. Konfidens: God

17. Ska de förväntade långsiktiga resultaten från insatsen åstadkommas direkt i samhälle/ natur eller i processer eller verksamheter?

I processer, organisationer, verksamheter som ska stärkas**,** byggas el förändras (som därefter direkt/indirekt kan påverka människor/ samhälle/natur)

Externt i berörda samhällen (direkt mot medborgaren, rättighetsbärare, i natur eller ekosystem)

I både och

Ej relevant/kan ej läsas ut

Precisering och begränsningar: slutsatsen baseras här på insatsens angivna mål/förväntade resultat och inte på vad som i praktiken utvärderats. Viss tolkning samt gränsfall har dock förekommit. Konfidens: God

18. Vilken av följande kategori motsvarar insatsens övergripande långsiktiga mål? (flera alternativ är möjliga)

Förändra el förbättra offentliga institutioners funktions-/arbetssätt eller kapacitet inom en sektor

Stärka implementering el efterlevnad av mänskliga rättigheterna inom ett/ flera områden (tematiskt el geografiskt)

Stärkt kapacitet att bedriva forskning

Stärkt användning av forskningsresultat

Skapa mobilisering, deltagande el åstadkomma demokratiskt efterfrågetryck

Minskad fattigdom i region, land eller landsdel

Ekonomisk empowerment för särskild grupp

Konfliktlösning, fred, fredliga samhällen och säkerhet

Kunskapssprida, utbilda, kompetensutveckla, informera om område/fråga

Nya och bättre jobb/arbetstillfällen

Förbättrade levnadsförhållanden eller ökande inkomster

Stärkt handel

Minskat våld i olika former, sammanhang eller gentemot olika grupper

Precisering och begränsningar: syftet är att genom en generisk ansats illustera övergripande målnivåer. En annan författare hade ev. formulerat kategorierna delvis annorlunda. Endast kategorier med fler än tre förekomster redovisas. Konfidens: Måttlig

19. Diskuterar, kommenterar, värderar eller analyserar utvärderarna hur insatsens mål eller förväntande resultat har formulerats (exempelvis avseende ambition, specificeringsgrad, mätbarhet, realism, acceptans eller tidsperspektiv)?

Ja

Nej

Om ja, ge exempel på vad som sägs: … (Öppet svar, klipp in de viktigaste slutsatserna)

Precisering och begränsningar: Frågan berör formuleringen och målens form inte substans. Frågan fångar här endast explicita diskussioner om mål/förväntade resultat och inte logiska ramverk el dyl. Utvärderarna kan därför i några fall ha velat uttrycka en underförstådd synpunkt på målformuleringar men som inte fångats. Vissa utvärderingar fokuserar också på fler än en insats. Konfidens: Måttlig

20. I vilken utsträckning har mål eller förväntade resultat uppnåtts så här långt i insatsen (enligt utvärderarna)?

Skalsteg:	Kriterier
Fullt uppfyllda (mål/ resultat)	• Med undantag för något eller några enskilda, el marginella, mål/resultat är utvärderarens bedömning att samtliga mål/resultat är uppfyllda eller på väg att nås/kommer att nås, och/eller: • Insatsen beskrivs nästan utan undantag som effektiv.
I huvudsak uppfyllda (mål/ resultat)	• Utvärderaren bedömer att majoriteten mål/förväntade resultat (alternativt de mål som utvärderaren angett som viktigast) är uppfyllda eller på väg att nås/kommer att nås, och/eller • Insatsen beskrivs huvudsakligen/främst som effektiv

Skalsteg:	Kriterier
I huvudsak ej uppfyllda (mål /resultat)	• Utvärderaren bedömer att majoriteten mål/förväntade resultat (alternativ de mål som utvärderaren angett som viktigast) inte är uppfyllda eller är på väg att nås/kommer att nås, och/eller • Insatsen beskrivs i huvudsak som varande ineffektiv
Inte uppfyllda (mål /resultat)	• Endast något el några enstaka, eller för insatsen marginella, mål/resultat, bedöms av utvärderaren ha uppnåtts eller på väg att nås/kommer att nås, och/eller • Insatsen beskrivs nästan utan undantag som ineffektiv
Går ej att läsa ut i rapporten/slutsatser för otydliga för att kunna skalstegbestämmas	• Information enligt ovan kan ej läsas ut i rapporten • Slutsatserna är mycket otydliga • Alt Utvärderingen har inte gjort något bedömning (trots att man har det i uppdrag)

Precisering och begränsningar: Slutsatserna i utvärderingarna är ofta otydliga eller sammansatta och studierna konkluderar sällan tydligt. Konfidensen stärks delvis av kriterna för individuella skalsteg som utgör en fyrgradig skala motsvarande 25% vardera i fördelningen. Kriterna har styrt genomgången inte skalstegets beteckning. Reproducerbarheten sannolikt lägre vid gränsfall mellan skalsteg. Konfidens: Måttlig

21. Beskriver utvärderarna specifikt och utförligt hur måluppfyllelsefrågan besvarats? (%)

Ja

Nej

Precisering och begränsningar: Det är delvis en bedömningsfråga vad som är specifikt och tydligt. Fokus har legat på Inception report samt metodkapitlet med särskilt fokus på ev utvärderingsmatris. Utgångspunkter har varit att hellre fria än fälla i bedömningen av gränsfall. Konfidens: Måttlig

22. Vilken utvärderingsdesign hänvisar utvärderaren till vid måluppfyllelseanalysen? (%)

Contribution analysis
Theory based evaluation (generell)
Outcome harvesting
Outcome mapping
Most significant change
Process tracing
Realist evaluation
Systems based Evaluation
QCA (Qualitative Comparitative Analysis)
QUIP (The Qualitative Impact Assessment Protocol)
Regressionsbaserad analys (någon form)
Kvasiexperiment (någon form)
Randomiserad kontrollstudie
Kombination av två alternativ
Övriga/Oklart/inget

Precisering och begränsningar: Analysen gjordes deduktivt med sökfunktionen. Endast designalternativ som explicit och tydligt nämns i texten fångas. En svaghet är analysens förmåga att fånga övriga/andra designalternativ som inte listats. Konfidens: Mycket god

23. Prövar utvärderaren en preciserad förändringsteori mot empiriska data i måluppfyllelseanalysen?

Ja

Nej

Oklart/kan ej läsas ut

Precisering och begränsningar: Bedömningen fångar bara det som sägs explicit i rapportens metodkapitel och Inception report. Vissa rapporter har saknat Inception report. Vid gränsfall har vi valt ”hellre fria än fall”. Vissa rapporter har saknat Inception report. Konfidens: God

24. Beskriver utvärderarna hur de bedömer och försökt hanterna kausalitet samt potentiellt påverkande externa faktorer vid analysen av måluppfyllelse? (%)

Ja

Nej

Oklart/kan ej läsas ut

Precisering och begränsningar: Bedömningen fångar här bara vad som sägs explicit i rapportens metodkapitel samt Inception report. Vissa rapporter har saknat Inception report. Vid gränsfall har vi valt ”hellre fria än fall”. Konfidens: God

25. Förekommer begreppen kontrafakticitet, additionalitet eller kausalitet i utvärderingen?

Ja + vilken/vilka

Nej

Precisering och begränsningar: Analysen genomfördes med sökfunktion i Adobe baserat på de tre begreppens ordstam och olika ändelser. Konfidens: Mycket god

26. Tar utvärderingens metodkapitel upp några metodologiska begräsningar avseende specifikt bedömningen av effektivitetsskriteriet och/eller kausalitet?

Ja för måluppfyllelse och för kausalitet
Ja för måluppfyllelse men inte för kausalitet
Ja för kausalitet men inte för måluppfyllelse
Nej
Rapportens metodkapitel saknar avsnitt om begränsningar
Om ja, vad sägs:

Precisering och begränsningar: Frågan fokuserar på vad som sägs explicit under rubriken begränsningar (Limitations) i metodkapitlet. I princip samtliga utvärderingar har haft en sådan rubrik. Det har samtidigt inneburit tolkning att bedöma vilka begränsningar som berör måluppfyllelse. Reproducerbarheten i förhållande till kausalitetsfrågan bedöms vara något bättre än i förhållande måluppfyllelse generellt. Konfidens: Måttlig

27. Ska utvärderingen enligt TOR analyserar eller diskutera (negativa el positiva) oförutsedda resultat eller effekter?

Ja

Nej

Precisering och begränsningar: Frågan fokuserar endast på vad som explicit efterfrågas i rapportens uppdragsbeskrivning/Terms of Reference. Utvärderingen kan mycket väl ha fokuserat på oförutsedda effekter utan att detta efterfrågats där. Konfidens: Mycket god

Bilaga 3 – Analyserade utvärderingar

Decentraliserade utvärderingar (80 stycken)

2017

1. *Midterm Review of GESTERRA Capacity Building Programme on Land Management and Administration within DINAT*. Final Report. Ian Christoplos, Dale Doré, André Calengo, Svend Erik Sørensen. Sida Decentralised Evaluation 2016:15.
2. *Evaluation of Sida's Support to the Gorée Institute (2012-2015). Final Report*. Jérôme Gouzou, Olivia Lazard. Sida Decentralised Evaluation 2017:1.
3. *Evaluation of Sida's Support to Tostan (2010-2016). Final Report*. Kimiko Hibri Pedersen, Alassane Binta Diop, Helena Neves, Olivia Lazard, Jérôme Gouzou. Sida Decentralised Evaluation 2017:2.
4. *Evaluation of the Sida supported RFSL projects "LGBT Voices" and "Rainbow Leaders"*. Final Report. Annika Nilsson, Jessica Rothman. Sida Decentralised Evaluation 2017:4.
5. *Final Evaluation of the Natural Resource Management Facility At Act!* Final Report. Eric Buhl-Nielsen, Winnie Wangari Wairimu. Sida Decentralised Evaluation 2017:6.
6. *Evaluation of Sida Support to ECPAT International*. Final Report. Louise Mailloux, Agneta Gunnarsson. Sida Decentralised Evaluation 2017:5.

7. *Evaluation of Swedish government research cooperation with Eduardo Mondlane University, Mozambique 2011-2016*. Synthesis Report. Stein-Erik Kruse, Inge Tvedten, Matti Tedre, Carmeliza Soares da Costa Rosário. Sida Decentralised Evaluation 2017:9.
8. *Evaluation of the OECD's 'Sector Competitiveness Strategy' Project in Ukraine. Final Report*. Andrea Spear, Vera Devine, Oleh Myroshnichenko. Sida Decentralised Evaluation 2017:10.
9. *Mid-Term Evaluation of the National Employment Programme, Rwanda*. Final Report. Sarah Gray, Charles Twesigye-Bakwatsa, David Muganwa, Marcienne Umubyeyi. Sida Decentralised Evaluation 2017:11.
10. *Evaluation of the Regional Statistics Cooperation on the Western Balkans 2013 – 2016*. Final Report. Leif Danielsson, Sabina Ymeri, Pier Giorgio Ardeni. Sida Decentralised Evaluation 2017:12.
11. *Evaluation of the Sida-funded Partnership in Statistics: A cooperation project between Statistical Office of the Republic of Serbia (SORS) and Statistics Sweden (SCB)*. Final Report. Pier Giorgio Ardeni, Leif Danielsson, Sabina Ymeri. Sida Decentralised Evaluation 2017:13.
12. *Evaluation of Core Support to Georgian Foundation for Strategic and International Studies (GFSIS)*. Final Report. Pierre Walther, Nelly Dolidze. Sida Decentralised Evaluation 2017:15.
13. *Evaluation of the Health Economics and HIV and AIDS Research Division (HEARD)*. Final Report. Silke Graeser, Anne L. Buffardi, Lennie Bazira Kyomuhangi. Sida Decentralised Evaluation 2017:17.
14. *Completion Evaluation of Sida Support to Environment and Climate Change Component of NREP*. Final Report. Leif Danielsson, Denis Rugege, Richard Ngendahayo. Sida Decentralised Evaluation 2017:18.

15. *Review of the Statistical Cooperation project between the National Institute of Statistics of Albania and Statistics Sweden, Phase IV, October 2014 – December 2017*. Final Report. Leif Danielsson, Sabina Ymeri. Sida Decentralised Evaluation 2017:20.
16. *Evaluation of Swedish Support to the One UN in Albania for gender equality work 2012–2017*. Final Report. Jonas Lövkrona, Sabina Ymeri. Sida Decentralised Evaluation 2017:21.
17. *Mid Term Review of the Agadir Technical Unit and the Swedish International Development Agency, Sida funded project "Support Quality Infrastructure in Agadir Countries"*. Final Report. Ali Dastgeer, Bourhan Kreitem, Jens Andersson, Stefano Ciarli. Sida Decentralised Evaluation 2017:22.
18. *Evaluation of ITP 299 Strategies for Chemicals Management*. Final Report. Ali Dastgeer, Abigail Hansen, Adriana Jalba, Alicia Borges Månsson. Sida Decentralised Evaluation 2017:23.
19. *Evaluation of Sida's research cooperation with Bolivia, 2007–2016*. Synthesis Report. Ananda S. Millard, Matti Tedre, Erik W. Thulstrup, Manuel Ramiro Munoz, Pamela Velasco. Sida Decentralised Evaluation 2017:24.
20. *End of Strategy Evaluation of the Zanzibar Legal Services Centre. Final Report*. Greg Moran, Clarence Kipobota, Johanna Lindgren Garcia. Sida Decentralised Evaluation 2017:26.
21. *Evaluation of EuroMed Rights Promotion of Human Rights in the Middle East and North Africa*. Final Report. Jérôme Gouzou, Kimiko Hibri Pedersen, Kristoffer Engstrand. Sida Decentralised Evaluation 2017:27.
22. *Evaluation of three projects on transboundary water management in the Middle East and North Africa region. Final Report*. Nicklas Svensson, Ludo Prins, Aly Kerdany. Sida Decentralised Evaluation 2017:28.

23. *End-Term Evaluation of the Swedish Education Support to Zanzibar 2010–2017*. Final Report. Sheila Reed, Angela Arnott, Idrissa Yussuf Hamad, Kristeen Oberlander Chachage. Sida Decentralised Evaluation 2017:32.
24. *End-Term Evaluation of the Global Partnership for Education (2014–2016) in Zanzibar*. Final Report. Sheila Reed, Angela Arnott, Idrissa Yussuf Hamad, Kristeen Oberlander Chachage.
25. *Evaluation of 3rd Call off of civil society support through umbrella organisations 2013–2017*. Final Report. Kathleen Webb, Agneta Gunnarsson, Camille Pellerin, Abdulatif Idris, Bezawit Bekele. Sida Decentralised Evaluation 2017:34.

2018

1. *Evaluation of the Sida-USAID/DCA Guarantee to Zanaco*. Final Report. Sarah Gray Emma Sitambuli, Jens Albråten. Sida Decentralised Evaluation 2018:29.
2. *Review of Ratmalana/Moratuwa and Ja-Ela/Ekala Wastewater Disposal Project in Sri Lanka*. Final Report. Åke Nilsson, Arumugham P.K., Nalin Wikramanayake, Kusum Athukorala. Sida Decentralised Evaluation 2018:30.
3. *Evaluation of the National Legal Aid Clinic for Women's Access to Justice Programme in Zambia*. Final Report. Greg Moran, Dr. Lungowe Matakala, Lovisa Arlid, Johanna Lindgren Garcia. Sida Decentralised Evaluation 2018:1.
4. *Evaluation of the Sida supported research capacity and higher education development program in Rwanda, 2013–2017*. Final Report. Inge Tvedten, Arthur Byabagambi, Johanna Lindström, Matti Tedre. Sida Decentralised Evaluation 2018:3.

5. *Evaluation of CIVICUS – World Alliance for Citizen Participation*. Final Report. Bente Topsøe-Jensen, Ingrid Obery, Khilesh Chaturvedi, Raphaëlle Bisiaux. Sida Decentralised Evaluation 2018:4.
6. *Evaluation of Programme Work Methods of The Raoul Wallenberg Institute of Human Rights and Humanitarian Law*. Final Report. Nicklas Svensson, Lone Lindholt, Henrik Alffram. Sida Decentralised Evaluation 2018:5.
7. *Evaluation of Afrobarometer's Regional Programme 2011–2017*. Final Report. Ian Christoplos, Peter da Costa, Dena Lomofsky, Faten Aggad, Linnea Hincks, Raphaëlle Bisiaux. Sida Decentralised Evaluation 2018:6.
8. *Evaluation of the Sida supported programmes "CLACSO Knowledge for a Sustainable World 2013–2016" and "South-South Tricontinental 2013–2015"*. Final Report. Ian Christoplos, Francisco Sagasti, Mario Bazan. Sida Decentralised Evaluation 2018:7.
9. *Mid-term Evaluation of Swedish government funded Civil Society Support through the AGIR II Programme in Mozambique 2014–2020*. Final Report. Stein-Erik Kruse, Inge Tvedten, José Jamie Macuane. Sida Decentralised Evaluation 2018:8.
10. *The key role of ICTJ in the designing of Colombia's complex Transitional Justice System*. Final Report. Jocke Nyberg, Heidi Abuchaibe Abuchaibe, Daniela Martínez Pérez. Sida Decentralised Evaluation 2018:9.
11. *Mid Term Review of SymbioCity Kenya.The Sustainable Urban Development Programme in Kenya – 2015-2018*. Final Report. Leif Danielsson, Tim Greenhow, Lucy Njigua. Sida Decentralised Evaluation 2018:10.

12. *Evaluation of AAU's Core Programmes and projects under the Core Programme 2013–2017, with particular focus on Sida's institutional and program support 2013–2017*. Final Report. Adam Pain, Catherine Ngugi, Michael Cross, Jennifer Sesabo. Sida Decentralised Evaluation 2018:11.
13. *Evaluation of CODESRIA's programme cycle "Forty Years of Social Research and Knowledge Production: Consolidating Achievements and Reaching New Frontiers 2012–2016"*. Final Report. Adam Pain, Stephen Webber, Jennifer Sesabo, Jerome Gouzou. Sida Decentralised Evaluation 2018:12.
14. *Evaluation of Swedish International Training Programmes (ITP); Quality Infrastructure – Technical Barriers to Trade (304) and Sanitary/Phytosanitary (305) 2013–2017*. Final Report. Eric Buhl-Nielsen, Mimi Groenbech, Stefano Sedola, Gemunu Wijesena, Jonas Noren. Sida Decentralised Evaluation 2018:13.
15. *Evaluation of Swedish International Training Programmes (ITP); Private Sector Development (Private Sector Growth Strategies and Strategic Business Management) 2011–2016*. Final Report. Sida Decentralised Evaluation 2018:14.
16. *Evaluation of the African Organisation of English-speaking Supreme Audit Institutions (AFROSAI-E)*. Final Report. Greg Moran, Marion Baumgart dos Santos, Kjeld Elkjaer. Sida Decentralised Evaluation 2018:15.
17. *Evaluation of Swedish Civil Society Support in Albania*. Final Report. Sida Decentralised Evaluation 2018:16.
18. *Evaluation of the Pungwe Basin Transboundary Integrated Water Resources Management and Development Programme (PP2)*. Final Report. Svend Erik Sørensen,Troels Kolster, Wellington Dzvairo, Tjitske Leemans, Klas Sandström. Sida Decentralised Evaluation 2018:17.

19. *Evaluation of the Sida supported programme"International Science Programme 2014–2018".* Final Report. Adam Pain, Trish Silkin, Gonçalo Carneiro. Sida Decentralised Evaluation 2018:18.
20. *Mid-term Evaluation of Small-holder Agriculture Reform through Enterprise Development (SHARED) Project, iDE.* Final Report. Eric Buhl-Nielsen, Mimi Groenbech, Stephen Tembo, Emelie Pellby. Sida Decentralised Evaluation 2018:19.
21. *Mid-term Review of Diakonia's Human Rights and Democracy Programme in Cambodia 2017–2019.* Final Report. Annika Nilsson, Kimsrun Chhiv. Sida Decentralised Evaluation 2018:20.
22. *Mid-Term Evaluation of "Regional Economic Integration through the Adoption of Competition and Consumer Policies in the Middle East and North Africa (COMPAL GLOBAL-MENA)"* Final Report. Kimiko Hibri Pedersen, John Lawrance, Richard Moody. Sida Decentralised Evaluation 2018:21.
23. *Mid-term Review of the EMPOWER Private Sector Project.* Final Report. Sarah Gray, Ulf Färnsveden, Art Shala. Sida Decentralised Evaluation 2018:22.
24. *Evaluation of four NGO implemented programmes in Georgia.* Final Report. Pierre Walther, Vera Devine, Nino Partskhaladze, Mikheil Kukava. Sida Decentralised Evaluation 2018:23.
25. *The Evaluation of the Dairy Hub and Dairy Academy Development Project in Bangladesh.* Final Report. Bo Tengnäs, Shibaji Roy. Sida Decentralised Evaluation 2018:24.
26. *Mid-Term Review of the Project 'Promoting Agriculture, Climate and Trade Linkages in the East African Community 2 (PACT EAC 2)'* Final Report. Greg Moran, Flor E Healy, Julia Leiss. Sida Decentralised Evaluation 2018:25.
27. *Evaluation of Health Guarantee to Centenary Rural Development Bank in Uganda.* Final Report. Ananda S. Millard, Ulrika Enemark, Basil Kandyomunda. Sida Decentralised Evaluation 2018:26.

28. *Mid-Term Review of Sida's regional core support (2014–2019) to the Eastern African Grain Council promoting grain trade in the East African region "Strengthening Regional Grain Markets II"* Final Report. George Gray, Sarah Gray. Sida Decentralised Evaluation 2018:27.
29. *Evaluation of the trade policy training centre in Africa (trapca).* Final Report. Jonas Lövkrona, Mark Pearson. Sida Decentralised Evaluation 2018:28.

2019

1. *Evaluation of the Sida supported research capacity building programme* "International Foundation for Science". Final Report. Pierre Walther, Bassirou Bonfoh, Thi Van Ha Nguyen. Sida Decentralised Evaluation 2019:1.
2. *Evaluation of the Strategic Plan 2014–2018 of the International Lesbian, Gay, Bisexual, Trans and Intersex Association (ILGA)* Final Report. Jocke Nyberg, Julie Thaarup, Larissa Arroyo Navarrete, Anna Lidström. Sida Decentralised Evaluation 2019:2.
3. *Evaluation of Swedish Leadership for Sustainable Development.* Final Report. Cecilia M Ljungman, Jens Andersson, Jonas Norén, Emelie Pellby. Sida Decentralised Evaluation 2019:3.
4. *Evaluation of the support to UNICEF to strengthen Child Protection in Abkhazia 2015–2018*. Final Report. Sida Decentralised Evaluation 2019:4.
5. *Evaluation of the Global Partnership for Education (GPE) – Literacy and Numeracy Education Support (LANES) program in Tanzania (2014–2018).* Final Report. Criana Connal, Annelie Strath, Khalid Dihenga. Sida Decentralised Evaluation 2019:5.

6. *Evaluation of ILO Road to Jobs (R2J), Afghanistan.* Final Report. Francis Watkins, Kimiko Hibri Pedersen, Naweed Rahmani. Sida Decentralised Evaluation 2019:6.
7. *Mid Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and The Mali Finance for Food Security and Women Entrepreneurs (FFSWE) programme.* Final Report. Frans van Gerwen, Knud-Erik Rosenkrantz, Hammou Haidara, Nadia Masri-Pedersen. Sida Decentralised Evaluation 2019:10.
8. *Evaluation of UN Women Country Programme in Afghanistan.* Final Report. Francis Watkins, Sippi Azarbaijani-Moghaddam. Sida Decentralised Evaluation 2019:14.
9. *Evaluation of the Guttmacher Institute programme on the use of SRHR evidence to inform action (2015–2018).* Final Report. Annica Holmberg, Philip Wambua, Abena Acheampong. Sida Decentralised Evaluation 2019:15.
10. *Mid-term Review of the Strategy for Sweden's Development Cooperation with Ethiopia, 2016–2020.* Final Report. Svend Erik Sørensen, Kristina Mastroianni, Imeru Tamrat, Matilda Svedberg. Sida Decentralised Evaluation 2019:16.
11. *Evaluation of IBON International and the CSO Partnership for Development Effectiveness (CPDE) Project.* Final Report. Annika Nilsson, Catherine Ngugi, Ellen Girerd Barclay, Ica Fernandez, Jonas Norén, Niels Dabelstein. Sida Decentralised Evaluation 2019:17.
12. *Evaluation of ISET Policy Institute Promoting Georgia's Development through Independent Policy Analysis, Trainings and Civil Society Engagement.* Final Report. Pierre Walther, Nino Parthaladze. Sida Decentralised Evaluation 2019:18.

13. *Evaluation of the Capacity Building for Improvement of Land Administration and Procedures in Bosnia and Herzegovina (CILAP project)*. Final Report. Svend Erik Sörensen, Bozena Lipej, Catharina Schmitz. Sida Decentralised Evaluation 2019:19.
14. *Review of Scandinavian Consortium Technical Assistance to Strengthening the Institutional Capacity of the Mozambican National Statistical System (2008–2017)*. Review Report. Frans van Gerwen, Norah Teopista, Wamayi Madaya. Sida Decentralised Evaluation 2019:20.
15. *Mid-term Evaluation of Women in Financial Inclusion and Capabilities Project in Zambia*. Final Report. Bente Topsøe-Jensen, Bethel Nakaponda, Stephen Tempo, Frida Ericsson. Sida Decentralised Evaluation 2019:21.
16. *Evaluation of Confidence in Georgia*. Final Report. Ian Christoplos, Joseph Salukvadze, Thomas Dubois. Sida Decentralised Evaluation 2019:22.
17. *Evaluation of Sida's Support to the Project "Organic Trade and Value Chain Development in East Africa" 2014–2019*. Final Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:23.
18. *Organising parliamentarians against armed violence Evaluation of Parliamentary Forum on Small Arms and Light Weapons (PFSALW) 2014–2018*. Final Report. FCG Swedish Development AB. Sida Decentralised Evaluation 2019:24.
19. *Evaluation of the International Disability Alliance (IDA) 2015–2018*. Final Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:25.
20. *Evaluation of the Sida-supported development cooperation with the Kenya Revenue Authority*. Final Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:26.

21. *Evaluation of Twaweza Strategic Plan 2015–2018, Programme Support in Tanzania. Final Report*. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:27.
22. *Evaluation of the Zimbabwe Election Support Network Project 2016–2018*. Final Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:28.
23. *Midterm review of the Sida core support to KCSF*. Final Report. FCG Swedish Development AB. Sida Decentralised Evaluation 2019:29.
24. *Evaluation of the Improving Court Efficiency and Accountability of Judges and Prosecutors in BIH Phase 2*. Evaluation Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:31.
25. *Evaluation of the regional development programme "Towards a Non-Toxic Environment in South-East Asia", Phase II*. Final Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:32.
26. *End of Project Evaluation of the BBC Media Action Radio Waves and Tikambe projects in Zambia*. Final Report. NIRAS Sweden AB. Sida Decentralised Evaluation 2019:33.

Strategiska utvärderingar (6 stycken)

1. *Evaluation of Policy Dialogue as an Instrument - the case of Gender Equality*. Dana Peebles, Jonas Lövkrona, Nadine Jub. Sida Evaluation 2015:01.
2. *Support to Capacity Development – Identifying Good Practice in Swedish Development Cooperation*. Sida Evaluation Report for the Joint Scandinavian Evaluation of Support to Capacity Development. Gonçalo Carneiro, Kristina Boman, Birgitte Woel, Arne Nylund. Sida Evaluation 2015:02.

3. *Evaluation of Sida's use of guarantees for market development and poverty reduction*. Hans Slegtenhorst, Mart Nugteren, Alwin de Haas, Rien Strootman, Marie Heydenreich, Paulo Luswata, Nino Serdarevic, Anders Grettve and Bart Schaap, Sida Evaluation 2016:01.
4. *Evaluation of Sida's ITP approach for Capacity Development*, Ingela Ternström, Elizabeth Bamwine, Marina Buch Kristensen, Japhet Makongo, Björn Ternström, Paul Balogun.Sida Evaluation, 2017:01.
5. *Evaluation of Sida's Global Challenge Funds. Lessons From a Decade Long Journey*. David Smith, Matthew Kentridge, Martin Wright, Juliette Seibold, Katharine May and Theo Sands,Sida Evaluation 2018:01.
6. *Evaluation of Sida's Support to Peacebuilding in Conflict and Post Conflict Contexts*. Erik Bryld, Julian Brett, Nadia Masri-Pedersen, Cécile Collin, Sida Evaluation 2019:01. + Underlagsrapporter.

Bilaga 4 – Måluppfyllelsekriteriet

En av flera tolkningsfrågor om måluppfyllelsekriteriet i biståndsutvärdering berör relationen till de två OECD -kriterierna Impact och Sustainability. Tabell B4:1 nedan anger definitioner innan år 2020 (som är fokus för denna rapport).

Tabell B4:1: Definitioner av tre utvärderingskriterier

Effectiveness	Impact	Sustainability
"The extent to which a development intervention has achieved its objectives, taking their relative importance into account."	"Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended."	"The continuation or longevity of benefits from a development intervention after the cessation of development assistance."

Källa: OECD/DAC.

En granskning av dessa definitioner visar att kriterierna är överlappande. Min slutsats är att Impact och Sustainability till effektivitetskriteriets förutsedda resultaten lägger oförutsedda (postiva och negativa) resultat och bärkraft när biståndet dras tillbaka. Övriga dimensioner finns i effektivitetskriteriet**.** Tidsperspektivet är inte alltid särskiljande då målsättningar naturligt formuleras på både lång och kort sikt beroende på vad man förväntat sig att åstadkomma. Inom biståndet är målen ofta (inte alltid) långsiktiga.

Om måluppfyllelseanalyser enbart undersöker kortsiktiga eller mellanliggande resultat så är det (som regel) inte en utttömande eller helt tillförlitlig analys av måluppfyllelse som görs. Värdet för beslutsfattare av en sådan analys är begränsat då de inte självklart säger något om verksamhetens effektivitet. Det går då inte heller i nästa steg att göra en fullständig bedömning av insatsens kostnadseffektivitet.

Många biståndsinsatser har lång tidshoriont och naturligt, men i olika utsträckning, långsiktiga effektmål. Effektivitetsskriteriet kräver liksom Impact bedömningar av kausalitet. Analyser av resultatens hållbarhet efter att bistånd avslutas kräver analys av resultat (som kräver bedömning av kausalitet). Rapporterad måluppfyllelse som inte tar hänsyn till resultatet bärkraft har troligen ett delvis lägre värde.

Vedung (2012) beskriver relationen mellan måluppfyllelse och bieffekter på följande sätt:

> ”Att bieffektsmodellen ändå är logiskt baserad på mål framgår av uttrycket sidoeffekter. Sidoeffekter kan endast finnas i relation till huvudeffekter, vilka pekas ut av målen för insatsen.” (Ibid).

Dvs. effekter (Impact) preciseras som regel i form av mål. Utöver detta finns det endast oförutsedda resultat.

Hur bör man se på dessa överlapp? En risk är att de främjar en felaktig bild bland utvärderare och beställare där man antar att man kan bedöma måluppfyllelse (effektivitet) utan att beakta kausalitet eller bedöma mot övergripande överenskomna mål. Alternativt att man antar att effektivitets- och impact-kriterierna i utvärdering motsvarar genuint separata analysmoment eller att det, exempelvis, är lättare att utvärdera måluppfyllelse än effekter.

För studien innebär diskussionen att vi strikt utgått ifrån effektivitetskriteriets definition som det preciserats innan 2020 (se ovan). Dess fokus på om insatserna nått sina mål har i linje med detta dock gjort det viktigt att söka svar på grundfrågan (om måluppfyllelse) i rapporterna som helhet och inte bara i kapitel som explicit eller i rubriken hänvisar till effektivitetskriteriet. Med överlappande kriterier blir det alltså nödvändigt att också analysera vad som sägs vid bedömningar av exempelvis impact i rapporterna för att kunna bedöma måluppfyllelse. Jag har behövt gå igenom

utvärderingarna som helhet, både texter om effectiveness, texter om Impact och rapporten som helhet för att kunna se vad rapporterna samlat har att säga om måluppfyllelse.

Samtidigt har effektivitetskriteriet formulerats om från 2020.

Tabell B4:2: Ny och äldre definitioner av effektivitet

Tidigare effektivitetskriterium (använs i denna rapport)	Nytt effektivitetskriterium (sedan 2020)
"The extent to which a development intervention's objectives were achieved, or is expected to be achieved, taking into account their relative importance"	"The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups".

Källa: OECD/DAC.

Förändringen motiveras enligt följande:

> ""Effektivitet och effekt (Impact) är kompletterande kriterier som fokuserar på olika nivåer i resultatkedjan. Effektivitet tar hänsyn till att resultaten uppnås i förhållande till målen för en intervention, nämligen på produktions- och utfallsnivån medan effekterna fokuserar på högnivåresultat resultatnivåer, nämligen vad deklarerade högre nivåresultaten är och vad som bidrar till dessa. I allmänhet bör insatsansvariga och utvärderingar se till att en tydlig distinktion görs mellan de olika resultatnivåerna (dvs. input, output, resultat och effekt) och att det är tydligt vilka aspekter som kommer att utvärderas under varje kriterium. "

En svårighet med denna lösning är att effektivitetskriteriet då inte handlar om effektivitet. Effektivitet bedöms *per definition* i förhållande till det man avsett att uppnå med intsatsen (som motiverar att man finansierat den) och den förväntan återfinns sällan på nivån ouput eller outcome i svenska och internationella biståndsprojekt. Det nya effektivitetskriteriet tycks även gå på tvärs mot hur effektivitet traditionellt definierats, förståtts och tillämpast inom utvärderingsfältet (Eliasson (2006), Vedung (2012), Faugert och Sandberg (2012). Syftet med förändringen är enligt OECD DAC att komma bort ifrån överlapp gentemot Impact-kriteriet men det finns en risk att det nya kriteriet samtidigt gör utvärderingar mindre relevant för beslutfattare genom ett ökat rapporterande om kortsiktiga resultat som säger lite eller inget om instsens effektivitet.

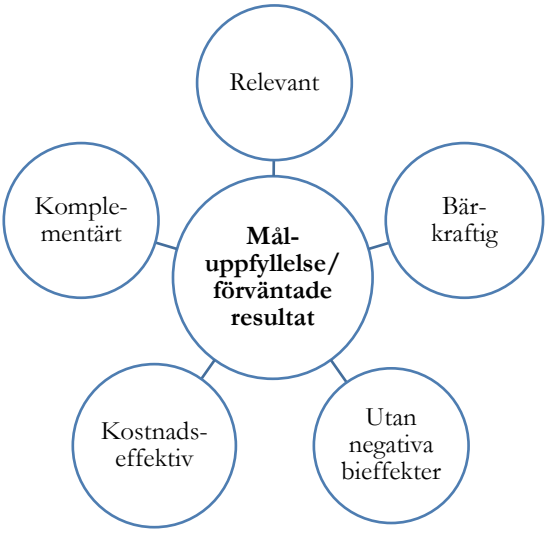
Ett bättre alternativ för framtiden hade eventuellt varit att övriga kriterier tolkats som kvalificeringar av måluppfyllelseanalysen/ bedömning mot förväntade resultat. En poäng med OECD/DAC kriteriernas isärhållande kvarstår om kriterierna tydliggör dimensioner eller kvaliteter som inte missas när utvärdering initieras.[80]

OECD/DACs övriga kriterier är också viktiga. Det är slöseri med skattepengar om biståndsinsatser inte når sina långsiktiga målsättningar men det är också slöseri om insatser (som når sina mål) inte är relevanta för fattiga och förtryckta (jmf Riddell 2007) eller som åstadkommits till en jämförelsevis hög kostnad. Sammantaget öppnar det för en tolkning av kritierierna där måluppfyllelse läggs i centrum för bedömningen (istället för att vara ett av fem kriterier) och som för helhetsperspektiv kompletteras med anayser av relevans, bärkraft, kostnadseffektivitet och avsaknad av negativa

[80] I Sidas tidigare utvärderingshandbok (Molund och Schill, 2007) menade man att överlappen är oproblematiska då de bidrar med olika dimensioner/ingångar och som är viktiga vid bedömningar av biståndsinsatser.

sidoeffekter.[81] Måluppfyllelsekriteriet är centralt även genom att kostnadseffektivitet och bärkraft (främst) bedöms i steget efter effektiviteten. På samma sätt behöver positiva el negativa sidoeffekter förstås i relation till ordinarie måluppfyllelse (Figur B4:1).

Figur B4:1: Måluppfyllelsekriteriet är centralt



Modellen återknyter till tanken att måluppfyllelse och förväntade resultat utgör kärnpunkt i utvärderingsuppdraget (Jmf Nachmias citerad i Vedung 2012). Men att insatsen *även* måste vara relevant, kostnadseffektiv, bärkraftig och inte leda till skadliga bieffekter i samhällen där biståndet verkar.

[81] På motsatt sätt skulle det nog inte vara lämpligt att lägga relevans eller koherens i centrum av analysen då det innebär en processfokusering av utvärderingsverksamheten.

Tidigare EBA-rapporter

2021:1, *Data Science Methods in Development Evaluation: Exploring the Potential*, Gustav Engström and Jonas Nóren

2020:07 *Effects of Swedish and International Democracy Aid*, Miguel Niño-Zarazúa, Rachel M. Gisselquist, Ana Horigoshi, Melissa Samarin and Kunal Sen

2020:06 *Sextortion: Corruption and Gender-Based Violence,* Åsa Eldén, Dolores Calvo, Elin Bjarnegård, Silje Lundgren and Sofia Jonsson

2020:05 *In Proper Organization we Trust – Trust in Interorganizational Aid relations,* Susanna Alexius and Janet Vähämäki

2020:04 *Institution Building in Practice: An Evaluation of Swedish Central Authorities' Reform Cooperation in the Western Balkans*, Richard Allen, Giorgio Ferrari, Krenar Loshi, Númi Östlund and Dejana Razić Ilić

2020:03 *Biståndets förvaltningskostnader För stora? Eller kanske för små*?, Daniel Tarschys

2020:02 *Evaluation of the Swedish Climate Change Initiative, 2009–2012,* Jane Burt, John Colvin, Mehjabeen Abidi Habib, Miriam Kugele, Mutizwa Mukute, Jessica Wilson

2020:01 *Mobilising Private Development Finance: Implications for Overall Aid Allocations*, Polly Meeks, Matthew Gouett and Samantha Attridge

2019:09 *Democracy in African Governance: Seeing and Doing it Differently,* Göran Hydén with assistance from Maria Buch Kristensen

2019:08 *Fishing Aid – Mapping and Synthesising Evidence in Support of SDG 14 Fisheries Targets,* Gonçalo Carneiro, Raphaëlle Bisiaux, Mary Frances Davidson, Tumi Tómasson with Jonas Bjärnstedt

2019:07 *Applying a Masculinities Lens to the Gendered Impacts of Social Safety Nets*, Meagan Dooley, Abby Fried, Ruti Levtov, Kate Doyle, Jeni Klugman and Gary Barker

2019:06 *Joint Nordic Organisational Assessment of the Nordic Development Fund (NDF)*, Stephen Spratt, Eilís Lawlor, Kris Prasada Rao and Mira Berger

2019:05 *Impact of Civil Society Anti-Discrimination Initiatives: A Rapid Review*, Rachel Marcus, Dhruva Mathur and Andrew Shepherd

2019:August *Migration and Development: the Role for Development Aid*, Robert E.B. Lucas (joint with the Migration Studies Delegation, Delmi, published as Delmi Research overview 2019:5)

2019:04 *Building on a Foundation Stone: the Long-Term Impacts of a Local Infrastructure and Governance Program in Cambodia*, Ariel BenYishay, Brad Parks, Rachel Trichler, Christian Baehr, Daniel Aboagye and Punwath Prum

2019:03 *Supporting State Building for Democratisation? A Study of 20 years of Swedish Democracy Aid to Cambodia,* Henny Andersen, Karl-Anders Larsson och Joakim Öjendal

2019:02 *Fit for Fragility? An Exploration of Risk Stakeholders and Systems Inside Sida*, Nilima Gulrajani and Linnea Mills

2019:01 *Skandaler, opinioner och anseende: Biståndet i ett medialiserat samhälle*, Maria Grafström och Karolina Windell

2018:10 *Nation Building in a Fractured Country: An Evaluation of Swedish Cooperation in Economic Development with Bosnia and Herzegovina 1995–2018*, Claes Lindahl, Julie Lindahl, Mikael Söderbäck and Tamara Ivankovic

2018:09 *Underfunded Appeals: Understanding the Consequences, Improving the System,* Sophia Swithern

2018:08 *Seeking Balanced Ownership in Changing Development Cooperation Relationships*, Nils Keizer, Stephan Klingebiel, Charlotte Örnemark, Fabian Scholtes

2018:07 *Putting Priority into Practice: Sida's Implementation of its Plan for Gender Integration*, Elin Bjarnegård, Fredrik Uggla

2018:06 *Swedish Aid in the Era of Shrinking Space – the Case of Turkey*, Åsa Eldén, Paul T. Levin

2018:05 *Who Makes the Decision on Swedish Aid Funding? An Overview*, Expertgruppen för Biståndsanalys

2018:04 *Budget Support, Poverty and Corruption: A Review of the Evidence*, Geske Dijkstra

2018:03 *How predictable is Swedish aid? A study of exchange rate volatility,* Númi Östlund

2018:02 *Building Bridges Between International Humanitarian and Development Responses to Forced Migration*, Alexander Kocks, Ruben Wedel, Hanne Roggemann, Helge Roxin (joint with the German Institute for Development Evaluation, DEval)

2018:01 *DFIs and Development Impact: an evaluation of Swedfund,* Stephen Spratt, Peter O'Flynn, Justin Flynn

2017:12 *Livslängd och livskraft: Vad säger utvärderingar om svenska biståndsinsatsers hållbarhet?* Expertgruppen för biståndsanalys

2017:11 *Sweden's Financing of UN Funds and Programmes: Analyzing the Past, Looking to the Future*, Stephen Browne, Nina Connelly, Thomas G. Weiss

2017:10 *Seven Steps to Evidence-Based Anticorruption: A Roadmap*, Alina Mungiu-Pippidi

2017:09 *Geospatial analysis of aid: A new approach to aid evaluation*, Ann-Sofie Isaksson

2017:08 *Research capacity in the new global development agenda*, Måns Fellesson

2017:07 *Research Aid Revisited – a historically grounded analysis of future prospects and policy options*, David Nilsson, Sverker Sörlin

2017:06 *Confronting the Contradiction – An exploration into the dual purpose of accountability and learning in aid evaluation*, Hilde Reinertsen, Kristian Bjørkdahl, Desmond McNeill

2017:05 *Local peacebuilding – challenges and opportunities*, Joakim Öjendal, Hanna Leonardsson, Martin Lundqvist

2017:04 *Enprocentmålet – en kritisk essä*, Lars Anell

2017:03 *Animal health in development – it's role for poverty reduction and human welfare,* Jonathan Rushton, Arvid Uggla, Ulf Magnusson

2017:02 *Do Anti-Discrimination Measures Reduce Poverty Among Marginalised Social Groups?* Rachel Marcus, Anna Mdee, Ella Page

2017:01 *Making Waves: Implications of the irregular migration and refugee situation on Official Development Assistance spending and practices in Europe*, Anna Knoll, Andrew Sherriff

2016:11 *Revitalising the policy for global development,* Per Molander

2016:10 *Swedish Development Cooperation with Tanzania – Has It Helped the Poor?* Mark McGillivray, David Carpenter, Oliver Morrissey, Julie Thaarup

2016:09 *Exploring Donorship – Internal Factors in Swedish Aid to Uganda,* Stein-Erik Kruse

2016:08, *Sustaining a development policy: results and responsibility for the Swedish policy for global development* Måns Fellesson, Lisa Román

2016:07 *Towards an Alternative Development Management Paradigm?* Cathy Shutt

2016:06 *Vem beslutar om svenska biståndsmedel? En översikt,* Expertgruppen för biståndsanalys

2016:05 *Pathways to change: Evaluating development interventions with Qualitative Comparative Analysis (QCA),* Barbara Befani

2016:04 *Swedish responsibility and the United Nations Sustainable Development Goals,* Magdalena Bexell, Kristina Jönsson

2016:03 *Capturing complexity and context: evaluating aid to education,* Joel Samoff, Jane Leer, Michelle Reddy

2016:02 *Education in developing countries what policies and programmes affect learning and time in school?* Amy Damon, Paul Glewwe, Suzanne Wisniewski, Bixuan Sun

2016:01 *Support to regional cooperation and integration in Africa – what works and why?* Fredrik Söderbaum, Therese Brolin

2015:09 *In search of double dividends from climate change interventions evidence from forest conservation and household energy transitions,* G. Köhlin, S.K. Pattanayak, E. Sills, E. Mattsson, M. Ostwald, A. Salas, D. Ternald

2015:08 *Business and human rights in development cooperation – has Sweden incorporated the UN guiding principles?* Rasmus Klocker Larsen, Sandra Atler

2015:07 *Making development work: the quality of government approach,* Bo Rothstein and Marcus Tannenberg

2015:06 *Now open for business: joint development initiatives between the private and public sectors in development cooperation,* Sara Johansson de Silva, Ari Kokko and Hanna Norberg

2015:05 *Has Sweden injected realism into public financial management reforms in partner countries?* Matt Andrews

2015:04 *Youth, entrepreneurship and development*, Kjetil Bjorvatn

2015:03 *Concentration difficulties? An analysis of Swedish aid proliferation*, Rune Jansen Hagen

2015:02 *Utvärdering av svenskt bistånd – en kartläggning*, Expertgruppen för biståndsanalys

2015:01 *Rethinking Civil Society and Support for Democracy*, Richard Youngs

2014:05 *Svenskt statligt internationellt bistånd i Sverige: en översikt*, Expertgruppen för biståndsanalys

2014:04 *The African Development Bank: ready to face the challenges of a changing Africa?* Christopher Humphrey

2014:03 *International party assistance – what do we know about the effects?* Lars Svåsand

2014:02 *Sweden´s development assistance for health – policy options to support the global health 2035 goals,* Gavin Yamey, Helen Saxenian, Robert Hecht, Jesper Sundewall and Dean Jamison

2014:01 *Randomized controlled trials: strengths, weaknesses and policy relevance,* Anders Olofsgård

Denna studie utgår från den övergripande frågan om det svenska biståndet når sina mål. I rapporten analyseras bedömningar av måluppfyllelse och effekter i Sidas decentraliserade utvärderingar under åren 2017–2019. I dessa rapporteras över lag god måluppfyllelse. Rapportförfattaren menar dock att analysen av orsakssamband kan stärkas betydligt, samt att övergripande mål med insatser bör bedömas i högre grad. Slutsatser, som självklart inte behöver vara felaktiga, brister annars i tillförlitlighet.

EBA
www.eba.se

Expertgruppen för biståndsanalys (EBA) är en statlig kommitté som oberoende analyserar och utvärderar svenskt internationellt bistånd.

The Expert Group for Aid Studies (EBA) is a government committee with a mandate to independently analyse and evaluate Swedish international development aid.