

Nota DGD over evaluatie (NGSOC 28 maart 2019)

Evaluatie in het KB van 11 september 2016 :

Deze nota handelt over het gebruik van de terminologie rond evaluatie, mede naar aanleiding van de verwarring die er zou kunnen ontstaan omwille van het door de Dienst Bijzondere Evaluatie gehanteerde onderscheid tussen de termen “interne” en “externe” evaluatie.

Daarnaast wordt er ook verduidelijkt welke de verwachtingen zijn ten aanzien van de verschillende evaluaties die de actoren moeten (laten) uitvoeren volgens het KB.

1. Terminologie en types evaluatie

Definitie van interne en externe evaluatie (Wet OS, KB DBE)

Interne evaluaties worden volgens die opvatting aangestuurd door de actoren van de Belgische ontwikkelingssamenwerking, onafhankelijk van het feit of die uitbesteed worden aan een extern bureau of een externe expert. **Externe evaluaties** worden dan weer gevoerd door de Dienst Bijzondere Evaluatie en maken deel uit van diens programma.

Definitie van onafhankelijke evaluatie en externe evaluator (KB 11/11/16)

In het KB van 11 september 2016 wordt gesproken van **onafhankelijke evaluatie** en van het uitbesteden van evaluatie aan **externe evaluatoren**. Om verwarring te vermijden over wat hiermee bedoeld wordt, worden hieronder de termen verduidelijkt.

Onafhankelijke evaluatie (art. 44, §1):

Een onafhankelijke evaluatie houdt in dat de evaluator kan werken zonder dat die onder invloed staat van de personen betrokken bij de planning of uitvoering van de interventies of acties die het voorwerp uitmaken van de evaluatie. Deze evaluator is iemand die zelf ook geen belangen heeft bij de interventie(s) en dus niet betrokken geweest is in de formulering bvb.

Externe evaluator (art. 44, §2)

Een externe evaluator staat buiten de organisatie en heeft geen organisatorische band met de actor waarvan hij of zij interventies of acties evalueert. Deze evaluator is iemand die zelf ook geen belangen heeft bij de interventie(s) en dus niet betrokken geweest is in de formulering, planning of uitvoering van de interventies of acties die het voorwerp uitmaken van de evaluatie. Voor het aanstellen van een externe evaluator wordt er een beroep gedaan op de markt.

Conclusie

De begrippen van onafhankelijke evaluatie en externe evaluator zijn dus niet noodzakelijk in contradictie met het onderscheid dat gemaakt wordt tussen interne evaluatie en externe evaluatie. Om het parallel circuleren van een verschillende interpretatie van dezelfde termen te vermijden, is het dus misschien interessanter om in het kader van de evaluaties van de programma's te spreken van interne en externe *evaluatoren*.

Dat leidt tot de volgende situatie:

Elke gesubsidieerde organisatie realiseert onafhankelijke evaluaties voor de opvolging van de interventies, om lessen te trekken en om verantwoording af te leggen. Die onafhankelijke evaluaties worden toevertrouwd aan een externe of een interne evaluator volgens de volgende onderverdeling:

- Externe evaluator: mid-term evaluatie van programma's van 5 jaar en eindevaluaties van alle programma's.
- Interne evaluator: mid-term evaluatie van programma's van minder dan 5 jaar, hoewel het hier ook mogelijk is om een externe evaluator aan te duiden (cf. definitie hierboven).

2. Doelstelling en timing van de verschillende evaluaties

Elke actor zal jaarlijks een evaluatieplanning overmaken aan DGD, die ook besproken kan worden tijdens de institutionele dialoog. Deze planning kan aan het begin van het programma meer algemeen en indicatief zijn, en preciezer voor die jaren waarin de evaluatie(s) aan de orde is (zijn). Zoals in het KB vermeld wordt, worden ook de referentietermen van de evaluaties gecommuniceerd aan DGD en heeft de administratie de mogelijkheid om er evaluatievragen aan toe te voegen.

Het KB heeft het over twee soorten evaluaties: mid-term en eindevaluaties. Daarbij wordt ook een onderscheid gemaakt tussen evaluaties binnen een vijfjarenprogramma en die voor subsidiedossiers met een kortere looptijd. Dat wordt hieronder toegelicht, net zoals er aangegeven wordt welke timing idealiter gevolgd wordt, en wat er verwacht wordt van deze evaluaties (welk doel moeten ze dienen).

Mid-term en eindevaluaties: externe of interne evaluator?

De volgende evaluaties worden uitgevoerd door een externe evaluator:

- Eindevaluaties van alle subsidiedossiers
- Mid-term evaluaties van subsidiedossiers van 5 jaar

Er moet een mid-term evaluatie uitgevoerd worden voor alle interventies die langer duren dan 3 jaar. Dat betekent dat voor interventies van 3 jaar of meer, maar die korter duren dan 5 jaar, de mid-term evaluatie door een interne evaluator uitgevoerd kan worden.

Elke interventie moet echter sowieso een eindevaluatie krijgen, die dus telkens uitgevoerd wordt door een externe evaluator zoals gedefinieerd hierboven.

Voor interventies (= specifieke doelstellingen) waaraan meerdere actoren samenwerken, lijken gemeenschappelijke evaluaties aangewezen.

Timing

De timing van de mid-term evaluaties is flexibel en kan dus aangepast worden in functie van de finaliteit ervan (cf. volgend punt: bijsturen, voorbereiden volgend subsidiedossier, collectief leren, ...). Dit kan besproken worden tijdens de institutionele of strategische dialoog.

De eindevaluatie heeft plaats aan het einde van de interventie. Organisaties kunnen starten met de eindevaluatie op 1/05/2021 mits de activiteiten ver genoeg gevorderd zijn om hierover een uitspraak te doen.

Doel / scope van de evaluaties?

Mid-term evaluaties moeten niet het geheel van het programma afdekken, eindevaluaties wel. Een mid-term evaluatie mag zich dus richten op een deelaspect van het programma, zoals bvb een land, een thema, of gewoon één interventie (specifieke doelstelling). Op die manier kan de mid-term evaluatie aangewend worden om de interventie bij te sturen, maar bijvoorbeeld ook om een volgend subsidiedossier voor te bereiden. Ook gezamenlijke evaluaties die het collectief leerproces dienen, kunnen passen in de optiek van een mid-term evaluatie.

Een eindevaluatie apprecieert voor alle outcomes van een subsidiedossier of die bereikt zijn of niet. In die zin is de primaire doelstelling van deze eindevaluatie om verantwoording af te leggen over de resultaten die behaald werden.

Het is echter niet onbelangrijk om op te merken dat door de (late) timing van de eindevaluatie (die plaatsvindt op het einde van de interventie), de mid-term evaluatie een belangrijke dubbele rol gaat spelen. Ze kan immers dienen om bij te sturen, maar zal tegelijkertijd ook een rol spelen bij de opstelling en beoordeling van een nieuw subsidiedossier.

3. Publicatie van evaluaties

Conform het KB worden de evaluaties gepubliceerd binnen 3 maanden na ontvangst van het eindrapport. Dit gebeurt minimum langs de website van de gesubsidieerde organisatie, waar ook de management response van de evaluatie gepubliceerd kan worden.

De gesubsidieerde organisatie kan beslissen om het hele evaluatierapport of delen ervan niet te publiceren. In dat geval moet er transparantie zijn over de redenen waarom het rapport niet publiek gemaakt wordt. De publicatie van het eindrapport en de bijhorende management response is dus niet verplicht, maar transparantie erover wel.

4. Uitvoering (eind)evaluaties in de praktijk

Art 44 § 1: De gesubsidieerde organisatie voert onafhankelijke evaluaties om haar interventies op te volgen, er lessen uit te trekken en rekenschap af te leggen. Ze besteedt hiervoor minstens 1% van haar directe kosten.

Een eindevaluatie oordeelt over het halen van alle outcomes op het einde van de interventie.

...

Art. 44 § 2:

...

De eindevaluaties, alsmede de tussentijdse evaluaties voor de subsidiedossiers met een looptijd van 5 jaar, worden uitgevoerd door externe evaluatoren.

De centrale vereisten voor eindevaluaties worden beschreven in artikel 44 § 1 en 44 § 2. In artikel 44 § 1 werd vastgelegd dat eindevaluaties dienen te 'oordelen over het halen van alle outcomes'. Artikel 44 § 2 voegt daaraan toe dat die eindevaluaties uitgevoerd dienen te worden door externe evaluatoren.

De interpretatie van deze vereisten dient in lijn te zijn met de drie doelstellingen van evaluaties zoals beschreven in artikel 44 § 1. Het uitvoeren van onafhankelijke evaluaties van gesubsidieerde interventies heeft volgende doelstellingen:

- *Accountability*: rekenschap afleggen t.a.v. de subsidiërende overheid/administratie (DGD)
- *Interventies opvolgen en lessen trekken*: Op basis van evaluaties kunnen organisaties hun programma's bijstellen en versterken om maximale impact te bereiken.

Opdat het budget dat CMO's en IA's besteden aan hun evaluaties (minimum 1% van hun directe kosten) effectief zinvol besteed wordt en betekenisvolle evaluaties oplevert voor alle betrokken actoren, is het belangrijk dat CMO's en IA's tussentijdse en eindevaluaties kunnen uitvoeren die samen bovenstaande doelstellingen bereiken.

Naast de ambitie van het realiseren van zinvolle evaluaties, staat ook de ambitie van realistische evaluaties voorop. In de interpretatie van de vereisten die gesteld worden aan (eind)evaluaties in artikel 44 § 1 en 44 § 2, is het dan ook nodig om een zekere flexibiliteit in te bouwen.

De nodige flexibiliteit in het evalueren van de outcomes kan worden ingebouwd op volgende manieren:

1. Relatie deel/geheel:

1.1 Combinatie evaluatie op terrein en desk study over de verschillende outcomes.

Een evaluatie kan op verschillende informatiebronnen gebaseerd zijn (monitoring gegevens van een organisatie, verzamelde informatie op het terrein, documentaire studie, ...). Voor zover mogelijk dienen databronnen getrianguleerd te worden met andere bronnen. Een terreinstudie vormt daarbij een adequate manier om monitoringdata van een organisatie te trianguleren. Binnen de evaluatie(s) die kijkt/kijken naar de outcomes van het programma kan een organisatie er echter voor kiezen om bepaalde outcomes via een desk study te evalueren gebaseerd op de monitoring gegevens van de organisatie. In het maken van de afweging welke outcomes uitgebreider (bijvb. via een evaluatie op het terrein) geëvalueerd worden, vormen de performantiescores een nuttige leidraad. Een desk study is meer aangewezen voor de gevallen waar beargumenteerd kan worden dat er geen problemen waren met de outcome of waar er weinig leerpotentieel is (bijvoorbeeld door te verwijzen naar de performantiescores of andere onderbouwde informatie), terwijl een terreinevaluatie aangewezen lijkt daar waar er meer problemen met een bepaalde interventie waren of daar waar het grootste leerpotentieel is.

1.2 Gebruik van cases binnen een outcome.

Een organisatie kan ervoor kiezen om een of enkele interventies te evalueren en bevindingen uit deze evaluaties te extrapoleren naar het niveau van de outcome. De geselecteerde interventies moeten dus representatief zijn voor de behaalde resultaten op het niveau van de outcome. Het kan hier bijvoorbeeld gaan over een bepaald type partnerschap, een bepaalde interventiestrategie, ... die geëvalueerd wordt binnen een outcome en waarvoor een aantal representatieve cases worden geselecteerd om een meer globale uitspraak te doen.

1.3 Uit de bovenstaande twee punten volgt dat evaluaties **verschillende vormen** kunnen aannemen om te 'oordelen over het halen van alle outcomes'.

- De aard van het programma speelt in die afweging een belangrijke rol. Bijvoorbeeld, voor een programma waar alle outcomes een gelijkaardige interventielogica of theory of change hebben, lijkt de flexibiliteit op het vlak van combinatie evaluatie op het terrein en desk study haalbaar. Eén of meerdere outcomes worden dan via een terrein studie geëvalueerd en de andere via een desk study (zie 1.1). In een programma waar de gelijkenis eerder te vinden is binnen een outcome is het interessant om te werken met cases binnen de outcome (zie 1.2).
- Om tot een afweging te komen over welke vorm de evaluatie(s) zullen aannemen doet de actor in samenspraak met de evaluator(en) een voorstel.
- Wanneer de referentietermen gecommuniceerd worden aan de DGD, dient duidelijk aangegeven te worden hoe met deze flexibiliteit omgegaan wordt. Daarvoor dient de organisatie een globaal beeld te geven aan de dossierbeheerder hoe in de evaluatie(s) de verschillende outcomes aan bod zullen komen. Indien het principe om een terreinmissie te doen niet gevolgd wordt, dient dit aangegeven te worden. Indien de DGD aanwijzingen heeft dat een terreinmissie toch nodig is, kan ze vragen om toch een terreinmissie te organiseren.

1.4 De vereiste om te oordelen over het 'halen van alle outcomes' geldt alleen voor de **eindevaluatie**. Bijgevolg is de interpretatie over de relatie deel/geheel ook alleen relevant voor eindevaluaties. De mid-term evaluatie mag zich richten op een deelaspect van het programma.

Gezamenlijke evaluatie:

- Waar organisaties het nuttig achten, kunnen gemeenschappelijke evaluaties uitgevoerd worden die een uitspraak doen over één of meerdere outcomes van de betreffende organisaties. Dit kan binnen een individueel of binnen een gemeenschappelijk programma. Een gemeenschappelijke evaluatie moet nuttig ervaren worden en kan niet verplicht worden (ook niet in het kader van gemeenschappelijke programma's, GSK's, ...).
- Dergelijke gemeenschappelijke evaluaties kunnen gekaderd worden binnen een GSK maar dit is zeker niet noodzakelijk.
- Een organisatie is verantwoordelijk voor haar eigen outcomes. Binnen een gezamenlijke evaluatie moeten er bijgevolg uitspraken zijn over de verschillende outcomes die er in aan bod komen. Binnen gemeenschappelijke programma's dragen organisaties voor sommige outcomes een gezamenlijke verantwoordelijkheid. Deze outcomes kunnen dan ook gezamenlijk geëvalueerd worden.
- Dit geldt zowel voor de mid-term als de eindevaluatie.

Participatieve externe evaluaties:

- Organisaties hebben de mogelijkheid om de evaluatie deels intern te organiseren en uit te voeren, mits organisaties een goed monitoringsysteem hebben uitgebouwd en een externe evaluator het proces van evaluatie valideert en als dusdanig garandeert dat de evaluatie op een onafhankelijke en betrouwbare manier wordt uitgevoerd. De externe evaluator treedt op als team leader van de evaluatie. De verantwoordelijkheid voor de kwaliteit van het eindproduct ligt dus bij de externe evaluator.
- Dit geldt zowel voor de mid-term als de eindevaluatie.

OESO-DAC criteria

- In het KB staat dat “De eindevaluatie gebeurt op basis van de criteria bepaald door het DAC”. Alle criteria dienen aan bod te komen, maar organisaties kunnen focussen op de criteria die het meest relevant zijn in de context van een bepaalde outcome. In samenspraak met de evaluator wordt bepaald op welke evaluatiecriteria ingezet zal worden.

Organisaties worden volgens het KB geacht de referentietermen van hun evaluaties te communiceren aan DGD. In die context leggen organisaties ook hun voorstellen over hoe zij bovenstaande richtlijnen plannen toe te passen voor aan hun dossierbeheerder van de DGD. Dit kan gaan over de argumentatie waarom men voor een bepaalde aanpak kiest (bijvoorbeeld: performantiescores bij keuze evaluatie op terrein versus desk study; interne leervragen als grond voor participatieve aanpak; etc.).

Het KB bepaalt dat de DGD evaluatievragen kan toevoegen aan de referentietermen. Dergelijke vraag tot toevoeging gebeurt steeds met respect voor de door de organisatie bepaalde scope en focus van de evaluatie zoals voorzien door de organisatie in de referentietermen.

DGD zal via de institutionele dialoog een globaal zicht hebben op de planning van de evaluaties van de organisatie voor het komende jaar. De referentietermen kunnen via een ad hoc overleg met DGD besproken worden of via mail afgestemd worden. In het laatste geval zal DGD de ontvangen referentietermen in de mate van het mogelijke becommentariëren binnen de twee weken na ontvangst ervan. Indien de termijn van twee weken niet gehaald kan worden, geeft DGD dit op voorhand aan. De referentietermen kunnen zowel becommentarieerd worden door de dossierbeheerder van het programma als de landenverantwoordelijke.

Aangezien het over een nieuw KB gaat, moeten de ervaringen met het uitvoeren van intermediaire en eindevaluaties als een leerproces gezien worden, zowel wat betreft de inhoud als het beheer en de uitvoering van de evaluaties. Daarbij houden de federaties de vinger aan de pols over de effectieve haalbaarheid van de vereisten die in het KB gesteld worden aan de (eind)evaluaties.