

(Zelf-)evaluatie van ESF-projecten

Leidraad voor het plannen van (zelf-)evaluaties

1. Inleiding	4
2. Wat wordt geëvalueerd?.....	5
a) Uitvoering van het project.....	5
b) Doorwerking van project	7
c) Output of outcome gerichte projecten.....	10
d) Mainstream of innovatie	10
3. Doelen van evaluatie en evaluatiecriteria.....	11
a) Doelen en timing van de evaluatie	11
Vroeg in het project.....	12
Kruissnelheid.....	12
Einde van het project.....	13
b) De evaluatiecriteria	13
4. Soorten evaluatie	16
a) Inleiding	16
b) Evaluatie van de prestatie	18
Getrouwheidsevaluatie	18
Prestatiegedreven effectenevaluatie	19
c) Ondersteunende evaluatie	20
Deblokkering van een project.....	20
Beoordeling van de evalueerbaarheid van het project.....	21
d) Theoriegedreven evaluatie	22
Theoriegedreven procesevaluatie	22
Theoriegedreven effectenevaluatie.....	22
5. Methodologische keuzes voor evaluatie.....	23
a) Inleiding	23
b) Wetenschappelijk evaluatie onderzoek	23
Algemeen.....	23
Kwantitatieve opzetten	25
Kwalitatieve opzetten	28
Andere vormen van wetenschappelijk onderzoek incl. “mixed methods”	29
c) Niet wetenschappelijke opzetten	30
Algemeen.....	30
Informatie inwinnen	31
Prioriteiten stellen.....	32
Coherentie verzekeren	33
d) Welke methodologie past bij welke benadering?.....	33
Op basis van benadering en evaluatievragen.....	33
Op basis van benodigde middelen	35
6. Organisatie van evaluatie	35
a) Plannen van de evaluatie als component van een projectvoorstel.....	35
b) Wie voert de evaluatie uit?.....	36

c) Wie volgt de evaluatie op en hoe?.....	37
7. Afstemming met evaluatie op programmaniveau.....	38
8. Opnemen van de evaluatieaanpak in een projectvoorstel.....	38
9. Literatuur	39

Lijst van afbeeldingen

Figuur 1: conceptualisering van uitvoering van een project.....	6
Figuur 2: voorbeeld van een schema van effecten	8
Figuur 3: voorbeeld van doelstellingen project, afgeleid van een effectenschema	9
Figuur 4: timing en doelen van de evaluatie	11
Figuur 5: evaluatiecriteria	14
Figuur 6: logische samenhang evaluatiecriteria	15
Figuur 7: keuze van evaluatiebenadering	17
Figuur 8: criterium en aard van de evaluatievraag naargelang evaluatiebenadering.....	18
Figuur 9: basismodel experiment	26
Figuur 10: overzicht keuze methode	34

Deze leidraad werd geschreven door Benedict Wauters (ESF-Agentschap Vlaanderen), met dank aan Marjolein Van Den Broeck, Veerle Moens en Maarten Christianen (ESF-Agentschap Vlaanderen) en Dries Verlet (Studiedienst Vlaamse Regering) voor het nalezen.

1. Inleiding

Deze leidraad beoogt de projectmanagers van ESF-projecten in Vlaanderen te oriënteren wat betreft het tussentijds en finaal evalueren van hun projecten.

Of er al dan niet geëvalueerd moet worden hangt echter van oproep tot oproep af. De onderhavige leidraad levert enkel een algemeen kader. Per oproep kan meer gespecificeerd worden wat al dan niet verwacht wordt van projectmanagers.

Om verwarring te vermijden dient eerst duidelijk gemaakt te worden wat evaluatie alvast niet is:

- ESF-kwaliteitsaudit (om het ESF-label te verkrijgen en te behouden) van de uitvoerende organisatie: de kwaliteitsaudit gebeurt niet op project- maar op organisatieniveau. Het uitgangspunt is dat organisaties die continu naar verbetering streven een vruchtbaarder bodem vormen voor ESF-projecten. Dit neemt echter niet weg dat de ESF-projecten zelf ook nog beoordeeld moeten worden. Een kwaliteitsvolle organisatie is immers geen sluitende garantie voor succesvolle projecten.
- Rapportering op het project: rapportering is een onderdeel van het projectmonitoren. “Monitoren is een continue functie die op systematische wijze informatie verzamelt betreffende indicatoren en mijlpalen om het management en belanghebbenden van een lopend project toe te laten de voortgang van het project, het bereiken van doelen en het gebruik van de ter beschikking gestelde fondsen op te volgen”¹.
- Controle (on desk en ter plaatse) van het project: bij controle gaat het om het nagaan (door het ESF-Agentschap, de Europese Commissie of een andere relevante instantie) of de projecten volgens de regels en waarheidsgetrouw (met betrekking tot projectvoorstel en rapportering) uitgevoerd werden.

Evaluatie is daarentegen “de systematische en objectieve beoordeling van een lopend of beëindigd project, met inbegrip van het ontwerp, de uitvoering en de effecten. De bedoeling is om de relevantie, efficiëntie, effectiviteit, impact en/of duurzaamheid vast te stellen. Een evaluatie moet informatie opleveren die geloofwaardig en nuttig is, wat zich uit in het integreren van het geleerde in de besluitvorming.”²

¹ OESO, 2002, “Glossary of Key Terms in Evaluation and Results-Based Management.” Paris: OECD/DAC.

² OESO, 2002

Bij evaluatie gaat het dus om een beoordeling met een zekere mate van objectiviteit en volgens een zekere systematiek. Evaluatie kan dus nooit gelijkgesteld worden aan een zelfreflectie door het projectmanagement en/of de projectuitvoerders. Deze leidraad zal dan ook methodieken aanreiken die gemeenschappelijk hebben dat andere belanghebbenden op een min of meer systematische wijze betrokken worden bij het beoordelen van een project. Dit levert minstens een inter-subjectief (meningen van een aantal mensen en vanuit verschillende perspectieven) oordeel op wat alleszins objectiever is dan loutere zelfreflectie.

Evaluatie kan zowel tijdens als na afloop van het project gebruikt worden om te oordelen, terwijl monitoren een projectactiviteit is en dus per definitie enkel tijdens het project plaats vindt.

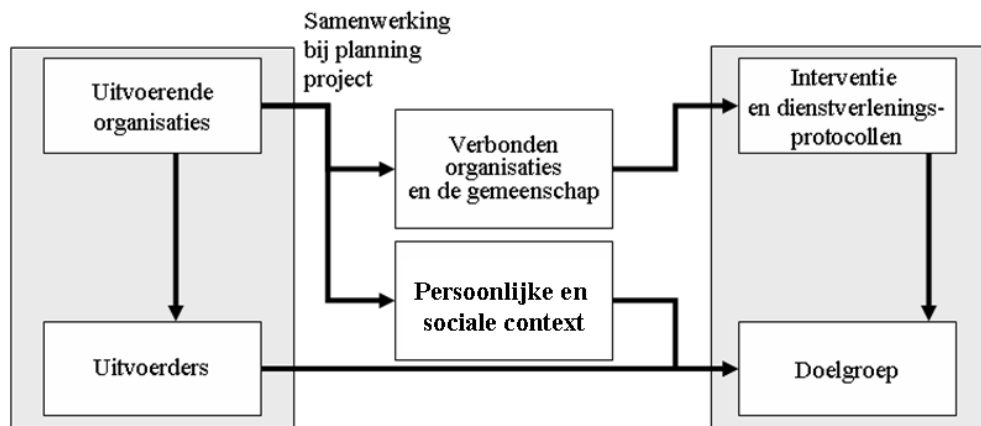
2. Wat wordt geëvalueerd?

De evaluatie van een project kan zich richten op de uitvoering van het project of op de uitkomst (effecten) van die uitvoering.

a) Uitvoering van het project

De conceptualisering van de uitvoering van een project is opgebouwd uit de elementen in Figuur 1. De logica van het schema loopt van links naar rechts en volgt de pijlen. Met deze pijlen wordt duidelijk gemaakt dat als er een probleem was in de uitvoering van één element, dit een blokkerend effect heeft op alle elementen waar het met een pijl naar verwijst. Het goed functioneren van elementen links is dus een voorwaarde voor het goed werken van elementen rechts. Als de uitvoerders bijvoorbeeld niet gemotiveerd zijn dan zal dit zijn weerslag vinden op de mate waarin aan de noden van de doelgroep voldaan wordt.

Figuur 1: conceptualisering van uitvoering van een project



Hieronder worden de elementen meer in detail besproken:

1. Beschrijving van de uitvoerende organisaties en meer bepaald hoe ze:
 - technologie / materiaal leveren;
 - uitvoerend en ander personeel recruterend, trainen en superviseren (inclusief prestatie meting en feed-back mechanismen);
 - middelen toewijzen (incl. financieel, externe expertise);
 - plannen en coördineren.
2. Beschrijving van het uitvoerend personeel (dat ervoor zorgt, dat het project iets beschikbaar stelt aan de doelgroep qua diensten, praktijken, attitudes, competenties, etc.), meer bepaald hun:
 - Kwalificaties;
 - Competenties;
 - Engagement en enthousiasme.
3. Verbonden organisaties en partners uit de gemeenschap:

Die organisaties wiens steun (soms wettelijk) vereist is om een actie te doen. Zij worden meestal al geïdentificeerd bij de analyse van belanghebbenden die een essentieel onderdeel is van ESF-projecten. Hier moet weer beschreven worden wat van hen verwacht en hoe met hen samengewerkt wordt (eventueel als partner binnen het project).

4. Persoonlijk en sociale context

Dit beschrijft de noodzakelijke voorwaarden om de doelgroep te laten deelnemen. Het gaat hier om:

- sociale, psychologische en/of materiële steun van nabije familie, vrienden en buurt; materiële steun in de omgeving (o.a. transportmogelijkheden, kinderopvang, financiële steun);
- heersende normen en cultuur.

5. Interventie- en dienstverleningsprotocol

Dit bevat:

- de filosofie van het project (gaat het over preventie of over oplossen van problemen, gaat het over ontwikkeling of over het uitrollen van een oplossing);
- hoe de doelgroep gerekruteerd, gescreend en beoordeeld zal worden;
- wat door het project qua output zal gerealiseerd worden voor die doelgroep;
- welke activiteiten voor die output zullen zorgen en wat hun timing is;
- wie wat doet, waar dit gedaan wordt en welke kanalen voor communicatie met de doelgroep worden gebruikt.

6. De doelgroep

- Kenmerken om in aanmerking te komen: geografisch, persoonskenmerken, grootte van de nood, etc.;
- Beschrijving van de noden;
- Mate waarin de doelgroep klaar is om deel te nemen: zowel mentaal (willen ze het probleem zien/erkennen, motivatie om oplossing te accepteren) als fysiek (bvb. gezondheid).

b) Doorwerking van project

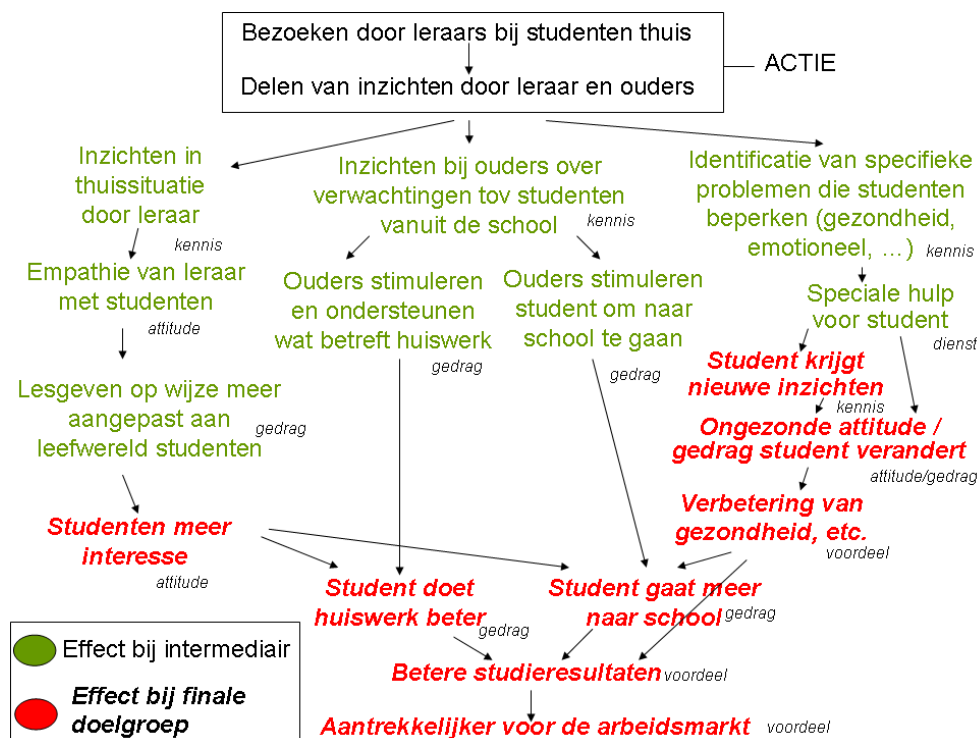
Projecten proberen veranderingen (effecten/outcomes) te realiseren in de maatschappij³. Veranderingen komen tot stand indien het project, via de gerealiseerde outputs (extensies van afgewerkte activiteiten), rechtstreeks inwerkt op een beperkt aantal sleutelfactoren in de externe omgeving bvb. op vakspecifieke kennis van werkzoekenden. Een directe verandering (bvb. een verbeterde kennis) zet dan weer verandering in gang van andere, “stroomafwaartse”, factoren (bvb.

³ We hanteren voor projecten gefinancierd door het ESF Vlaanderen definities die in het verlengde liggen van “Developing a logic model” (2008), Ellen Taylor-Powell et al. , University of Wisconsin. Vanuit de Europese Commissie wordt op programmaniveau voor de structuurfondsen, incl. het Europees Sociaal Fonds, echter de term “outcome” niet gebruikt. De EC spreekt eerder over “resultaten” en “impact”. Zowel de OESO als de meeste multi-laterale instanties gebruiken echter de term “outcome” en reserveren “resultaten” voor de verzameling van outcomes en outputs.

aantrekkelijkheid van een werkzoekende voor een werkgever). Uiteindelijk moet duidelijk worden welke ketting van oorzakelijk verbonden factoren aan de noden van de doelgroep tegemoet komt.

Dit wordt gevisualiseerd in een **schema van effecten**: het voorbeeld in Figuur 2 betreft een actie waarbij leerkrachten de ouders van studenten thuis bezoeken. Een schema van effecten kan meer bevatten dan de keten van veranderingen die door het project in gang gezet werd. Ook versterkende of belemmerende factoren waar het project zelf niet op inspeelt kunnen weergegeven worden!

Figuur 2: voorbeeld van een schema van effecten



De enige actie die gefinancierd wordt in het voorbeeld zijn de thuisbezoeken waarbij leraars en ouders met elkaar inzichten delen. Daarom vormen de bezoeken de output van dit project. Alles wat er daarna uit voortvloeit is een effect. Het was bijvoorbeeld ook mogelijk geweest dat er na de bezoeken ook nog een debriefing van de leraar met een gespecialiseerde begeleider gefinancierd zou worden waarbinnen inzichten gekristalliseerd worden en op papier gezet in een actieplan. Dan zouden dergelijke actieplannen ook output van het project zijn. De inzichten van de leraar zoals weergegeven in het actieplan zouden tevens geen effect meer zijn aangezien ze vervat zijn in een output van het project (dat er inzichten zullen zijn valt m.a.w. binnen de controle van de projectuitvoering).

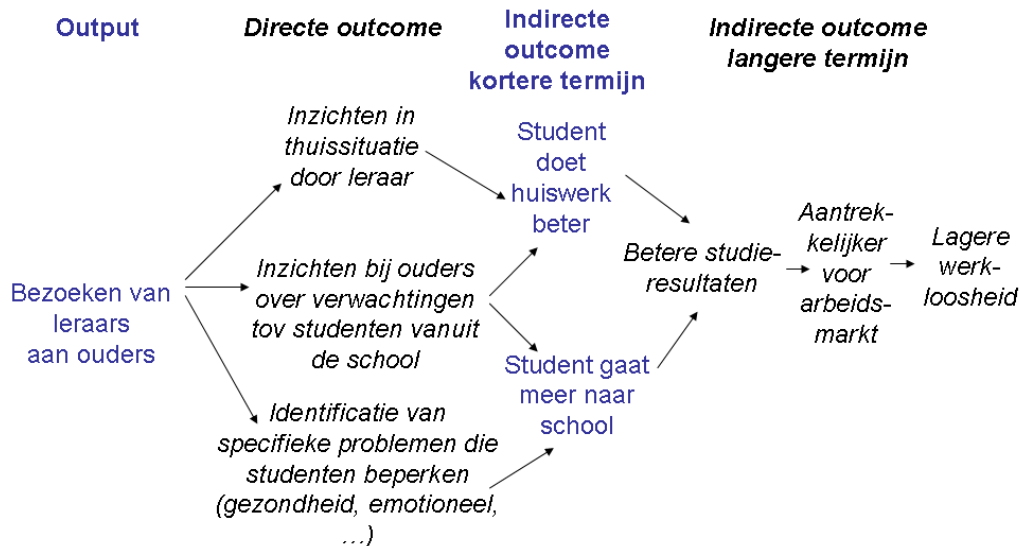
Het schema geeft een ketting van op elkaar inwerkende effecten weer. Het maakt hierin ook een onderscheid tussen effecten op het niveau van de intermediaire doelgroepen (leerkrachten, ouders) en de uiteindelijke doelgroep (studenten).

Er zijn tevens verschillende soorten effecten: het kan gaan om verbeteringen wat betreft kennis/vaardigheden, attitudes, gedrag en praktijken van zowel intermediairen als de finale doelgroep maar ook om een verbeterde dienstverlening (waar de dienstverlening niet als dusdanig gefinancierd wordt in het project maar een dienstverlening wel dankzij het project op verbeterde wijze ter beschikking staat).

Ten slotte moet een project een voordeel opleveren voor de finale doelgroep. Hierbij gaat het om iets wat een intrinsieke waarde heeft voor de doelgroep omdat het een fundamentele nood representeert.

Dit schema van effecten vormt de basis voor de selectie van de doelstellingen van een project. Dit betekent dat de doelen van een project niet het volledige schema reflecteren maar inzoomen op die effecten die als meest belangrijk ervaren worden.⁴

Figuur 3: voorbeeld van doelstellingen project, afgeleid van een effectenschema



De directe outcomes evenals de indirecte outcomes op korte termijn (gereflecteerd in operationele doelen) worden in principe zichtbaar op het moment dat het project nog loopt (de deelnemers staan nog in contact met het project), terwijl de indirecte outcomes (specifieke en globale doelen) op langere termijn soms pas geruime tijd na het project zichtbaar worden.

⁴ Bij outcome-gerichte ESF-projecten wordt door middel van probleem –en doelenbomen⁴ bij het formuleren van een projectvoorstel reeds een goede aanzet gegeven voor een effectschema dat bij bepaalde soorten evaluatie meer uitgewerkt kan worden. De probleem –en doelenboom dient ook om de doelen van een project op te maken maar zal zich beperken tot de indirecte effecten/outcomes waarbij de indirecte outcomes op korte termijn de operationele doelen zijn en de indirecte outcomes op langere termijn de specifieke doelstelling (intrinsiek voordeel voor de doelgroep bvb. aantrekkelijkheid arbeidsmarkt) of de globale doelstelling (wat betreft de uiteindelijke beleidsdoelstellingen bvb. lagere werkloosheid) van het project weerspiegelen:

c) Output of outcome gerichte projecten

Het Vlaamse ESF programma 2007-2013 hanteert over het algemeen twee soorten projectoproepen. “Outcome” gerichte oproepen richten zich op het verwezenlijken van een bredere, beleidsrelevante doelstelling. Oplossingen binnen het relevante beleidsdomein worden niet door de oproep bepaald. Indieners kunnen dus zelf bedenken en beslissen hoe ze tot de beleidsdoelstelling zullen bijdragen, in tegenstelling tot “output” gerichte oproepen waar de oproep zelf vastlegt wat het project moet opleveren (activiteitsgerelateerde outputs zoals trainingen, trajecten voor specifieke doelgroepen).

Bij outputgerichte projecten zullen min of meer (een) gestandaardiseerde acties (outputs) voor meerdere cohorten deelnemers uit de doelgroep herhaald worden (bvb. begeleidingstrajecten voor werkzoekenden naar werk). Het uitvoeringsconcept situeert zich dus reeds volledig op het niveau van deze cycli.

Binnen outcomegerichte projecten bouwen verschillende activiteiten voort op elkaar (bvb. reorganisatie van een bedrijf ten bate van de kwaliteit van de arbeid) met het oog op het realiseren van outcomes. Het uitvoeringsconcept situeert zich dus op het niveau van het hele project.

In de praktijk is het natuurlijk wel mogelijk dat bij outputgerichte projecten er ook een voortbouwen van activiteiten op elkaar is (bvb. na een output als training volgt een andere output zoals werkervaring), terwijl bij outcome gerichte projecten er ook cohorten van deelnemers kunnen zijn (bvb. binnen een project rond arbeidsorganisatie worden dezelfde opleidingen over zelfsturende teams voor alle medewerkers gegeven) binnen een breder uitvoeringsconcept. We zullen dus in deze leidraad eerder spreken over projecten die vooral werken met cohorten versus diegene waar dit geen substantiële component is van de uitvoering.

Het onderscheid is van belang voor evaluatie omdat bij projecten met cohorten gekozen kan worden om de evaluatie toe te spitsen op een beperkt aantal cohorten van deelnemers, na deelname aan (een reeks van) acties, (zie hoofdstuk 3. Doelen van evaluatie).

d) Mainstream of innovatie

Van innovatieve projecten worden nog niet meteen effecten verwacht op het niveau van de finale doelgroepen. Er wordt binnen het project immers ook nog niets ter beschikking gesteld aan deze doelgroepen. Dit zal pas, na afloop van het innovatief project, gebeuren door middel van “mainstreaming” (het inbedden van de innovatie in de reguliere werking van dienstverleners op de arbeidsmarkt).

Het ESF-Agentschap zelf speelt een belangrijke rol bij het evalueren van innovaties door middel van “validering”. Voor het evalueren van dit soort projecten wordt dan ook verwezen naar de richtlijnen betreffende de validering. Deze richtlijnen handelen ook over de opvolging door hert ESF-Agentschap nadat de valideringsprocedure beëindigd werd.

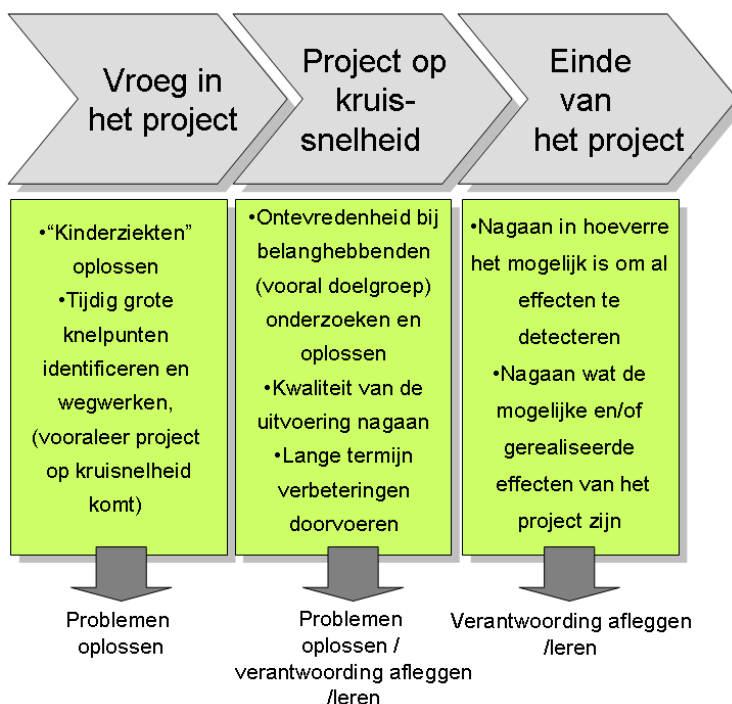
Niettemin kan bijkomende zelf-evaluatie door de promotor zeker nog nuttig zijn. De oproepen voor voor (transnationale) innovatieve projecten verplichten de promotor echter niet om zelf te evalueren.

3. Doelen van evaluatie⁵ en evaluatiecriteria

a) Doelen en timing van de evaluatie

De doelstellingen van evaluatie kunnen verschillend zijn naargelang de timing in het project (zie Figuur 4). De doelstellingen en de timing van de evaluatie zullen de keuze van de evaluatiebenadering bepalen (zie hoofdstuk 4).

Figuur 4: timing en doelen van de evaluatie



⁵ Bewerking op basis van Chen, 2005, Practical Program Evaluation, Sage en Evaluating socio-economic development (Evalsed), EC website. Er wordt bij deze laatste abstractie gemaakt van het “networking/institutional strengthening” doel aangezien dit altijd speelt.

Vroeg in het project

Vroeg in het project komt het er op aan zo snel mogelijk de operationele “kinderziekten” (bvb. afstemming tussen verschillende partners) op te lossen die zich in het begin van een project aandienen (hierbij gaat het dan ook eerder om de conceptualisering van de uitvoering –zie 2. a)). Ook grotere problemen, die met de logica en de aannames van het project betreffende doorwerking (zie 2. b)) te maken hebben, kunnen opgespoord en opgelost worden. Achterliggende doelstelling is dus **probleemoplossing**.

Voor projecten met cohorten is dit typisch na het aflopen van de deelname van een cohort aan een actie. Voor andere projecten gaat het typisch om 3 tot 6 maanden, afhankelijk van de complexiteit en totale looptijd van het project.

In principe kan hierover dus reeds binnen het evaluatieliek van de eerste tussentijdse rapportering gerapporteerd worden.

Kruissnelheid

Eens het project de beginperiode afgerond heeft en de projectuitvoering op kruissnelheid begint te draaien, kan op regelmatige wijze feedback over het project van de deelnemende doelgroep en andere belanghebbenden verkregen worden zowel over uitvoering als doorwerking (zie 2. a) en b)). Indien deze feedback negatief is, dan zal de evaluatie aanleiding geven tot het verder onderzoeken van de oorzaken en oplossingsrichtingen aanbrengen. Achterliggende doelstelling is dus **probleemoplossing**.

Tevens kan nagegaan worden of de projectuitvoering wel verloopt volgens de vooraf vastgelegde normen beschreven in de conceptualisering (zie 2.a)) van de uitvoering (kwaliteit van de uitvoering). Achterliggende doelstelling hierbij is het afleggen van **verantwoording** ten opzichte van de belanghebbenden.

Hoe langer het project loopt, hoe meer fundamentele lessen kunnen getrokken worden betreffende de conceptualisering van de uitvoering (zie 2.a)) van het project en een mogelijke optimalisering naar de toekomst toe. Het gaat dan niet meer over het ad hoc oplossen van knelpunten maar over het grondig herdenken van de projectuitvoering. Achterliggend is er dus een **leerdoelstelling**.

Bij projecten gericht op cohorten kan men bovenstaande soorten evaluatie op regelmatige basis uitvoeren (na elke deelname van een cohort aan een afgeronde actie). Bij andere projecten zal dit eerder gebeuren na het (geplande) bereiken van belangrijke mijlpalen waarvan het vervolg van het project afhangt.

Deze fase van het project loopt tot het effectieve einde van het project. Er kan dus over gerapporteerd in het evaluatielook van de tussentijdse rapporteringen én in de finale rapportering.

Einde van het project

Evaluatie kan zich uitspreken over de tijdshorizon waarin (onderscheiden) effecten verwacht worden zich te manifesteren. Naarmate het einde van het project nadert zal men immers meer zicht beginnen krijgen op de effecten van het project (zie “2. b) Doorwerking van project”). Verschillende effecten manifesteren zich echter binnen verschillende termijnen. Sommige effecten zullen pas na afloop van het project voelbaar worden en andere kunnen zich reeds ten volle manifesteren tijdens de looptijd van het project.

Na het volgen van een opleiding kan het bijvoorbeeld nog wel enkele maanden tijd kosten vooraleer een deelnemer een aangepaste job vindt. Het is echter bij het evalueren van de doorwerking van een project wel onontbeerlijk een idee te kunnen geven van die termijn en te bespreken of de doelgroep tijdens de projectuitvoering al in de juiste richting evolueert (bvb. actief solliciteert en positieve feedback krijgt van werkgevers).

Bij projecten die met cohorten werken zal het meestal wel mogelijk zijn om binnen de looptijd van het project volledig gerealiseerde cycli op effecten te evalueren. Indien een project met een totale looptijd van twee jaar bijvoorbeeld al in de eerste 6 maanden voor twee cohorten een trainingscyclus heeft uitgevoerd en er van uitgaat dat het gemiddeld 6 maanden duurt om de effecten te zien, dan kan een (zij het tot twee cohorten beperkte) evaluatie van effecten al afgerond worden na een jaar.

De achterliggende doelstelling van het evalueren van effecten is meestal het afleggen van **verantwoording** ten opzichte van belanghebbenden met betrekking tot de vooraf beoogde effecten (dit kan zonder volledig effectenschema, door enkel te focussen op die effecten die in de doelen van de interventie gereflecteerd worden (zie

Figuur 3). Anderzijds schuilt er veelal ook een **leerdoelstelling** achter, namelijk nagaan op welke wijze een proces van verandering voor de finale doelgroep in gang gezet wordt na deelname (op basis van een volledig effectenschema (zie Figuur 2).

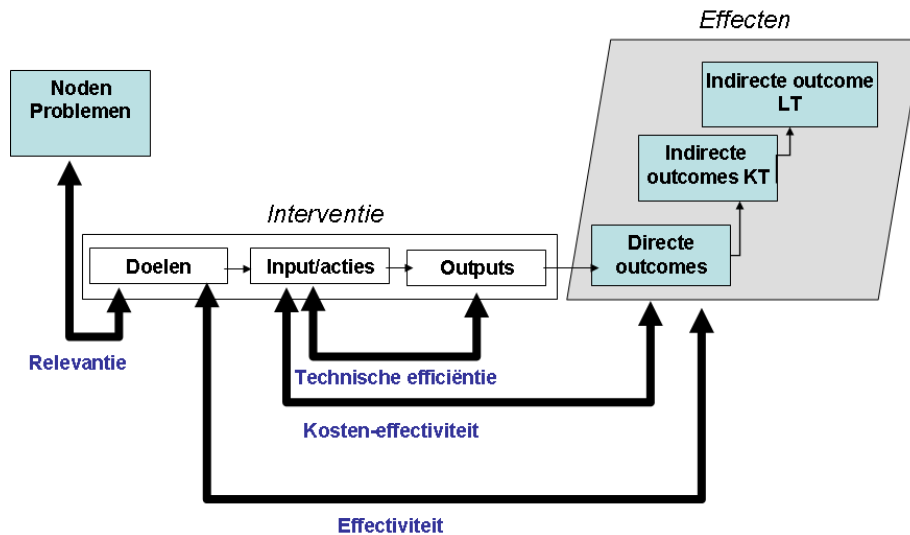
b) De evaluatiecriteria

Naast de doelen en timing van evaluatie worden ook generieke evaluatiecriteria gehanteerd om vorm te geven aan de evaluatie. Waar doelen en timing de keuze voor de evaluatiebenadering zullen

bepalen, zullen de evaluatiecriteria de aard van de evaluatievragen binnen de gekozen evaluatiebenadering bepalen.

In de algemene inleiding werd al verwezen naar de meest gebruikte evaluatiecriteria namelijk relevantie, efficiëntie, effectiviteit, impact en duurzaamheid.

Figuur 5: evaluatiecriteria



De evaluatiecriteria worden op basis van de bovenstaande figuur als volgt beschreven:

- **Efficiëntie:**
 - enerzijds gaat het om kosten-effectiviteit: is er een goede verhouding tussen de ingezette middelen en de effecten (in de eerste plaats directe outcomes en de indirecte outcomes op kortere termijn)?
 - anderzijds gaat het om technische efficiëntie: de best mogelijke verhouding tussen input en outputs
- **Effectiviteit:** leidt de uitgevoerde actie tot het behalen van de beoogde effecten (in de eerste plaats de directe en indirecte outcomes op kortere termijn gereflecteerd in operationele doelen, evenals het behalen van een intrinsiek voordeel –een outcome op de langere termijn - voor de doelgroep, gereflecteerd in het specifiek doel)?
- **Impact:** het onderscheid tussen impact en effectiviteit bestaat er enkel uit dat impact gaat om het realiseren van de indirecte outcomes op langere termijn (meestal beleidsdoelen) die voortvloeien uit het verkrijgen van een duurzaam voordeel voor de doelgroep (zie effectiviteit).
- **Relevantie:** beantwoorden de doelstellingen (inclusief het duurzaam voordeel) daadwerkelijk aan de (eventueel gewijzigde) probleemsituatie en noden?

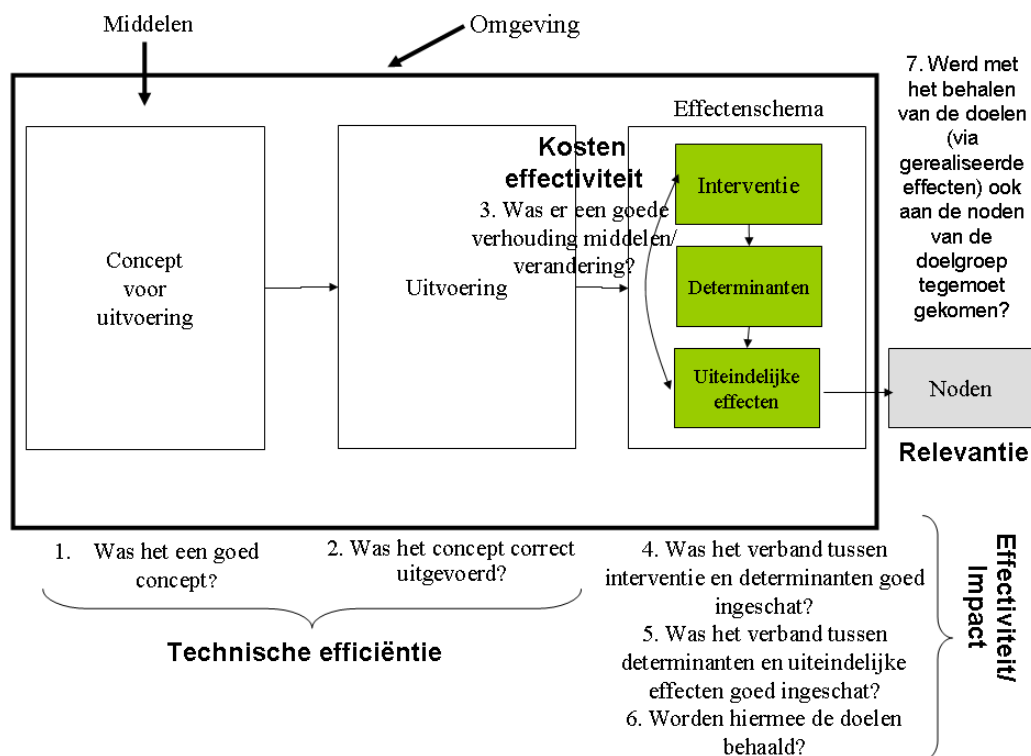
Duurzaamheid wordt in het schema niet vermeld. Het gaat hierbij om de vraag of de effecten (zowel blijvend zullen zijn, of wegvallen met het stopzetten van het project.

Al deze criteria zijn onderling ook verbonden.

Idealiter is de logica van de criteria immers als volgt (zie onderstaande figuur): indien een goed concept (1 in de figuur) voor de uitvoering bedacht is (waarbij de verhouding tussen inputs en opututs geoptimaliseerd werd: geplande technische efficiëntie) en dit wordt ook daadwerkelijk zo uitgevoerd (gerealiseerde technische efficiëntie) (2), dan werkt de interventie in op de realiteit van de doelgroep via een effectenschema van determinanten en uiteindelijke effecten. Sommige van die effecten worden vooropgesteld als doelen en als die gehaald worden dan is de interventie effectief (4, 5 en 6). Als de doelen gehaald worden zonder de vooropgestelde, billijke, allocatie van middelen te overschrijden, dan is de interventie kosteneffectief (3). Als de behaalde doelen er in slagen om de noden van de doelgroep te vervullen dan is de interventie ook relevant (7).

Uit dit schema wordt dan ook duidelijk hoe de evaluatiecriteria met elkaar verbonden zijn. Indien bijvoorbeeld het concept wel goed was, maar de geplande middelen bleven uit, met als gevolg een gebrekkige implementatie, dan heeft dit niet enkel een weerslag op de efficiëntie (minder middelen besteed maar hierdoor ook de geplande output niet in de vooropgestelde mate opgeleverd) maar waarschijnlijk ook op de (kosten-)effectiviteit.

Figuur 6: logische samenhang evaluatiecriteria



Uit het schema blijkt ook duidelijk dat enkel de invloed van de interventie als geheel op de determinanten en effecten beschouwd wordt. Het is immers niet zinvol om causale verbanden te proberen leggen tussen geïsoleerde componenten van het uitvoeringsconcept en de effecten.

4. Soorten evaluatie⁶

a) Inleiding

Afhankelijk van de doelstellingen van de evaluatie evenals het tijdstip, zullen er verschillende evaluatiebenaderingen gekozen worden. Het betreft hier:

- Evaluatie van de prestatie:
 - Getrouwheid
 - Effecten
- Ondersteunende evaluatie:
 - Deblokkerend
 - Evalueerbaarheid
- Theoriegedreven evaluatie
 - Proces
 - Effecten

De keuze voor een evaluatiebenadering hangt af van de doelstelling en het tijdstip van de evaluatie. Verschillende benaderingen leggen hun focus ook verschillend op concept van uitvoering en/of doorwerking (zie 2.a) en b)) . Het overzicht wordt geboden in de onderstaande tabel.

⁶ Bewerking op basis van Chen, 2005, Practical Program Evaluation, Sage, waarbij, om redenen van vereenvoudiging, de “troubleshooting”, “development facilitation” en “development partnership” benaderingen onder de gezamenlijke noemer “deblokkering” gebracht werden.

Figuur 7: keuze van evaluatiebenadering

Timing	Begin	Kruissnelheid	Einde
Primair doel			
Problemen oplossen	Ondersteunende evaluatie (deblokkering) (focus op uitvoeringsconcept en/of doorwerking)	Ondersteunende evaluatie (deblokkering) (focus op uitvoeringsconcept en/of doorwerking)	Ondersteunende evaluatie (evalueerbaarheid) (focus op uitvoeringsconcept en/of doorwerking)
Verantwoording afleggen		Getrouwheidsevaluatie (focus op uitvoeringsconcept)	Prestatiegedreven effectenevaluatie (focus op doorwerking)
Leren		Theoriegedreven procesevaluatie (focus op uitvoeringsconcept)	Theoriegedreven effectenevaluatie (focus op doorwerking)

Afhankelijk van de evaluatiecriteria zal dan weer de aard van de evaluatievragen binnen de gekozen evaluatiebenadering bepaald worden. De verschillende criteria moeten immers geconcretiseerd worden met behulp van specifieke evaluatievragen. Deze evaluatievragen verschillen qua aard naargelang het criterium. Sommige criteria passen niet bij een bepaalde benadering. De keuze voor een benadering impliceert dus ook de keuze voor een criterium.

Er zijn vier verschillende types vragen:

- het bevragen van knelpunten (knelpuntgerichte vragen);
- vragen rond causale verbanden tussen determinanten en uiteindelijke gerealiseerde effecten (causale vragen-;
- een beschrijving van de feiten (descriptieve vragen);
- een vergelijking van de feiten met een norm (normatieve vragen).

De onderstaande figuur geeft weer welke de aard van de vragen is naargelang het evaluatiecriterium gegeven de gekozen evaluatiebenadering.

Figuur 8: criterium en aard van de evaluatievraag naargelang evaluatiebenadering

	Technische efficiëntie	Kosten-effectiviteit	Effectiviteit/impact	Relevantie
Deblokkering	Knelpuntgericht	Knelpuntgericht	Knelpuntgericht	Knelpuntgericht
Getrouwheid	Vgl. met normen			
Prestatiegedreven effecten		Vgl. met normen	Causaal	
Theoriegedreven effectenevaluatie			Causaal	
Theoriegedreven procesevaluatie	Beschrijvend / vgl. met normen			

Alle bovenstaande begrippen worden per hieronder per benadering verduidelijkt.

b) Evaluatie van de prestatie

Getrouwheidsevaluatie

Dit soort evaluatie wordt uitgevoerd als het **project op kruissnelheid** is (wanneer een zicht bestaat op het grootste deel van het uitvoeringsconcept) en beperkt zich tot het nagaan of het geplande concept voor uitvoering ook effectief zo geïmplementeerd werd en of eventuele afwijkingen door het projectmanagement hadden kunnen vermeden worden. Er wordt niet nagegaan of het geplande concept van uitvoering wel het meest gepaste was. De doelstelling is louter het **afleggen van verantwoording**. Indien ernstige afwijkingen geconstateerd worden kan dit wel aanleiding geven tot een andere soort evaluatie (gericht op probleemoplossing of leren). De focus ligt dus op het **uitvoeringsconcept** (zie 2.a)).

Het criterium waarop geëvalueerd wordt is **technische efficiëntie**.

De belangrijkste evaluatievraag die hier gesteld wordt moet dan ook zijn: **“in hoeverre komt de uitvoering van het project overeen met wat gepland was?”**. Het geplande (inclusief de inzet van, o.a. financiële, middelen) verwijst naar de meest recente en in voege zijnde versie van de interventie, niet naar de oorspronkelijke versie. Het is wel belangrijk dat, indien er wijzigingen gemaakt werden, deze effectief gedocumenteerd en gerechtvaardigd werden. Dit vormt immers het startpunt van dit soort evaluatie.

Deze vraag kan uiteraard opgesplitst worden in meer gedetailleerde vragen betreffende (een selectie van) verschillende aspecten van het uitvoeringsconcept (zie 2.a)) bvb.:

- Worden alle vooropgestelde activiteiten uitgevoerd?

- Worden de activiteiten en de stappen binnen deze activiteiten uitgevoerd op de wijze zoals vooropgesteld (aantallen en kwaliteit) en aan de vooropgestelde billijke kost?"
- Wordt voldoende rekening gehouden met randvoorwaarden om deel te nemen, onder andere door deelnemers door te verwijzen naar complementaire oplossingen?
- Is het aantal personen dat door het project bediend in overeenstemming met wat vooropgesteld werd (kwantiteit)?
- Is het aantal personen dat bediend wordt in voldoende mate afkomstig uit de oorspronkelijk beoogde (moeilijkere) doelgroep?
- Zijn de deelnemers tevreden over hun deelname?
- Etc...

Het gaat hier steeds om **vragen betreffende het halen van vooropgestelde normen** (vergelijking van de feiten met een geplande norm).

Dit soort evaluatie lijkt erg op projectmonitoring met dat verschil dat het om een omvattend én objectief oordeel (dus niet dat van de projectmanager of projectmedewerkers –zie eerder in hoofdstuk 1) gaat betreffende de mate dat het projectmanagement tot op dat moment het concept van het project heeft weten uitvoeren op het verwachte kwalitatieve (bvb. zijn deelnemers in de beoogde mate tevreden) en kwantitatieve niveau (bvb. werd het vooropgestelde aantal deelnemers gehaald). Het is dan ook duidelijk dat deelname en zelfs tevredenheid van deelnemers geen deel uit maken van een effectenevaluatie (zie hieronder) aangezien tevredenheid een beoordeling van de uitvoering is.

Prestatiegedreven effectenevaluatie

Hierbij wordt, **naar het einde van het project toe**, enkel nagegaan of de vooraf beoogde effecten wel degelijk gerealiseerd werden door de interventie. Er wordt geen inzicht verschaft in de wijze waarop deze effecten tot stand kwamen noch in eventuele onvoorziene effecten. Ook hier is de doelstelling louter **verantwoording**. Het verschil met monitoring op de projectindicatoren zit erin dat een effectenevaluatie hard probeert te maken dat de veranderingen die gerealiseerd worden wel degelijk aan het project toe te wijzen zijn (oorzakelijk verband tussen effecten en interventie) en niet aan andere factoren. De focus ligt dus op **doorwerking** (zie 2.b)).

Het criterium is hier **effectiviteit**.

Hier is de hoofdvraag dan ook: **“in hoeverre worden de door de interventie vooropgestelde doelen behaald?”**. Ook hier kan men de vraag verder detailleren naargelang de (meest recent) beoogde uiteindelijke doelstellingen. Dit soort evaluatie wordt effectenevaluatie genoemd omdat men via de gerealiseerde effecten nagaat of de doelen behaald worden. Men is echter, in tegenstelling tot de theoriegedreven effectenevaluatie (zie infra), niet geïnteresseerd in de wijze waarop deze effecten tot stand gekomen zijn (welke determinanten invloed hebben uitgeoefend), enkel in het toewijzen aan de

interventie van de behaalde niveaus van de geselecteerde effecten. Er wordt dus geen gebruik gemaakt van een volledig effectenschema (zie Figuur 2) enkel van de vooropgestelde doelen (zie Figuur 3).

Het gaat hier primair om **een causale vraag** aangezien men de gerealiseerde effecten aan de interventie wil kunnen toeschrijven. De effecten die men waarneemt kunnen immers ook te wijten zijn aan een complementaire actie, een ontwikkeling in de omgeving die niets met het ESF project te maken heeft etc.. Die alternatieve verklaringen zal men willen uitsluiten.

Verder kan ook gevraagd worden naar de **kosteneffectiviteit**: **“in hoeverre had men dezelfde effecten kunnen bereiken met minder middelen of meer met dezelfde middelen?”**. Dit is een **vraag betreffende het halen van vooropgestelde normen** waarbij de norm meestal (vooraf of later) op basis van een “benchmark” gesteld wordt (ofwel dezelfde soort interventie in het verleden, ofwel een gelijkaardige interventie elders ofwel op basis van het oordeel van belanghebbenden en/of experts).

c) Ondersteunende evaluatie

Deblokkering van een project

Dit soort ondersteunende evaluatie richt zich louter op het oplossen van knelpunten bij de uitvoering van het project. In dit geval ligt de focus dus op het **uitvoeringsconcept** (zie 2.a)). Anderzijds is het ook mogelijk dat een project geblokkeerd raakt omdat er onenigheid aan de oppervlakte komt tussen belanghebbenden wat betreft de prioriteiten van het project (aard van de doelen) of de samenhang tussen doelen. Dan ligt de focus op **doorwerking** (zie 2.b)).

De doelstelling is dus louter **probleemoplossend**. Meestal gebeurt dit soort evaluatie dus in het **beginstadium of wanneer het project op kruissnelheid** begint te komen. Alleszins moet dit soort evaluatie vroeg genoeg gebeuren vooraleer er door het project schade aangericht wordt bij de doelgroep of andere belanghebbenden.

In projecten met cohorten kan dit soort evaluatie ook nog na elke afgerond deelname van een cohort gebeuren in de mate dat er effectief problemen waren. In andere projecten kan het gebeuren na kritische mijlpalen, eveneens naargelang er problemen ondervonden werden die het verdere verloop van het project in gevaar brengen.

De vragen kunnen betrekking hebben op **eender welk evaluatiecriterium** en zijn eerder praktisch van aard: **“hoe lossen we best een bepaald knelpunt op?”**. Het aantal vragen zal dus afhangen van het aantal concrete knelpunten en hun samenhang. Het kan betrekking hebben op:

- doelen die niet goed geformuleerd werden of waarover onenigheid is ontstaan tussen belanghebbenden bvb. omdat het onderliggende effectenschema niet helder is (zie 2b));

- problemen met de conceptualisering van de uitvoering (zie 2.a)).

Beoordeling van de evalueerbaarheid van het project

Naar het einde van het project toe kan een ondersteunende evaluatie gebruikt worden om na te gaan in hoeverre

- de **conceptualisering** van de uitvoering (zie 2.a) nog steeds in overeenstemming is met de praktijk
- doelen van een project (nog steeds) welomschreven zijn en wanneer effecten zouden kunnen geobserveerd worden (focus op **doorwerking** zie 2.b))
- relevante informatie verzameld kan worden om een toekomstige evaluatie uit te voeren
- beoogde gebruikers van een toekomstige evaluatie het eens kunnen worden over doelen en vragen van een toekomstige evaluatie

Dit soort evaluatie is dus eigenlijk eerder een voortraject voor een latere evaluatie (van prestaties of theoriegedreven). Ze kan dus ook gebruikt worden om een keuze te maken voor het soort evaluatie dat zal volgen naar het einde van het project toe.

In de praktijk zal dit soort evaluatie enkel nodig zijn indien belanghebbenden gedurende de implementatie van het project sterk van mening begonnen te verschillen over de doelen en de aard van het project of wanneer hierover eigenlijk nooit overeenstemming bereikt was. De doelstelling is dan ook eerder **probleemoplossend** in de zin dat het project anders niet geëvalueerd kan worden omdat men het er niet over eens is waaruit het project nu eigenlijk bestaat.

Dit soort evaluatie **valt buiten het kader van de generieke criteria**. Het is er immers op gericht om de slaagkansen te verhogen van een evaluatie van de prestatie of een theoriegedreven evaluatie. De centrale vraag is dan ook **“in hoeverre is het mogelijk om een evaluatie van de prestatie of een theoriegedreven evaluatie van het project uit te voeren?”**.

Voorbeeldvragen zijn:

- in hoeverre is er overeenstemming tussen belanghebbenden betreffende de conceptualisering en/of de beoogde effecten van het project (zie 2.a en b)?
- in hoeverre is er overeenstemming tussen belanghebbenden betreffende doelen en benadering van een toekomstige evaluatie?
- in hoeverre is er data beschikbaar of kan deze nog verzameld worden om een eventuele toekomstige evaluatie uit te voeren?

Bovenstaande vragen hebben allen betrekking op het **exploreren** van de precondities om te evalueren.

d) Theoriegedreven evaluatie

Theoriegedreven procesevaluatie

Dit soort evaluatie start vanuit een welomschreven **conceptualisering van de uitvoering** van het project (zie 2.a). De evaluatie is er op gericht om na te gaan of deze conceptualisering wel degelijk aangepast was aan de reële situatie. Ze wordt uitgevoerd als het project reeds op **kruissnelheid** zit. De doelstelling is in de eerste plaats **leren** hoe het concept van projectuitvoering (zie 2.b)) in de praktijk gewerkt heeft, of het oorspronkelijke concept wel het beste was en welke aanpassingen eventueel aangewezen zijn indien men dit concept wil herhalen.

Een theoriegedreven procesevaluatie gaat verder dan een getrouwheidsevaluatie. Ze stelt zich immers ook de vraag “in hoeverre het concept a priori werkbaar was en in hoeverre een ander concept beter had kunnen werken?”. De concrete vragen over specifieke elementen van het concept zullen dus beschrijvend (bvb. welke knelpunten waren er) zijn of betreffen het halen van vooropgestelde normen (bvb. hoe snel verloopt de screening van deelnemers in vergelijking met andere projecten of met een beoogd streefcijfer). Het criterium is technische efficiëntie.

Een veel gemaakte fout is causale vragen op te nemen binnen een procesevaluatie waarbij gevraagd wordt wat de invloed is van een specifiek element uit het uitvoeringsconcept op de doorwerking. Effecten kunnen echter enkel zinvol verbonden worden met een interventie als geheel (zie ook bespreking van Figuur 6).

Theoriegedreven effectenevaluatie

Net als bij de theoriegedreven procesevaluatie heeft dit soort evaluatie als primair doel **leren** en wordt ze uitgevoerd naar het **einde toe van het project**. Net als de prestatiegedreven effectenmeting beoogt ze hard te maken in hoeverre effecten aan het project kunnen toegewezen worden. Het verschil met deze laatste is dat er ook inzicht verworven wordt in de wijze waarop de **veranderingen doorwerken** (zie 2.b)). Bij de theoriegedreven effectenevaluatie is het uitgangspunt, i.t.t. de prestatiegedreven effectenmeting, dus een volledig schema van effecten (zie Figuur 2) dat de tussenliggende **causale** verbanden tussen interventie en uiteindelijke effecten, via determinanten, blootlegt.

Een theoriegedreven evaluatie kan, naast leren, uiteraard ook gebruikt worden om verantwoording af te leggen en om op langere termijn problemen duurzaam op te lossen. Het primaire doel is echter leren waarom en hoe het project al dan niet werkt. Indien deze doelstelling ontbreekt, is een andere benadering aangewezen (namelijk een ondersteunende of prestatie-evaluatie).

Het gaat hier dus, net zoals bij de prestatiegedreven effectenevaluatie, om **effectiviteit**. Specifieke vragen worden afgeleid uit het effectenschema: “**in welke mate is er een oorzaak-gevolg relatie**

tussen de interventie en directe beïnvloedbare determinanten, tussen determinanten onderling of tussen determinanten en de uiteindelijk beoogde effecten”.

5. Methodologische keuzes voor evaluatie

a) Inleiding

Dit hoofdstuk is niet bedoeld om de lezer in staat te stellen alle mogelijke methoden correct toe te passen. Het is wel bedoeld om de lezer te helpen een keuze te maken uit de ruime waaier aan methoden en te oriënteren bij het kiezen van de juiste professionele ondersteuning.

b) Wetenschappelijk evaluatie onderzoek

Algemeen

Binnen de wetenschappelijke benadering bestaat een grote verscheidenheid aan onderzoeksontwerpen die elk meer of minder geschikt zijn naargelang de aard van de vraag. Een ontwerp wordt gedefinieerd door de aard van de informatie die men zal verzamelen (bvb. gesloten vragen met een beperkt aantal antwoordmogelijkheden), de instrumenten die men gebruikt voor dataverzameling (bvb. een vragenlijst) en analyse (bvb. statistische regressieanalyse) in combinatie met de selectie van de databronnen (bvb. toevalsgewijze steekproef van respondenten). Ook van belang is de timing van de selectie van databronnen en het verzamelen van de data en het uitvoeren van de analyse (bvb. alles op voorhand bepalen versus ad hoc).

Alvorens dieper in te gaan op wetenschappelijke methoden is het van belang te herhalen dat sommige evaluatievragen peilen naar een oorzakelijk verband (causaal), terwijl andere een beschrijvend karakter hebben (descriptief) ofwel een vergelijking maken met een norm (zie hoofdstuk 4.a)). Indien er causale vragen zijn, impliceren deze echter meestal ook dat er eerst iets beschreven moet worden. Dit neemt niet weg dat de opzet die men moet kiezen er één moet zijn die causaliteit kan aantonen. Deze opzet kan in principe ook het beschrijvende deel voor haar rekening nemen.

Iedere vorm van wetenschappelijk evaluatief onderzoek heeft echter wel volgende kenmerken:

- er wordt een onderzoeksopzet uitgewerkt met de bedoeling een hoge betrouwbaarheid en validiteit van de conclusies te borgen;
- er wordt een, aan de gekozen opzet, aangepast instrumentarium gebruikt voor dataverzameling (bvb. een vragenlijst) en analyse (bvb. statistisch);
- er wordt een aangepaste steekproefstrategie uitgewerkt;

- er wordt expliciet getoetst aan criteria van wetenschappelijke betrouwbaarheid en validiteit⁷.

Er bestaan verschillende soorten wetenschappelijke onderzoeksopzetten. Het vraagstuk van de validiteit van het onderzoek kan niet losgekoppeld worden van het overkoepelende ontwerp van het onderzoek bvb. een experimentele opzet dient om een hoge mate van interne validiteit (conclusies betreffende causaliteit) te borgen door middel van vergelijkingsgroepen. Dataverzamelmethode evenals analysemethoden moeten dus beoordeeld worden op hun wetenschappelijkheid binnen een overkoepelend ontwerp.

Binnen de wetenschappelijke opzetten wordt een eerste onderscheid gemaakt tussen kwantitatieve en kwalitatieve benaderingen.

Alle kwantitatieve opzetten hebben het volgende gemeenschappelijk:

- de opzet (selectie van steekproef, dataverzameling –en analyse methoden) staat van bij het begin vast;
- de informatie die verzameld zal worden, is op voorhand vastgelegd;
- er wordt gebruik gemaakt van gesloten vragen en/ of observatie (met vooraf bepaalde antwoord/observatiemogelijkheden) en instrumenten voor dataverzameling die hiervoor geschikt zijn (gestructureerde interviews, gesloten vragenlijsten, testen, checklists,...);
- de data wordt statistisch geanalyseerd;
- steekproeven worden dusdanig bepaald om statistisch naar de populatie te kunnen veralgemenen (ofwel wordt de volledige populatie gemeten);
- de realiteit wordt gereduceerd tot een beperkt aantal variabelen.

Kwalitatief onderzoek verschilt van kwantitatief onderzoek op alle vlakken:

- het ontwerp ligt niet volledig vast in het begin: er is flexibiliteit om andere dataverzameling en analysemethoden in te zetten naargelang nodig;
- ook de benodigde informatie ligt niet volledig vast; tijdens het onderzoek kan duidelijk worden dat nieuwe informatie nodig is;
- er wordt gewerkt met open vragen en/of observatie waarbij de onderzoeker zelf als dataverzameling instrument functioneert;
- analyse van tekst, beelden, audio, etc. gebeurt door het zoeken naar categorieën die de data omvatten en naar verbanden tussen deze categorieën;
- steekproeven worden bepaald volgens het principe van doelmatigheid (waarbij eerst die bronnen aangeboord worden die het snelst het meeste nieuwe informatie opleveren) en verzadiging (als verdere dataverzameling geen nieuwe elementen meer bevat dan wordt de dataverzameling afgesloten);

⁷ Zie website ESF Agentschap www.esf-agentschap.be voor meer informatie omtrent betrouwbaarheid en construct, conclusie, interne en externe validiteit.

- de realiteit kan in al haar complexiteit gevat worden.

In zowel kwalitatief als kwantitatief onderzoek kan er met al bestaande gegevens gewerkt worden (secundaire data). In vele gevallen zullen er echter ook ten behoeve van de evaluatie nieuwe gegevens verzameld moeten worden (primaire data).

Indien gekozen wordt voor de wetenschappelijke benadering dan is het aangewezen om deze soort evaluatie te laten uitvoeren door personen met een gepaste onderzoeksachtergrond. Dit geldt zowel voor kwalitatief als kwantitatief onderzoek. Ofwel vindt men deze intern (bvb. personen met een doctoraat in de menswetenschappen of met een achtergrond in de onderzoeksmethodologie, ...) ofwel extern (doctorale onderzoekers aan een universiteit, Master studenten die een thesis onderwerp zoeken,...). De volgende hoofdstukken zijn bedoeld om projectmanagers in staat te stellen om een eerste planning te maken wat betreft evaluatie en een geïnformeerde discussie over een eventuele evaluatieopdracht te voeren met een interne of externe deskundige. In de literatuurlijst op de ESF website worden meer gedetailleerde bronnen opgelijst.

Kwantitatieve opzetten

Experimentele en quasi-experimentele opzetten

De bedoeling van experimentele en quasi-experimentele opzetten is het onderzoeken van causaliteit. De gangbare noodzakelijke en voldoende voorwaarden voor causaliteit tussen twee factoren worden als volgt samengevat: A is de oorzaak van B als:

- er een (probabilistische) relatie is tussen A en B (veranderingen in A gaan meestal samen met een vermindering of vermeerdering in B, wat neerkomt op een negatieve of positieve samenhang)
- A in de tijd voor B komt
- er geen derde variabele C bestaat die de oorzaak is van zowel A als B en deze beide doet co-variëren zonder dat deze eigenlijk rechtstreeks met elkaar in verband staan.

Indien "A" de interventie is en "B" een effect dan is het duidelijk dat het een uitdaging is om het effect toe te schrijven aan de interventie. Het is namelijk mogelijk dat een andere externe oorzaak "C" het gewenste effect te weeg heeft gebracht. Omgekeerd is het mogelijk dat een gewenst effect dat effectief door de interventie werd gecreëerd negatief werd gecompenseerd door een externe oorzaak "C".

Een voorbeeld kan bovenstaande duidelijker maken: indien na een onderzoek blijkt dat in organisaties waar meer dan gemiddeld vrouwen aan het werk zijn er ook minder dan gemiddelde stressniveaus zijn, dan zou men verkeerdelijk kunnen besluiten dat de oververtegenwoordiging van vrouwen de oorzaak is van het lagere stressniveau. Er is echter enkel voldaan aan voorwaarde 1 wat betreft causaliteit. Indien men ook aan voorwaarde 2 wil voldoen dan moet men over meerdere metingen

over de tijd heen beschikken. Op die wijze kan men nagaan of de stressniveaus voorafgaan aan de proportie vrouwelijke medewerkers of omgekeerd. Nochtans is ook dit niet voldoende om een oorzakelijk verband aan te tonen. Er kan immers nog een derde factor in het spel zijn bvb. een gezinsvriendelijk beleid dat bepaalde organisaties hanteren waardoor ze én meer vrouwen aantrekken (omdat deze toch gemiddeld meer voor gezinstaken instaan) én aan de basis liggen van de lagere stressniveaus. Om dit aan te tonen is het, binnen het kwantitatief wetenschappelijk onderzoek, nodig een (quasi-)experimentele opzet te hanteren.

Al de hieronder besproken (quasi-)experimentele opzetten zijn deductief van aard. Ze gaan dus uit van een theoretisch inzicht dat getest zal worden (bvb. een interventie heeft een bepaald effect).

Experimentele opzetten kennen VOOR de interventie van start gaat toevalsgewijs een representatief staal personen uit de beoogde doelgroep toe aan een groep deelnemers en niet-deelnemers. Op basis van deze toekenning wordt dan aangenomen dat de twee groepen GEMIDDELD in alles (zowel invloeden van buitenaf als persoonseigen kenmerken) gelijk zijn.

Het basismodel is het volgende:

Figuur 9: basismodel experiment

	VOOR	Project	NA
Groep deelnemers	01	X	02
Controle groep	03		04

0 staat voor een observatie (meting). X staat voor het deelnemen aan het project. De opzet laat toe om na te gaan in hoeverre de situatie ten opzichte van de baseline (VOOR deelname aan het project) verbeterd is (NA deelnamen aan het project). Het verschil tussen 03 en 04 is immers de invloed van allerlei andere factor die afgetrokken moet worden van het verschil tussen 02 en 04 (het verschil tussen de groepen) om het netto-effect van de interventie in te schatten.

Zwakkere experimentele ontwerpen laten één of meerdere van de bovenstaande metingen vallen (bvb. geen voormeting van de controlegroep) of verschuiven metingen (bvb. geen voormeting van de controlegroep maar wel een meting tijdens het project).

Het is mogelijk om bovenstaande ontwerp nog rigoureuzer te maken op basis van meerdere voor –en na metingen die toelaten om op betere wijze de autonome trend in te schatten (evolutie van de controlegroep).

In de praktijk is de volledig toevalsgewijze toekenning echter bijna niet realiseerbaar o.a. omdat:

- de meeste interventies ofwel personen selecteren op basis van een niet toevalsgewijs criterium (bvb. hoogste nood) ofwel iedereen laten deelnemen;

- geselecteerde personen kunnen weigeren om deel te nemen aan de interventie;
- niet-geselecteerde personen kunnen weigeren deel te nemen aan het onderzoek;
- het aantal deelnemers te klein is om het toevalsmechanisme effectief te laten spelen.

Het is dus meestal nodig om gebruik te maken van *quasi-experimentele opzetten*. Deze proberen op een of andere manier te zorgen dat groepen met elkaar over de tijd heen vergeleken kunnen worden zonder toevalsgewijze toekenning. Hierin bestaan een aantal varianten.

In de eerste plaats gaat het om *opzetten met "niet equivalente groepen"*. Hierbij kunnen groepen gemaakt worden waarbij men op basis van karakteristieken waarvan men weet dat ze belangrijk zijn (mede de metingen beïnvloeden bvb. motivatie) een evenwichtige verdeling van respondenten over de groepen probeert te verkrijgen (bvb. evenveel ongemotiveerde als sterk gemotiveerde personen in zowel de controle als de deelnemersgroep). Deze procedure wordt "matching" genoemd⁸.

"Regression discontinuity" opzetten werken dan weer anders. Hierbij is het mogelijk om diegenen die de hoogste nood kennen te laten deelnemen aan het project en de anderen in een controle groep te plaatsen. Vervolgens wordt ook hier een VOOR en NA meting gebruikt voor de twee groepen. Voor elk van de twee groepen wordt vervolgens een regressieanalyse gemaakt. De eventuele "discontinuïteit" tussen de twee regressies is het effect van deelname.

Niet-experimentele opzetten

Het gaat hierbij om statische (metingen van respondenten op één punt in de tijd) ofwel om longitudinale opzetten (meerder punten in de tijd).

Niet-experimentele opzetten zijn in principe niet geschikt om causale vragen te beantwoorden. Zoals vermeld in het bovenstaande voorbeeld over het verband tussen de vertegenwoordiging van vrouwen en stress in organisaties, kunnen statische opzetten maar aan één van de drie voorwaarden van causaliteit voldoen nl. een probabilistische relatie vaststellen. Longitudinale opzetten kunnen ook de tijdsopvolging meenemen maar zijn ook niet in staat om op sluitende wijze de invloed van achterliggende variabelen uit te sluiten.

⁸ De meest vooruitstrevende techniek waarvan gebruik gemaakt wordt is propensity score matching. Deze techniek gebruikt beschikbare info over potentiële deelnemers om via statistische analyse uit te maken hoe verschillend/gelijkaardig deze zijn betreffende een gamma relevante variabelen en ze daarna in te delen in quasi-equivalente groepen. De statistische analyse gebeurt dan op basis van variantie analyse. Een alternatief wordt gevormd door de zogenaamde "difference in difference" ontwerpen waarbij statistische regressieanalyse toelaat om effecten te berekenen zonder matching. Een bijkomend voordeel is dat het niet nodig is om evenveel respondenten in elke groep te hebben.

Indien (quasi)-experimentele opzetten echter niet (meer) mogelijk zijn om causaliteit te onderzoeken, wordt soms toch gebruik gemaakt van longitudinale en zelfs statische niet-experimentele opzetten. Deze kunnen dan immers toch een tipje van de sluier oplichten over één of twee van de voorwaarden betreffende causaliteit. Hiertoe wordt een breed gamma aan statistische analyse technieken gebruikt⁹.

De niet-experimentele opzetten zijn wel aangewezen om, op basis van een steekproef uit de hele populatie van respondenten, de feiten in kaart te brengen (descriptieve wie, waar, wanneer, hoeveel,... vragen) of om deze te vergelijken met een norm. Deze opzetten kunnen uitgevoerd worden nadat het project begonnen is (in die mate dat respondenten de informatie kunnen geven die gevraagd wordt).

Kwalitatieve opzetten

Casestudy onderzoek

Evaluatief casestudy onderzoek betreft meestal meervoudig casestudy onderzoek waarbij verschillende cases doelgericht geselecteerd worden conclusies te kunnen trekken die gelden voor een ruimere context dan enkel die van één specifieke case. Binnen de context van de evaluatie van een ESF-project gaat het echter bijna per definitie om een enkelvoudige casestudy (tenzij projecten samenwerken wat betreft evaluatie).

Van belang is wel dat een causale casestudy niet louter beschrijvend is (wat gebeurt er wanneer) maar ook de causale dynamiek kan onderbouwen (wat gebeurt er wanneer en vooral, waarom). Theoretische inzichten over causale verbanden worden hierbij aan de feiten getoetst. Uiteraard is het zo dat de inzichten van de casestudy niet voorbij de specifieke context van de case zullen toegepast kunnen worden (i.t.t. de meervoudige casestudy).

Casestudies worden meestal toegepast als de focus ligt op het onderzoeken van een hedendaags fenomeen (bvb. het verloop van een organisationeel proces, veranderingen betreffende attitudes van deelnemers, etc.) binnen een dagdagelijkse context, vooral wanneer de grens tussen context en fenomeen niet zo helder is.

⁹ Een uitweiding hierover valt buiten de scope van deze handleiding. We verwijzen naar de uitgebreide literatuur op het vlak van univariate en multivariate statistische analyse. De meest recente ontwikkelingen betreffende de analyse met behulp van structurele vergelijkingsmodellen van zowel statische als longitudinale data (bij deze laatste meer specifiek cross-lagged panel analyse en analyse van groeipaden) evenals de analyse van tijdsreeksen (indien men beschikt over een groot aantal meetpunten) verdienen hierbij een hoge mate van aandacht.

Andere vormen van wetenschappelijk onderzoek incl. “mixed methods”

Binnen het wetenschappelijk onderzoek dient een onderscheid gemaakt te worden tussen inductieve en deductieve benaderingen. De inductieve benadering vertrekt vanuit data om een theoretisch inzicht op te bouwen. De deductieve vertrekt vanuit een theoretisch inzicht om met data het inzicht te testen (ofwel wordt het inzicht weerlegd ofwel wordt het onderbouwd). Het gaat dus om twee fundamenteel verschillende benaderingen. De inductieve benadering is echter niet geschikt om een inzicht te weerleggen of te onderbouwen. Het deductieve proces is dan weer niet geschikt om een theorie te creëren. Het eindproces van inductie is het startpunt van deductie.

De hierboven geschetste benaderingen werd vanuit een deductief perspectief besproken. Indien voor de start van een deductief evaluatie onderzoek echter blijkt dat niet duidelijk is welke inzichten eigenlijk onderbouwd moeten worden dan is het aangewezen om eerst een ander proces op te starten. Dit proces kan gebaseerd zijn op een synthese van reeds bestaande theoretische inzichten door middel van een literatuurstudie ofwel op een inductieve onderzoeksbenadering waarbij men data verzamelt om op basis daarvan een theoretisch inzicht op te bouwen. Men spreekt dan van een exploratieve studie.

Case studies kunnen niet enkel gebruikt worden om op deductieve wijze een theoretisch inzicht te testen (zie supra). Ze kunnen ook dienen om op inductieve wijze vorm te geven aan theoretische inzichten. Andere kwalitatieve inductieve benaderingen zijn de etnografie, narratief onderzoek, “grounded theory” en de fenomenologie. Deze benaderingen hebben soms zeer specifieke richtlijnen wat betreft de inzet van dataverzameling –en analysemethoden en de focus van het onderzoek (bvb. etnografie is gericht op het onderzoeken van cultuur op basis van langdurige onderdompeling in de te bestuderen context).

Er bestaan ook kwantitatieve analyse technieken zoals cluster analyse en exploratieve factor analyse die passen binnen een inductieve opzet.

In de gangbare evaluatiepraktijk wordt zelden gebruikt gemaakt van één van de varianten van een inductieve wetenschappelijke evaluatiebenadering. Meestal wordt er ten behoeve van exploratie niet-wetenschappelijk te werk gegaan of beperkt men zich tot een literatuurstudie. Nochtans kan een goed opgezette wetenschappelijke inductieve studie zeer nuttig zijn (bvb. bij een ondersteunende evaluatie rond de evalueerbaarheid van een project –zie 6.d) infra).

De combinatie van een eerste inductieve opzet (meestal kwalitatief), gevolgd door een hierop voortbouwende deductieve opzet (meestal kwantitatief) valt onder de noemer ‘mixed methods’ waarbij verschillende opzetten gecombineerd worden in één onderzoek.

Er zijn echter nog andere “mixed methods” benaderingen

- Triangulering van onderzoekopzetten: een kwalitatieve en kwantitatieve benadering (bvb. een casestudy en een experiment) worden naast elkaar gebruikt en de resultaten worden met elkaar vergeleken.
- Deductieve sequentiële opzet: deductieve kwalitatieve en kwantitatieve benaderingen volgen elkaar in de tijd op met de bedoeling elkaars conclusies verder te onderbouwen of toe te lichten.

c) Niet wetenschappelijke opzetten

Algemeen

Het onderscheid met wetenschappelijk onderzoek ligt erin dat bij niet-wetenschappelijk onderzoek:

- een omkaderende onderzoekopzet ontbreekt die er op gericht is om betreffende een bepaalde soort vraag de hoogste mate van wetenschappelijke betrouwbaarheid en validiteit te waarborgen;
- er meestal geen expliciete benadering is wat betreft de steekproef (noch statistisch, noch doelmatig)
- toetsing aan de criteria van “wetenschappelijke betrouwbaarheid en validiteit” ondergeschikt is aan:
 - de dynamiek die deze opzetten op gang kunnen brengen (bvb. door middel van doorgedreven participatie van belanghebbenden);
 - de snelheid waarmee meer klaarheid over knelpunten kan verkregen worden;
 - de grotere toegankelijkheid voor leken.

Het is duidelijk dat de grens tussen wel en niet-wetenschappelijk vaag is. Dit geldt vooral voor de inductieve kwalitatieve benaderingen waarbij in een evaluatiecontext in vele gevallen abstractie gemaakt wordt van de omkaderende opzet. Nochtans is de omkaderende opzet, ook voor kwalitatieve inductieve benaderingen wel van belang. Indien men bvb. in een etnografische studie gebruik zou maken van het instrument “directe observatie”, dan gebeurt dit meestal over een lange tijdsperiode waarbij de onderzoeker zich zoveel mogelijk onderdompelt in de te bestuderen realiteit. Dit is toch iets anders dan een evaluator die één of twee keer ter plekke komt observeren. Nochtans gaat het telkens om hetzelfde dataverzamelingsinstrument nl. “directe observatie”.

Indien gekozen wordt voor een niet-wetenschappelijke opzet dan is de kans groot dat men de benodigde vaardigheden binnen de eigen organisatie kan vinden of deze vaardigheid relatief snel kan opbouwen (bvb. op basis van de gedetailleerde fiches die op de website van het ESF-Agentschap te vinden zijn). Uiteraard kan men ook beroep doen op externe ondersteuners, wat vooral bij het faciliteren van groepen een troef kan zijn.

De methoden die gebruikt worden in een niet-wetenschappelijke benadering verschillen naargelang de gebruiker één van volgende doelen nastreeft:

- informatie inwinnen
- prioriteiten stellen
- coherentie verzekeren wat betreft de doelen van een project

De onderstaande opsomming is uiteraard niet exhaustief.

Informatie inwinnen

Bij het inwinnen van informatie zijn volgende methoden aangewezen:

- Sleutelinterviews
- Focusgroepen
- Vragenlijsten
- Observatie
- “Mystery” shoppers

Sleutelinterviews worden gebruikt om in een individueel gesprek bij een relatief beperkt aantal personen te peilen naar hun zicht op een situatie of idee. Deze interviews kunnen semi-gestructureerd (met een aantal sleutelvragen of onderwerpen) zijn of volledig ongestructureerd (vrije conversatie met enkel een breed onderwerp als startpunt). Het gebruik van volledig gestructureerde open vragen (die bij alle respondenten op exact dezelfde wijze voorgelezen worden) is eerder zeldzaam. Men zal dan eerder beroep doen op gesloten vragen (zie infra).

Een *focusgroep* is een discussie onder een kleine groep belanghebbenden (8-12 personen) begeleid door een gespreksleider. Net zoals met semi-gestructureerde interviews is er vooraf reeds bepaald over welke vragen men zal discussiëren. Het aantal vragen is relatief beperkt (2-5 hoofdvragen). De meeste focusgroepen streven niet naar het bereiken van consensus maar naar het peilen van de reikwijdte wat betreft verschillende percepties en meningen.

De meest gebruikte methodiek is waarschijnlijk die van de *gesloten vragenlijsten*. Een voorbeeld van een aanpak gestoeld op een vragenlijst is het zelfevaluatie-instrument ontwikkeld door het ESF-Agentschap in het kader van het EQUAL-programma¹⁰. Andere voorbeelden zijn vragenlijsten die gebruikt worden na het einde van een evenement (bvb. na een training) of om de tevredenheid van klanten te bevragen betreffende de dienstverlening. Kenmerken voor gesloten vragenlijsten is dat respondenten keuze hebben tussen een beperkt aantal antwoordmogelijkheden. Meestal worden de vragen ook op een gestandaardiseerde wijze (telefonisch, persoonlijk of via brief/web).

¹⁰ Zie website van het ESF-Agentschapi

Met directe observatie kan men rechtstreeks een situatie of een proces waarnemen. Het is heel nuttig in het geval er twijfel is of respondenten wel eerlijk op, bijvoorbeeld, een vragenlijst willen (het gaat misschien over een gevoelig onderwerp) of kunnen (het gaat misschien over onbewust gedrag) antwoorden. Het is ook nuttig als techniek om meer inzicht te krijgen in situaties die niet goed begrepen worden vooraleer men andere methoden gaat gebruiken.

“Mystery” shoppers zijn personen die de rol van “klant” opnemen om vervolgens (op basis van een vragenlijst of via een interview) te rapporteren over hun ervaringen. De kwaliteit van de dienstverlening kan op die manier beoordeeld worden. Deze personen worden hiertoe ook opgeleid. Het gaat eigenlijk om een specifieke vorm van observatie.

Prioriteiten stellen

Bij het stellen van prioriteiten zijn volgende technieken aangewezen:

- Nominale groep
- Expert panels
- Delphi

Bij de *nominale groep methode* is het de bedoeling om van een groep deelnemers zoveel mogelijk ideeën met betrekking tot een centrale vraag/probleemstelling te bekomen. Er wordt gewerkt met rondes waarin de deelnemers individueel hun idee opschrijven. Na elke ronde worden de resultaten gepresenteerd maar niet bediscussieerd. Deze presentatie bevordert het bedenken van nieuwe ideeën. Pas als alle ideeën uitgeput zijn, is er een discussieronde. De sessie eindigt met een (meestal geheime) stemming over welke de beste ideeën zijn (men komt tot een ranking).

Expert panels worden gebruikt wanneer er een hoog niveau van expertise vereist is en/of de complexiteit van het onderwerp zeer groot is. Het panel zal het opgegeven onderwerp bestuderen en in een schriftelijke rapportage conclusies en aanbevelingen naar voren schuiven. Er worden meerdere bijeenkomsten georganiseerd waarbij men er in een eerste sessie voor zorgt dat het panel ondergedompeld wordt in de materie en inzicht krijgt in het proces. Het is mogelijk dat besloten wordt dat het onderzoek van het onderwerp op wetenschappelijke leest moet gebeuren. Dan is het expert panel uiteraard te situeren binnen de reeds besproken wetenschappelijke aanpak.

Bij *Delphi panels* wordt een groep deskundigen in verschillende rondes op afstand bevraagd. Vervolgens krijgen ze elkaars antwoorden te zien met een focus op die antwoorden die significant afwijken van de meerderheid. Dit wordt herhaald tot een consensus bereikt wordt. Een variant is dat men niet tot consensus probeert te komen wat betreft het antwoord op de centrale vraag, maar wel komen tot een consensus wat betreft de redenen en het belang van de afwijkende standpunten.

Coherentie verzekeren

Bij het verzekeren van coherentie wat betreft de doelen van een project kunnen volgende methoden gebruikt worden:

- Faciliteren van conceptualisatie
- Kwantitatieve concept mapping

Met *het faciliteren van conceptualisatie* worden een reeks technieken bedoeld waarbij ideeën/concepten in een groepsessie geïnventariseerd worden en vervolgens op één of andere wijze geclusterd worden tot meer abstracte concepten. Er bestaan vele variaties op deze techniek gaande van mind mapping tot metaplan workshops. Ook het maken van probleembomen in een workshop behoort tot deze technieken. Al deze variaties hebben gemeenschappelijk dat er, via een gefaciliteerd groepsgebeuren, structuur wordt aangebracht in een complex onderwerp.

Kwantitatieve concept mapping is dan weer een techniek waarbij brainstormsessies gehouden worden om uitspraken van belanghebbenden op te leveren. Deelnemers zullen tevens de uitspraken een rating geven (bvb. op een schaal van 1 tot 5) en de uitspraken clusteren. Vervolgens wordt gebruik gemaakt van exploratieve statistische analyse¹¹ om deze output van de brainstormsessies te analyseren en een conceptuele kaart voort te brengen. Het gaat dus eigenlijk om een zeer specifieke vorm van faciliteren van conceptualisatie.

d) Welke methodologie past bij welke benadering?

Op basis van benadering en evaluatievragen

Idealiter wordt de keuze voor methodologie enkel gedreven door de aan de doelstelling aangepaste evaluatiebenadering en de daarmee gepaard gaande evaluatievragen.

Onderstaand wordt samengevat weergegeven welke methoden passen bij welk type evaluatie. Tussen haakjes staat de aard van de vraag aangegeven (bij de wetenschappelijke benaderingen). Dit wordt dan in de tekst verder toegelicht.

¹¹ Multi-dimensional scaling en cluster analyse

Figuur 10: overzicht keuze methode

Methode Type	Niet- wetenschappelijk	Wetenschappelijk		
		(Quasi)- experimenteel	Niet-experimenteel	Kwalitatief
Deblokkering	X			
Evalueerbaarheid	X		X (exploratief)	X (exploratief)
Getrouwheid			X (Vgl. met norm)	
Prestatiegedreven effectenevaluatie		X (Causaal: wat)		
Theoriegedreven effectenevaluatie		X (Causaal: hoe en wat)		X (Causaal: hoe en wat)
Theoriegedreven procesevaluatie			X (Vgl. met norm/ descriptief)	

Ondersteunende evaluaties doen meestal geen beroep op wetenschappelijke methoden. Deze evaluaties zijn immers bedoeld om relatief snel knelpunten op te lossen. Een uitzondering kan de beoordeling van de evalueerbaarheid van het project zijn indien een exploratieve wetenschappelijke (meestal kwalitatieve) onderzoeksopzet gebruikt wordt om op inductieve wijze een effectenschema in kaart te brengen.

Evaluaties van de prestaties maken meestal wel gebruik van wetenschappelijke methoden omwille van de doelstelling om verantwoording af te leggen aan belanghebbenden. De voorkeur gaat hierbij uit naar kwantitatieve methoden. In het geval van een getrouwheidsevaluatie is dit meestal een niet-experimentele opzet om de feiten te vergelijken met een norm terwijl dit in het geval van de effectenevaluatie eerder een (quasi)experimentele opzet zal zijn omwille van de interesse in causaliteit.

Theoriegedreven evaluaties zullen ook eerder gebruik maken van wetenschappelijke methoden. Niet-wetenschappelijke methoden zijn immers minder geschikt om een complexe situatie op een omvattende en objectieve wijze te doorgronden.

Idealiter wordt bij een theoriegedreven effectenevaluatie gebruik gemaakt van een mix van kwalitatieve en kwantitatieve methoden. Wat betreft deze theoriegedreven effectenevaluatie is de vraagstelling primair causaal.

Wat betreft de theoriegedreven procesevaluatie is deze gericht op het vergelijken van feiten met een norm of louter het beschrijven van de feiten (zonder vergelijking). De toegepaste methode is bij uitstek een niet-experimentele methode.

Op basis van benodigde middelen

De kostprijs van de methodologie speelt echter ook een grote rol in de keuze van een methode. Het is evident dat kleinere projecten verhoudingsgewijs minder middelen aan evaluatie kunnen besteden. Soms is het dan ook aangewezen om een minder geschikte methodologie te gebruiken die echter realistisch is wat betreft de benodigde middelen.

Algemeen gesproken zijn wetenschappelijke methoden duurder dan niet-wetenschappelijke¹².

6. Organisatie van evaluatie

a) Plannen van de evaluatie als component van een projectvoorstel

In de meeste gevallen zal een ESF oproep een getrouwheidsevaluatie als minimumvereiste voor evaluatie vooropstellen. Indien het grotere projecten betreft is het ook mogelijk dat een prestatiegedreven effectmeting verplicht is. De optiek die voor het ESF-Agentschap primeert is immers dat er verantwoording afgelegd wordt aan alle belanghebbenden. Dit betekent echter niet dat er financiële gevolgen verbonden kunnen worden aan de evaluatie indien blijkt dat er problemen bij de uitvoering waren. Indien uit een evaluatie echter blijkt dat er sprake is van een schending van de ESF-reglementering dan zal dit uiteraard wel gevolgen hebben.

Andere soorten evaluatie worden meestal niet opgelegd, maar wel aanbevolen.

Bij het plannen van evaluatie is het nuttig om met volgende principes rekening te houden:

1. De evaluatiefunctie moet zo vroeg mogelijk in het project worden opgenomen, ongeacht de benadering. De latere evaluatieactiviteiten zullen daardoor gemakkelijker en ook succesvoller verlopen.
2. Bij het aanvangen van hun opdracht moeten de (interne of externe) evaluatoren er voor zorgen dat een haalbaar evaluatieplan op tafel ligt. Ze moeten toelichten hoe de resultaten van

¹² Indicaties van kostprijs worden op de website van het ESF-Agentschap meegegeven o.a. op basis van informatie uit de evaluatierapporten van projecten.

de evaluatie aangewend kunnen worden. Ze moeten ook voor ogen houden dat evaluatie totaal zinloos is, wanneer de bevindingen ervan niet worden (mee)gedeeld aan besluitvormers.

3. Belanghebbenden moeten, in de mate van het mogelijke, vanaf het vroegste stadium bij de evaluatie worden betrokken. Dit zorgt ervoor dat in het evaluatieontwerp en in het evaluatieplan rekening wordt gehouden met hun prioriteiten en agenda's. Het is dan ook aangewezen een stuurgroep samen te stellen (zie infra). Anderzijds kan het noodzakelijk zijn om selectief op te treden ten aanzien van de "stemmen" die de evaluatieagenda uiteindelijk bepalen, zodat voldoende gericht gewerkt kan worden en de evaluatie ook daadwerkelijk beheersbaar is.
4. De evaluatievragen in een evaluatieontwerp zijn uitermate belangrijk. Zoniet komt men al snel in de verleiding om grote hoeveelheden gegevens te verzamelen en soms technisch gesofisticeerde analyses te produceren, die slechts heel weinig bijdragen tot de praktijk of het beleid. Uiteraard is het niet gemakkelijk om evaluatievragen zo te formuleren dat ze ook worden beantwoord. Hoewel dit een technische aangelegenheid is, bestaat hier ook een overkoepelende zorg betreffende het gebruik van de antwoorden.

In hoofdstuk 3 werd reeds de timing van de evaluatie in samenhang met de evaluatiedoelen besproken. De rapportering naar het ESF toe wordt hierdoor, zoals vermeld in hoofdstuk 3, grotendeels bepaald.

b) Wie voert de evaluatie uit?

De evaluatie kan intern of extern ondersteund worden.

Er zijn een aantal voor- en nadelen verbonden aan het gebruik van interne of externe evaluatoren. Een interne evaluator heeft in principe een beter zicht op en toegang tot relevante informatiebronnen en heeft meer voeling met de context (inclusief interne gevoeligheden). Langs de andere kant is het mogelijk dat er een zekere belangenvermenging is. Dit kan zowel leiden tot eenzijdig positieve of negatieve beoordelingen afhankelijk van de interne dynamiek in de organisatie.

Externe evaluatoren hebben dan weer het voordeel van een onbevangen en ongebonden blik op de interventie. Hierdoor kunnen ze gemakkelijker negatieve conclusies formuleren en een nieuwe wind laten waaien.

Indien het gaat om verantwoording af te leggen aan externe belanghebbenden dan is het dan ook geloofwaardiger als de evaluatie extern uitgevoerd wordt. Bij knelpunten oplossen stelt zich dit probleem van geloofwaardigheid minder. Dan zijn contextgevoeligheid en toegang tot informatie weer belangrijker.

Wat betreft leren is meestal doorslaggevend of de benodigde expertise al dan niet intern of extern te vinden is en, indien dit het geval is, of die expertise beschikbaar kan worden gesteld binnen de beoogde termijn. Ook de vraag of externe disseminatie van lessen of internalisering in de organisatie van groter belang is, speelt mee om respectievelijk externe of interne expertise aan te werven.

c) Wie volgt de evaluatie op en hoe?

Een interactieve benadering, bijvoorbeeld via regelmatige bijeenkomsten van een stuurgroep, zorgt voor bewustwording rond de keuzes die tijdens de looptijd van het project moeten worden gemaakt. Verder zorgt dit ervoor dat de evaluatieteams voldoende ondersteuning krijgen, dat de informatie toegankelijk is en dat wordt gebriefd over mogelijke veranderingen van de context.

Deelnemers aan een stuurgroep moeten aan volgende criteria voldoen:

- Beschikbaar zijn (voldoende tijd kunnen besteden)
- Expertise wat betreft het onderwerp van de evaluatie
- Bereid om eigen perspectief naar voren te brengen én open te staan voor andere perspectieven
- Bereid om samen te werken met het evaluatieteam en eventuele vertrouwelijkheid te respecteren.

Eén van de taken van de stuurgroep is om de kwaliteit van het de evaluatie te bewaken. De kwaliteit kan continu bewaakt worden aan de hand van volgende criteria:

- Kwaliteitscriteria betreffende het evaluatieproces:
 - Goed opgestelde opdracht: Een interne of externe evaluatieopdracht dient minimaal volgende elementen te bevatten:
 - Beschrijving van het kader (Europese en Vlaamse context, ESF oproep, etc.)
 - Reikwijdte van de evaluatie (het hele project of slechts onderdelen)
 - Voornaamste gebruikers van en belanghebbenden bij de evaluatie
 - Evaluatievragen
 - Beschikbare informatie
 - Voornaamste methodes of technieken die moeten worden gebruikt
 - Tijdsschema
 - Indicatief budget
 - Vereiste kwalificaties van het team
 - Vereiste structuur van het door de evaluatoren in te dienen voorstel
 - Regels voor indiening en criteria voor selectie en toewijzing
 - Degelijk selectieproces bij het aanduiden van evaluatoren
 - Effectieve dialoog en feedback doorheen het evaluatieproces

- Goede beschikbaarheid van informatie om mee te werken
- Een goed management en dito coördinatie door het evaluatieteam
- Kwaliteitscriteria betreffende de resultaten van het evaluatiewerk:
 - Tegemoet komend aan de noden zoals bepaald in de opdracht
 - Relevante reikwijdte en dekking van de projectactiviteiten
 - Verdedigbaar ontwerp en methodes
 - Betrouwbare gegevens
 - Goede analyse
 - Geloofwaardige resultaten met betrekking tot analyse en gegevens
 - Onpartijdige conclusies zonder vooringenomenheid en die getuigen van gezond verstand
 - Duidelijk verslag met managementsamenvatting en ondersteunende gegevens in bijlagen.
- Effectieve verspreiding van verslagen/onderzoeksoutputs aan de stuurgroep en andere belanghebbenden

Wat dit laatste betreft moeten de analyses en conclusies van de evaluatie(s) ook bij de relevante (tussentijdse) rapportering voor het ESF-Agentschap gevoegd worden.

7. Afstemming met evaluatie op programmaniveau

Het ESF-Agentschap laat zelf op regelmatige wijze evaluaties uitvoeren over verschillende projecten en/of oproepen heen. Indien reeds gepland dan zal dit ook in de oproepfiche vermeld staan, inclusief eventuele verwachtingen naar promotoren toe.

8. Opnemen van de evaluatieaanpak in een projectvoorstel

Volgende componenten moeten bij het opstellen van een projectvoorstel in acht genomen worden:

- Doel evaluatie: de mogelijke doelen van de evaluatie werden besproken in hoofdstuk 3. Indien meerdere doelen gekozen worden moet voor elk doel een aparte aanpak met alle onderstaande componenten opgesteld worden.
- Soort evaluatie: de soorten evaluatie naargelang doel en timing van de evaluatie werden besproken in hoofdstuk 4. Het overzicht wordt geboden in Figuur 7.
- Evaluatievragen: de verschillende soorten evaluatievragen naargelang evaluatiecriteria en soort evaluatie werden ook besproken in hoofdstuk 4. Het overzicht wordt geboden in Figuur 8.
- Opzet evaluatie: de verschillende soorten opzet werden besproken in hoofdstuk 5. Het overzicht naargelang soort evaluatie en evaluatievragen wordt geboden in Figuur 10

- Externe of interne evaluatoren: de keuze van de evaluatoren werd toegelicht in hoofdstuk 6b).
- Stuurgroep evaluatie: de samenstelling en rol van de stuurgroep werd besproken in hoofdstuk 6c) .
- Tijdstip evaluatie en deliverables van de evaluatie (link met rapportering): hier moet een planning van de evaluatieactiviteiten opgesteld worden incl. bijeenkomsten van stuurgroepen en het opleveren van rapporten. Dit werd toegelicht in hoofdstuk 6a).
- Budget: tenslotte moet een globaal budget opgesteld worden voor de evaluatieactiviteiten. Dit bestaan uit een tijdsbesteding voor internen (aantal dagen) en/of externe loonkosten.

9. Literatuur

Op de website van het ESF (www.esf-agentschap.be) wordt wordt zo veel mogelijk ondersteunende literatuur (ingedeeld per hoofdstuk van de leidraad) gratis digitaal ter beschikking gesteld.