



KONINKRIJK BELGIË

Federale Overheidsdienst

Buitenlandse Zaken,

Buitenlandse Handel en

Ontwikkelingssamenwerking

Studie van de evalueerbaarheid van interventies van de Belgische Ontwikkelingssamenwerking



Inleiding

Deze studie situeert zich tegen de achtergrond van het toenemend belang van evaluaties in de ontwikkelingssamenwerking en de vaststelling dat tot nu toe de kwaliteit en het gebruik van evaluaties om verschillende redenen onder de verwachtingen zijn gebleven. De studie heeft in totaal 40 interventies in 4 landen (België, Benin, DRC en Rwanda) bekeken. Het ging om interventies in verschillende sectoren, uitgevoerd via een brede waaier van ontwikkelingsactoren met (mede-)financiering van de Belgische federale overheid.

Wat is evalueerbaarheid?

De studie heeft de OESO/DAC definitie gehanteerd die evalueerbaarheid definieert als *“de mate waarin een activiteit of project op een betrouwbare en geloofwaardige manier kan worden geëvalueerd”*. Evalueerbaarheidsbeoordeling wordt dikwijls in één adem genoemd met evalueerbaarheid en omschreven als *“een instrument dat toelaat te bepalen of een evaluatie aangewezen is in een bepaalde situatie”*. Een belangrijk onderscheid dat verder wordt gemaakt is dat tussen theoretische evalueerbaarheid en praktische evalueerbaarheid. Theoretische evalueerbaarheid verwijst naar evalueerbaarheid in principe, zoals die kan worden afgeleid uit het ontwerp van de interventie, zonder de praktijk in aanmerking te nemen. Praktische evalueerbaarheid neemt die praktijk wel in aanmerking en zal bijvoorbeeld nagaan of de gegevens over de voortgang van een interventie daadwerkelijk worden verzameld en of M&E systemen daadwerkelijk worden gebruikt, zodat die een bijdrage aan de verbetering van het beheer en de resultaten mogelijk maken.

Waarom is evalueerbaarheid belangrijk?

De bepaling van de evalueerbaarheid van een interventie dient niet om de (ontwikkelings-) waarde van die interventie te bepalen. Het is perfect mogelijk dat erg waardevolle interventies moeilijk evalueerbaar zijn; er bestaat met andere woorden geen verband tussen de mate van evalueerbaarheid en de (ontwikkelings-) waarde van een interventie. Een Evalueerbaarheidsmeting kan er evenwel voor zorgen dat de kwaliteit van evaluaties verbetert en evaluaties een grotere meerwaarde opleveren in verhouding tot hun kostprijs en ervoor zorgen dat de evaluatieresultaten effectiever worden gebruikt. Een goede evalueerbaarheidsbeoordeling kan daarbij uitgevoerd worden aan een kost die slechts een fractie vertegenwoordigt van het totale evaluatiebudget, maar verhoudingsgewijs een groot verschil kan maken. Verder is het niet zo dat een evalueerbaarheidsbeoordeling enkel nuttig is in de voorbereiding van een evaluatie en kan zulke beoordeling in alle fasen van een interventie bijdragen tot een beter beheer.

Nut en belang van een beoordeling van de evalueerbaarheid:

Fase interventie-cyclus	Doel van de beoordeling	Focus van de beoordeling	Beoogd resultaat
Formulering van de interventie	Verbetering van het interventieontwerp	Kwaliteit van het interventieontwerp	Verbeteringen en aanvullingen aan het interventieontwerp
Opstart van de interventie	Input leveren voor uitbouw M&E systeem	Beschikbaarheid en kwaliteit van de informatie en data-verzameling	Verbetering van het M&E systeem (inhoud en proces)
Uitvoering en afronding	<ul style="list-style-type: none"> Beslissen of evaluatie nu of later plaatsvindt Nagaan of interventie goed is aangepast aan ontwikkelingen 	<ul style="list-style-type: none"> Beschikbaarheid, actualiteit en kwaliteit gegevens Omgeving 	<ul style="list-style-type: none"> Evaluatie mogelijk op dit ogenblik? Alternatieven qua timing en inhoud? Bijsturing interventie ontwerp en M&E systeem
	Input leveren voor geplande evaluatie		<ul style="list-style-type: none"> Inzicht in evaluatiedoelen, hoofdvragen en aanpak
Post-interventie	Evaluatie mogelijk op betrouwbare en geloofwaardige wijze	<ul style="list-style-type: none"> Beschikbaarheid en kwaliteit van gegevens Haalbaarheid van contactname met betrokken partijen 	<ul style="list-style-type: none"> Inzicht in moeilijkheidsgraad en wenselijkheid van uitvoering evaluatie Mogelijk om impact en duurzaamheid te evalueren?

Studiekader

Uitgaande van de studiedoelstellingen is getracht om een realistisch studiekader te ontwerpen, waarmee per interventie in totaal in vier dagen, inclusief het veldbezoek, de evalueerbaarheid kon bepaald worden.

Het kader bestaat uit drie delen :

- de analyse van het interventie-ontwerp (inclusief de onderliggende veranderingstheorie) waarmee vooral (maar niet uitsluitend) de theoretische evalueerbaarheid in rekening wordt gebracht. De analyse van het interventie-ontwerp (in strikte zin) werd aangewend om ook de kwaliteit van de geplande M&E systeem te beoordelen;
- de analyse van de praktijk m.b.t. de uitvoering en beheer van de interventie en de context (inclusief gebruik van menselijke en financiële middelen);

- de analyse van contextuele factoren; deze factoren kunnen een rol spelen zowel op het vlak van het interventie-ontwerp als bij de dataverzameling, als bij de (toekomstige) uitvoering van de evaluatie zelf.

DAC Evaluatiecriteria	Relevantie	Effectiviteit	Efficiëntie	Impact	Duurzaamheid
Dimensies/Componenten	Aantal items dat werd gescoord				
1. Analyse van het interventie-design (7)					
1.1 Onderliggende analyse	(7)	(4)	(0)	(0)	(0)
1.2 De interventielogica en veranderingstheorie (8)	(1)	(6)	(5)	(5)	(3)
1.3 Het voorgestelde M&E systeem (9)	(5)	(7)	(9)	(7)	(7)
1.4 De consistentie en adaptatie van de interventielogica en veranderingstheorie (3)	(3)	(3)	(3)	(3)	(3)
2. Praktijk mbt. uitvoering en beheer van de interventie en van de context					
2.1 Basisinformatie mbt. de interventie-uitvoering (11)	(5)	(10)	(11)	(10)	(8)
2.2 Het M&E systeem in de praktijk (12)	(11)	(12)	(12)	(12)	(12)
3. De evaluatie-context					
3.1 Houding van de sleutelactoren (9)	(9)	(9)	(9)	(9)	(9)
3.2 De bredere context (3)	(3)	(3)	(3)	(3)	(3)
3.3 Praktische elementen (2) (°)					
Geaggregeerde score	(44)	(54)	(52)	(49)	(45)

Weinig gekend en gebruikt

De studie heeft vastgesteld dat de concepten “evalueerbaarheid” en “evalueerbaarheidsbeoordeling” tot nu toe weinig gekend en gebruikt worden in de Belgische ontwikkelings-samenwerking. Elementen van een evaluateerbaarheidstoets worden hier en daar wel in de praktijk gebracht zonder ze evenwel als zodanig te benoemen; verder is er nergens sprake van een systematische toepassing. Uit de analyse bleek verder dat interventies vanuit de optiek van evaluateerbaarheid in het algemeen een aantal sterke punten hebben, maar ook nog een groot aantal werkpunten als ze hun evaluateerbaarheid willen verbeteren.

Het M&E systeem scoort zwak in de ontwerpfase, de uitvoering scoort beter

De onderdelen van het interventie-beheersysteem die zwak scoren zijn 'het voorgestelde M&E systeem', 'de kwaliteit van de basisinformatie m.b.t. de uitvoering' en 'de kwaliteit van de interventielogica en veranderingstheorie'. Sterke aspecten zijn onder meer de 'de houding van de sleutelactoren t.a.v. evaluatie' en 'de invloed van de (politieke) en institutionele context'. Verder blijken onvolkomenheden in het interventieontwerp in vele gevallen door te werken tijdens de uitvoering en daardoor de evalueerbaarheid te beïnvloeden. Een goede ontwerpfase is in dat opzicht veelal de 'voorbode' van een kwalitatief goed interventiebeheer.

Er blijkt voorts een verschil te bestaan in de gevolgen van een zwakke ontwerpfase voor monitoring enerzijds en evaluatie anderzijds. Het is mogelijk om op het (operationele) niveau van monitoring relatief eenvoudig correcties aan te brengen, zoals blijkt uit het feit dat de kwaliteit van het M&E systeem in de praktijk aanmerkelijk hoger ligt dan de kwaliteit van het voorgestelde M&E systeem, al blijven lacunes uit de beginfase doorwerken in de uitvoering. Lacunes in het interventieontwerp hebben echter zwaardere gevolgen voor de evaluatiefunctie omdat correcties daar moeilijker zijn aan te brengen en impliceren dat bepaalde belangrijke aspecten van een interventie (zoals bv het effectief bereiken van initiële doelgroepen) niet of bijna niet geëvalueerd kunnen worden.

Item	Relevantie	Effectiviteit	Efficiëntie	Impact	Duurzaamheid	Evalueerbaarheidsindex
Het voorgestelde M&E systeem	2,45	2,83	2,85	1,95	2,46	2,51
M&E systeem in de praktijk	3,03	3,30	3,50	2,53	2,85	3,04

'scores van 1 tot 5'

Focus op effectiviteit en efficiëntie.

De globale evalueerbaarheidsscores per (OESO/DAC) evaluatiecriterium geven lagere scores voor duurzaamheid en vooral impact. Het feit dat hetzelfde score-patroon bijna steeds terugkomt, heeft uiteraard te maken met de verschillende moeilijkheidsgraad - in alle fasen van de interventiecyclus - van de evaluatie van deze criteria: duurzaamheid en vooral impact zijn moeilijker te evalueren zijn dan de drie andere criteria. Impact-evaluatie stelt hoge methodologische (en ook financiële) eisen terwijl de moeilijkheid bij de evaluatie van duurzaamheid meestal te maken heeft met de uitdaging om gefundeerde uitspraken te doen m.b.t. een situatie die zich pas in de toekomst zal voordoen. Anderzijds scoren effectiviteit en vooral efficiëntie aanmerkelijk hoger, wat een indicatie vormt van de kwaliteit van het interventiebeheer, met name wat betreft monitoring en evaluatie. In dit kader is het ook belangrijk de invloed van de financierende overheid te onderkennen: DGD is in eerste instantie bekommerd om een correcte aanwending van de ter beschikking gestelde publieke fondsen en heeft daartoe belangrijke voorwaarden (via procedures, formats, ...) opgelegd voor het beheer van de gefinancierde interventies.

Item	Relevantie	Effectiviteit	Efficiëntie	Impact	Duurzaamheid	Evalueerbaarheidsindex
Dimensie 1 (interventieplan)	3,11	3,22	3,29	2,52	3,15	3,16
Dimensie 2 (uitvoeringspraktijk)	2,99	3,25	3,51	2,33	2,73	2,96
Dimensie 3 (context)	3,89	3,98	3,98	3,84	3,89	3,91
Globale score evalueerbaarheid	3,26	3,38	3,52	2,70	3,13	3,23

Wisselende kwaliteit tussen de verschillende ontwikkelingsactoren

De comparatieve analyse van de evalueerbaarheid per land en interventietype (complexe versus minder complexe veranderingstheorie) leverde geen significante verschillen op. Verder valt op dat er op het vlak van kwaliteit van beheer nog grote verschillen bestaan onder de Belgische actoren en types interventies, ondanks een sturing van de financierende overheid. De evalueerbaarheidsscores per financieringskanaal (type Belgische actor) geven grotere verschillen aan met name tussen 'bilateraal/NGOs/ vakbonden' aan de ene kant en de 'andere actoren' aan de andere kant (betere scores voor de eerste groep), al zijn er ook in deze laatste groep voorbeelden van *good practices*.

Voornaamste aanbevelingen

De aanbevelingen hierna kunnen pas goed worden toegepast als de beheerslast voor de betrokken actoren op een aantal vlakken kan worden verminderd door - overeenkomstig de visie van de Strategienota Ontwikkelingsresultaten - een *andere invulling* te geven aan de eisen m.b.t. de uitwerking van interventievoorstellen, uitvoeringsrapportage, e.d.m.. De focus moet daarbij verlegd worden naar de ontwikkelingsresultaten (outcomes, impact) ten koste van de operationele niveaus (middelen, activiteiten, outputs).

De voornaamste aanbevelingen betreffen:

1. **De systematische integratie, door alle actoren in hun werking, van evalueerbaarheid en evalueerbaarheidsbeoordeling** waarbij deze als een middel voor een meer performante ontwikkelingssamenwerking worden aanzien en niet als een hefboom voor bureaucratische controle of sturing door donor en/of binnen organisatie).
2. **De invoering van een coherente evalueerbaarheidsbeoordeling als belangrijk hulpmiddel bij een ex ante analyse van elke evaluatie** om de potentiële baten van de evaluatie te analyseren en aan te tonen, en om op die wijze tot een gefundeerde beslissing te komen m.b.t. het al of niet laten uitvoeren van een evaluatie.
3. **Een (geleidelijke) verbetering van het voorbereidingstraject van interventies**, waarbij meer moet worden ingezet op kwaliteit en minder op routine, wat DGD via een aangepast kader (met incentives en via een vermindering van de administratieve eisen en regelgeving) kan ondersteunen.
4. **Een hernieuwde aandacht en opwaardering van de outcome en impact niveaus doorheen de gehele interventiecyclus** (interventieontwerp, M&E, ...) via een duidelijke definiëring van deze basisnoties en hoe ze worden ingevuld in interventie-voorstellen en -rapporten.
5. **Een verdere ontwikkeling van de dikwijls reeds goed uitgewerkte M&E systemen en praktijken met de ambitie een goede evalueerbaarheid van effectiviteit, impact en duurzaamheid te bereiken.** Zulke ontwikkeling gebeurt best op een geleidelijke wijze, waarbij middelen, instrumenten, capaciteit en ervaring overhands toenemen. In dit verband is het ook belangrijk dat een kader wordt gecreëerd waarbinnen deze (ambitieuze) veranderingen niet alleen worden gefaciliteerd, maar ook gestimuleerd en positief gewaardeerd. DGD speelt in dit verband een cruciale rol en zou in samenspraak met de andere sleutelactoren (a) de huidige regelgeving, instrumenten en procedures verder kunnen herzien en vereenvoudigen zodanig dat deze meer gericht zijn op de (beoogde) ontwikkelingseffecten, (b) incentives ontwikkelen om de M&E functie (en dan vooral de evaluatiefunctie) verder te ontwikkelen, (c) een fonds oprichten ter financiering van studies en evaluaties op effect- en impactniveau waarvan het initiatief bij voorkeur uitgaat van het geheel van Belgische ontwikkelingsactoren.
6. **De geplande certificering van M&E systemen wordt best niet losgekoppeld van een bredere en meer geïntegreerde benadering, zoals bv. de screening van de NGAs.** Eerder dan een formele en gestandaardiseerde certificering lijkt het aangewezen dat vanuit DGD goede incentives worden ontwikkeld om de kwaliteit van het beheer van interventies en meer specifiek M&E systemen te verhogen.



KONINKRIJK BELGIË

Federale Overheidsdienst

Buitenlandse Zaken,

Buitenlandse Handel en

Ontwikkelingssamenwerking

Dienst Bijzondere Evaluatie van de Belgische Ontwikkelingssamenwerking