

Leren van Evalueren

*Onderzoek naar het gebruik van evaluatieonderzoek
bij het ministerie van Buitenlandse Zaken*

Voorwoord

De kwaliteit en effectiviteit van buitenlands beleid gaat iedereen aan het hart. Er staat immers veel op het spel voor Nederland. Of het nu gaat over overwegingen van politieke aard of over veiligheid, het klimaat of economische ontwikkeling. Beleidsevaluatie is belangrijk om inzicht te krijgen in de resultaten van het beleid dat door het ministerie van Buitenlandse Zaken (BZ) is gevoerd, de inspanningen die daarvoor geleverd zijn en de middelen die daarmee gemoeid zijn. Het ministerie streeft naar goed gebruik van evaluatie. Evaluaties waar geen gebruik van wordt gemaakt zijn zonde van de tijd maar vooral zonde van het geld. Goed gebruik impliceert dat evaluatieonderzoek wordt benut om te komen tot een kwalitatief betere beleidsontwikkeling en uitvoering én betere beleidsverantwoording. Beide functies zijn van belang, mede in het licht van de rijksbrede Regeling Periodiek Evaluatieonderzoek en beleidsinformatie 2006 (RPE 2006).

In Nederland is sprake van toenemende aandacht voor beleidsverantwoording. Een belangrijk uitgangspunt voor de RPE 2006 is dat vakministers verantwoordelijk zijn voor de doeltreffendheid en de doelmatigheid van het beleid dat aan hun begroting ten grondslag ligt. Om te kunnen bepalen of het al dan niet bereiken van de gestelde doelstellingen toe te schrijven is aan het gevoerde beleid, de doeltreffendheid van het beleid, is evaluatieonderzoek noodzakelijk. Een tweede belangrijk uitgangspunt is dat vakministers ook verantwoordelijk zijn voor het periodiek evalueren van de doeltreffendheid en de doelmatigheid van het beleid. Met behulp van evaluatieonderzoek dienen vakministers zich tegenover het parlement te verantwoorden over het gevoerde beleid. Dit is ook vastgelegd in de Comptabiliteitswet (CW). Naast verantwoorden is leren van evalueren cruciaal omwille van kwalitatief betere beleidsontwikkeling. Het gaat daarbij steeds vaker om samen leren. In een internationale werkomgeving is samenwerking met anderen van belang. Evaluatiediensten werken ook steeds meer samen. Het werkterrein van het ministerie is complex en reikt ver. Evaluaties geven informatie over of we de juiste dingen doen, of we op de juiste weg zijn en of er betere manieren zijn om hetzelfde te doen. Evaluaties geven een beoordeling van de relevantie, effectiviteit, duurzaamheid en/of efficiëntie van de inspanningen van dit ministerie. Bij leren van evalueren gaat het niet om een modieuze trend. Het gaat gewoon om beter beleid. De bevindingen en aanbevelingen van het voorliggende onderzoek zijn daarbij van nut.

De Directie Financieel Economische Zaken (FEZ) heeft dit onderzoek uitgevoerd op verzoek van het Audit Committee. De doelstelling van dit onderzoek was om meer inzicht te krijgen in het huidige gebruik van evaluaties in de beleidsontwikkeling en uitvoering alsmede beleidsverantwoording met het oog op de identificatie van knelpunten en het aandragen van oplossingen. Het onderzoeksteam bestond uit hoofdonderzoekers Frans L. Leeuw van de Universiteit in Maastricht en Ger H. C. van Gils van BeleidsOnderzoek & Advies (BOA) en onderzoeksmedewerkers Marie Gildemyn en Liesbeth Inberg van het Instituut voor Ontwikkelingsbeleid en -beheer van de Universiteit van Antwerpen. De begeleidingsgroep van dit onderzoek bestond uit de heer H. Jorritsma van de Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie (IOB) en de heer H. Pelgrøm van de Directie Effectiviteit en Coherentie (DEC). De groep stond onder voorzitterschap van Margret Verwijk (FEZ). De leden van de groep hebben waardevol commentaar gegeven op eerdere versies van dit rapport, waarvoor onze dank. Dank gaat ook naar iedereen die als respondent heeft meegewerkt aan dit onderzoek. De namen staan vermeld bij de geraadpleegde bronnen.

| 4 |

De onderzoekers zijn verantwoordelijk voor de inhoud van dit rapport.

Peter Arnoldus
Directeur Financieel Economische Zaken

Inhoud

Voorwoord	3
Lijst van afkortingen	7
Hoofdbevindingen en aanbevelingen	9
1. Inleiding	16
1.1. Aanleiding voor het onderzoek (en doelstelling)	17
1.2. Doel van het onderzoek en onderzoeksopzet	17
1.3. Inhoud van het rapport	21
2. Achtergrondinformatie evaluaties en evaluatiebeleid	22
2.1. Evaluatiebeleid rijksbreed	23
2.2. Hoezo evalueren?	24
2.3. Evaluatiebeleid bij het ministerie	25
3. Inventarisatie van evaluaties (2004-2007)	28
3.1. Inleiding	29
3.2. Typen van evaluatieonderzoek	29
3.3. Aanleiding voor de evaluatie	30
3.4. Opdracht en uitvoering	30
3.5. Beoogde gebruikers	32
3.6. Presentatie en communicatie evaluatieresultaten	33
4. Gebruik van evaluaties: een beknopt literatuuroverzicht	34
4.1. Inleiding	35
4.2. Vormen van gebruik van evaluatieonderzoek	36
4.3. Factoren van belang voor benutting	38
5. Analyse van gebruik van 30 evaluaties	44
5.1. Inleiding	45
5.2. Gebruik van evaluaties	45
5.3. Invloed op beleid	49
5.4. Factoren gebruik evaluatieresultaten	53

6. Leren van evaluaties: een beknopt literatuuroverzicht en resultaten uit het empirisch onderzoek	62
6.1. Inleiding	63
6.2. Enige theoretische noties over leren	63
6.3. Leren van evaluaties bij het ministerie van Buitenlandse Zaken	66
7. Conclusies, aanbevelingen en discussie	72
7.1. Conclusies	73
7.2. Aanbevelingen	81
Bijlage	
Bijlage 1. Terms of Reference	86
Bijlage 2. Overzicht evaluaties in de steekproef	96
Bijlage 3. Instrumenten en faciliteiten voor kennisuitwisseling	100
Bijlage 4. Geraadpleegde personen	108
Bijlage 5. Geraadpleegde bronnen	110

Lijst van afkortingen

AMAD-OS	Ambassadeur in Algemene Dienst voor OS-aangelegenheden
BOA	BeleidsOnderzoek & Advies
BZ	Ministerie van Buitenlandse Zaken
CW	Comptabiliteitswet 2001
CNV	Christelijk Nationaal Vakverbond
DEC	Directie Effectiviteit en Coherentie
DG	Directeur-Generaal of Directoraat Generaal
DGES	Directoraat-Generaal Europese Samenwerking
DGIS	Directoraat-Generaal Internationale Samenwerking
DGPZ	Directoraat-Generaal Politieke Zaken
DGRC	Directoraal-Generaal Regiobeleid en Consulaire Zaken
EDP	<i>Enabling Development Policy</i>
EU	Europese Unie
FEZ	Directie Financieel Economische Zaken
IOB	Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie
IPDET	<i>International Programme for Development Evaluation Training</i>
KOS	Kennis en Onderzoeksstrategie
KREO	Kenniscentrum Rechtstaat en Ontwikkelingssamenwerking
MDTF	<i>Multi Donor Trust Funds</i>
MSSM	<i>Maryland Scale of Scientific Methods</i>
NFP	<i>Nederlands Fellowship Programme</i>
NIMF	Nederlands Investerings Matchings Fonds
NONIE	<i>Network of Networks Impact Evaluation Initiative</i>
NORAD	<i>Norwegian Agency for Development Cooperation</i>
NPT	<i>Netherlands Programme for the Institutional Strengthening of Post- Secondary Education and Training Capacity</i>
OECD	<i>Organisation for Economic Co-operation and Development</i>
OESO DAC	Organisatie voor Economische Samenwerking en Ontwikkeling - <i>Development Assistance Committee</i>
ORET	Ontwikkelingsrelevante Exporttransacties
ORIO	Ontwikkelingsrelevante Infrastructuur Ontwikkeling
OS	Ontwikkelingssamenwerking
PlvDGRC	Plaatsvervangend Directeur-Generaal Regiobeleid en Consulaire Zaken
RPE	Regeling Periodiek Evaluatieonderzoek en beleidsinformatie 2006

SRGR	Seksuele en Reproductieve Gezondheid en Rechten
TMF	Thematisch Medefinancierings Programma
TOR	<i>Terms of Reference</i>
VBTB	<i>Van Beleidsbegroting Tot Beleidsverantwoording</i>
VMP	<i>Vakbonds Medefinancierings Programma</i>
WFP	<i>United Nations World Food Programme</i>
WODC	Wetenschappelijk Onderzoek en Documentatie Centrum (ministerie van Justitie)
WRA	Wetenschappelijk Raadadviseur (ministerie van Buitenlandse Zaken)

Hoofdbevindingen en Aanbevelingen

Inleiding

De doelstelling van het onderzoek was om meer inzicht te krijgen in het huidige gebruik van en leren door evaluaties in de beleidsontwikkeling en uitvoering, alsook bij de beleidsverantwoording, om zowel knelpunten te identificeren als oplossingen aan te dragen. Het onderzoek heeft betrekking op evaluaties die deel uitmaken van de evaluatieprogrammering van het ministerie van Buitenlandse Zaken in de jaren 2004 t/m 2007.

De centrale vragen voor het onderzoek waren:

1. of evaluatieonderzoeken worden gebruikt voor de verbetering van beleid en/of beleidsuitvoering (de leerfunctie);
2. of evaluatieonderzoeken worden gebruikt voor het verstrekken van informatie over effectiviteit en doelmatigheid van het gevoerde beleid (de verantwoordingsfunctie); en
3. welke factoren en omstandigheden het gebruik van evaluatieonderzoek beperken dan wel bevorderen en hoe beperkingen zijn op te heffen.

191

De bevindingen zijn gebaseerd op een uitgebreide analyse van het gebruik via een steekproef van 30 evaluaties. Het onderzoek kent twee belangrijke informatiebronnen: interviews met beleidsmedewerkers ofwel de beoogde gebruikers van de evaluaties en documentenonderzoek. De analyse is mede gebaseerd op een verkenning van de (internationale) literatuur over de relatie tussen onderzoek en beleid.

Hoofdbevindingen

1. Evaluaties bij het ministerie van Buitenlandse Zaken laten diverse vormen van gebruik zien

Er zijn verschillende vormen van gebruik van evaluaties aangetroffen. In de eerste plaats worden evaluaties gebruikt als communicatiemiddel: intern vooral ter ondersteuning van de overdracht bij functie(wisselingen). Evaluaties lenen zich daarvoor omdat ze vaak het enige document zijn dat

een overzicht van het beleid bevat. Deze vorm van gebruik vindt alleen op individueel initiatief plaats; er wordt vanuit de organisatie niet op gestuurd. Een enkele keer worden evaluaties gebruikt om een intern debat te voeren. In de externe communicatie worden evaluaties vooral gebruikt om een standpunt of beleid te legitimeren, of als basis voor overleg met organisaties die regelingen uitvoeren of subsidies ontvangen. Sommige respondenten onderscheiden ook de promotiefunctie van evaluaties. Ze bieden een kans om jezelf weer eens in de schijnwerpers te plaatsen. In een enkel geval heeft een evaluatie een rol gespeeld als een middel om beleid systematisch vast te leggen, zodat verschillende partijen zich er aan kunnen binden of gebonden worden. Evaluaties worden gebruikt ter verantwoording van beleid door onder meer toezending van evaluaties van de Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie (IOB), voorzien van een beleidsreactie, aan de Tweede Kamer, door publicatie van rapporten op de website van het ministerie en door gebruik als informatiebron voor het jaarverslag van het ministerie en de resultatenrapportages over ontwikkelingsamenwerking.

| 10 |

2. De relatie tussen evaluaties bij het ministerie van Buitenlandse Zaken en beleidswijzigingen is complex en veelvormig

Veel evaluaties binnen het Directoraat-Generaal Internationale Samenwerking (DGIS) gaan gepaard met een herziening van het geëvalueerde beleid. Hoewel dit niet zonder meer wil zeggen dat de evaluaties verantwoordelijk zijn voor de aanpassingen of er in belangrijke mate toe bijdragen, is de samenhang op geaggregeerd niveau opvallend. Bij de andere directoraten gaan de evaluaties meestal gepaard met beperktere aanpassingen, en ook wel met het achterwege blijven van enige actie. Het is moeilijk om in afzonderlijke gevallen een lineair en causaal verband aan te tonen. In een aantal gevallen lijken de conclusies en aanbevelingen van de evaluatie de erop volgende beleidswijziging niet helemaal te kunnen rechtvaardigen. Daarnaast komt het voor dat beleidsaanpassingen vooruit lopen op conclusies en aanbevelingen van de evaluatie. Overigens is dat een bevinding, die vaker in de literatuur over gebruik van evaluaties is te vinden. De geschetste relatie tussen evaluaties en beleidsaanpassingen, die eruit voortvloeien, is als volgt verklaarbaar. Ten eerste wordt een deel van de conclusies en aanbevelingen van evaluatieonderzoek al tijdens het onderzoek via deelname aan referentie-

groepen bekend en ter harte genomen. Een aantal respondenten vindt een referentiegroep dan ook de best denkbare leeromgeving om iets van een evaluatie op te steken. Voorts zijn belangrijke ideeën voor beleidswijzigingen uit meerdere bronnen afkomstig. Evaluaties uit de evaluatieprogrammering van het ministerie van Buitenlandse Zaken maken deel uit van een beleidsdebat, waaraan verschillende informatiebronnen een bijdrage leveren. Een beleidsaanpassing, die volgt op een aanbeveling in een evaluatie, kan ook gebaseerd zijn op een eerdere aanbeveling uit een andere bron.

Niettemin is er over het algemeen sprake van een aanzienlijke mate van overeenstemming tussen de inhoud en teneur van evaluaties en die van het beleidsdebat en de beleidswijzigingen. De meest bescheiden conclusie is dat in elk geval een deel van de door evaluaties gesignaleerde problemen en voorgestelde oplossingen op de beleidsagenda staat.

3. Het type kennis dat evaluaties opleveren is een belangrijke factor bij benutting en leren van evaluaties

| 11 |

Een belangrijke factor van invloed op gebruik van evaluaties is het type kennis dat evaluaties produceren. Evaluaties leveren weinig substantiële kennis op over hoe resultaten te bereiken. De meeste evaluaties hebben een bestuurskundig en procesmatig karakter. Indien effecten worden bepaald, worden deze in verband gebracht met de vraag of het beleid is uitgevoerd zoals oorspronkelijk bedoeld, met een reconstructie van beleidsintenties en afspraken.

Dat betekent niet dat deze procesmatige evaluaties geen gebruikswaarde hebben. Ze bevatten waardevolle informatie over hoe beleidsprocessen volgens recente inzichten in te richten. Ze maken echter niet duidelijk op welke manier een gewenst resultaat is te bereiken of waarom de processen volgens de instructies dienen te worden ingericht. De wet van de afnemende meeropbrengsten kan voor dit type evaluatie mogelijk in werking zijn getreden. Bovendien heeft substantieve kennis over hoe gedrag te beïnvloeden een grotere generaliseerbaarheid en biedt daarom meer aanleiding en gelegenheid tot leren. Er is ook sprake van enige onduidelijkheid over de aard van de kennis die wordt geleverd. Cruciale evaluatieve begrippen als 'doelmatigheid' en 'doeltreffendheid' worden niet altijd consistent gebruikt. De evaluaties in de steekproef kennen een duidelijk aanwijsbare doelgroep en hebben doorgaans een beperkte algemene geldigheid.

Voorts dragen extra inspanningen voor de verspreiding van evaluaties niet bij tot meer gebruik of tot gebruik in bredere kring.

Er is geen evaluatie-vijandige opstelling aangetroffen; medewerkers zijn vooral geïnteresseerd in bruikbare resultaten. Bij de directoraten, die vooral diplomatieke inspanningen plegen, heerst meer twijfel over het nut van evaluaties vanuit de vraag of het mogelijk is om resultaten van diplomatieke inspanningen te meten. Vrijwel nergens bestaat een spontane voorstelling van het leren van (gedragswetenschappelijk georiënteerde) evaluaties over de doeltreffendheid van interventies en beleid. Verder is het niet gebruikelijk om actief kennis uit evaluaties te zoeken. Men beperkt zich tot de evaluaties van beleid en programma's betreffende de eigen portefeuille en zoekt niet naar lessen, die elders zijn te leren.

4. Definitie van en visie op leren bepalend voor beoordeling van leren van evaluaties

| 12 |

Er zijn meerdere definities van en visies op leren mogelijk. Welke wordt aangehangen is in belangrijke mate bepalend voor de beoordeling van leren van evaluaties binnen het ministerie. Er zijn voorbeelden van leerprocessen naar aanleiding van evaluaties, maar het zijn geen standaardpraktijken. De gebruikers van de evaluaties geven vaak zelf desgevraagd aan dat zij niet het idee hebben iets nieuws te leren van evaluaties. Dat is in verband te brengen met de aard van de kennis, die de evaluaties doorgaans produceren. Voorts is er geen sprake van een echt leergierige cultuur. Het is niet gebruikelijk om actief kennis uit evaluaties te zoeken. Tot slot is het duidelijk dat de ambities met betrekking tot leren binnen het ministerie verder reiken dan wat tot nu toe wordt gerealiseerd. Dat blijkt uit klachten over het gebrek aan leren, maar ook uit diverse initiatieven om het ministerie een lerende organisatie te maken. Deze initiatieven vormen echter nog geen duidelijk en samenhangend beleid, gestoeld op een visie over wat leren voor de organisatie moet en kan betekenen. Er zijn wel noemenswaardige noties en aanzetten, die bouwstenen voor een dergelijke visie en beleid leveren. Belangrijk is dat het verschil tussen individueel en institutioneel leren binnen het departement door beleidsbepalende medewerkers wordt onderkend. Een andere interessante aanzet is de kennisagenda waarmee met een veelheid aan kleine stapjes, die elk afzonderlijk gemakkelijk worden geaccepteerd, wordt geprobeerd belangrijke veranderingen in de

leercultuur te bewerkstelligen. Dat geldt ook voor de notie van voortdurende her- en bijscholing om kennis op peil te houden en dit te stimuleren via het beoordelings- en beloningsbeleid. Een laatste interessante notie betreft de manier waarop onenigheid over de conclusies van evaluatieonderzoek kan leiden tot leren omdat daarbij lessen niet noodzakelijkerwijs veronachtzaamd worden. Er ontbreekt echter ontbreekt een visie op leren.

Aandachtspunten

1. **Schep meer duidelijkheid over wat bij het ministerie onder (gebruik en) leren (van evaluaties) verstaan wordt**

Het ontbreekt aan een theorie of visie over wat leren is, waarom het belangrijk is dat er geleerd wordt, wie leert en dient te leren, of afleren ook leren is, wat er gebeurt als er niet geleerd wordt, wat de factoren zijn die bepalend zijn om tot een lerende omgeving en cultuur te komen, welk type kennis het meest leerzaam is en of er beperkingen gelden of gesteld worden aan leren. De laatste vraag is vooral van belang in verband met een eventueel aan leren aangepast personeelsbeleid. Het is daarom belangrijk om, in lijn met (buitenlandse) literatuur en ervaringen, tot een theorie te komen over wat leren op verschillende niveaus inhoudt en hoe organisatie en medewerkers geprikkeld kunnen worden tot leren (van evaluaties).

| 13 |

2. **Zorg voor digitale (webgebaseerde) en andere voorzieningen om gebruik van kennis en leren te bevorderen**

Het ministerie beschikt in beperkte mate over webgebaseerde- en digitale voorzieningen om het leren van onderzoeken te bevorderen. Faciliteiten zoals *repositories* (kennisschatkamers), die eenvoudig toegankelijk zijn, voor het opslaan, beheren en gebruiken van digitale informatie waaronder de resultaten van evaluaties, zijn van groot belang en verdienen verdere ontwikkeling. Het consequent instellen van zorgvuldig samengestelde referentiegroepen waaraan in elk geval de beoogde gebruikers van evaluaties deelnemen, kan een belangrijk effect hebben op het gebruik en leren van evaluaties.

3. Heb meer oog voor het type kennis dat evaluaties opleveren in relatie tot het gebruik ervan en de behoefte aan kennis

In de eerste plaats is het van belang dat studies voorzien in de specifieke behoeften aan evaluatieve kennis van verschillende directoraten met uiteenlopende operationele doelstellingen om gebruik en leren van evaluaties (verder) te bevorderen. Daarnaast gaat het om de gerichtheid op de reconstructie en toetsing van interventietheorieën. Waarbij meer aandacht nodig is voor de analyse van sociaal- en gedragswetenschappelijke mechanismen, die aan het succes of het falen van interventies en programma's ten grondslag liggen.

4. Doe meer aan syntheses maar dan *state of the art*¹

Er zou meer werk gemaakt moeten worden van systematische of wetenschappelijke reviews, die er via synthese van bevindingen op gericht zijn een antwoord te geven op beleidsvragen. Daarbij is het aan te bevelen om meer in lijn te werken met de standaarden uit de internationale evaluatie-literatuur over syntheses en ook gebruik te maken van bevindingen en onderzoek van derden.

| 14 |

5. Zorg dat verantwoording kan worden afgelegd over de aard en mate van het gebruik van evaluatieonderzoek

Zonder veel moeite is het mogelijk om het gebruik van evaluatieonderzoek in het kader van de beleidscyclus bij te houden. Daarbij kan geleerd worden van wat elders in de Rijksdienst voorhanden is.

¹ volgens standaard kwaliteitsnormen voor het maken van syntheses

1

Inleiding

1.1. Aanleiding voor het onderzoek (en doelstelling)

Jaarlijks laat het ministerie van Buitenlandse Zaken een aanzienlijk aantal evaluaties uitvoeren. Het ministerie streeft naar een goed gebruik van deze evaluaties. Goed gebruik betekent dat er van evaluatieonderzoek gebruik gemaakt wordt ter verbetering van zowel het beleid als de verantwoording over het beleid. Het ministerie heeft dit onderzoek laten uitvoeren om meer inzicht te krijgen in het gebruik van evaluaties in de praktijk van de beleidscyclus.

Het onderwerp van onderzoek is in de tijd beperkt tot evaluaties die deel uitmaken van de evaluatieprogrammering en die zijn afgerond in de kalenderjaren 2004 t/m 2007 onder (mede-) verantwoordelijkheid van het ministerie. Dit betekent dat ruim 70 afgeronde evaluaties onderwerp van dit onderzoek zijn geweest.

1.2. Doel van het onderzoek en onderzoeksopzet

| 7 |

De centrale vraag voor het onderzoek is of evaluatieonderzoeken daadwerkelijk worden benut voor (a) verbetering van beleid en/of beleidsuitvoering (de leerfunctie) en (b) voor informatieverstrekking over effectiviteit en doelmatigheid van het gevoerde beleid (de verantwoordingsfunctie en de kwaliteit van de beleidsinformatie). Beantwoording van deze vraag dient inzicht te leveren in hoe gebruik van evaluaties kan worden geoptimaliseerd.

Direct verbonden met deze centrale vraag zijn de volgende deelvragen:

- Is het gebruiksdoel (van de evaluaties) helder gespecificeerd en zijn de evaluatie-uitkomsten tijdig beschikbaar?;
- Sluiten onderzoeksvragen, bevindingen, conclusies en aanbevelingen in de te onderzoeken evaluaties op elkaar aan?; en
- Zijn de gepresenteerde aanbevelingen praktisch uitvoerbaar en liggen deze binnen het bereik van de ontvangers?

Resterende deelvragen zijn meer inventariserend van karakter:

- Welke typen van evaluatieonderzoek zijn in de periode 2004-2007 in het kader van de evaluatieprogrammering uitgevoerd (tijdstip van de evaluatie, aard en omvang van het onderwerp van evaluatie, etc.)?;

- Wat was de aanleiding voor de evaluatie?;
- Hoe is de evaluatie uitgevoerd?;
- Wie waren de ontvangers van de onderzoeksresultaten?;
- Wanneer en op welke wijze waren de belanghebbenden (beleidsmakers) betrokken bij de uitvoering van de evaluatie?; en
- Op welke wijze zijn de evaluatieresultaten gepresenteerd en gecommuniceerd?

Analytische benadering

De aard en omvang van het onderzoek dat is verricht laat een (statische) causale analyse van relaties tussen geobjectiveerde kenmerken van evaluaties en gebruik niet toe. In plaats daarvan is gekozen voor een meer kwalitatieve benadering. Deze is gebaseerd op de veronderstelling dat gebruikers van evaluaties belangrijke redenen voor het al of niet gebruik zelf kunnen onderkennen en een inschatting maken van het relatieve belang van verschillende redenen en factoren. Interviews met betrokkenen, de beoogde gebruikers van de evaluaties, zijn daarom een eerste belangrijke informatiebron. In deze interviews zijn zij ondervraagd over de evaluaties, de aard en mate van benutting en achtergronden. Er wordt van uit gegaan dat gebruikers van evaluaties belangrijke redenen voor al of niet gebruik kunnen onderkennen en ook een inschatting kunnen maken van het relatieve belang van verschillende redenen en factoren.

| 18 |

Dit uitgangspunt strookt met een belangrijk kenmerk van de ‘*realist evaluations-benadering*’, namelijk dat deelnemers aan beleidsprocessen via eigen inzicht er mede voor zorg dragen dat interventies, in dit geval evaluatieonderzoek, invloed kunnen hebben (Pawson and Tilley 2007). Dit betekent uiteraard niet dat het gebruik van evaluaties eenvoudig is te reconstrueren. Een deel van het proces verloopt weinig bewust of is mede afhankelijk van zaken die onvolgende problematisch zijn voor gebruikers en daarom niet de aandacht krijgen die ze verdienen. Een uitgebreide analyse van bijpassende documentatie is daarom de tweede belangrijke informatiebron voor het onderzoek.

De afgelopen twee decennia lag de nadruk in onderzoek van het gebruik van evaluaties en bij maatregelen om gebruik te bevorderen doorgaans op het management van onderzoek, op een goede communicatie van onderzoeksresultaten en op een zekere standaardisering van kwaliteit door middel van het gebruik van *checklists*. Over het algemeen werd er minder aandacht besteed aan

de vraagzijde van onderzoek, ofwel aan kenmerken en wensen van de beoogde en potentiële gebruikers van beleidsonderzoek. Uit recent wetenschappelijk onderzoek blijkt echter dat het effect van evaluaties meer afhangt van inhoudelijke kenmerken van evaluaties, zoals het soort kennis dat ze opleveren, hun methodologische kwaliteit en van de beleidscontext waarin ze tot stand komen of gebruikt dienen te worden, dan van gestandaardiseerde kwaliteitseisen of de marketing van rapporten met behulp van communicatie-activiteiten (Johnson et al 2009). Zowel de wensen van beoogde gebruikers van onderzoek, de methodologische kwaliteit van evaluaties en het soort kennis dat ze leveren stonden daarom in dit onderzoek centraal.

Voor het bepalen van de methodologische kenmerken en kwaliteit worden de gangbare standaarden gebruikt zoals die van de Wereldbank, de *Organisation for Economic Co-operation and Development* (OECD), *Network of Networks Impact Evaluation Initiative* (NONIE) en het ministerie van Buitenlandse Zaken. Deze kunnen ook inzicht geven in relevante indelingen van de soorten kennis die evaluaties kunnen leveren.

Bijzondere aandachtspunten met betrekking tot de vraagzijde van evaluaties zijn:

- de opleiding en ervaring van de beleidsmakers met het oog op vertrouwde-heid met en kennis van onderzoek, het vermogen onderzoek te beoordelen en gebruiken, de vraag hoe men in het algemeen nieuwe kennis verwerft en gebruikt en wat de houding en cultuur is ten aanzien van evaluaties en het verwerven van kennis;
- de verhouding tussen vraag en aanbod van onderzoek op strategisch en operationeel niveau: krijgen beleidsmakers het soort kennis uit evaluaties waaraan zij behoefte hebben (bijvoorbeeld kennis over naleving van afspraken en procedures, kennis over processen, kennis over effecten, kennis over oorzaken);
- de relatie tussen beleidsmedewerkers en onderzoek, onderzoekmanagement en onderzoekers, zowel institutioneel als informeel;
- de (cognitieve en politieke) beleidscontext van de evaluatie, ofwel de stand van het beleidsdebat, vragen die er leefden, gebruikte argumenten, zekerheden die bestonden en elementen van het geëvalueerde beleid die al of niet ter discussie stonden. Deze context is van belang om het gebruik van evaluatieresultaten te beoordelen, en kan zowel het gebruik van evaluaties bevorderen of belemmeren.

Uitvoering van het onderzoek

Het onderzoek kende twee onderdelen.

Het eerste deel bestond uit een korte verkenning van de beleidscyclus en de evaluatiepraktijk van het ministerie, uit een korte literatuurstudie ter voorbereiding van het onderzoek en tot slot uit een inventarisatie van de evaluaties die in de periode 2004 tot en met 2007 zijn uitgevoerd. De literatuurstudie diende ter inventarisatie van elementen die van belang kunnen zijn voor het bepalen van gebruik van evaluaties en factoren die daarbij een rol spelen. De inventarisatie is verricht op basis van bovenstaande resterende deelvragen.

Het tweede deel van het onderzoek bestond uit een analyse van een selectie van 30 evaluaties. Het gebruik van deze evaluaties is onderzocht inclusief de factoren die daarbij een rol spelen. Daarbij is gebruik gemaakt van informatie uit evaluatierapporten, de *Terms of Reference* (TOR) voor de evaluaties, de beleidsreacties op de evaluaties en informatie uit diverse overige documentatie en interviews.

| 20 |

Om een evenwichtig beeld van het gebruik van de evaluaties te waarborgen, is een steekproef samengesteld op basis van de volgende criteria:

- verdeling over IOB-onderzoeken en niet-IOB-onderzoeken naar verhouding van het aantal evaluatieonderzoeken per directoraat-generaal in de evaluatieprogrammering;
- verdeling over de verschillende directoraten-generaal naar verhouding van het totaal aantal evaluatieonderzoeken per directoraat-generaal in de evaluatieprogrammering; en
- verdeling over verschillende typen evaluatieonderzoek, waarbij in elk geval een impactevaluatie en een beleidsdoorlichting in de selectie opgenomen dienden te zijn. Dit zijn nieuwe typen onderzoek die in de periode waarop dit onderzoek betrekking heeft, zijn geïntroduceerd. Bovendien bestaat de verwachting dat deze typen bijzondere gebruiksmogelijkheden bieden.

Resultaten van onderzoek kunnen ook nog lang na publicatie langs vele omwegen en op diverse manieren kleine of grotere bijdragen leveren aan beleid. In zowel interviews als documentatie is daar eveneens aandacht aan besteed. Voor de interviews zijn beleidsmedewerkers benaderd die direct

verantwoordelijk zijn voor het beleid waarop de evaluatie gericht was. In een enkel geval is een interview gevoerd met een leidinggevende. Vanwege de grote mobiliteit van medewerkers was het voor de meeste evaluaties niet mogelijk om één aanspreekpunt te vinden, die het gehele proces had meegemaakt, van initiatief voor de evaluatie tot en met de periode na oplevering. Bijlage 4 bevat een overzicht van geïnterviewde personen.

De interviews waren half gestructureerd. Er is gewerkt met lijsten met gespreksonderwerpen waarbij er voldoende ruimte bleef bestaan om de argumentatie van de respondenten te volgen. Verder is gebruik gemaakt van een uitgebreidere lijst met aandachtspunten om te kunnen doorvragen op factoren en argumenten die in het interview ter sprake kwamen. Een deel van de vragen was gericht op een reconstructie van het gebruik van evaluaties en andere vragen waren gericht op beperkende danwel bevorderende factoren voor gebruik.

1.3. Inhoud van het rapport

Hoofdstuk 2 geeft achtergrondinformatie over het beleid inzake evaluaties. Hoofdstuk 3 behandelt de onderzoeksvragen. Hoofdstuk 4 beschrijft en analyseert de literatuur over gebruik van evaluaties en gaat in op de vraag welke vormen van gebruik van evaluaties worden onderscheiden en welke factoren van belang zijn voor een effectief gebruik. De eerste onderzoeksvraag, of evaluaties daadwerkelijk worden gebruikt, is beantwoord door een uitgebreide en gedetailleerde analyse van een steekproef van 30 evaluaties. De nadruk lag daarbij op de afzonderlijke evaluaties, op de reacties en acties van de gebruikers of het achterwege blijven daarvan. Daarbij is gekeken naar de beleidscontext van de evaluatie zoals boven omschreven. Hoofdstuk 5 bevat de verkorte versie van deze analyse. In dit hoofdstuk wordt ook de vraag beantwoord of gebruik ook betekent dat evaluaties van invloed zijn op het beleid. Dit hoofdstuk behandelt daarnaast de factoren die in de evaluatiepraktijk van het ministerie het meest relevant zijn voor een effectief gebruik van evaluaties. Hoofdstuk 6 gaat in op de vraag of gebruik ook leidt tot leren van evaluatieresultaten. Hoofdstuk 7, tot slot, bevat de conclusies en aanbevelingen van het onderzoek. De bijlagen informeren over de onderzoeksopzet en de gebruikte bronnen.

2

Achtergrondinformatie evaluaties en evaluatiebeleid

2.1. Evaluatiebeleid rijksbreed

Evaluaties van overheidsbeleid komen er in de kern op neer dat onderzocht wordt hoe beleid tot stand is gekomen, is ingevoerd en welke effecten en/of neveneffecten dit beleid heeft of heeft gehad. Hoewel de geschiedenis van evaluatieonderzoek in Nederland enkele decennia langer terug gaat, bestaat er sinds 1991 een kabinetsstandpunt dat het evaluatiebeleid van de Nederlandse centrale overheid beschrijft (Leeuw 2009). Later is dit aangevuld met richtlijnen die te vinden zijn in de Comptabiliteitswet (2001) en in lagere regelgeving (RPE 2006).

Evaluatiebeleid omvat regels, afspraken en principes die een organisatie heeft geformuleerd om als richtsnoer te dienen voor besluiten over en handelingen ten aanzien van het doen van evaluaties. Recent is in een brief van de Minister van Financiën over dit onderwerp (18 december 2008) een aantal aspecten van het evaluatiebeleid nogmaals op een rijtje gezet. Een belangrijk uitgangspunt voor de Regeling Periodiek Evaluatieonderzoek en beleidsinformatie (RPE 2006) is dat vakministers verantwoordelijk zijn voor de doeltreffendheid en de doelmatigheid van het beleid dat aan hun begroting ten grondslag ligt. Dit is vastgelegd in artikel 20 van de Comptabiliteitswet (CW). Om te kunnen bepalen of het al dan niet bereiken van de gestelde doelstellingen toe te schrijven is aan het gevoerde beleid – de doeltreffendheid van het beleid – is evaluatieonderzoek noodzakelijk. Met deze evaluaties dienen vakministers zich tegenover het parlement te verantwoorden over het gevoerde beleid. Daarbij wordt onderscheid gemaakt tussen verschillende soorten onderzoek: ex ante, ex post, proces evaluaties, effect (soms ook impact genoemd) evaluaties en beleidsdoorlichtingen.

Onderdeel van het rijksbrede evaluatiebeleid is, naast het aspect van de onafhankelijkheid van het onderzoek en de periodiciteit ervan, vooral ook de bruikbaarheid en het gebruik. Immers, evaluatieonderzoek is niet onderzoek omwille van onderzoek. Evaluatieonderzoek staat in het teken van het benutten van de onderzoeksresultaten en het leren daarvan, opdat toekomstig beleid effectiever en efficiënter wordt. Dat betekent ook dat er op goede gronden kan worden afgezien van (nieuw) beleid, omdat uit onderzoek blijkt dat andere maatschappelijke krachten beter in staat zijn om de gestelde maatschappelijke doelen te realiseren.

2.2. Waarom evalueren?

Het meest eenvoudige antwoord is omdat het moet. Zoals hierboven is toegelicht, hebben overheden zichzelf de taak opgelegd kosten, proces, effecten en neveneffecten van hun beleid (en organisatie) te (laten) evalueren, een activiteit die in het verlengde ligt van het oudere controle-onderzoek naar de rechtmatigheid van uitgaven en ontvangsten van overheden door onder andere de Rekenkamers. Echter, ook voordat er van een evaluatiebeleid sprake was, waren tal van landen al actief op dit terrein, te beginnen bij de Verenigde Staten en Canada. Inmiddels is evaluatie-onderzoek een wereldwijde activiteit geworden, zoals de *International Evaluation Atlas*, die de situatie in ruim 20 landen beschrijft, laat zien (Furubo, Rist and Sandahl 2002). Deze atlas markeert de tientallen beroepsverenigingen, de verschillende *systems of evaluation*, die eigen handboeken, leerstoelen, organisaties en geldstromen kennen, maar ook voorschriften en evaluatie-*frameworks* (Leeuw and Furubo 2008).

| 24 |

Welke andere overwegingen zijn er om werk te maken van beleidsevaluaties?

- Overheden dienen transparant te zijn en over hun activiteiten, doelen, ingezette middelen en doelbereiking verantwoording af te leggen;
- Waar er *market failures* zijn, zijn er ook *government failures*. Om die te vinden, respectievelijk te herkennen en de oorzaken ervan te begrijpen, is evaluatieonderzoek onmisbaar;
- Het stimuleert overheden om te leren van ervaringen en gemaakte fouten (Leeuw, Rist and Sonnichsen 1994). In plaats van zoals door von Mises verwoord fout op fout te maken, kan overheidsbeleid preciezer zijn in het detecteren van voor welke problemen en maatschappelijke actoren welk beleid relevant en bruikbaar is (von Mises 1976). Soms vertoont overheidsbeleid trekken van beleidshomeopathie, of te wel van beleid dat niet gebaseerd is op werkzame (gedrags)mechanismen². Een van de taken van evaluatieonderzoek is ook dát te onderkennen;

² Leeuw, Frans L. 2008. Oratie Gedragsmechanismen achter overheidsinterventies en rechtsregels. Universiteit Maastricht.

- Het vergroot het wetenschappelijk inzicht en levert een bijdrage aan theorievorming. Beleid kan daar de vruchten van plukken. Langs die weg kunnen evaluaties mogelijk bijdragen aan het realiseren van een meer *evidence-based* overheidsbeleid ofwel beleid dat werkt.

2.3. Evaluatiebeleid bij het ministerie

Bij het ministerie van Buitenlandse Zaken staat evalueren in dienst van het afleggen van verantwoording over bereikte resultaten en het leren van ervaringen uit het verleden. Dat betekent dat evaluaties een rol spelen in het streven naar een beleidsverantwoording, zoals vastgelegd in de nota 'Van Beleidsbegroting Tot Beleidsverantwoording' (VBTB). Daarbij dienen de evaluaties te voldoen aan de kwaliteitseisen zoals vastgelegd in de RPE 2006 die integraal onderdeel van de Comptabiliteitswet is. De leerfunctie van evaluaties speelt op diverse niveaus. Het gebruik van of het geven van follow-up aan individuele evaluatieonderzoeken is de verantwoordelijkheid van de betreffende Directeur-Generaal of directie. Een belangrijk element van het gebruik van evaluaties is de opstelling van een beleidsreactie door of namens de betreffende bewindspersoon. In het geval van IOB-evaluaties en in het geval van beleidsdoorlichtingen worden het evaluatierapport en de beleidsreactie altijd aan de Kamer aangeboden. Bij andere evaluatieonderzoeken wordt van geval tot geval bezien of aanbieding van het evaluatierapport aan de Tweede Kamer wenselijk is. Door het geven van een beleidsreactie op het rapport geeft de betrokken bewindspersoon aan welke lessen uit de bevindingen worden getrokken. Tijdens de uitvoering van evaluaties ontvangen ambtenaren vaak tussentijds al informatie, bijvoorbeeld via concept-rapporten. Hierdoor komen feitelijke bevindingen al eerder beschikbaar voor de bij de uitvoering betrokken ambtenaren en instellingen.

IOB is de belangrijkste uitvoerende en coördinerende actor binnen het ministerie van Buitenlandse Zaken waar het evaluatieonderzoek betreft, maar er zijn ook andere actoren die een rol spelen. De evaluatieprogrammering van het ministerie bestrijkt een periode van zes jaar en wordt jaarlijks bijgesteld op basis van een voortrollende programmering. De Directie Financieel Economische Zaken (FEZ) is samen met IOB verant-

woordelijk voor de coördinatie van de evaluatieprogrammering inclusief de advisering aan beleidsdirecties over evaluatie en evaluatiebeleid. Een deel van de evaluatieonderzoeken wordt uitgevoerd door IOB. Een ander deel van de evaluatieonderzoeken wordt uitgevoerd door of in opdracht van beleidsdirecties.

Naast IOB en FEZ is een aantal andere actoren en voorzieningen bij het ministerie van belang voor het verwerven en gebruik van kennis uit onderzoek of andere bronnen. Zo is de Directie Effectiviteit en Coherentie (DEC) onder meer verantwoordelijk voor sturing op resultaat om enerzijds het functioneren van het DGIS op een kwalitatief hoger niveau te brengen en anderzijds de effecten van samenwerking en hulpverlening te vergroten. De directie richt zich daarbij op landen- en thema-overstijgende onderwerpen. Daarnaast speelt de ontwikkeling van kennis- en onderzoeksstrategieën (KOS) een rol. In het kader van training en opleiding zijn er mogelijkheden als deelname aan het *International Program for Development Evaluation Training* (IPDET) vanuit de Wereldbank en Carleton University. Overige leer- en ontwikkelmogelijkheden zijn er via de Hoofddirectie Personeel en Organisatie en FEZ-leerportalen. Daarnaast zijn te noemen de Academie voor Internationale Samenwerking (IS-Academie), de Wetenschappelijk Raadsadviseur (WRA) en de Ambassadeur in Algemene Dienst voor OS-aangelegenheden (AMAD-OS). Hoofdstuk 6 gaat kort in op genoemde functies en faciliteiten.

3

Inventarisatie van evaluaties (2004-2007)

3.1. Inleiding

Alle evaluaties, geprogrammeerd voor de jaren 2004 tot en met 2007, zijn opgenomen in een inventarisatie aan de hand van zes deelvragen (zie TOR). Voor een steekproef van 30 evaluaties zijn bovendien drie additionele deelvragen beantwoord, die verbonden zijn met de centrale vraag. Dit hoofdstuk geeft een kort overzicht van de belangrijkste gegevens. De hier gepresenteerde resultaten zijn niet alleen op de inventarisatie gebaseerd, maar ook op gegevens uit het onderzoek van de steekproef.

3.2. Typen van evaluatieonderzoek

In de evaluatieprogrammering van 2006 is er nog geen onderscheid gemaakt tussen effectenonderzoek en overig evaluatieonderzoek. In de evaluatieprogrammering van 2008 is dit onderscheid wel gemaakt. De categorie beleidsdoorlichting is duidelijk te onderscheiden van overige vormen van evaluatieonderzoek. Zowel de steekproef als de groep van evaluaties buiten de steekproef bevatten een beleidsdoorlichting. Het specifieke karakter van beleidsdoorlichtingen komt bij de bespreking van dit type evaluatie aan de orde.

| 29 |

De evaluatietypen effectenonderzoeken (*ex post*) plus effectenonderzoek *ex post* en overig evaluatieonderzoek zijn nader gecategoriseerd in termen van de *Maryland Scientific Methods Scale* (MSMS). Dit is een vijfpuntsschaal die een uitspraak kan doen over de interne validiteit van effectevaluaties (Farrington en Welsch 2005). Op de MSMS is score 3 (quasi-experimenteel *design*) de minimale onderzoeksopzet die nodig is om verantwoorde conclusies te trekken over de effectiviteit van een maatregel. Niveau 4 is een quasi-experiment met omvangrijke statistische controles terwijl niveau 5 de *randomized controlled trial* is. Lagere niveaus dan 3 betreffen onderling samenhangende studies.

Van de 24 effectonderzoeken in de steekproef hebben 20 rapporten een niveau 1 op de MSMS, twee rapporten een niveau 2 en één rapport deels niveau 1 en deels niveau 2. Het evaluatieonderzoek 'De impactevaluatie drinkwater en sanitaire voorzieningen in Tanzania' is een twijfelgeval. De evaluatie maakte weliswaar gebruik van een controlegroep, maar er is geen

echte nulmeting gebruikt. Er is een nulmeting gereconstrueerd op basis van historische gegevens. Met deze kanttekening is de evaluatie te beschouwen als een evaluatie van het niveau 3: een quasi-experimenteel design.

3.3. Aanleiding voor de evaluatie

De aanleiding voor een evaluatie wordt in veel rapporten niet specifiek aangegeven, maar valt wel af te leiden uit de tekst. De beschrijvingen van de aanleidingen zijn heel divers. Er is getracht om een onderscheid te maken tussen de aanleiding voor een evaluatie en het gebruiksdoel ervan ofwel de inhoudelijke reden om de evaluatie uit te voeren. De aanleiding is meer de reden, gegeven het gebruiksdoel, om de evaluatie nu en niet eerder of later uit te voeren. Het is niet helemaal gelukt om dit onderscheid aan te brengen. De aanleiding voor de evaluaties zoals omschreven in de rapporten en ToRs is vaak niet duidelijk te onderscheiden van het gebruiksdoel.

| 30 |

Betrekkelijk duidelijk te onderscheiden als aanleidingen zijn het verstrijken van een termijn voor periodiek te ondernemen evaluaties en toezeggingen aan de Tweede Kamer. Bij vijf evaluaties in de steekproef en vijf van de overige evaluaties is sprake van een toezegging aan de Kamer als een aanleiding voor evaluatie. Bij 13 evaluaties in de steekproef wordt melding gemaakt van het verstrijken van de tijd en/of de verplichting tot een periodieke evaluatie. Hetzelfde geldt voor 16 van de overige evaluaties.

3.4. Opdracht en uitvoering

Opdrachtgevers zijn het ministerie van Buitenlandse Zaken (soms in samenwerking met andere ministeries) en andere instellingen (vaak internationale organisaties of daaraan gelieerde donoren, en soms de Algemene Rekenkamer). Doorgaans is IOB uitvoerder, al of niet in samenwerking met externe consultants en/of bureaus. In een enkel geval betreft het buitenlandse zusterinstellingen. Verder is er een onderscheid tussen uitvoering door externen en (zelfstandige) uitvoering door directies. Evaluaties, opgenomen in de evaluatieprogrammering van het ministerie, zijn in meerderheid een opdracht van het ministerie. Deze evaluaties zijn zowel door IOB als door externe partijen uitgevoerd. Er staan ook externe

evaluaties in de programmering. Dat zijn evaluaties, die in opdracht van andere partijen dan het ministerie van Buitenlandse Zaken zijn uitgevoerd. Een klein aantal betreft een gezamenlijke opdracht van het ministerie en een externe instelling. Doorgaans zijn deze externe instellingen internationale organisaties, ministeries in andere landen of zusterinstellingen van IOB. De volgende tabel verschaft een overzicht van het opdrachtgeverschap.

Tabel 1: Opdrachtgevers van alle evaluaties uit de evaluatieprogrammering (2004 - 2007)

Evaluaties	Opdrachtgever				
	BZ	BZ m.b.v. externen	Extern	Onbekend	Totaal
Onderdeel van steekproef	21	3	5	1	30
Buiten Steekproef	21	6	12	3	42
Totaal	42	9	17	4	72

Gebruiksdoel

In de meeste rapporten (28) in de steekproef wordt een gebruiksdoel aangegeven, al dan niet expliciet. Het doel is over het algemeen omschreven in termen van verantwoord en/of leren waarbij het specifieke beleidsveld of -onderdeel waarvoor de resultaten van belang zijn staat vermeld.

Zijn evaluatie uitkomsten tijdig beschikbaar?

Van slechts drie evaluaties waren uitkomsten volgens planning beschikbaar; 15 evaluaties waren niet tijdig beschikbaar en voor 12 evaluaties was geen planning gespecificeerd in het evaluatierapport en/of de *Terms of Reference*. Het is echter moeilijk om de tijdige oplevering van een evaluatie op basis van de openbaar beschikbare gegevens te beoordelen. Vaak worden resultaten al op een andere wijze gecommuniceerd. Van slechts één evaluatie uit de steekproef hebben de gebruikers aangegeven dat de resultaten te laat beschikbaar waren voor een effectief gebruik.

Hoe ziet een evaluatierapport er uit?

De evaluatie bestaat doorgaans uit een eindrapport. Indien er deelstudies zijn uitgevoerd is er een synthese-rapport plus deelstudierapporten. In een enkel geval zijn de deelstudies alleen op cd-rom beschikbaar.

3.5. Beoogde gebruikers

Voor de meeste evaluaties (24 van de evaluaties in de steekproef en 34 van de overige evaluaties) worden in de rapporten en/of de TOR de ontvangers van de onderzoeksresultaten of de beoogde gebruikers niet vermeld. Wel is in de programmering, rapportage of TOR vaak de verantwoordelijke directie vermeld.

De volgende tabel geeft een overzicht van de evaluaties in de steekproef per directoraat-generaal.

| 32 |

Tabel 2: Evaluaties¹ in steekproef naar verantwoordelijk Directoraat Generaal (evaluatieprogrammering 2004 - 2007)

Directoraat - Generaal	Aantal evaluaties in steekproef
Directoraat - Generaal Politieke Zaken	5
Directoraat - Generaal Internationale Samenwerking	15
Directoraat - Generaal Regiobeleid en Consulaire Zaken	5
Directoraat - Generaal Economische Samenwerking	5
Totaal	30

Van de evaluaties in de steekproef zijn 22 rapporten aangeboden aan de Tweede Kamer, waarvan 19 met een beleidsreactie. Eén rapport is eveneens aan de Eerste Kamer aangeboden. Van de evaluaties buiten de steekproef zijn er 19 aangeboden aan de Tweede Kamer, waarvan 14 met een beleidsreactie.

¹Tabel 2: Twee evaluatierapporten kwalificeren niet als evaluaties volgens zowel medewerkers als onderzoekers en dienen voornamelijk ter verantwoording

Betrokkenheid belanghebbenden

Bij de meeste evaluaties, die in opdracht van het ministerie zijn uitgevoerd, is er sprake van een referentie-, stuur-, of klankbordgroep, die de evaluatie begeleidt. Deze bestaat doorgaans uit medewerkers van de betrokken beleidsdirectie(s), vaak aangevuld met externe experts en, indien van toepassing, met vertegenwoordigers van andere departementen of maatschappelijke organisaties. Bij drie evaluaties in de steekproef is de betrokkenheid van belanghebbenden niet aangegeven. Bij twee rapporten is de vraag niet van toepassing. Het betreft een rapport van de Algemene Rekenkamer en een rapport van de Adviescommissie Vreemdelingenzaken. Bij 16 evaluaties, die niet tot de steekproef behoren, is er geen informatie over de betrokkenheid van belanghebbenden.

3.6. Presentatie en communicatie evaluatieresultaten

| 33 |

Voor een groot deel van de evaluaties worden resultaten alleen via een rapport gepresenteerd. Aanvullend wordt van de volgende communicatiemiddelen gebruik gemaakt:

Tabel 3: Presentatie en communicatie evaluatie resultaten
(evaluatieprogrammering 2004 - 2007)

	Evaluatie in steekproef	Overige evaluaties
Mondelinge presentatie (o.a. gesprekken en consultaties)	6	3
Workshop, seminar, conferentie	8	6
Studiedagen	1	-
Website	3	2
Cd-rom/ dvd	5	2
Publicatie in academisch tijdschrift / working paper	-	2

Bij vier evaluaties in de steekproef werden de concept-rapportages al op conferenties gepresenteerd aan een breed forum van belanghebbenden.

4

Gebruik van evaluaties: een beknopt literatuuroverzicht

4.1. Inleiding

Evaluatieonderzoek dat niet gebruikt wordt, omdat het te laat komt met het oog op besluitvorming, de kwaliteit ervan onder de maat is of omdat evaluatoren te hoge verwachtingen creëerden, draagt doorgaans niet bij aan een doelmatige besteding van publieke onderzoeksmiddelen. Daar is ook sprake van als het onderzoek wel voldoet, maar de bevindingen beleidsmakers niet bevallen en de rapporten daardoor op de plank terechtkomen of in de prullenbak verdwijnen. Immers, evaluatieonderzoek beoogt beleidsmakers, politici en maatschappelijke organisaties aan te zetten te leren van het verleden en langs die weg bij te dragen aan het realiseren van een meer *evidence-based* overheidsbeleid dat effectief en efficiënt is.

Dit streven wordt bemoeilijkt door een spanningsveld dat inherent is aan een dergelijk onderzoek. Beleid komt immers mede tot stand op grond van andere, vaak sterkere, krachten uit de politiek en de media dan uitsluitend op grond van kennis uit onderzoek. Sommige onderzoekers klagen dan ook over het onbenut laten maar soms ook misbruik van resultaten van onderzoek, door bijvoorbeeld het zoeken naar uit hun verband gehaalde bevindingen, die staand beleid ondersteunen ofwel *'cherry picking'* (Walters 2008 en Shulha and Cousins 1997). Schertsend is dit ook wel aangeduid als *'policy-based evidence making'*³.

| 35 |

Tegelijkertijd wordt in de wetenschappelijke literatuur geschreven over een soms te groot vertrouwen in de bijdrage van evaluatieonderzoek aan de beleidsonderbouwing. Weiss et al. (2008) zijn van oordeel dat de groeiende eisen, die overheden stellen aan de onderbouwing en effectiviteit van beleidsmaatregelen, maken dat aan resultaten uit evaluatieonderzoek te veel gewicht wordt toegekend. Evaluatieonderzoek is daardoor vaak van aanzienlijke betekenis bij de besluitvorming over het al dan niet continueren van programma's, respectievelijk het financieren ervan. Dit fenomeen van 'verplichte' benutting wordt wel aangeduid met de term *targeted of imposed-use*. Zij werpen daarbij de vraag op of evaluatieonderzoek in staat is om aan alle verwachtingen te voldoen.

³ Groenewegen, Peter. 2008. Persoonlijke mededeling. Nivel.

In dit hoofdstuk wordt ingegaan op factoren uit de literatuur over de relatie tussen onderzoek en beleid, die het gebruik van evaluatieonderzoek medebepalen. Eerst wordt nader ingegaan op wat er verstaan wordt onder gebruik van evaluatieonderzoek.

4.2. Vormen van gebruik van evaluatieonderzoek

In de literatuur is het meest gemaakte onderscheid in de wijze van gebruik die tussen instrumenteel en conceptueel gebruik van onderzoek ofwel direct en indirect gebruik (Swanborn 1999, Horton et al. 2003, Brandt 2007 en Fabian 2009).

Direct gebruik slaat op de directe toepassing van onderzoeksbevindingen in beleid. Volgens Eijk en Hartkamp (1992, 105) komt instrumentele of directe benutting vrijwel uitsluitend voor bij redelijk eenvoudige en relatief onbelangrijke zaken van beleid. Indirect gebruik kan verschillende vormen aannemen zoals:

- de aanpassing van beleidsmaatregelen of van de organisatie erachter;
- legitimatie van bestaand beleid; en
- systematisering van beleid (Fabian 2009).

Indirect gebruik is ook wel aangeduid als gebruik ter ondersteuning van de ontwikkeling van kennis (*enlightment*) van bestuurders, uitvoerders, politici en de samenleving. Onderzoeksresultaten fungeren dan primair om het veld inzicht te verlenen in thema's, om problemen te verhelderen, oude waarheden open te breken, enzovoorts, dit ter verbreding van de oordeelsvorming. Overigens moet het onderscheid tussen instrumenteel en conceptueel gebruik niet worden beschouwd als een tweedeling. In de praktijk zijn de verschillen subtiel. Een variant hierop is dat evaluaties dienen als communicatiemedium binnen grotere organisaties. Bij complex beleid, dat verschillende organisatieonderdelen en hiërarchische lagen betreft, zijn evaluatierapporten soms de enige platforms van interne educatie en communicatie. Zij systematiseren het beleid in kwestie, de achtergronden en de beweegredenen ervan en leggen de ontwikkeling van het beleid in de loop der jaren binnen een organisatie vast.

Naast de begrippen instrumentele en conceptuele benutting zijn er nog legio andere aanduidingen en typeringen van het gebruik en negeren van onderzoek. Zo wordt er ook gesproken over ‘symbolisch’ gebruik. Dit betekent bijvoorbeeld dat een onderzoeksbevinding slechts ter illustratie dient dat een bepaald onderwerp aandacht heeft gehad. Van aanpassing (*improvement*), respectievelijk verbetering (*embetterment*) is daarentegen sprake als het maatschappelijk probleem, dat onderwerp is van het evaluatieonderzoek, na herziening van het beleid op basis van de evaluatie in meerdere of mindere mate effectief is aangepakt dan wel daadwerkelijk is verbeterd. Bij verbetering wordt niet primair gekeken of het beleid aangepast is, maar of het maatschappelijk probleem als zodanig is verminderd of verdwenen als gevolg van het geëvalueerde beleid. Zie voor enkele voorbeelden van dit soort studies de lijst van geraadpleegde bronnen (World Bank 2004). Tegenover deze vormen van gebruik kan het eerder beschreven niet-gebruik dan wel misbruik van onderzoek geplaatst worden. Daarnaast wordt nog gesproken in de literatuur over evaluaties die politieke standpunten bekrachtigen of wel weerleggen en evaluaties die nuttig zijn vanwege het proces dat daarbij doorlopen wordt. Fabian (2009) verwijst in navolging van Stamm (2003) ook nog op niet-gebruik in de vorm van het opwerpen van blokkades gericht op het voorkomen van enig gebruik, omdat er een aanzienlijke weerstand tegen de onderzoeksbevindingen is.

| 37 |

Over gebruik van evaluatieonderzoek bestaan dus verschillende beelden. Een beeld is dat van de evaluatie als *the dusty report on the shelf*; een ander beeld is dat van de evaluator als *the stranger with the ax*. Bij dat laatste beeld gaat het erom dat het evaluatieonderzoek verantwoordelijk is voor de stopzetting van een bepaald beleid: het werkt niet, dus weg er mee. In onderzoek dat Dahler-Larsen (2000) naar het gebruik van evaluaties bij Deense gemeenten deed, bleken beide beelden niet gesteund te worden door de feiten. Een ander beeld is wel in de praktijk herkenbaar. Het betreft het gebruik van evaluatie als management ondersteuning waarbij, omwille van snelle acties om programma’s te verbeteren, de nadruk meer ligt op tijdigheid en relevantie in plaats van betrouwbaarheid (Hellstern 1985 en Dahler-Larsen 2000).

Er bestaan vandaag de dag nog andere beelden van evaluaties en benutting. Zo wijzen Frey en Osterloh (2006) op *evaluiitis*, ofwel een overdaad aan evaluaties waarvan gehoopt wordt dat beleidsmakers er iets van kunnen

leren, terwijl dat lang niet altijd het geval is. Elders wordt naar een vergelijkbaar fenomeen verwezen: *death by evaluation*. De kern daarvan is dat enerzijds door het volume aan evaluaties, en anderzijds door de aard van de kennis die deze studies produceren gebruik ervan door onder andere beleidsmakers gering is of achterwege blijft. Het betreft dan meestal processtudies en studies gericht op naleving, beheersdoelmatigheid en tevredenheidsmetingen, die proceskennis en kennis over naleving van afspraken en procedures opleveren. Immers, de kennis is vaak weinig spannend, is meer van hetzelfde en is vaak moeilijk te gebruiken, met andere woorden, levert geen of een beperkt inzicht in hoe en waarom het beleid werkt en hoe effect is te bereiken. Bovendien kan dit type evaluatie aanleiding geven tot bureaucrativering. Power (1999) heeft in dit verband gewezen op de ontwikkeling van een *audit society*. Groeiende aantallen audit rapporten, die naar naleving van procedurele beleidsregels vragen zonder veel nieuwe inzichten of verklaringen te genereren. Of grote aantallen procesevaluaties die steeds maar de knoppen beschrijven waarop gedrukt had moeten worden zonder dat inzichtelijk wordt wat de maatschappelijke effecten daarvan eigenlijk zijn (Leeuw en Cooksy 2004 en Leeuw en Vaessen 2009). Dit maakt dat de toegevoegde waarde van nog meer studies van dit zelfde type afneemt en daarmee vermoedelijk ook het gebruik. Er wordt in dit kader daarom wel over de wet van de afnemende meeropbrengsten van evaluaties en audits gesproken.

| 38 |

Niettemin geeft een rapport van het *Center for Global Development* (2005) aan dat er, ondanks het grote aantal evaluaties op het terrein van ontwikkelings-samenwerking, een *evaluation gap* bestaat. De *gap* komt er op neer dat nog maar weinig bekend is over de impact van ontwikkelingsinterventies (Clements, Chianca and Sasaki 2008).

4.3. Factoren van belang voor benutting

De internationale literatuur over benutting van onderzoek en evaluatie-onderzoek in het bijzonder onderscheidt een groot aantal factoren en kenmerken van onderzoeken, die kunnen bijdragen tot gebruik van onderzoeksresultaten. In de volgende paragrafen wordt kort een aantal belangrijke factoren besproken. Daarbij wordt bijzondere aandacht besteed aan de vraagzijde, de beleidscontext, de methodologische kwaliteit en het

soort kennis dat evaluaties leveren (Leviton and Hughes 1981, Nutley, Walter and Davies 2003, Leeuw 2009 en Swanborn 1999).

Intrinsieke kwaliteitswaarborgen

Een onderzoek staat of valt bij de geloofwaardigheid van wat een onderzoeksrapport ons leert en hoe het ons wordt geleerd (Chelimsky 1995 en Swanborn 1999). De waarde daarvan baseert zich onder andere op het gebruik van een wetenschappelijk verantwoord onderzoeksontwerp. Andere belangrijke indicatoren voor wetenschappelijkheid en daarmee de geloofwaardigheid van evaluaties zijn validiteit, betrouwbaarheid, objectiviteit en onafhankelijkheid (Chelimsky 1995).

Aard van de kennis die geproduceerd wordt

Evaluaties produceren verschillende typen kennis. Zo maken Nutley, Walter en Davies (2003) een onderscheid tussen vijf typen (evaluatieve) kennis, te weten:

- * *knowledge - about problems;*
- * *knowledge - about what works;*
- * *knowledge how to put it into practice;*
- * *knowledge who to involve; and*
- * *knowledge - why.*

Een ander onderscheid op grond van een inhoudsanalyse van tientallen rapporten over (Rekenkamer)doelmatigheidsonderzoek is ook mogelijk (Leeuw 1998). Doelmatigheidsonderzoeken produceren enerzijds kennis over procedureel-organisatorische voorwaarden (POV-evaluaties) waaraan voldaan moet worden om tot een effectief en efficiënt beleid te kunnen komen zoals bijvoorbeeld de planning van evaluaties, de coördinatie en planning van beleid. Anderzijds produceren zij substantieve kennis over wat de effecten en de werkzame bestanddelen van overheidsbeleid zelf zijn, voor wie, hoe lang en waarom. Een andere typering van kennis uit evaluatieonderzoek heeft betrekking op de aard van het onderzoeksontwerp dat gebruikt wordt. Sommige onderzoeksontwerpen staan het doen van uitspraken over de causale relatie tussen een interventie en een uitkomst niet toe, dit omdat het attributievraagstuk niet geadresseerd is, bijvoorbeeld omdat nagelaten is op meerdere momenten te meten en te randomiseren.

Tot slot is de aard van de kennis, die evaluaties opleveren over de effecten van beleid belangrijk. Er circuleren in de evaluatieliteratuur veel definities van begrippen als effectiviteit, efficiëntie, doelmatigheid en doeltreffendheid. Daarbij wordt niet altijd duidelijk aangegeven welk begrip wat beoogt en welk deel van de relatie tussen inspanning en resultaat of kosten en baten precies geëvalueerd is. Zo komt het in praktijk regelmatig voor dat uitspraken over de doelmatigheid of doeltreffendheid van beleid worden gebaseerd op een evaluatie van de vraag naar de randvoorwaarden daarvoor in plaats van een vaststelling van effecten. Ook komt het voor dat doelmatigheidsonderzoeken alleen betrekking hebben op de kosten of uitgaven voor een bepaald beleid, zonder onderzoek naar (eventuele) baten. Dit terwijl voor doelmatigheidsonderzoeken altijd kosten én baten bekeken moeten worden.

| 40 |

Het maken van onderscheid in de aard van de geproduceerde kennis is vooral belangrijk omdat er een mismatch kan zijn tussen wat beleidsmakers, bestuurders en politici verwachten, hopen of wensen dat aan kennis geproduceerd wordt en wat de evaluator daadwerkelijk biedt. Uiteenlopende verwachtingen over wat aan inzichten over de impact van een interventie geproduceerd wordt, kunnen een belangrijke factor zijn bij het niet gebruiken van de resultaten. Als gehoopt wordt dat uitspraken over impact gedaan worden, en over het soort instrumenten waarmee effecten al of niet te bereiken zijn en vervolgens informatie geproduceerd wordt over de vraag of het beleid is uitgevoerd conform de procedurele en organisatorische voorwaarden in het beleidsontwerp, dan kan dit leiden tot het niet gebruiken van evaluatieonderzoek en niet leren.

Het maken van onderscheid in de aard van de geproduceerde kennis is ook van belang vanwege de toenemende profilering van zogenaamde *evaluation machines*, of systemen van evaluatieonderzoek met eigen handboeken (Dahler-Larsen 2000), claims op een bepaald type geproduceerde kennis, leerstoelen, professionele verenigingen, websites, enzovoorts. Het is daarom belangrijk bij opdrachtgeving te weten welke school van denken om een onderzoek wordt verzocht (Leeuw and Furobo 2008).

Extrinsieke kwaliteit waarborgen

De wat wel genoemd wordt extrinsieke kwaliteit van evaluatieonderzoek speelt ook een belangrijke rol bij het gebruik van onderzoeksbevindingen.

Evaluaties moeten daarvoor in lijn worden gebracht met de context, zoals de beleidspraktijk, de beleidsgeschiedenis en de institutionele omgeving. Daarnaast moet er rekenschap zijn van de fase van het beleidsproces waarvoor onderzoeksresultaten worden gebruikt. Deze aanpassing aan tijd en context van evaluaties beperkt echter wel de houdbaarheidsdatum van onderzoeksbevindingen.

Wat betreft evaluaties en hun relevantie noemt Swanborn (1999) verder het belang van tijdigheid van de oplevering en de handvatten voor beleidsmakers. Ook wordt aanvaardbaarheid door Swanborn onderscheiden als kernaspect van de relevantie van onderzoek. Daarmee wordt bedoeld op het inlevingsvermogen van de onderzoeker in de wereld van de beleidsmaker. Bij grote afwijkingen van de gebruikte theorievorming van de huidige beleidslijn dient de onderzoeker extra zorg te besteden aan de onderbouwing.

Onderzoek uitdragen en verspreiden

De wijze waarop de resultaten van onderzoek worden verspreid krijgt vaak een prominente plaats als een factor van belang voor gebruik. Onderzoekers kunnen bijvoorbeeld een persbericht uitgeven over een nieuw verschenen onderzoek. Daarnaast kunnen onderzoekers bewust meer doen aan kennisverspreiding, door bijvoorbeeld lezingen of workshops over hun onderzoek te verzorgen voor de opdrachtgever of via moderne wegen zoals wiki's of blogs en *crowdsourcing*⁴ (Sunstein 2006).

Veel van de zojuist aangehaalde (wetenschappelijke) literatuur over gebruik van evaluatieonderzoek richt zich op de beïnvloedingsmogelijkheden van de onderzoeker: hoe kan de onderzoeker ervoor zorgen dat er iets of meer met zijn onderzoek wordt gedaan? Echter, hoe en in welke mate onderzoeksresultaten worden gebruikt hangt ook in sterke mate af van de houding, respectievelijk het gedrag van de opdrachtgever, aanvrager en het onderwerp van de evaluatie.

⁴ het vooral via internet gebruikmaken van een grote groep niet vooraf gespecificeerde individuen (professionals, vrijwilligers, geïnteresseerden) voor consultancy, innovatie, beleidsvorming en onderzoek.

Persoon, houding en cultuur

Allereerst is de persoonlijke bereidheid en het vermogen tot handelen op basis van onderzoek bij de opdrachtgever cruciaal. Sommige beleidsmedewerkers of ambtenaren hebben meer belangstelling voor onderzoek dan andere. Bovendien zijn sommigen van hen beter in staat of bereid dan anderen om onderzoek te verbinden met het denken over beleid.

De houding en cultuur die binnen afdelingen, directoraten of een ministerie bestaat ten aanzien van het belang van evaluaties en de bereidheid er iets mee te doen, speelt ook een rol bij gebruik van resultaten van evaluaties. In verschillende kringen is te beluisteren dat de cultuur van het ministerie van Buitenlandse Zaken niet gericht zou zijn op leren en het delen van kennis. Zo merken De Gaay Fortman en De Ruijter in hun commentaar op de Afrika-evaluatie bijvoorbeeld op dat zij geen overtuigende aanwijzingen gevonden hebben dat er op het ministerie een cultuur is, die gericht is op leren van het verleden. Zij vragen zich af of er op het ministerie een op leren georiënteerd instrumentarium en mentaliteit bestaan (Gaay Fortman en de Ruijter 2008). IOB voegt daar in een reactie aan toe dat het een *understatement* is voor een hardnekkig probleem, waar ook door anderen al vaker op gewezen is. Het gemis van een dergelijke cultuur zou te maken hebben met verschillende kenmerken van het werk bij het ministerie, zoals een diplomatencultuur waarbij men vooral leert via persoonlijke contacten, een hoog verloop in functies en een gebrek aan specialistische kennis.

| 42 |

Samenwerking

Het zal duidelijk zijn dat gebruik in de praktijk vaak het resultaat zal zijn van een wisselwerking tussen enerzijds de specifieke context en anderzijds het gedrag van zowel onderzoekers als beleidsmakers. Daarom wordt meer en meer gewezen op het belang van het denken in termen van trajecten rondom het gebruik van evaluaties door beleid. Een bekende suggestie is om de uitvoering van evaluatieonderzoek en het denken over beleid, waar mogelijk, meer aan elkaar te koppelen. Enerzijds kunnen begeleidingscommissies daaraan bijdragen, anderzijds tussentijdse andere overlegsituaties. Wat een complicerende factor is, is de mate waarin er in een dergelijke situatie sprake is van *partnerial evaluations* (Pollitt 1999). Dat zijn evaluaties, die niet meer onafhankelijk worden uitgevoerd, maar co-producties zijn van beleids- en uitvoeringsorganisaties, waarbij als het ware de slager een deel van zijn of haar eigen vlees aan het keuren is (RPE 2006).

5

Analyse van gebruik van 30 evaluaties

5.1. Inleiding

Dit hoofdstuk bevat een beknopte analyse van het gebruik van 30 evaluaties in de steekproef waarbij de volgende onderzoeksvraag wordt beantwoord:

Worden evaluatieonderzoeken daadwerkelijk gebruikt voor verbetering van beleid en/of beleidsuitvoering (de leerfunctie) en voor informatieverstrekking over effectiviteit en doelmatigheid van het gevoerde beleid (de verantwoordingsfunctie)?

Deze vraag wordt uitgesplitst in diverse vormen van gebruik, invloed op beleid en het leren van evaluaties. In het tweede deel van dit hoofdstuk worden de factoren besproken, die op grond van deze analyse van belang worden geacht voor het gebruik van evaluaties door het ministerie van Buitenlandse Zaken. De vraag naar het leren van evaluaties komt in hoofdstuk 6 aan de orde.

| 45 |

5.2. Gebruik van evaluaties

Interne communicatie

Evaluaties kunnen een rol spelen in de interne of in de externe communicatie van het ministerie, de beleidsdirectie of het Directoraat Generaal. Bij de evaluaties van de EDP (*Enabling Development Policy* binnen het *World Food Programme*) en het NIMF (*Nederlands Investerings Matching Fonds*) speelde de evaluatie een rol in het interne debat om een standpunt te bepalen. Bij de EDP-evaluatie zijn sporen van dat debat aan te treffen in interne notities en correspondentie en bij de NIMF-evaluatie richtte de directie een interne beleidsreactie aan de DG om haar standpunt uiteen te zetten. In een ander voorbeeld vonden felle discussies plaats waarbij controversiële punten werden doorgesproken. Dit was nodig voor de voortgang van de gedachtebepaling en besluitvorming. De evaluatie ontleende een deel van haar belang hieraan.

Een veel voorkomende vorm van interne communicatie, waarbij evaluaties als medium fungeren, is de overdracht van dossiers bij de start in een nieuwe functie. Bij meerdere evaluaties wordt deze vorm van gebruik gemeld. Bij één voorbeeld (ORET) vergrootte de evaluatie de kennis van een

deel van de zittende medewerkers over een bepaald programma. Een evaluatie is vaak het meest complete overzicht van een beleid. Een beleidsdoorlichting biedt op basis van beschikbare onderzoeksgegevens het meest complete overzicht van het beleid inzake een of meer operationele doelstelling(en), de legitimatie van en de ratio achter het beleid. Bij het voorbeeld van de beleidsdoorlichting, die in de steekproef is opgenomen, wordt echter geen melding gemaakt van deze vorm van gebruik. Een van de evaluaties, die een aantal kenmerken heeft van een beleidsdoorlichting, de evaluatie Energie en Ontwikkelingssamenwerking, vervulde de rol van overdrachtsdocument wel. Het gebruik van evaluaties als overdrachtsdocument bij functiewisseling is dan ook veeleer een zaak van individueel initiatief van de betrokken medewerkers dan van beleid van het ministerie. Beleid inzake het gebruik van evaluaties als institutioneel geheugen is afwezig.

| 46 |

Het komt ook voor dat de onderzoekers, die betrokken zijn bij een evaluatie, een rol spelen bij de overdracht van kennis bij verandering van functie. Bij een van de evaluaties worden zowel de evaluatie als de evaluator geregeld geconsulteerd bij lopende zaken, met name waar het gaat om overleg met ontvangers van subsidies. In een ander voorbeeld vervult een inventariserend rapport een committerende en disciplinerende rol in een arena van beleidsbepalers van verschillende directies en departementen. Het document legt de gedeelde standpunten vast en kan op een later moment gebruikt worden om de uitgezette lijn vast te houden en druk tot oprekken van de agenda te weerstaan. Weer een andere evaluatie heeft als belangrijkste functie te dienen als agenda en 'to-do-lijst' voor wenselijke aanpassingen waaraan BZ-medewerkers in internationaal verband dienen te werken. Men kan deze rapporten beschouwen als *cognitive organizers* van beleid zoals in hoofdstuk 4 wordt beschreven.

Sommige respondenten onderscheiden daarnaast de promotiefunctie van evaluaties. Ze bieden een kans om jezelf weer eens in de schijnwerpers te plaatsen. Deze functie wordt vooral beschreven als een vorm van interne communicatie, omdat het in de eerste plaats gaat om het vragen van aandacht van een DG en/of een minister. Wanneer het rapport ook naar de Kamer wordt gestuurd, is het daarnaast een voorbeeld van externe communicatie.

Externe communicatie

In de externe communicatie met partijen buiten het ministerie is legitimatie een belangrijke vorm van gebruik. Legitimatie is in dit verband het verantwoord en toelichten van een reeds ingenomen standpunt. Dit in tegenstelling tot overleg waarbij een standpunt nog gevormd dient te worden. Een opmerkelijke vorm van legitimatie in het kader van dit onderzoek, is de aanspraak op het gebruik van evaluaties als basis van beleidsaanpassingen. Drie voorbeelden zijn de evaluaties van de programma's energie en ontwikkelingssamenwerking, van ORET en van het Afrikabeleid.

Geregeld dienen evaluaties als aanleiding en agenda voor overleg met de doelgroep en achterban van het geëvalueerde beleid, bijvoorbeeld bij de evaluaties van NPT/NFP, ORET en het onderzoeksbeleid. Het onderscheid tussen overleg en legitimatie is dan niet altijd even scherp te maken. De beschrijvingen van de contacten geven in een aantal gevallen de indruk dat men met een redelijk open agenda het overleg aanging en er sprake was van uitwisseling van informatie en standpunten. In het geval van de TMF-evaluatie getuigen verslagen van vergaderingen met de branche-organisatie daarvan. De ORET-evaluatie werd daarentegen onder meer gebruikt om het bedrijfsleven er van te overtuigen dat gebonden hulp niet meer paste in het programma.

| 47 |

De evaluatie van het onderzoeksbeleid neemt een aparte positie in. Deze evaluatie had betrekking op een voorgaande versie van het onderzoeksbeleid, maar werd toch aangegrepen als aanleiding voor een openbaar debat over de inhoudelijke en beleidsmatige aspecten van ontwikkelingsonderzoek. Dat debat ging over meer algemene aspecten van ontwikkelingsonderzoek en beleid inzake onderzoek.

Verantwoording

De verantwoordingsfunctie van evaluaties richt zich op de informatie-behoefte van het parlement en is vastgelegd in de RPE 2006 en de Comptabiliteitswet. Verantwoording wordt afgelegd door toezending van evaluaties en beleidsreacties aan het parlement, door publicatie van de evaluatieprogrammering en de evaluatierapporten op de website van het ministerie en door gebruik van evaluaties als informatiebron voor het departementale jaarverslag en de resultatenrapportages over ontwikkelings-

samenwerking. Volgens de interne notitie 'Naar een effectiever gebruik van evaluaties door het DGIS' (Ministerie van Buitenlandse Zaken 2008, 2)⁵ is de verantwoordingsfunctie duidelijk vastgelegd en wordt deze in de praktijk ook adequaat ingevuld. De regel is dat alle evaluaties, uitgevoerd door IOB, aan het parlement worden toegezonden en worden begeleid door een beleidsreactie van de bewindspersoon. Veel evaluaties zijn in de Tweede Kamer op de agenda geplaatst. Voor zover na te gaan, zijn er over geen van de evaluaties in de steekproef rechtstreeks schriftelijke Kamervragen gesteld⁶.

De eerste resultatenrapportage over ontwikkelingssamenwerking, 'Resultaten in Ontwikkeling 2004', bevat nog geen verwijzingen naar evaluaties. Er zijn alleen reguliere prestatiegegevens gebruikt. De gegevens voor de rapportage worden ontleend aan het rapportagesysteem *Piramide* en aan internationale rapporten van onder meer de VN en de Wereldbank. De tweede resultatenrapportage, 'Resultaten in Ontwikkeling 2005-2006', gaat uitvoerig in op de voorlopige resultaten van de tot dan uitgevoerde impactevaluaties, te weten de impactevaluatie van drinkwatervoorziening in Tanzania en van primair onderwijs in Oeganda en Zambia. In de laatste resultatenrapportage worden aanzienlijk meer en ook een grotere verscheidenheid aan evaluaties aangehaald. Daartoe behoren onder meer de volgende evaluaties uit de steekproef: ORET, VMP, Afrikabeleid en de impactevaluatie drinkwatervoorziening in Tanzania. Ook de andere impactevaluaties worden opnieuw gebruikt voor de rapportage.

| 48 |

In de jaarverslagen van het ministerie wordt verantwoording afgelegd over de programmering en uitvoering van evaluaties. De vaststelling van de begrotingsstaten bevat overzichten van geplande en afgeronde evaluatieonderzoeken. Hierin zijn de belangrijkste evaluatieonderzoeken per beleidsartikel opgenomen, geordend naar operationele doelstelling, waarbij conform de Rijksbegrotingsvoorschriften onderscheid wordt gemaakt tussen beleidsdoorlichtingen, effectenonderzoek en overig evaluatieonderzoek. Voor meer gedetailleerde informatie over geplande evaluatieonderzoeken wordt verwezen naar de evaluatieprogrammering van het ministerie en naar de website. In het jaarverslag over 2007 is uitvoering

⁵ Ministerie van Buitenlandse Zaken. 2008. Naar een effectiever gebruik van evaluaties door het DGIS. DEC.

⁶ Zoekactie Parlando

gegeven aan een Kamermotie om in het jaarverslag systematisch aandacht te besteden aan de uitkomsten van evaluaties. In eerdere jaarverslagen gebeurde dat slechts incidenteel.

5.3. Invloed op beleid

In deze paragraaf wordt de volgende vraag behandeld:

Betekenen de verschillende vormen van gebruik nu ook dat evaluaties leiden tot aanpassingen van het beleid en invloed hebben op nieuw beleid?

Veel evaluaties binnen het DGIS en ook een aantal binnen andere DG's gaan gepaard met een herziening van het geëvalueerde beleid. In de periode, die het onderzoek bestrijkt, is er binnen het DGIS een aantal nieuwe beleidsnotities en beleidskaders tot stand gebracht, die aanpassingen van het beleid introduceerden. Bij de andere DG's gaan de evaluaties meestal gepaard met beperkte aanpassingen en een aantal malen blijft actie geheel achterwege. De aanpassingen binnen het DGIS zijn te onderscheiden in een aantal typen. Een belangrijke reden om verschillende programma's aan te passen, was dat zij vooral gebaseerd waren op zelfstandige projecten en niet meer goed pasten in de nieuwe sectorbenadering voor ontwikkelings-samenwerking. Een tweede overweging is te typeren als het opvangen van de gevolgen van de sectorbenadering. Deze heeft in een aantal gevallen geleid tot een verminderde nadruk op bestrijding van armoede en gender-gelijkheid en tot minder aandacht voor de productieve sectoren en ontwikkeling van het platteland in partnerlanden. Een aantal programma's werd op de schop genomen om deze zaken recht te zetten.

| 49 |

Deze twee typen beleidsaanpassingen komen het vaakst voor en leiden over het algemeen ook tot de meest ingrijpende wijzigingen. Maar er zijn ook andere overwegingen. In het geval van de evaluatie van het beleid inzake Hiv/Aids en seksuele en reproductieve gezondheid en rechten (SRGR) was de wens tot een verdere integratie van beide onderdelen in het buitenlandse beleid de reden voor het opstellen van één nieuwe en geïntegreerde beleidsnotitie. Het programma VMP sloot al voldoende aan op belangrijke

accenten binnen de millenniumdoelen en werd vooral om beheersmatige redenen herzien en om prioriteiten, doelstellingen en gewenste resultaten beter af te bakenen.

De voorbeelden laten zien dat de verhouding tussen evaluaties en beleid veelvormig en gecompliceerd is. Er zijn twee voorbeelden van een betrekkelijk eenduidige relatie. In het ene concludeert de Algemene Rekenkamer dat het ministerie sinds een eerder Rekenkamerrapport een groot aantal verbeteringen heeft doorgevoerd in zowel beleid, uitvoering als verantwoording van het betreffende beleid gericht op de begeleiding van Nederlandse gedetineerden in het buitenland. In het geval van de evaluatie van het Hiv/Aids en SRGR beleid is er sprake van een betrekkelijk duidelijke lijn tussen evaluatie en beleidsaanpassing. Deze evaluatie was ook bedoeld als voorbereiding van een nieuwe geïntegreerde beleidsnotitie voor beide beleidsvelden. Voorts werd enige tijd na het verschijnen van het evaluatierapport een nieuwe beleidsnotitie gepresenteerd. Deze verwees meermalen naar het evaluatierapport als bron en vertoont ook inhoudelijke overeenkomsten met het rapport.

| 50 |

Andere voorbeelden van aanpassingen laten daarentegen een aanzienlijk inhoudelijk verschil zien tussen evaluatie en beleid. Bij de meeste voorbeelden was de voorbereiding van een nieuwe beleidsnotitie of van een ingrijpende beleidsaanpassing niet in de opdracht voor de evaluatie vervat. Dit geldt o.a. voor NPT/NFP, ORET en NIMF. Ook de resultaten van deze evaluaties geven geen aanleiding tot ingrijpende aanpassingen. Soms is er een verschil in kritische toon tussen de evaluaties en de reactie van de beleidsdirectie op de evaluatie. In een aantal gevallen wijzen de respondenten er op dat al tijdens de begeleiding van het onderzoek een verschil in kritische toon werd ervaren. Een duidelijk voorbeeld daarvan is het beleid inzake ontwikkelingssamenwerking en hoger onderwijs. Het beleid werd na het uitbrengen van de evaluatie meer in overeenstemming gebracht met de nieuwe 'architectuur voor ontwikkelingssamenwerking', inclusief de sectorbenadering en de principes van de *Paris Declaration*. De aanbevelingen van de evaluatie gaan ook in die richting, maar de beleidsreactie en medewerkers van de beleidsdirectie wijzen vooral op het feit dat de evaluatie niet aanstuurde op fundamentele veranderingen en vooral technisch van aard was. Volgens een van de beleidsmedewerkers was het rapport zelfs redelijk braaf en heeft de beleidsdirectie de onderzoekers

uitgedaagd om kritischer te werk te gaan en strategisch te denken in plaats van operationeel. De relatie tussen evaluatie en de er op volgende beleidswijziging is dan ook niet duidelijk aan te geven.

Een tweede voorbeeld betreft de vervanging van het programma ORET door ORIO. Daarbij lijkt er eerder sprake te zijn van een andere waardering van bevindingen uit de beschikbare evaluaties in het licht van nieuwe inzichten over gewenste resultaten, dan van nieuwe bevindingen. De regeling werd meer in overeenstemming gebracht met gegroeide inzichten over ontwikkelingssamenwerking, met het millenniumdoel 'bestrijding van armoede' en met de OESO-DAC afspraken over de ontbinding van hulp van 2001. De evaluatie van het NIMF laat zien dat een inhoudelijk verschil tussen evaluatie en de erop volgende beleidsaanpassing mogelijk te verklaren is door het gebruik van andere informatiebronnen incl. evaluaties. Dan wordt met name gedacht aan de PSOM-evaluatie van december 2005, aan het AIV-advies inzake het belang van de particuliere sector voor ontwikkelingsamenwerking van oktober 2006 en aan adviezen van de Wereldbank en het *DCD/DAC Investment Committee*. Samen konden deze publicaties leiden tot de conclusie dat het NIMF, zoals het tot dan toe had gefunctioneerd, niet voldeed aan de behoefte aan een meer op armoedebestrijding gericht programma.

| 51 |

In andere gevallen is er wel sprake van duidelijke inhoudelijke overeenkomsten tussen evaluatie en nieuw beleid, maar lopen de beleidswijzigingen al vooruit op de bevindingen en aanbevelingen van de evaluaties. Een aantal belangrijke aanbevelingen uit de TMF-evaluatie is bijvoorbeeld te herkennen in een nieuw MFS-beleidskader, dat in concept al gereed was bij de start van de evaluatie. Betrekkelijk veel elementen uit de evaluatie van het Vakbondsmedefinancieringsprogramma (VMP) zijn terug te vinden in het nieuwe VMP-beleidskader voor de periode 2009-2012. Maar een aantal concrete elementen is gedurende de looptijd van de evaluatie, of al eerder, in andere bronnen te herkennen.

De evaluatie sectorbenadering wijst op een aantal problemen, die volgens de beleidsreactie en de respondenten al langer de aandacht hadden van het ministerie. Het is dan ook niet altijd duidelijk welke van de voorgestelde aanpassingen van de sectorbenadering op de evaluatie zijn gevolgd en welke stammen van vóór het IOB-onderzoek. Een intern stuk noemt diverse

evaluaties, contacten met internationale partners, ervaringen van medewerkers van ambassades en het ministerie in Den Haag als de bronnen voor de lessen, waarvan geleerd is. Ook de Afrika-evaluatie sneed thema's en problemen aan, die voor velen inmiddels herkenbaar waren als gevolg van het beleid van het voorgaande decennium, zoals problemen met begrotingssteun en goed bestuur. Er werden voorafgaande aan de evaluatie en gedurende de looptijd al initiatieven genomen om problemen aan de orde te stellen en aan te pakken.

De voorbeelden laten zien dat er veel inhoudelijke overeenkomsten zijn tussen de aanbevelingen en conclusies uit evaluaties en beleidswijzigingen. Dit betekent echter niet dat een evaluatie leidt tot een daarop volgende beleidsaanpassing. De evaluatie van het onderzoeksbeleid toont het duidelijkst aan, dat het beleid en de organisatie van beleid binnen het DGIS een krachtige zelfstandige dynamiek kennen, die onafhankelijk van de resultaten van een evaluatie kunnen leiden tot hervormingen. Maar ook de inhoudelijke kloof tussen betrekkelijk mild gestemde evaluaties en omvangrijke hervormingen die er op volgen, tonen dat evaluaties zelden hiervoor verantwoordelijk waren de afgelopen jaren. Inhoudelijke overeenkomsten tussen evaluatie en beleid berusten in een aantal gevallen op beleidsaanpassingen die voorafgaan aan de evaluatie. Deze onduidelijke en gecompliceerde verhouding tussen evaluaties en beleid is ook regelmatig beschreven in de literatuur over gebruik van onderzoek.

| 52 |

Voor deze verhouding tussen evaluaties en beleidsaanpassingen zijn twee verklaringen. In de eerste plaats is bekend dat een deel van de conclusies en aanbevelingen van evaluatieonderzoek al tijdens het onderzoek via deelname aan de referentiegroep bekend worden en ter harte worden genomen. De beleidsreactie op de evaluatie van het beleid inzake uitbreiding van de Europese Unie maakt bijvoorbeeld melding van aanpassing van het Matra-programma op basis van evaluatieresultaten, die tijdens de looptijd van evaluatie reeds werden besproken. Een aantal respondenten vindt dan ook deelname aan een referentiegroep de best denkbare leeromgeving om over een evaluatie geïnformeerd te raken.

De tweede verklaring is dat evaluaties, uit de evaluatieprogrammering van het ministerie, deel uitmaken van een beleidsdebat waaraan evaluaties uit

eerdere perioden, evaluaties van buitenlandse instellingen, adviezen, internationale verdragen en -verklaringen, maar ook politieke programma's en beleidsnotities bijdragen. Noties over het belang van doelstellingen als armoedebestrijding en emancipatie, maar ook over het gebruik van concrete instrumenten als contextanalyses en (varianten van) monitoring-systemen, zijn onderdeel van dat debat. Een beleidsaanpassing die volgt op een aanbeveling in een evaluatie kan gebaseerd zijn op een eerdere aanbeveling of conclusie in een andere bron. Dat maakt het moeilijk om de invloed van een bepaald rapport op een beleidsvoorstel aan te tonen. Niettemin zijn de overeenkomsten tussen uitkomsten van een evaluatie en beleidsaanpassingen niet zonder belang voor de discussie over gebruik van evaluaties. De meest bescheiden conclusie is dat in elk geval een deel van de door evaluaties gesignaleerde problemen wordt aangepakt.

5.4. Factoren gebruik evaluatieresultaten

| 53 |

In deze paragraaf worden de factoren behandeld die van invloed zijn op gebruik van evaluatieresultaten. Het gaat om onderzoeksvraag 4:

Welke factoren en omstandigheden zijn beperkend of juist bevorderlijk voor gebruik van evaluatieonderzoek?

Hierbij gaat het om factoren, die in zowel de deelvragen als inventariserende onderzoeksvragen aan de orde komen en in het bijzonder potentieel belangrijke factoren, die in de internationale literatuur worden behandeld (zie hoofdstuk 4).

Type kennis

Het type kennis dat evaluaties opleveren speelt een belangrijke rol bij het in meer of mindere mate gebruiken van resultaten uit evaluaties. Veel respondenten constateren dat de evaluaties, ook wanneer ze op een van de bovenstaande manieren werden gebruikt, geen instrumentele kennis leverden, dat wil zeggen kennis over hoe iets te doen of te regelen en een gewenst resultaat te bereiken. Evaluaties leveren over het algemeen te weinig substantieve kennis over effecten en over de werkzame bestanddelen van beleid die tot deze effecten leiden.

Evaluaties laten vooral een bestuurskundige invalshoek zien. In veel gevallen komt weliswaar ook de vraag aan de orde of de gestelde doelen zijn bereikt, maar er wordt geen verband gelegd met een mogelijke (gedrags-wetenschappelijke) verklaring voor het al of niet behalen van resultaten, de vraag naar wat werkt en wat niet werkt. De evaluaties hebben vooral betrekking op een adequate inrichting van het beleidsproces en de doelmatigheid van activiteiten in plaats van het verwerven van inzicht in de causale factoren en processen die menselijk gedrag reguleren. De evaluaties starten doorgaans met een beschrijving of reconstructie van het beleid. Vaak wordt ook getoetst of dit beleid voldoet aan de eisen, die aan een rationeel ontwikkeld beleid gesteld mogen worden. Vervolgens wordt nagegaan of het beleid daadwerkelijk volgens dit gereconstrueerd spoorboekje is uitgevoerd. In overeenstemming hiermee bestaan de aanbevelingen dan ook doorgaans uit nieuwe prioriteiten ofwel nieuwe regels, procedures en afspraken met verschillende doelen, zoals een scherpere afbakening van doelgroepen en van doeleinden, verscherpte criteria voor de toekenning van middelen en een duidelijker plaats voor planning, monitoring en evaluatie. In de evaluaties ontbreekt echter een gedrags-wetenschappelijke invalshoek waarin de vraag wordt gesteld of de veronderstellingen, waar het beleid, het programma of de interventie vanuit gaat, inderdaad leiden of kunnen leiden tot de gewenste resultaten.

154

Een belangrijke evaluatie als die van het sectorbeleid beoogt antwoord te geven op de vraag wat er van de intenties van dit beleid terecht komt. De evaluatie steunt op een interessante beleidsreconstructie van veranderingen in de beleidsintenties van ontwikkelingssamenwerking voor en na 1998. De evaluatie gaat na of het beleid ook conform deze intenties is uitgevoerd, of de beoogde toename van meer zeggenschap over de hulp bij de ontvangende landen is gerealiseerd en wat de gevolgen ervan zijn voor het perspectief op armoedebestrijding⁷. De beleidsreconstructie is echter geen reconstructie van de beleids- of interventietheorie, zoals dat in *state-of-the-art* ofwel kwalitatief goede evaluaties is aan te treffen (Leeuw and Vaessen 2009). Het is vooral een bestuurskundige reconstructie zonder aandacht voor mechanismen die mensen, groepen en bestuurders aan kunnen zetten tot gewenst gedrag. Er zou bijvoorbeeld aandacht kunnen

⁷ IOB. 2006. Van Projecthulp naar Sectorsteun: evaluatie van de sectorale benadering 1998 – 2005.

worden besteed aan de theorie op grond waarvan *local ownership* relevant wordt geacht voor ontwikkeling en armoedebestrijding en welke vormen van *ownership* (bezit, eigendom, gedeeld belang, zeggenschap) daarmee bedoeld worden met welk effect? Bij de evaluatie van het Nederlands voorzitterschap van de Europese Unie (EU) wordt bijvoorbeeld voorbij gegaan aan kennis uit de sociale psychologie die het inzicht in de bestuurde processen kan vergroten. Deze evaluatie is vooral gericht op formele processen en op papier vastgelegde doelen en afspraken. Er is één voorbeeld van een gedragswetenschappelijke benadering en vraagstelling aangetroffen. Het betreft de constatering van IOB in de evaluatie van het VMP dat volgens de literatuur ter zake, coachen, adviseren en managementontwikkeling doeltreffender zijn dan training. Deze aanbeveling werd in het nieuwe VMP-beleidskader overgenomen.

Kennis over netto-effecten, ook wel impact genoemd, van beleid en over de vraag welke elementen van het beleid en mechanismen verantwoordelijk zijn voor die impact, of het ontbreken daarvan, wordt nog nauwelijks geproduceerd. Mogelijk dat de geplande reeks impactevaluaties daar verandering in zal brengen. De impactevaluatie van drinkwatervoorziening en sanitatie in Tanzania die in de steekproef is opgenomen, levert nog maar ten dele de hier bedoelde kennis. Er worden gegevens over impact geleverd, maar de evaluatie bevat geen uitgewerkte reconstructie van de aannames achter de water- en sanitatieprogramma's en belangrijke elementen daarvan, zoals het genderaspect en het functioneren van gebruikersgroepen. Bovendien zijn de gebruiksmogelijkheden van impactevaluaties niet goed aan de hand van dit voorbeeld te beoordelen. De impactevaluatie in Tanzania was een nieuw type evaluatie in een nieuwe constellatie. Programmasteun in Tanzania is sinds de evaluatie vervangen door sectorsteun met gedelegeerde zeggenschap. Nederland heeft als *silent partner* geen rechtstreekse invloed meer op de uitvoering van water- en sanitatiebeleid in Tanzania. Het is daarom moeilijker te bepalen of en hoe aanbevelingen uit de evaluatie en toezeggingen in de beleidsreactie tot concrete acties hebben geleid.

In dit beeld van relatief eenzijdig procesgerichte evaluaties zijn ook geregeld opmerkingen te horen over de voorspelbaarheid van de evaluaties. Elke volgende evaluatie biedt volgens deze geluiden meer van hetzelfde en deskundige medewerkers vernemen daardoor weinig nieuws uit de rapporten. Een van de respondenten maakte er aanspraak op dat hij

ongezien al de helft van de bevindingen van elke willekeurige evaluatie kon voorspellen: een beetje versnippering, een beetje onvoldoende aandacht voor monitoring en evaluatie, gender vergeten, een beetje gebrekkige relevantie en weer is bijna een evaluatie klaar. Het feit dat dit soort geluiden geregeld zijn te horen kan er op wijzen, dat de in hoofdstuk 4 aangehaalde wet van de afnemende meeropbrengsten in werking is getreden voor het type evaluatie dat bij het ministerie doorgaans wordt uitgevoerd.

Aansluiting onderzoeksvragen, bevindingen, conclusies en aanbevelingen

Er zijn weinig duidelijke discrepanties tussen onderzoeksvragen, bevindingen, conclusies en aanbevelingen aangetroffen. In één voorbeeld is er een duidelijk onderscheid tussen de opgeworpen probleemstelling en onderzoeksvragen en de uiteindelijke bevindingen en aanbevelingen waartoe het rapport komt. Het betreft een evaluatie, die nadrukkelijk de ambitie had om een lerende evaluatie te zijn, maar daar door gebrek aan geschikte gegevens niet in slaagde. De evaluatieprogrammering bevat een aantal rapporten, die op eigen initiatief van een DG zijn opgesteld, die meer het karakter hebben van een inventarisatie van maatregelen of activiteiten, dan van een evaluatie. Deze rapporten bevatten geen onderzoeksvragen en nauwelijks of geen aanbevelingen. In geen van deze gevallen heeft een onduidelijk verband tussen onderzoeksvragen, bevindingen, conclusies en aanbevelingen gevolgen gehad voor het gebruik van de rapporten.

| 56 |

Wel valt op dat cruciale begrippen als doelmatigheid en doeltreffendheid in de evaluaties, en soms zelfs binnen één evaluatie, niet altijd consistent worden gebruikt. Dat leidt tot verwarring over de aard van de uitspraken die gedaan worden en de kennis die de rapporten leveren. Hoewel bijvoorbeeld in de evaluatie van het Nederlands Investerings Matching Fonds (NIMF) meermalen over effectiviteit (doeltreffendheid) van beleid wordt gesproken, blijkt het bij nader inzien vooral over de doeltreffendheid van het beheer van het NIMF te gaan. Dit evaluatierapport is ook om een andere reden interessant. In de studie wordt aangegeven dat ook de doelmatigheid van het beheer geëvalueerd is. Doelmatigheid wordt echter beperkt opgevat. Er wordt vooral gekeken of men zich heeft gehouden aan bepaalde criteria. Van een vergelijking van kosten en baten, zoals bij een beoordeling van doelmatigheid te verwachten, is geen sprake. Daarmee is dit onderzoek op dit punt geen doelmatigheidsonderzoek, maar eerder een *compliance* studie: er is nagegaan of men zich aan de regels heeft gehouden.

In de evaluatie van het mensenrechtenbeleid wordt aangegeven dat in de deelstudies het accent op efficiëntie en effectiviteit ligt, terwijl het daarbij meestal gaat om voorwaardenscheppende omstandigheden.

De evaluatie van het Nederlands voorzitterschap van de EU geeft ook enige aanleiding tot verwarring over de aard van de uitspraken die gedaan worden. Te lezen is dat het doel van deze evaluatie, naast verantwoording afleggen over de wijze van voorbereiding en uitvoering van het voorzitterschap, vooral het trekken van lessen is uit de ervaringen met het voorzitterschap. Voorts worden in de rapportage herhaaldelijk conclusies over effectiviteit getrokken, terwijl de opzet van het onderzoek daar geen basis voor biedt. Wat gebeurd zou zijn wanneer het Nederlandse voorzitterschap niet had plaatsgevonden of op een andere manier was vormgegeven dan door IOB is beschreven, is immers onbekend. De wijze waarop de methodologie wordt toegelicht kan ook onbedoeld hebben bijgedragen aan onduidelijkheid over de aard van de uitspraken. Zo wordt gesteld dat de beantwoording van de onderzoeksvragen vraagt om een combinatie van een procesevaluatie en van een productevaluatie. De processen betreffen dan het interdepartementale coördinatieproces ten behoeve van een Nederlandse inzet, de onderhandelingsprocessen in de Raad en het organiseren van diverse evenementen in het kader van het voorzitterschap. De producten bestaan dan uit een inzet van het Nederlandse voorzitterschap, de uitkomst van de onderhandelingen, een positief Nederlands imago, en evenementen. De gebruikte termen zijn verwarrend. Doorgaans zijn productevaluaties metingen van de daadwerkelijke gevolgen van interventies of processen, terwijl het in deze evaluatie gaat over inputs, outputs en neveneffecten. De onduidelijkheid neemt niet af als nagegaan wordt hoe deze resultaten gemeten zijn. Ook een in evaluatieonderzoek geoefende lezer weet, na zekere tijd, niet meer wat nu de aard van de uitspraken is: gaat het om de wijze waarop processen verliepen, gaat het om de doelmatigheid daarvan (inclusief de uitgaven), of gaat het over de doeltreffendheid van het voorzitterschap?

Houding ten aanzien van evaluaties en de evaluatiecultuur

Bij aanvang van het onderzoek is getracht op grond van een aantal oriënterende gesprekken en documenten een beeld te vormen van de cultuur binnen het ministerie, als mogelijke verklaring voor wat aan mate van gebruik en leren zou worden aangetroffen. Zo wordt regelmatig op de

gesloten cultuur binnen het departement geweest, die past bij het vak van diplomaat. Die cultuur zou niet gericht zijn op het delen van kennis en van bronnen van kennis, maar eerder op het ontwikkelen van een netwerk van persoonlijke contacten, die op het juiste moment informatie kunnen leveren of kunnen doorverwijzen. Verder werkt men bij het ministerie volgens velen te kort in een functie om verdieping van specialistische kennis mogelijk en profijtelijk te maken. Dat is een van de redenen waarom de typische medewerker er naar zou streven om generalist te zijn en geen specialist. De tweede reden is dat men vooral bedreven is in het vaardig hanteren van procedures in plaats van te kiezen voor een inhoudelijk specialisme. Er wordt ook gewezen op het feit dat binnen een dergelijke cultuur de noodzaak of prikkels voor gebruik van evaluaties ontbreken.

| 58 |

Dit profiel wordt door een groot aantal geïnterviewden herkend, maar niet of nauwelijks op de eigen situatie toegepast. Het is niet mogelijk om in dit onderzoek de houding van medewerkers ten aanzien van evaluaties of de cultuur binnen het ministerie ten aanzien van leren te beschrijven. Maar de resultaten van de interviews laten wel zien dat een andere visie en een andere cultuur binnen de organisatie mogelijk en zelfs tot op zekere hoogte aannemelijk is. Die visie houdt in dat er geen sprake is van een evaluatievijandige opstelling, maar van een houding die varieert van geïnteresseerd tot nieuwsgierig. Over het algemeen weet men in grote lijnen ook aan te geven wat bruikbare resultaten zouden zijn en gebruikt men ook andere kennisbronnen, zoals rapporten van adviesraden of adviescommissies en van buitenlandse organisaties. Een aantal keren wordt de gedachte uitgesproken dat het ministerie of de betreffende directie geen kritische rapportages en opmerkingen kan negeren, omdat men in een glazen huis leeft. De beperkingen voor leren over de inhoud van beleid als gevolg van de frequente veranderingen van functie zijn mede daarom te relativiseren. Het gebruik van persoonlijke netwerken, passend bij een diplomatencultuur, hoeft het verwerven van nieuwe kennis niet uit te sluiten.

Bij het DGPZ, DGES en DGRC komt men een meer sceptische visie tegen op het nut van evaluaties. Hier neigt men er vaker toe om evaluaties als een vorm van corvee te beschouwen. Men acht de mogelijkheden om bruikbare evaluaties uit te voeren gering vanwege de beperkte invloed die de betrokken directies kunnen uitoefenen op het diplomatieke eindresultaat. Verder kent de diplomatie meer vertrouwelijke informatie. Deze meer

sceptische visie kan blijkens voorbeelden echter nog steeds samengaan met een zakelijke bereidwilligheid om aanbevelingen, die men bruikbaar acht, over te nemen.

Twee beperkingen vallen op in zijn algemeenheid. Vrijwel nergens komt men een spontane voorstelling tegen van het aanscherpen van beleidstheorieën en interventielogica's op grond van via evaluatieonderzoek verworven inzicht in de causale factoren en processen die menselijk gedrag reguleren. Verder is het niet gebruikelijk om actief kennis uit evaluaties te zoeken. Men beperkt zich tot de rechtstreeks aan de directie gerichte evaluaties van beleid en programma's in eigen portefeuille en zoekt niet naar lessen die elders zijn te leren. Dat geeft ook aan dat secundair gebruik van de evaluaties uit de evaluatieprogrammering weinig voorkomt.

Verspreiding van resultaten

Er is weinig bekend over de verspreiding van de resultaten van evaluaties buiten de kring van directe gebruikers. De rapporten en ToR bevatten er meestal weinig gegevens over en de respondenten zijn er over het algemeen niet goed van op de hoogte. Het is gebruikelijk dat IOB door middel van workshops en samenvattende publicaties resultaten bekend maakt in landen waar evaluatieonderzoek heeft plaats gevonden. De indruk bestaat dat bij evaluaties, waarbij geen substantieel onderzoek in ontwikkelingslanden heeft plaats gevonden, verspreiding van de resultaten beperkter plaatsvindt. Dit geldt ook voor de evaluaties van de andere DG's. De evaluaties, die betrekking hebben op aspecten van de sectorale benadering en de evaluatie van het sectorbeleid in het bijzonder, vormen een uitzondering. Vooral aan deze laatste wordt in bredere kring, zowel nationaal als internationaal, ruimere bekendheid gegeven. Deze evaluaties hebben dan ook een algemenere geldigheid dan de meeste andere evaluaties in het onderzoek. Een ander voorbeeld van verspreiding in brede kring geeft de evaluatie van coördinatie, complementariteit en coherentie in de Europese ontwikkelingssamenwerking. De verspreiding van de studie is zeer professioneel opgezet. Vooral beïnvloeding van politieke initiatieven en besluitvorming in Europese kring was van meet af aan doel van deze evaluatie.

De evaluaties in de steekproef kennen een duidelijk aanwijsbare doelgroep en hebben doorgaans een beperkte algemene geldigheid. Bijzondere

inspanningen voor verspreiding van de evaluaties in de evaluatieprogrammering dragen waarschijnlijk niet bij tot meer gebruik of tot gebruik in bredere kring. Dit geldt te meer voor IOB-evaluaties, waarvoor al een vorm van *imposed use* door beoogde gebruikers bestaat. Na het verschijnen van de evaluatie dient een beleidsreactie te worden opgesteld.

Deelname van beleidsmedewerkers aan de referentiegroep van een evaluatie is waarschijnlijk een belangrijk mechanisme om verspreiding en gebruik van de evaluatieresultaten te bevorderen. Twee respondenten vinden deelname aan een referentiegroep zelfs de best denkbare leeromgeving om over een evaluatie geïnformeerd te raken. Als belangrijkste voordelen wordt gezien dat er intensief wordt gediscussieerd over punten van belang en er meteen toelichting kan worden gevraagd aan de onderzoekers en experts. Volgens een andere respondent past deze vorm van leren ook goed bij de heersende diplomatencultuur op het ministerie. Men leert vooral via persoonlijke contacten.

| 60 |

Helderheid gebruiksdoel en tijdige beschikbaarheid evaluatie-uitkomsten

Bij de meeste onderzoeken is het gebruiksdoel omschreven in de ToR en/of het rapport van de evaluatie. Die omschrijvingen zijn vaak betrekkelijk algemeen en standaard. De evaluaties zijn in de meeste gevallen bedoeld om te verantwoorden en om van te leren. Dit onderzoek heeft laten zien dat het geregeld voorkomt dat de beleidswijzigingen die op de evaluatie volgen niet in de omschrijving van het gebruiksdoel van de studies is opgenomen. De data van geplande en daadwerkelijke oplevering zeggen weinig over de tijdige oplevering van een evaluatie. In referentiegroepen en door interim- en conceptrapportages worden veel evaluatieresultaten vaak al ruim voor de definitieve opleverdatum van het rapport bekend en komen ze beschikbaar voor gebruik. Slechts van één geval is bekend dat de resultaten van een onderzoek te laat beschikbaar waren om te worden gebruikt voor het bestemde doel.

Uitvoerbaarheid aanbevelingen

Bij twee evaluaties van diplomatieke inspanningen levert de beleidsdirectie kritiek dat de evaluatie weinig realistische veronderstellingen hanteerde over de geëvalueerde activiteiten en daarom ook weinig realistische aanbevelingen deed. In één van die gevallen neemt de directie een van deze aanbevelingen toch deels ter harte. Bij de overige evaluaties werd geen melding gemaakt van een gebrek aan uitvoerbaarheid van aanbevelingen.

Belang methodologische aspecten

Debatten waarin methodologische aspecten van evaluaties aan de orde komen vinden slechts beperkt plaats. Bij die voorbeelden blijkt dat methodologische of andersoortige controverses niet betekenen dat aanbevelingen in de wind worden geslagen. Eventuele controverses zitten dus het zakelijk beoordelen en opvolgen van aanbevelingen niet in de weg. Aan de onafhankelijkheid van het IOB wordt slechts zelden gerefereerd. Voor de meeste respondenten is die onafhankelijkheid een gegeven.

6

Leren van evaluaties: een beknopt literatuuroverzicht en resultaten uit het empirisch onderzoek

6.1. Inleiding

De beelden over wat leren van evaluaties inhoudt, verschillen. De vraag, die voorligt, is wat men onder leren moet verstaan en in hoeverre er van evaluaties geleerd wordt. Op grond van een beknopte analyse van enige wetenschappelijke literatuur en bevindingen uit het onderzoek wordt in dit hoofdstuk hiervan een beeld gegeven. Daarnaast wordt ook een aantal voorzieningen vermeld waarover een organisatie zou moeten beschikken om het leren van haar medewerkers te bevorderen en te faciliteren en daarmee te institutionaliseren.

6.2. Enige theoretische noties over leren

In haar onderzoek naar de vraag of van parlementaire onderzoeken geleerd wordt stelt de Algemene Rekenkamer (2007) dat leren te maken heeft met het vergroten van kennis en kunde. De Rekenkamer onderscheidt drie vormen van leren. Leren kan ten eerste gericht zijn op het voorkomen van fouten. Dit gebeurt meestal naar aanleiding van incidenten. Een tweede vorm van leren is gericht op het vinden van oplossingen voor meer structurele (beleids-)problemen. De derde vorm van leren is gericht op de voortdurende ontwikkeling van nieuwe oplossingen. De Rekenkamer merkt ook op dat leren er toe kan leiden dat mensen of organisaties iets (een handeling, een procedure enzovoort) dat al tot hun repertoire behoorde, beter gaan beheersen. Leren kan ook een meer fundamenteel karakter hebben en leiden tot een herbezinning op de uitgangspunten van het beleid. Verder kan ook van parlementair onderzoek geleerd worden doordat er meer kennis ontstaat over het betreffende beleidsterrein, op basis waarvan het beleid kan worden bijgesteld.

| 63 |

Waar de Rekenkamer tot op zekere hoogte op voortbouwt is het onderscheid tussen 'single-loop learning' en 'double-loop learning' (Argyris and Schön 1978, Argyris and Schön 1996 en Leeuw, Rist and Sonnichsen 1993). Van 'enkele lus leren' is sprake als van fouten geleerd wordt en die kennis ingezet wordt voor het verbeteren van wat de organisatie al doet. Van 'dubbele lus leren' is sprake als, naast het detecteren van uitvoeringsproblemen en -fouten, de organisatie ook bestaande procedures, beleid en doelstellingen ter discussie stelt. Ofwel de operationele normen zelf staan

voortdurend ter discussie en worden bijgesteld. Een belangrijk aspect van deze vorm van leren is dat de aandacht vooral uit gaat naar het reflecteren op de interventietheorie, de beleidstheorie(ën), respectievelijk het verband tussen probleem, beleid en beleidsdoelstelling achter beleidsprogramma's en andere interventies. Een derde vorm van leren (deutero leren) is dat organisaties weten wanneer zij, met het oog op kwaliteit van de dienstverlening of de effectiviteit van het beleid, moeten kiezen voor enkele of voor dubbele lus leren.

Het beeld over het leren van evaluatieachtig onderzoek is op grond van de verschillende Nederlandse studies niet zeer helder; er zijn voorbeelden van leren, maar ook het omgekeerde vindt plaats. Of leren ook afleren inhoudt is eveneens onduidelijk. Saskia Stuiveling (President Algemene Rekenkamer, 1999 – heden) presenteerde onlangs een interessante verklaring waarom overheden soms niet leren van onderzoeken en ervaringen. Ze zei: “De grootste leerblokkade voor overheidsorganisaties zie ik echter in het gegeven dat de politieke leiding op *damage control* en daarmee dus op defensief gedrag is ingesteld. Door deze houding krijgen organisaties tot in hun haarvaten vooral prikkels richting defensieve routines waardoor er min of meer ingebouwde leerbarrières ontstaan en in stand worden gehouden”⁸.

| 64 |

Verder is het van belang onderscheid te maken tussen het leren van individuen en het leren van organisaties, tussen individueel en institutioneel leren. Het is in eerste instantie het individu dat leert en dat nieuwe feiten, inzichten, verklaringen en theorieën moet recipiëren, verwerven en verwerken en vervolgens zijn gedrag aanpassen om zo nieuwe en betere resultaten te bereiken. Echter, individueel leren is niet meer dan een cruciale randvoorwaarde voor leren op organisatieniveau. March and Olsen (1975) deden een poging individueel en institutioneel leren met elkaar te verbinden. In hun benadering leiden individuele leerervaringen tot individuele actie, die, wanneer er sprake is van een accumulatie, tot acties op organisatieniveau leiden, die vervolgens weer feedback naar individuen en hun netwerken opleveren. Daarbij is Schön's concept van de *reflective*

⁸ Van Slingerlandt lezing “Hardop denken”. 2007. Toespraak 09-10-2007. http://www.rekenkamer.nl/Actueel/Toespraken/Van_Slingerlandt_lezing (laatst bezocht: 8 juni 2010).

practitioner van betekenis. Zo'n beleidsmaker is zowel in staat aan *reflection-in-action* te doen, als aan het meer traditionele *reflection-on-action*. Dat betekent dat:

“the practitioner allows himself to experience surprise, puzzlement, or confusion in a situation which he finds uncertain or unique. He reflects on the phenomenon before him, and on the prior understandings which have been implicit in his behaviour. He carries out an experiment which serves to generate both a new understanding of the phenomenon and a change in the situation”
(Schön 1983, 68).

Om het leren van haar medewerkers te bevorderen en te institutionaliseren, dient een organisatie een doelgericht beleid te voeren en voorzieningen te bieden, die het leren faciliteren. Wikipedia vult deze redenering in een artikel over *organizational learning* als volgt aan:

“A learning organization actively creates, captures, transfers, and mobilizes knowledge to enable it to adapt to a changing environment. (...) Capturing individual learning is the first step to making it useful to an organization. There are many methods for capturing knowledge and experience, such as publications, activity reports, lessons learned, interviews, and presentations. (...) Capturing includes organizing knowledge in ways that people can find it; multiple structures facilitate searches regardless of the user’s perspective (e.g., who, what, when, where, why and how). Capturing also includes storage in repositories, databases, or libraries to ensure that the knowledge will be available when and as needed. (...) Thus, a culture that learns and adapts as part of everyday working practices is essential. (...) Adapting an idea must be rewarded along with its initial creation. Sharing to empower the organization must supersede controlling to empower an individual.” (Wikipedia 2009⁹).

| 65 |

Samenvattend

Om te leren van evaluatieonderzoek is in eerste instantie nodig dat individuele afnemers van de aangeboden kennis, de kennis overnemen, opslaan en verwerken. Vervolgens moet die kennis zowel tijdens als na handelingen leiden tot reflectie daarop. Parallel daaraan is nodig dat

⁹ Wikipedia. 2009. http://www.en.wikipedia.org/wiki/Organizational_Learning (laatst bezocht: 2 augustus 2009).

meerdere actoren met elkaar van gedachten wisselen over wat zij individueel leren met het doel het geleerde vast te houden en in (handelings-)repertoires op te slaan. Deze kunnen vervolgens weer door anderen gebruikt worden bij het analyseren van bestaand beleid en het maken van nieuw beleid. Het opslaan van leerervaringen kan op verschillende manieren plaatsvinden. Behalve medewerkers zijn *repositories* ofwel kennisschatkamers, workshops en seminars, databases, wiki's en traditionele bibliotheken van belang.

6.3. Leren van evaluaties bij het ministerie van Buitenlandse Zaken

Evaluaties worden op verschillende manieren gebruikt. Rechtstreekse verbanden tussen afzonderlijke evaluaties en beleidsaanpassingen zijn echter moeilijk aan te tonen. Wel is er in de periode, die het onderzoek bestrijkt, sprake van een aanzienlijke mate van samenhang tussen de inhoud en teneur van evaluaties en die van het beleidsdebat en beleidswijzigingen. Daarnaast laten de bestudeerde casus ook een aantal voorbeelden zien van processen van verwerving en uitwisseling van kennis gericht op het aanpassen van beleid, die (mede) naar aanleiding van evaluaties in gang zijn gezet. Elementen uit de eerder gegeven typering van leerprocessen in organisaties zijn er in te herkennen. Het betreft diverse activiteiten als bijeenkomsten, internetdiscussies, reviews en experimenten met doelen, die overeenkomen met de hierboven aangehaalde omschrijvingen. Naar aanleiding van de evaluatie van mensenrechtenbeleid is bijvoorbeeld sprake van het introduceren van nieuwe instrumenten ter ondersteuning van een meer strategische en professionele benadering van het onderwerp mensenrechten. Daarvoor worden richtlijnen en voorbeelden voor ambassade medewerkers ter beschikking gesteld via terugkomdagen, intranet en interne publicaties. Naar aanleiding van de evaluatie van gedetineerdenbegeleiding in het buitenland zijn onderzoeken en experimenten gestart betreffende onderwerpen waarover aanbevelingen werden gedaan. Naar aanleiding van de evaluatie van algemene ambtberichten zijn expertbijeenkomsten georganiseerd om informatie in te winnen over ervaringen en opvattingen van de gebruikers van de berichten. Verder zijn er *pilots* gestart naar aanleiding van aanbevelingen van de evaluatie.

Naar aanleiding van de conceptrapportage over MDTF's (*Multi Donor Trust Funds*) is een conferentie belegd ter bespreking van *“ways to improve on this instrument as an effective tool for supporting countries in early recovery and reconstruction”*. Vervolgens werd ter voorbereiding van een nieuwe bijeenkomst in opdracht van het ministerie een synthese opgesteld van diverse evaluaties met betrekking tot hetzelfde onderwerp. Het doel van deze bijeenkomst was *“to share lessons learned and best practices in order to build consensus and identify concrete steps (...) to improve pooled funding mechanisms in post-conflict situations”*. De sectorbenadering is in brede kring besproken, mede naar aanleiding van de evaluatie en voor een belangrijk deel gefaciliteerd door het ministerie, via mail, web, terugkomdagen, publicaties en consultatiebijeenkomsten met interne en externe experts. Voorts is er een traject gepland voor herziening van de benadering waarmee de lessen, die de afgelopen 10 jaar zijn geleerd door diverse evaluaties, door contacten met internationale partners en door ervaringen van medewerkers ten goede moeten komen aan nieuw beleid.

Of deze activiteiten zijn te betitelen als leren is afhankelijk van de definitie van en visie op leren waar men van uit gaat. Er zijn meerdere definities en visies mogelijk. In elk geval is bij de beschreven activiteiten een aantal kwalificaties te plaatsen. In de eerste plaats gaat het eerder om incidentele voorbeelden dan om standaardpraktijken binnen de organisatie. In de tweede plaats geven de gebruikers van de evaluaties zelf desgevraagd aan dat zij niet het idee hebben iets nieuws te leren van de evaluaties. De evaluaties leveren hen, naar eigen zeggen, weinig instrumentele kennis op over hoe een gewenst resultaat te bereiken. Een aantal respondenten schrijft het gebrek aan nieuwe en leerzame inhoud toe aan het feit dat de evaluaties te veel zijn gericht op het verantwoorden van beleid en te weinig op leren. Dit gemis is al in verband gebracht met het procesgerichte karakter van evaluaties, die vooral de procedurele en organisatorische voorwaarden voor het voeren van een beleid zoals bedoeld in kaart brengen en toetsen, maar weinig kennis leveren over wat werkt. Wat ontbreekt is een belangrijk kenmerk van het dubbele lus leren zoals omschreven in paragraaf 6.2, en een reflectie op de interventietheorie, de beleidstheorie(ën), respectievelijk het verband tussen probleem, beleid en beleidsdoelstelling achter beleidsprogramma's. In de derde plaats zijn de cultuur en houding inzake het gebruik van evaluaties weliswaar niet evaluatie-vijandig, maar ook niet bepaald leergierig. Uit de gesprekken met respondenten bleek niet dat het gebruikelijk is om actief kennis uit evaluaties te zoeken. Men beperkt zich over het algemeen tot de evaluaties, die rechtstreeks aan de eigen directie worden

aangeboden en die betrekking hebben op programma's in eigen portefeuille. Men zoekt niet naar lessen die elders zijn te leren. Ook het feit dat het type procesgerichte evaluaties zo vaak is te herhalen, laat zien dat de organisatie op zijn minst enigszins hardleers is als het gaat om het op orde brengen van procedurele en organisatorische voorwaarden voor een adequate beleidsuitvoering.

Tot slot is het duidelijk dat de ambities met betrekking tot leren binnen het ministerie verder reiken dan wat tot nu toe wordt gerealiseerd. Dat blijkt uit klachten over het gebrek aan leren, maar ook uit diverse pogingen om van het ministerie een meer lerende organisatie te maken. In de Kamer gaf de minister een overzicht van initiatieven, die er toe moeten leiden dat onderzoek en evaluatie vaker worden ingezet voor het ministerie als lerende organisatie. Daarin werd onder meer gewezen op de rol van de eerder vermelde kennis- en onderzoeksstrategie (KOS). In het kader van de IS-academie werkt het ministerie via meerjarige samenwerkingsovereenkomsten samen met wetenschappelijke instellingen. In deze samenwerking participeren beleidsmedewerkers, wetenschappers, studenten en jonge onderzoekers. Het ministerie steunt verder het *Development Policy Review Network* van wetenschappers en beleidsmedewerkers, werkzaam op het terrein van ontwikkelingsamenwerking in Nederland en Vlaanderen. Er is voorts een aantal samenwerkingsverbanden met betrekking tot specifieke beleidsvragen. Voorbeelden hiervan zijn samenwerking op het gebied van islam en het thema migratie, en op het terrein van goed bestuur door middel van het Kenniscentrum Rechtsstaat en Ontwikkelingsamenwerking (KREO). Andere initiatieven om kennisverwerving en samenwerking tussen beleidsmakers en wetenschappers te bevorderen zijn:

- de aanstelling per 1 juli 2007 van een Wetenschappelijk Raadsadviseur als spil in de kennis- en strategienetwerken van het ministerie;
- het in het leven roepen van de functie van ambassadeur voor ontwikkelingsamenwerking (AMAD-OS), die zich onder meer bezig houdt met het ontwikkelen van kennis en deskundigheid in het DGIS, het onderhouden van contacten met kennisinstellingen en het realiseren van de moderniseringsagenda voor ontwikkelingsamenwerking;
- een in 2006 ingestelde promotieregeling voor ambtenaren van het ministerie om beleidsrelevant onderzoek te doen;
- de instelling van enkele BZ-leerstoelen, die bij toerbeurt worden bezet;
- de uitwisseling van ambtenaren en wetenschappers door detachering.

Een andere aanzet is de kennisagenda. Deze agenda is gebaseerd op een visie ten aanzien van kennismanagement en op de filosofie van de kleine stapjes aldus een van de geïnterviewden. Er wordt een veelheid aan kleine stapjes gezet, die elk afzonderlijk gemakkelijk worden geaccepteerd, maar tezamen een dynamiek moeten krijgen tot grote veranderingen. Er is geen specifiek plan of mandaat en er staat bijna niets op papier. Voorbeelden van dit soort stappen zijn de eerder vermelde initiatieven van de IS-Academie en de uitwisseling van ambtenaren en wetenschappers door detachering. Voorts is er het onder de loep nemen van het trainingsbeleid, met de vraag of en hoe het meer geïnteresseerd raken van medewerkers in kennis kan worden bevorderd.

Al deze initiatieven vormen echter nog geen duidelijk en samenhangend beleid dat gestoeld is op een visie over wat leren voor de organisatie moet betekenen. Wel is er een aantal vermeldenswaardige noties en zijn er aanzetten, die bouwstenen voor een visie en beleid kunnen leveren. Hieronder volgt een toelichting.

In de notitie van DEC over effectiever gebruik van evaluaties is, zij het voor één directoraat-generaal, een aanzet gepresenteerd van wat een gewenste samenhang tussen evaluaties, het afleggen van verantwoording en het daaruit leren van lessen, kan zijn. Een van de uitgangspunten van de notitie is dat evaluaties zowel een verantwoordings- als een leerfunctie hebben. Een ander uitgangspunt is dat evalueren beter dan monitoring in staat is lessen te trekken voor toekomstig beleid, omdat het oog heeft voor doelbereiking en verklarende factoren. Ten aanzien van het leren van evaluaties is een van de vertrekpunten dat leren in eerste instantie een intern karakter draagt. Leren komt voort uit de behoefte tot onderbouwing, regelmatige toetsing en, waar nodig, bijstelling van de interventielogica en bijbehorende strategieën, die in het kader van OS worden gebruikt. Ook wordt aangegeven dat evaluaties een functie hebben om vastroesten van beleid tegen te gaan. Vervolgens wordt over een externe leerfunctie gesproken, deels geïnitieerd vanuit het parlement, maar ook vanuit partnerlanden. Hoe deze externe functie vorm gegeven zou moeten worden en in hoeverre dit extern leren als iets anders wordt beschouwd dan het afleggen van verantwoording, wordt niet duidelijk.

Aan het einde van de notitie wordt een aantal aanbevelingen geformuleerd, onder meer de aanbeveling dat IOB de handreiking voor activiteiten-

evaluaties dient aan te scherpen, de posten hun leer- en informatiebehoefte beter dienen te omschrijven en worden er suggesties gedaan hoe de beleidsmatige follow up van evaluaties beter getraceerd kan worden. Daarbij wordt onder andere gewezen op het belang van internetvoorzieningen. Hoewel deze notitie een poging is om tot een visie te komen over de relatie tussen evalueren, verantwoorden en leren, wordt de vraag wat het ministerie verstaat onder leren niet beantwoord. Wie leert, wat leren eigenlijk is, wat het doel is van leren, hoe leren beklijft en hoe wat geleerd wordt zichtbaar is en gedeeld kan worden met anderen, of er beperkingen gelden of gesteld worden aan leren, blijft onduidelijk.

Tot slot volgen hieronder enkele ideeën over de rol van leren binnen het ministerie, die interessante elementen en aspecten van een meer overkoepelende visie kunnen uitmaken.

| 70 |

Het belang van het verschil tussen individueel en institutioneel leren wordt binnen het departement door beleidsbepalende medewerkers onderkend. Tevens wordt onderkend dat instrumenten om leren door individuen te bevorderen niet altijd hoeven te werken ten gunste van het leren door de gehele organisatie. Verder bestaat de gedachte dat wezenlijke onderdelen van leren in ieder geval zijn dat medewerkers *out-of-the-box* moeten kunnen denken, daartoe gestimuleerd moeten worden, innovatief dienen te zijn en er op afdelingsniveau een verband moet bestaan tussen het leren van individuen en de kennis en onderzoekstrategie. Van dit alles is, volgens een respondent, thans nog te weinig sprake. Ook hield deze respondent een pleidooi voor *education permanente*, op een nieuwe, niet vrijblijvende manier. Centraal daarbij staat de gedachte dat medewerkers er punten mee kunnen verdienen, punten die werken als prikkel tot verder leren en onderdeel vormen van het personeelsbeleid.

Een andere notie van een andere geïnterviewde betreft het zogenaamde '*contrecoeur* leren'. Daarvan lijkt de kern te zijn dat wanneer er institutionele, politieke druk is om uit een evaluatie of de discussie daarover lering te trekken, dat ook gebeurt. Aansluitend gaf deze geïnterviewde aan dat wanneer er, zoals in het geval van de evaluatie van het Europese Voorzitterschap, in termen van bevindingen en methode aanzienlijke onenigheid is over de evaluatie, dat niet noodzakelijkerwijze slecht is met het oog op de leerfunctie van zo'n rapport. Immers, er is dan ook flinke

aandacht voor de evaluatie, die gedurende zekere tijd wordt vastgehouden en er mogelijkwijze ook toe leidt dat onderliggende veronderstellingen ter discussie komen te staan. Grotendeels haaks op deze notie staat de notie dat druk van buiten leren juist afremt: *damage control* lijkt dan centraal te gaan staan en dat zou het leren van evalueren minder kansen geven. Een visie waarvan bijvoorbeeld de evaluatie van de Afrika-evaluatie getuigt (Gaay Fortman en de Ruijter 2008, 19). De casus laten in elk geval zien dat controversie geen belemmering hoeft te zijn voor gebruik van een evaluatie.

7

Conclusies, aanbevelingen en discussie

7.1. Conclusies

1. Evaluaties bij het ministerie van Buitenlandse Zaken laten diverse vormen van gebruik zien.

Er zijn verschillende vormen van gebruik van evaluaties aangetroffen.

In de eerste plaats worden evaluaties gebruikt als communicatiemiddel. Intern vooral ter ondersteuning van functie-overdracht. Evaluaties lenen zich daarvoor, omdat ze vaak het enige document zijn dat een overzicht van beleid bevat. Deze vorm van gebruik vindt, voor zover kon worden vastgesteld, alleen op individueel initiatief plaats en wordt vanuit de organisatie niet gericht gestuurd. Een enkele keer worden evaluaties gebruikt om een intern debat te voeren voorafgaande aan een beleidswijziging of alleen om een standpunt te bepalen, zonder dat een beleidswijziging volgt.

In de externe communicatie worden evaluaties vooral gebruikt om een standpunt of beleid te legitimeren, of als basis voor overleg met organisaties die regelingen uitvoeren of subsidies ontvangen. Sommige respondenten onderscheiden ook de promotiefunctie van evaluaties. Ze bieden een kans om jezelf weer eens in de schijnwerpers te plaatsen. Voor veel programma's is er doorgaans geen aanleiding om iets te melden, zolang het *business as usual* is. In een enkel geval heeft een evaluatie een rol gespeeld als *cognitive organizer* waarin een beleid is vastgelegd zodat verschillende partijen zich er aan kunnen binden of gebonden worden.

| 73 |

Evaluaties spelen geen belangrijke rol bij verantwoording van resultaatbereiking. Het lijkt er wel op dat die rol groeit binnen de resultatenrapportages en dat impactevaluaties daarbij een bijzonder gewicht krijgen.

2. De relatie tussen evaluaties bij het ministerie van Buitenlandse Zaken en beleidswijzigingen is complex en veelvormig.

Evaluaties gaan vaak gepaard met beleidsaanpassingen. Maar dat wil niet zonder meer zeggen dat ze verantwoordelijk zijn voor de aanpassingen of er in belangrijke mate toe bijdragen. De onderzochte casus van gebruik laten zien dat er veel inhoudelijke overeenkomsten zijn tussen de aanbevelingen

en conclusies uit evaluaties en beleidswijzigingen, maar dat dit niet noodzakelijk betekent dat een evaluatie leidde tot een er op volgende beleidshervorming. Vaak zijn er inhoudelijke overeenkomsten tussen conclusies en aanbevelingen van evaluaties en aanpassingen van programma's. In beleidsdocumenten wordt kennis genomen van conclusies en aanbevelingen van evaluaties, er wordt rekenschap van gegeven en ze worden gereproduceerd. Het is echter moeilijk om in afzonderlijke gevallen een lineair en causaal verband aan te tonen.

In een aantal gevallen is er sprake van een inhoudelijk verschil tussen evaluatie en de er op volgende beleidsaanpassing. De conclusies en aanbevelingen van de evaluatie lijken dan de erop volgende beleidswijziging niet helemaal te kunnen rechtvaardigen. Men zou kunnen zeggen dat er dan sprake is van overbenutting. Daarnaast komt het voor dat aanbevolen beleidsaanpassingen al tijdens de looptijd van de evaluatie of zelfs voorafgaande aan de evaluatie hebben plaats gevonden. Veel evaluaties binnen het DGIS bijvoorbeeld gaan gepaard met een herziening van het geëvalueerde beleid. Bij de andere drie DG's gaan de evaluaties meestal gepaard met beperktere aanpassingen of met het achterwege blijven van enige actie.

| 74 |

De aanpassingen binnen het DGIS komen tot stand op basis van twee overwegingen. Een eerste overweging is dat verschillende programma's niet meer goed passen in de nieuwe sectorbenadering. Een tweede overweging is juist min of meer te typeren als het opvangen van de gevolgen van de sectorbenadering, een verminderde aandacht voor de bestrijding van armoede en voor de productieve sectoren. Een aantal programma's werd op de schop genomen om dit gemis recht te zetten. Maar inhoudelijke overeenkomsten tussen de aanbevelingen en conclusies van evaluaties en beleidswijzigingen betekenen niet noodzakelijk dat een evaluatie leidde tot de er op volgende beleidsaanpassing. De evaluaties zijn zelden zelfstandig verantwoordelijk voor de grootscheepse programmawijzigingen bij het DGIS van de afgelopen jaren. De architectuur van het beleid voor ontwikkelingssamenwerking kent een krachtige zelfstandige dynamiek, die weinig behoefte heeft aan evaluaties om een hervorming door te voeren. Ook de gesignaleerde inhoudelijke kloof tussen betrekkelijk mild gestemde evaluaties en majeure hervormingen die er op volgen, laat zien dat evaluaties zelden zelfstandig verantwoordelijk waren voor de grootscheepse

programmawijzigingen bij het DGIS van de afgelopen jaren. Verder berusten inhoudelijke overeenkomsten tussen evaluatie en beleid in een aantal gevallen op beleidsaanpassingen, die voorafgaan aan de evaluatie. Deze onduidelijke en gecompliceerde verhouding tussen evaluaties en beleid is regelmatig beschreven in de literatuur over gebruik van onderzoek.

Voor deze verhouding tussen evaluaties en ermee gepaard gaande beleidsaanpassingen zijn twee verklaringen. In de eerste plaats is bekend dat een deel van de conclusies en aanbevelingen van evaluatieonderzoek al tijdens het onderzoek via deelname aan de referentiegroep bekend worden en ter harte worden genomen. Een aantal respondenten vindt deelname aan een referentiegroep de best denkbare leeromgeving om iets van een evaluatie op te steken. De tweede verklaring is dat belangrijke ideeën voor beleidswijzigingen uit meerdere bronnen afkomstig kunnen zijn. Evaluaties opgenomen in de programmering van het ministerie maken deel uit van een beleidsdebat waaraan evaluaties uit eerdere perioden, evaluaties van buitenlandse instellingen, adviezen, internationale verdragen en -verklaringen, maar ook politieke programma's en beleidsnotities bijdragen leveren. Noties over het belang van algemene doelstellingen als armoedebestrijding en gender-gelijkheid, van *harmonisation* en *alignment* om effectiviteit te verhogen, maar ook over het gebruik van concrete instrumenten als contextanalyses en verschillende monitoring- en evaluatiesystemen, figureren in dat debat. Een beleidsaanpassing die volgt op een aanbeveling in een evaluatie kan daarom ook gebaseerd zijn op een eerdere aanbeveling uit een andere bron. Dit gegeven maakt het moeilijk om in een concreet geval de invloed van een bepaald rapport op een beleidswijziging aan te tonen. Wel is er in de periode die het onderzoek bestrijkt sprake van een aanzienlijke mate van samenhang tussen de inhoud en teneur van evaluaties en die van het beleidsdebat en beleidswijzigingen. Deze samenhang is niet zonder belang voor de discussie over gebruik van evaluaties. De meest bescheiden conclusie is dat in elk geval een deel van de door evaluaties gesignaleerde problemen en voorgestelde oplossingen op de beleidsagenda staat en in aanpassingen resulteert.

3. De belangrijkste factoren van invloed op gebruik van evaluaties

Type kennis

Een belangrijke factor van invloed op het gebruik van de evaluaties van het ministerie van Buitenlandse Zaken betreft het type kennis dat de evaluaties produceren. Meerdere evaluaties kijken naar effecten en resultaatbereiking, maar er wordt geen relatie gelegd met daarvoor verantwoordelijke sociale, institutionele en gedragsmechanismen. De achterliggende beleids- en interventietheorie(ën) worden niet geproblematiseerd, gereconstrueerd en getoetst. Daardoor leveren de evaluaties geen of beperkt substantieve kennis over hoe ontwikkelingsresultaat te bereiken. De meeste evaluaties hebben in plaats daarvan eerder een bestuurskundig en procesmatig karakter. Indien effecten worden bepaald worden deze in verband gebracht met de vraag of het beleid is uitgevoerd zoals bedoeld, volgens een reconstructie van beleidsintenties en afspraken. Vergelijkbare opmerkingen met elementen uit deze diagnose zijn eerder door andere analyses van evaluaties uit de evaluatieprogrammering van het ministerie gemaakt. Zowel de inhoud van de rapporten als de analyses van respondenten leiden tot deze conclusies. De reacties van respondenten wijzen er op dat inmiddels de wet van de afnemende meeropbrengsten voor dit type evaluatie mogelijk in werking is getreden.

| 76 |

Overigens is niet gezegd, dat het type procesmatige evaluaties zoals beschreven geen gebruikswaarde hebben. Ze bevatten waardevolle informatie over hoe beleidsprocessen volgens recente inzichten in te richten. Ze maken echter niet duidelijk op welke manier een gewenst resultaat is te bereiken of waarom de processen volgens de instructies dienen te worden ingericht. Bovendien heeft de kennis die ze leveren over het algemeen een beperkte geldigheid. Substantieve kennis over hoe gedrag te beïnvloeden en zo resultaten te bereiken heeft een grotere generaliseerbaarheid en biedt daarom meer aanleiding en gelegenheid tot leren.

Verspreiding van evaluatieresultaten

In de literatuur wordt vaak verwezen naar het belang van diverse vormen van verspreiding van evaluatieresultaten. Over verspreiding van de resultaten van de evaluaties buiten de kring van beoogde gebruikers zijn er weinig gegevens. Het is gebruikelijk dat IOB door middel van workshops en samenvattende publicaties resultaten bekend maakt in landen waar

evaluatieonderzoek heeft plaats gevonden. De evaluaties voor de andere DG's kennen over het algemeen een beperkte circulatie. Aan de evaluaties van de sectorale benadering wordt in breder kring, zowel nationaal als internationaal, ruimere bekendheid gegeven. Deze evaluaties hebben vanwege het fundamentele belang van het onderwerp dan ook een algemenere geldigheid dan de meeste andere evaluaties in het onderzoek.

De evaluaties in de steekproef kennen een duidelijk aanwijsbare doelgroep en hebben doorgaans een beperkte algemene geldigheid. Bijzondere inspanningen voor verspreiding van de evaluaties in de evaluatieprogrammering zullen waarschijnlijk niet bijdragen tot meer gebruik of tot gebruik in bredere kring. Dit geldt te meer voor IOB-evaluaties waarvoor al de noodzaak bestaat om er iets mee te doen. Na het verschijnen van de evaluatie dient immers een beleidsreactie te worden opgesteld.

Deelname van beleidsmedewerkers aan de referentiegroep van een evaluatie is waarschijnlijk een belangrijk mechanisme om verspreiding en gebruik van de evaluatieresultaten te bevorderen. Er wordt intensief gediscussieerd over punten van belang en er kan meteen toelichting worden gevraagd. Volgens een andere respondent past deze vorm van leren goed bij de heersende diplomatencultuur op het ministerie. Men leert vooral via persoonlijke contacten.

| 77 |

Methodologische en inhoudelijke kwaliteiten

De respondenten zien voor gebruik van de evaluaties geen duidelijke rol voor andere in de literatuur genoemde factoren zoals validiteit en onafhankelijkheid van de evaluator. Deze methodologische kenmerken zijn over het algemeen weinig problematisch. Opvallend is dat controverses over methodologie van de evaluatie niet hoeven te leiden tot veronachtzaming van de resultaten van een evaluatie en een zakelijke waardering niet uitsluiten. Het blijkt ook dat dissensus over enkele rapporten juist bijdraagt aan het publieke debat.

Houding en cultuur

Er is geen evaluatie-vijandige opstelling aangetroffen, maar een houding, die varieert van in principe geïnteresseerd tot een zekere nieuwsgierigheid naar resultaten van evaluaties. Over het algemeen weten de respondenten in grote lijnen ook aan te geven wat bruikbare resultaten zouden zijn en

gebruikt men ook andere kennisbronnen, als leverancier van bruikbare ideeën, bijvoorbeeld rapporten van adviesraden of -commissies en van buitenlandse organisaties. Bij DGPZ en DGES komt men een meer sceptische visie tegen op het nut van evaluaties. Hier neigt men er vaker toe om evaluaties als een vorm van corvee te beschouwen. Men acht de mogelijkheden om bruikbare evaluaties uit te voeren beperkt, omdat het vermogen om eigen doelstellingen te kiezen en na te streven klein is in verhouding tot de rol van andere actoren. Diplomatieke inspanningen in multinationalaal verband kunnen geen eindresultaten bereiken, maar alleen beperkte en soms moeilijk herkenbare, meetbare bijdragen leveren, zo is de gedachte. Verder kent de diplomatie meer vertrouwelijke informatie die in evaluaties openbaar kan worden, dan ontwikkelings samenwerking, zo luidt de redenering.

| 78 |

Twee beperkingen vielen op. Vrijwel nergens komt men een spontane voorstelling tegen van het leren van gedragswetenschappelijke georiënteerde evaluaties van de doeltreffendheid van interventies en beleid; met andere woorden, evaluaties waarin de werkzaamheid van mechanismen van verandering een belangrijke rol speelt. In de tweede plaats is het niet gebruikelijk om actief kennis uit evaluaties te zoeken. Men beperkt zich tot de evaluaties van beleid en programma's betreffende de eigen portefeuille en zoekt niet naar lessen die elders zijn te leren.

Overige factoren

In het onderzoek zijn geen aanwijzingen gevonden dat kenmerken als:

- de aansluiting tussen onderzoeksvragen, bevindingen, conclusies en aanbevelingen in de evaluaties;
- de tijdigheid van de rapportages;
- de duidelijkheid van het gebruiksdoel van de studie; en
- de mate waarin aanbevelingen uitvoerbaar zijn voor betrokken medewerkers, bepalend zijn voor de aard en mate van gebruik van de resultaten.

Wel valt op dat cruciale evaluatieve begrippen als *doelmatigheid* en *doeltreffendheid* in de evaluaties, en soms zelfs binnen één evaluatie, niet altijd consistent worden gebruikt. Dat leidt tot verwarring over de aard van de uitspraken die gedaan worden en van de kennis die de rapporten leveren.

4. Wordt van evaluaties geleerd?

Er is een overeenkomst tussen de inhoud van evaluaties en die van het beleidsdebat, die relevant is voor een beoordeling van het leren dat naar aanleiding van evaluaties plaats kan vinden. Daarnaast laten de bestudeerde casus ook een aantal voorbeelden zien van processen van verwerving en uitwisseling van kennis die mede naar aanleiding van evaluaties in gang zijn gezet. Het betreft diverse activiteiten zoals bijeenkomsten, internetdiscussies, reviews en experimenten. Elementen van leerprocessen in organisaties zijn daarin te herkennen. Of deze activiteiten zijn te betitelen als leren, is afhankelijk van de definitie van en visie op leren waar vanuit wordt gegaan, er zijn meerdere definities en visies mogelijk. Een belangrijk element in die visies is het verschil tussen individueel en organisationeel leren. Om het leren van haar medewerkers te bevorderen en te institutionaliseren, dient een organisatie een doelgericht beleid te voeren en voorzieningen te bieden die het leren faciliteren. Het ministerie heeft een aantal voorzieningen en activiteiten in gang gezet die tot een lerende organisatie beogen te komen.

Echter, er is een aantal redenen om kanttekeningen te plaatsen. In de eerste plaats zijn de bovengenoemde voorbeelden van kennisuitwisseling naar aanleiding van evaluaties geen standaardpraktijk. In de tweede plaats geven de gebruikers van de evaluaties zelf desgevraagd aan dat zij niet het idee hebben iets nieuws te leren van evaluaties. De evaluaties leveren hen, volgens dit onderzoek, weinig substantieve kennis op over hoe met beleid resultaat te boeken. Dit gemis van respondenten is al in verband gebracht met het procesgerichte (bestuurskundige) karakter van evaluaties, die vooral de procedurele en organisatorische voorwaarden voor het voeren van een beleid zoals bedoeld in kaart brengen en toetsen, maar weinig kennis opleveren over gedrags- en andere mechanismen, die beleid effectief moeten maken. Ze missen daardoor een belangrijk kenmerk van dubbelus leren, namelijk de reconstructie en toetsing van de theorievorming achter beleidsprogramma's en andere interventies. In de derde plaats is de cultuur en houding inzake het gebruik van evaluaties niet bepaald een leergierige. Het is voor beleidsmedewerkers niet gebruikelijk om actief kennis uit evaluaties te zoeken. Men beperkt zich over het algemeen tot de evaluaties, die rechtstreeks aan de eigen directie worden aangeboden en die betrekking hebben op programma's in eigen portefeuille. Ook het feit dat het type procesgerichte evaluaties zo vaak is te herhalen, laat zien dat de organisatie op zijn minst gezegd enigszins hardleers is als het gaat om het

op orde brengen van procedurele en organisatorische voorwaarden voor een adequate beleidsuitvoering.

Tot slot is het duidelijk dat de ambities met betrekking tot leren binnen het ministerie verder reiken dan wat tot nu toe wordt gerealiseerd. Dat blijkt uit klachten over het gebrek aan leren, maar ook uit diverse initiatieven van het ministerie om een meer lerende organisatie te maken. Deze initiatieven vormen nog geen duidelijk en samenhangend beleid gestoeld op een visie over wat leren voor de organisatie moet betekenen. Wel is er een aantal vermeldenswaardige noties en aanzetten, die bouwstenen voor een dergelijke visie en beleid kunnen leveren.

Een notie is dat het verschil tussen individueel en institutioneel leren binnen het departement door beleidsbepalende medewerkers wordt onderkend. Tevens wordt onderkend dat maatregelen om leren door individuen te bevorderen, niet altijd hoeven te werken ten faveure van het leren door de gehele organisatie. Een andere interessante aanzet is de kennisagenda. Deze is gebaseerd op een visie ten aanzien van kennismanagement en op de filosofie van de kleine stapjes. Er wordt een veelheid aan kleine stapjes gezet, die elk afzonderlijk gemakkelijk worden geaccepteerd, maar tezamen een dynamiek moeten krijgen om tot grotere veranderingen te komen. Een aantal initiatieven, dat de organisatie meer ontvankelijk moet maken voor nieuwe kennis uit onderzoek, maakt deel uit van deze agenda. Een andere notie is *education permanente* voor medewerkers te stimuleren door het onderdeel te maken van het beoordelings- en beloningsbeleid. Een laatste interessante notie betreft het *contrecoeur* leren. Dit is de gedachte dat onenigheid over evaluaties niet noodzakelijkerwijs hoeft te leiden tot het verontachtzamen van de lessen die zijn te leren, maar juist, integendeel, kan leiden tot een vruchtbaar debat. De casus laten in elk geval zien dat controverserige geens belemmering hoeft te zijn voor gebruik van een evaluatie.

Niettemin ontbreekt in de diverse gedachten, aanzetten en initiatieven een visie op leren. Wie leert en dient te leren, wat leren eigenlijk is, wat het doel is van leren, welk type kennis het meest leerzaam is, of afleren ook leren is, hoe leren beklijft en hoe wat geleerd wordt zichtbaar is en gedeeld kan worden met anderen, of er beperkingen gelden of gesteld worden aan leren, zijn vragen die in het ongewisse blijven.

7.2. Aanbevelingen

De aanbevelingen hebben ten eerste tot doel het gebruik en leren van evaluaties verder te stimuleren en dat op een manier, die zichtbaar en duurzaam is en waarover verantwoording kan worden afgelegd. Er mee verbonden zijn aanbevelingen betreffende de aard van de kennis die evaluaties van het ministerie van Buitenlandse Zaken opleveren en hoe syntheses van onderzoeksbevindingen vorm te geven zijn. Ook zijn er enkele aanbevelingen ter zake van de verhouding van IOB onderzoeken tot vraagstellingen die niet het terrein van ontwikkelings samenwerking betreffen.

Aanbeveling 1: Schep meer duidelijkheid over wat onder gebruik en leren van evaluaties verstaan wordt

Regelmatig is in interviews en documenten het woord leren gebruikt. Doorgaans alsof het duidelijk is wat onder leren van te verstaan is, soms aangevuld met een poging het begrip wat te operationaliseren. Er blijken verschillende noties te bestaan over de mate waarin er van evaluaties geleerd wordt en in hoeverre het ministerie door een lerende cultuur gekenmerkt wordt. Waar het echter aan ontbreekt, is een theorie of visie over vragen als:

- wat leren is;
- waarom het belangrijk is dat er geleerd wordt;
- wie leert en dient te leren;
- wat er gebeurt als er niet-geleerd wordt;
- welke factoren bepalend zijn om tot een lerende omgeving en cultuur te komen;
- welke typen kennis het meest leerzaam zijn;
- of af-leren ook leren is; en
- of er beperkingen gelden of gesteld worden aan leren: een vraag vooral van belang in verband met een eventueel aan leren aangepast personeelsbeleid.

In lijn met buitenlandse literatuur en ervaringen is het daarom aan te bevelen tot een theorie te komen over wat leren op verschillende niveaus inhoudt en hoe mens en organisatie in BZ-verband geprikkeld kunnen

worden tot leren van evaluaties. Het zoeken van aansluiting bij concepten zoals instrumentele, conceptuele en reflexieve benutting van evaluaties, *single* en *double loop*-leren en afleren kan meer helderheid verschaffen. Dat geldt ook voor kennis, die uit studies afkomstig is hoe organisaties geleerd hebben en het vinden van indicatoren, die nodig zijn om te kunnen vaststellen wanneer mens en organisatie geleerd hebben. Daarom wordt aanbevolen om duidelijker te zijn over wanneer het ministerie tevreden is met bereikte leerniveaus.

Onderdeel van het komen tot een theorie over gebruik en leren dient ook een standpunt te zijn over de toepasbaarheid van het verschijnsel van ‘verplicht’ gebruik van de resultaten uit evaluatieonderzoek, zoals dat internationaal, maar in de vorm van bijvoorbeeld de Erkenningscommissie Gedragsinterventies Justitie, ook in Nederland, al bestaat. Dat is de strategie waarbij van overheidswege geëist wordt dat interventies of beleidsprogramma’s wetenschappelijk zijn beproefd en succesvol zijn, willen ze erkend en opnieuw gefinancierd worden. De goedgekeurde interventies komen dan op een lijst te staan die gebruikt moet worden bij het nemen van beslissingen over gedragsinterventies. Aan deze strategie kleven voor- en nadelen.

| 82 |

Aanbeveling 2: Zorg voor digitale (webbased) en andere voorzieningen om gebruik van kennis en leren te bevorderen

Het ministerie beschikt in beperkte mate over *webbased* en digitale voorzieningen om het leren van onderzoeken te bevorderen. Kennisschatkamers die elders wel al bestaan, die digitaal ontsluitbaar zijn en goed bezoekbaar, en onder andere gevuld zijn met resultaten van evaluaties van IOB en van andere onderzoeken en evaluaties, zijn niet aangetroffen. Databestanden die voor niet-onderzoekers toegankelijk zijn zoals StatLine evenmin. Elektronische leeromgevingen waarin wiki’s figureren, die betrekking hebben op resultaten van en leerervaringen met evaluatieonderzoeken, zijn op het terrein van evaluaties niet aangetroffen. Dit soort voorzieningen draagt bij aan een cultuur waarin leergierige en reflexieve medewerkers goed kunnen functioneren. Dit zal zeker gelden in een omgeving als bij het ministerie van Buitenlandse Zaken, dat gekenmerkt wordt door frequent wisselende functiebezettingen en door een decentrale uitvoering van beleid

door de posten. Elders wordt in dit verband gewezen op een zwakke rapportagecultuur van het ministerie, waardoor het moeilijk is om snel inzicht te krijgen in de details van een dossier over een langere termijn. De mogelijkheden van *open source internet* archivering blijven bij het ministerie onbenut.

Het gaat uiteraard niet alleen om digitale voorzieningen. Op ander vlak heeft het ministerie al aanzienlijk meer stappen gezet, waaronder de Academie voor Internationale Samenwerking (IS-Academie), een kennis-ambassadeur, een directie die zich bezig houdt met effectiviteit, participatie van medewerkers aan *IPDET* en in de toekomst wellicht aan andere master-opleidingen. Het is aan te bevelen om hier mee door te gaan en daarbij te opteren voor investeringen in activiteiten, die zich kenmerken door succesvol gebleken combinaties van *on-line* en *off-line* permanente educatie. Een betrekkelijk eenvoudig te realiseren voorziening met een mogelijk belangrijke impact op gebruik en leren van evaluaties is het consequent instellen van zorgvuldige samengestelde begeleidingscommissies of referentiegroepen, waarin de beoogde gebruikers van de evaluaties in elk geval zitting hebben. Te overwegen valt om deze praktijk uit te breiden met enkele experimenten waarin de uitvoering van evaluatieonderzoek en het denken over beleid, waar mogelijk, meer naast elkaar oplopen in trajecten.

Aanbeveling 3: Heb meer oog voor het type kennis dat evaluaties opleveren, voor de gevolgen daarvan voor het gebruik dat er van gemaakt wordt en voor de behoefte aan bepaalde (andere) typen kennis bij de organisatie

Er zijn verschillende typen evaluatieve kennis te onderscheiden. Zorgen bestaan er ten aanzien van de mate waarin (IOB-)evaluaties vooral bestuurskundige, procesgerichte en nalevingsgerichte informatie opleveren. In eerdere analyses van evaluaties op het terrein van ontwikkelingssamenwerking wordt ook op elementen van deze problematiek gewezen. Het is aan te bevelen meer helderheid na te streven over welke typen kennis meer en welke minder geproduceerd zouden moeten worden. Ten aanzien van impact-evaluaties zijn belangrijke stappen gezet, maar dat is niet voldoende. Ten eerste omdat het aantal impact-evaluaties nog maar zeer gering is en ten tweede omdat er ook behoefte is te onderkennen aan

andere typen kennis en studies. In de eerste plaats aan synthesestudies en in de tweede plaats aan studies, die de specifieke behoeften aan evaluatieve kennis binnen de verschillende Directoraten Generaal kunnen vervullen.

IOB stelt terecht dat het niet zo kan zijn dat essentiële delen van het beleid van het ministerie niet geëvalueerd zouden worden. Maar deze uitspraak legt tegelijkertijd een uitdrukkelijke verplichting op aan evaluatoren om met antwoorden te komen op de vraag hoe in de kennisbehoefte van alle DG's te voorzien is. Het zoeken van aansluiting bij de literatuur over *law & economics*, *institutional economics*, organisatiepsychologie en organisatie-economie zou daartoe interessante mogelijkheden kunnen bieden. Waar het de aandacht betreft, die uitgaat naar reconstructie en toetsing van interventietheorieën, respectievelijk beleidstheorieën kan geconstateerd worden dat flink wat evaluaties aandacht wijden aan beleidsreconstructies, maar dan vooral in termen van beleidsintenties, processen, procedures en institutionele voorzieningen en actoren. Meer aandacht voor de analyse van de gedrags-, sociale en institutionele mechanismen, die aan interventies, programma's en arrangementen ten grondslag liggen, is nodig om gebruik en leren van evaluaties verder te bevorderen.

| 84 |

Aanbeveling 4: Doe meer aan syntheses maar dan *state of the art*

De uitgevoerde en gerapporteerde evaluaties leveren per rapport een schat aan informatie op, die ook vaak per rapport op enigerlei wijze gebruikt wordt. Er werden echter veel minder syntheses aangetroffen van de geproduceerde kennis, die voor beleidsmedewerkers, de medewerkers op de posten, medewerkers van de Directie Financieel Economische Zaken, leidinggevend en anderen binnen en buiten het ministerie, toegankelijk en bruikbaar zijn. Het is daarom aan te bevelen om meer werk te maken van syntheses. Daarbij is het dienstig niet alleen eigen werk te synthesiseren maar vooral ook bevindingen van zusterorganisaties, OESO DAC, de Wereldbank en van extern wetenschappelijk onderzoek. Daarbij is het eveneens aan te bevelen meer te werken in lijn met de standaarden uit de internationale evaluatieliteratuur over syntheses.

Voor leken-geïnteresseerden in resultaten van evaluatieonderzoeken, is wellicht te overwegen een op de jaarlijkse rapportage van NORAD 'Results

Report' (2008) gelijkende vorm te kiezen. Interessant daaraan is dat de bevindingen van onderzoeken van deze organisatie niet alleen verbonden worden met thema's, maar juist expliciet aan bepaalde assumpties (lees: hypothesen).

Aanbeveling 5: Zorg dat verantwoording kan worden afgelegd over de aard en mate van benutting en leren op organisatieniveau

Tot begin jaren tachtig werden in Nederland evaluaties nog niet of nauwelijks verricht. Weinigen hoefden zich druk te maken over de ermee gepaard gaande rijksuitgaven, dan wel over het nut van dit soort werk. Die tijden liggen achter ons. Evaluaties zijn *booming*, maar tegelijkertijd neemt de vraag naar de doelmatigheid en doeltreffendheid ervan ook toe. Deze kritiek moet serieus genomen worden. Immers, zoals met elke menselijke activiteit kan ook evaluatieonderzoek te maken krijgen met de Wet van de Afnemende Meeropbrengsten. Daarom is het aan te bevelen tot een *tracing*-instrument te komen dat het gebruik van evaluatieonderzoek en van leerervaringen bijhoudt. Daarbij kan geleerd worden van wat elders in de Rijksdienst voorhanden is. Het is wel verstandig dit instrument niet al te zwaar aan te zetten, omdat de kosten anders wellicht hoger zijn dan de opbrengsten. Daartoe zijn mogelijkheden door gebruik te maken van de eerder vermelde *open source* internet archivering. Instituten als de Algemene Rekenkamer en het WODC maken al een langere tijd gebruik van relatief eenvoudige technieken om dit doel te bereiken.

Bijlage 1

Terms of Reference

Onderzoek naar het gebruik van evaluatieonderzoek in de beleidscyclus van het ministerie van Buitenlandse Zaken

1. Aanleiding tot, en doelstelling van het onderzoek

Het Ministerie van Buitenlandse Zaken (BZ) streeft naar goed gebruik van evaluatie, waarbij wordt aangesloten op de beleidscyclus voor de verschillende beleidsterreinen van het ministerie. Goed gebruik impliceert dat evaluatieonderzoek wordt benut om te komen tot een kwalitatief betere beleidsontwikkeling en uitvoering én betere beleidsverantwoording. Beide functies zijn van belang, mede in het licht van de rijksbrede Regeling Periodiek Evaluatieonderzoek en beleidsinformatie 2006 (RPE 2006).

In Nederland is sprake van toenemende aandacht voor beleidsverantwoording. Dit heeft onder meer geleid tot recente toezeggingen van de Ministers aan de Kamer dat (1) in jaarverslagen wordt aangegeven welke evaluaties het beleid op hoofdlijnen hebben beïnvloed en (2) een Raad van Advies adviseert over zowel het gebruik als de bruikbaarheid van IOB-evaluaties op het terrein van ontwikkelings samenwerking. Daarnaast bestaat de wens om het leren van evaluatieresultaten voor beleidsontwikkeling en uitvoering te bevorderen en beter te verankeren.

Op grond hiervan heeft het Audit Committee (AC) verzocht voorstellen te ontwikkelen voor een beter gebruik van evaluaties. Dit heeft geresulteerd in de notitie *'naar een effectiever gebruik van evaluaties door het DGIS'* (DEC, 2008). Voorts heeft het AC een onderzoek gelast naar het gebruik van evaluaties in de praktijk van de beleidscyclus bij het ministerie van Buitenlandse Zaken. De doelstelling van dit onderzoek is om meer inzicht te krijgen in het huidige gebruik van evaluatie in de beleidsontwikkeling en uitvoering alsmede beleidsverantwoording met het oog op de identificatie van knelpunten en het aandragen van oplossingen.

2. Beschrijving en afbakening van het onderwerp van onderzoek

Het onderzoek is beperkt tot beleidsevaluaties die deel uitmaken van de BZ-evaluatieprogrammering. Deze voortrollende meerjarige programmering bevat naast IOB-evaluaties, evaluaties die onder directe verantwoordelijkheid van beleidsdirecties worden uitgevoerd (in de regel via uitbesteding aan een consultant) evenals de zgn. derden evaluaties (uitgevoerd door een andere organisatie dan BZ).

De evaluatieprogrammering is een openbaar document en wordt - via de Memorie van Toelichting - ook aan de Kamer aangeboden. Ook de individuele evaluaties zijn - een enkele uitzondering daargelaten - openbaar en via internet toegankelijk. Voor alle IOB-evaluaties en voor sommige niet-IOB-evaluaties wordt door de betrokken bewindspersoon(onen) een beleidsreactie opgesteld. Deze beleidsreacties worden in de regel gelijktijdig met het evaluatierapport aan de Kamer aangeboden en zijn openbaar en via internet toegankelijk.

| 88 |

Het onderwerp van onderzoek is in de tijd beperkt tot evaluatieonderzoeken afgerond in de kalenderjaren 2004 t/m 2007 uit de BZ Evaluatieprogrammering. Het onderzoek beperkt zich verder tot evaluaties die (mede) onder verantwoordelijkheid van het Ministerie van Buitenlandse Zaken zijn uitgevoerd. Dit betekent dat circa 105 afgeronde en digitaal beschikbare evaluaties onderwerp van dit onderzoek zullen zijn.

3. Probleemstelling en onderzoeksvragen

De centrale vraag voor het onderzoek is of evaluatieonderzoeken daadwerkelijk worden benut voor (a) verbetering van beleid en/of beleidsuitvoering (de leerfunctie) en (b) voor informatieverstrekking over effectiviteit en doelmatigheid van het gevoerde beleid (de verantwoordingsfunctie, kwaliteit van de beleidsinformatie).

Om deze centrale onderzoeksvraag te kunnen beantwoorden dienen in het onderzoek de volgende deelvragen aan de orde te komen:

1. Welke typen van evaluatieonderzoek zijn in de periode 2004-2007 in het kader van de BZ Evaluatieprogrammering uitgevoerd (tijdstip van de evaluatie, aard en omvang van het onderwerp van evaluatie, etc.)?;
2. Wat was de aanleiding voor de evaluatie?;
3. Hoe is de evaluatie uitgevoerd?;
4. Wie waren de ontvangers van de onderzoeksresultaten?;
5. Wanneer en op welke wijze waren de belanghebbenden (beleidsmakers) betrokken bij de uitvoering van de evaluatie?;
6. Op welke wijze zijn de evaluatieresultaten gepresenteerd en gecommuniceerd?

Deze zes vragen hebben betrekking op alle evaluatieonderzoeken die onderwerp van dit onderzoek zijn.

De volgende vragen onderzoeken het gebruik van evaluatieresultaten en hebben betrekking op de steekproef ofwel een representatieve deelverzameling van circa 30 evaluaties:

7. Wat hebben de ontvangers gedaan met de evaluatieresultaten met het oog op leren en verantwoorden? En waarom?;
8. Is het gebruiksdoel helder gespecificeerd en zijn de evaluatie-uitkomsten tijdig beschikbaar?;
9. Sluiten in de te onderzoeken evaluaties onderzoeksvragen, bevindingen, conclusies en aanbevelingen helder en volledig op elkaar aan?;
10. Zijn de gepresenteerde aanbevelingen praktisch uitvoerbaar en liggen deze binnen het bereik van de ontvangers?;

Voor de vragen 8 t/m 10 gaat het zowel om de perceptie van de ontvangers als het oordeel van de onderzoekers.

4. Opzet van het onderzoek

Het onderzoek valt uiteen in twee fasen. De eerste voorbereidende fase omvat de volgende componenten:

- een beschrijving van de beleidscyclus en de evaluatiepraktijk bij het Ministerie van Buitenlandse Zaken;
- een inventarisatie van de uitgevoerde evaluaties op basis van de vragen 1 t/m 6.

Aan de hand van de inventarisatie zal een selectie van circa 30 evaluaties voor het onderzoek worden gemaakt, op basis van de volgende criteria:

- verdeling over DGIS/DGES/DGRC/DGPZ gewogen in het licht van het totaal aantal evaluatieonderzoeken per DG in de BZ-evaluatie-programmering;
- verdeling over typen evaluatieonderzoek;
- verdeling over IOB-onderzoeken en niet-IOB-onderzoeken gewogen in het licht van het aantal evaluatieonderzoeken per DG in de BZ-evaluatieprogrammering waarvoor IOB wel/geen verantwoordelijkheid draagt.

De tweede fase richt zich op onderzoek naar het feitelijk gebruik van evaluaties op basis van:

- interviews met belanghebbende beleidsmedewerkers en departementsleiding, en bij de evaluaties betrokken evaluatoren;
- dossieronderzoek (van relevante beleidsdossiers, kamerstukken, etc.);
- analyse planning- en beleidsdocumenten voor bewijsvoering gebruik evaluatie-uitkomsten;
- analyse presentatievorm en -stijl, en feedback mechanismen onderzoeksresultaten;

Deze fase van het onderzoek beoogt antwoord te geven op de vraag of er wel/geen/ten dele gebruik is gemaakt van de evaluatieresultaten en de factoren die hiervoor bepalend zijn.

Betrokkenheid van en communicatie over en weer met potentiële gebruikers en belanghebbende DG's gedurende het onderzoek verdienen bijzondere aandacht.

5. Organisatie van het onderzoek

Als opdrachtgever voor dit onderzoek treedt op het AC van het Ministerie van Buitenlandse Zaken. Voor de uitvoering zal op basis van onderhandse aanbesteding een externe consultant worden gecontracteerd. De consultant zal worden gevraagd binnen drie weken een plan van aanpak voor het onderzoek ter goedkeuring voor te leggen aan de opdrachtgever. Dit plan omvat onder meer een voorstel voor de selectie van evaluaties, een onderzoeksplan en een lijst van te houden interviews.

Rol belanghebbenden

De direct belanghebbenden bij het onderzoek zijn het AC, de DG's en de voor de kwaliteit van de beleidspraktijk en beleidsverantwoording verantwoordelijke directies evenals de IOB. Voor het AC zal de uitvoering van deze opdracht begeleid en ondersteund worden door een referentiegroep bestaande uit medewerkers van DEC, HDPO/OO, IOB en FEZ (trekker). Deze groep vervult ook de rol van klankbord voor het onderzoeksteam en kritische toetsing. Specifieke taken van de referentiegroep zijn het voeren van een startgesprek met de consultant, bespreking van het plan van aanpak en het becommentariëren van het concept eindrapport en het verlenen van decharge aan de consultant. De uiteindelijke verantwoordelijkheid voor de inhoud van het rapport ligt bij de consultant.

Evaluatieteam

Het onderzoek zal worden uitgevoerd door een team van een tweetal ervaren evaluatoren en een junior onderzoeker. Leden van het evaluatieteam dienen te beschikken over kennis van bestuur en overheidsmanagement, de evaluatiepraktijk en ervaring op het terrein van onderzoek. Bij voorkeur beschikken zij over ervaring met onderzoek naar het gebruik van evaluatieonderzoek en kennis op het terrein van de formulering en uitvoering van buitenlands beleid.

| 91 |

Rapportage

De consultant zal de opzet voor de rapportage en de voorlopige bevindingen met de referentiegroep bespreken. Het rapport zal in het Nederlands worden opgesteld. De hoofdtekst (exclusief bijlagen) zal in totaal niet meer dan 50 pagina's zijn.

Budget en planning

Het onderzoek zal binnen een periode van 4 maanden worden uitgevoerd, vanaf het moment van de start van de werkzaamheden van het evaluatieteam, tot aan het opleveren van het concept eindrapport. Uitvoering van het onderzoek zal onderhands worden aanbesteed. Voor dit onderzoek wordt een budget van max. Euro 130.000,- ter beschikking gesteld. De kosten komen ten laste van het FEZ bedrijfsvoeringbudget.

Follow-up

Het AC zal op grond van het rapport richting geven aan de verbetering van het gebruik van evaluatie-uitkomsten en concrete stappen ondernemen die er toe leiden dat evaluatieonderzoek wordt benut om te komen tot een kwalitatief betere beleidspraktijk én betere beleidsverantwoording.

Literatuur

Adviesraad voor Wetenschaps- en Technologiebeleid. 2005. *Kennis voor beleid – beleid voor kennis*. Rijswijk: Quantes.

Algemene Rekenkamer. 2003b. *Tussen beleid en uitvoering: lessen uit recent onderzoek*. Tweede Kamer der Staten-Generaal. Vergaderjaar 2002-2003, 28 831, nr. 2. 's-Gravenhage: Sdu Uitgevers.

Bamberger, M., Rugh, J and Mabry, L. 2006. *RealWorld Evaluation: Working under Budget, Time, Data and Political Constraints*. Thousand Oaks, CA Sage Publications.

Commission of the European Communities. 2002. *Communication from the Commission on Impact Assessment*. Brussels: COM(2002) 276 Final.

Covey, Stephen R. 1999. *The 7 Habits of Highly Effective People*. New York: Simon & Schuster.

David A. Kolb. 1984. Kolb's Theory of Adult Learning.
<http://www.learningandteaching.info/learning/experience.htm>

Government Accountability Office. 1992. *The Evaluation Synthesis*. GAO/PEMD-10.1.2. Washington D.C.

Hatry, H.P. 2006. *Performance Measurement: Getting Results*. Second edition. Washington D.C: The Urban Institute.

Independent Evaluation Group. 2004. *Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs*. World Bank.

Independent Evaluation Group. 2005. *Influential Evaluations: Detailed Case Studies*. World Bank.

Independent Evaluation Group. 2007. *Sourcebook for Evaluating Global and Regional Partnership Programs. Indicative Principles and Standards*. World Bank.

IOB Evaluations, No 298. 2004. *Quality Assessment of Evaluations completed in 2002*. Ministry of Foreign Affairs, The Netherlands.

IOB Werkdocument (Rita Tesselaar en Roland Rodts). 2005. *Onderzoek naar de kwaliteit van in 2003 afgerond decentrale evaluaties*. Ministerie van Buitenlandse Zaken, Den Haag.

Kusek, Jody Zall and Rist Ray C. 2004. *A Handbook for Development Practitioners: Ten Steps to a Results-Based Monitoring and Evaluation System*, Washington: the World Bank.

Ministerie van Financiën. 2004a. *Eindrapport VBTB-evaluatie – Lessen uit de praktijk*. Den Haag, Interdepartementaal Overlegorgaan Financieel Economische Zaken: 57.

| 93 |

Organisation for Economic Cooperation and Development. 2001. *Evaluation Feedback for Effective Learning*. Paris. OECD/DAC.

Patton, Michael Quinn. 1996. *Utilization-Focused Evaluation*, Third Edition, SAGE Publications.

Patton, Michael Quinn. 2002. *Utilization-Focused Evaluation (U-FE) Checklist*. <http://www.wmichi.edu/evalctr/checklists>.

Rand Europe. 2005. *Working Paper: richting een beleidsevaluatie cultuur in Nederland*.

Riet, O.A.W.T. Van de. 2003. *Policy Analysis in Multi-Actor Policy Settings – Navigating Between Negotiated Nonsense & Superfluous Knowledge*. Eburon Publishers, Delft.

Rist, R.C. 1990. *Program evaluation and the management of government. Patterns and prospects across eight nations*. New Brunswick and London, Transaction Publishers.

RPE richtlijnen. Uit: Staatscourant 28 april 2006, nr. 83 / pag. 14.

Swanborn, P.G. 1999. *Evalueren*. Amsterdam, Boom: 364.

Vaughan, R.J. & Buss, T.F. 1998. *Communicating Social Science Research to Policymakers*. Thousand Oaks, CA, Sage Publications.

Weiss, C.H. 1977. "Introduction" pp. 1-22. *Using Social Research in Public Policymaking*. Edited by Carol Weiss. Lexington MA. D.C. Health.

Wholey, J.E., Scanlon, J.W., Duffy, H.G., Fukumotu, J.S. and Vogt, L.M. 1970. *Federal Evaluation Policy: Analyzing the Effects of Public Programs*. Washington D.C. The Urban Institute.

Bijlage 2

Evaluaties in de steekproef

Directoraat-generaal Europese Samenwerking

BZ. 2004. *Tampere-agenda.*

BZ. 2004. *Terugblik op het Nederlands voorzitterschap van de Raad van Europa 6 november 2003 tot en met 13 mei 2004*

IOB. 2005. *Een uitgebreid Europabeleid, Evaluatie van het Nederlands beleid inzake de toetreding van Midden-Europese landen tot de Europese Unie 1997-2003.*

BZ. 2005. *Verder denken over Thinking Forward - Evaluatie van het culturele programma in het kader van het Nederlandse EU-voorzitterschap.*

IOB. 2007. *Coördinatie, complementariteit en coherentie (de drie C's) van het verdrag van Maastricht.*

IOB. 2008. *Nederlandse EU-voorzitterschap.*

| 97 |

Directoraat-generaal Internationale Samenwerking

BZ. 2005. *Joint evaluation of effectiveness and impact of the Enabling Development Policy of the World Food Programme (WFP).*

CERES. 2005. *Rural Development Project Performance: A Review of 46 Evaluation Studies of DGIS-funded Projects in the Themes Agricultural Research, Natural Resource Management, Water Management and Area Development and the Implications for the Sector Wide Approach.*

IOB. 2005. *Aid for Trade? An evaluation of Trade-Related Technical Assistance.*

IOB. 2006. *Sectorale benadering.*

BZ. 2006. *Evaluation of the Theme-Based Co-Financing Programme (TME).*

BZ. 2006. *Evaluatie Nederlands Investerings Matching Fonds (NIMF) 2002-2006.*

IOB. 2006. *ORET/MILIEV in China.*

BZ. 2007. *Evaluation of the international education programmes NPT and NFP managed by Nuffic.*

IOB. 2007. *Evaluatie van de vernieuwing van het Nederlandse onderzoeksbeleid 1992-2005.*

IOB. 2007. *Impactevaluatie water en sanitaire voorzieningen in Tanzania.*

BZ. 2006. *ORET/MILIEV programmabreed 1999 - 2004.*

BZ. 2007. *Evaluatie Inventarisatie Energie & Ontwikkelingssamenwerking 1996 - 2006*

IOB. 2007. *Clean Development Mechanism.*

IOB. 2008. *Beleidsdoorlichting seksuele en reproductieve gezondheid en rechten en hiv/aids*

IOB. 2008. *Het Nederlandse Afrikabeleid 1998-2006 / Evaluatie van de bilaterale samenwerking*

| 98 |

IOB. 2008. *VakbondsMedefinancieringsProgramma (VMP)*

Directoraat-generaal Politieke Zaken

IOB. 2004. *Gemeentelijke internationale samenwerking en pre-accessie beleid.*

BZ. 2004. *Vredesfonds 2001-2003.*

IOB. 2006. *Mensenrechtenbeleid.*

BZ. 2007. *Evaluatie Stabiliteitsfonds 2004 en 2005.*

Directoraat-generaal Regiobeleid en Consulaire Zaken

BZ. 2006. *Algemene Rekenkamer – Rapport van de Gedetineerdenbegeleiding 2005.*

BZ. 2006. *Matra Politieke Partijen Programma 2000-2005.*

BZ. 2006. *Transparant en toetsbaar, een advies over landeninformatie in het vreemdelingenbeleid.*

Alle evaluaties zijn voorhanden via de website van het ministerie van Buitenlandse Zaken: www.minbuza.nl

Bijlage 3

Instrumenten en faciliteiten voor kennisuitwisseling

Belangrijk voor het leren door organisaties is dat meerdere actoren met elkaar van gedachten wisselen over wat zij individueel leren met het doel het geleerde vast te houden en in (handelings)repertoires op te slaan. Anderen nemen hier vervolgens kennis van en gebruiken het geleerde bij het analyseren van bestaand en het maken van nieuw beleid. Bij dat opslaan van leerervaringen zijn verschillende vormen van *storage* belangrijk. Deze bijlage geeft een kort overzicht van mogelijkheden om deze voorzieningen en instrumenten in te zetten om leren binnen een organisatie te institutionaliseren.

Een eerste manier is om kennisoverdracht en kennis-interactie tot stand te brengen via workshops, seminars, hei-sessies en soortgelijke arrangementen. Dit is een relatief vaak gebruikte aanpak en er zijn zeker voorbeelden van succesvol gezamenlijk leren. Deze manier kent echter een schaduwzijde en die heeft deels te maken met groepsdynamische processen. Groepsessies kunnen *groupthink* veroorzaken (gezamenlijk geloven in één bepaalde richting, zeker als tegendenken niet gestimuleerd wordt), *cascade-vorming* (een mildere vorm van *groupthink*, waarbij er zich in een discussie slechts enkele hoofdlijnen aftekenen waar steeds meer mensen zich mee verbinden, waardoor de overblijvenden - die niet noodzakelijkerwijze onzin vertellen - steeds meer gemarginaliseerd worden) en '*de-karavaan-trekt-verder-gedrag*'. Omdat mensen bedreven zijn in praatgedrag, dreigt het risico dat na een seminar van een dag de indruk gevestigd is dat het morgen anders gaat. Als de morgen daar is, en de overmorgen ook, blijkt van het leren geen sprake te zijn. Immers, morgen regeert niet het praatgedrag zoals tijdens het seminar plaatsvond, maar of ander praatgedrag dan wel '*doegedrag*' (Kogel 2008).

| 101 |

Een tweede manier is te werken via kennischatkamers (*repositories*) om langs die weg leren op individueel, maar vooral op organisatieniveau, te bevorderen. Dit zijn digitaal toegankelijke, goed gestructureerde *portals* waarin relevante evaluatieonderzoeken en hun resultaten zijn samengevat. Het gaat doorgaans niet om afzonderlijke (project) evaluaties maar om *systematic reviews*, meta-analyses, meta-evaluaties en syntheses. Onderstaande box (Leeuw en Vaessen 2009), licht het een en ander toe:

- **Systematic reviews** are syntheses of primary studies that, from an initial explicit statement of objectives, follow a transparent, systematic, and replicable methodology of literature search, inclusion and exclusion of studies according to clear criteria, and extracting and synthesizing information from the resulting body of knowledge.

- **Meta-analyses**, a common type of systematic review, quantitatively synthesize “scores” for the impact of a similar set of interventions from a number of individual studies across different environments. They follow a strict procedure to search for and select appropriate evidence, typically using a hierarchy of methods, with more quantitatively rigorous (experimental) studies being ranked higher as sources of evidence.
- **Narrative reviews** are descriptive accounts of intervention processes and/or results covering a series of interventions, relying on a common analytical framework and template to extract data from the individual studies and summarizing the main findings in a narrative account and/or tables and matrices representing key aspects of the interventions.
- **Realist syntheses** are theory based and do not use a hierarchy of methods. They collect earlier research findings by placing the policy instrument or intervention that is evaluated in the context of other similar instruments and describe the intervention in terms of its context, social and behavioral mechanisms (what makes the intervention work), and outcomes (the deliverables).

| 102 |

In de literatuur over gebruik van en leren door evaluatieonderzoek wordt voor wat betreft kennischatkamers ook verwezen naar het concept *second-order knowledge-producing organizations*. Vooral omdat de instellingen, die verantwoordelijk zijn voor *portals*, meestal niet zelf evaluaties uitvoeren maar evaluaties *reviewen* en, na scheiding van het kaf van het koren, de bevindingen bundelen en toegankelijk maken (zie ook Hanne Foss Hansen and Olaf Rieper). Soms gaat het om de bundeling van experimentele en quasi-experimentele evaluaties (Cochrane Collaboration & Campbell Collaboration), soms over de bundeling van evaluaties die meer dan alleen dit type onderzoek toelaten (evidencenetwerk.org). Een ander relevant onderscheid is dat de kennis*portals* geordend worden via de aard van de evaluaties van interventies en beleidsmaatregelen (zoals bij de Campbell Collaboration), maar soms ook via de gedragsmechanismen, die aan beleidsmaatregelen, de beleidslogica en de interventietheorie ten grondslag liggen (vgl. *realist synthesis* (Pawson and Tilley 1997); *browsing for evidence* (Kleemans et al. 2007)).

Om van deze schatkamers te kunnen leren, moeten zij die willen of moeten leren kunnen beschikken over adequate digitale zoekmachines. Achteraan deze bijlage vindt u een overzicht van een tiental van dergelijke voorzieningen (in *figure 1*). In het rapport *When will we ever learn* (Center for Global Development, 2005) is een aantal andere te vinden, die direct relevant zijn

voor evaluaties op het gebied van ontwikkelingsamenwerking. Datzelfde geldt in toenemende mate voor 3ie (<http://www.3ieimpact.org/>).

Twee voorbeelden: Cochrane Collaboration (<http://www.cochrane.org/>) en de Campbell Collaboration (<http://www.campbellcollaboration.org>). De oudste is de Cochrane Collaboration. Deze op evaluaties van medische interventies en therapieën gerichte organisatie *published more than 4,000 reviews and protocols (which are plans for ongoing work) since its establishment in 1993. The idea had its offspring in the 1970s in England where Archie Cochrane spoke for the point that results of all relevant studies concerning given interventions ought to be summarized in reviews* (Rieper & Foss Hansen, in druk: 33).

Van een meer direct nut voor de sociaal-wetenschappelijke wereld is de Campbell Collaboration *which was modeled after Cochrane's Collaboration, [and] was established in the late 1990s. Since then it has published a little less than 100 reviews and protocols. By publishing a publication with the title "Randomized Experiments for Planning and Evaluation" Robert F. Boruch in 1997 had followed in the footsteps of Donald Campbell who in the 1960s argued in favor of the importance of designing new social interventions as experiments in order to be able to evaluate the effects.*

| 103 |

Een derde manier van kennisoverdracht, benutting en leren is via databases. Voorbeelden zijn *Statline* (<http://statline.cbs.nl/statweb/>) van het CBS, waarbij bezoekers hun eigen tabellen kunnen maken, *Repris* (<http://www.wodc.nl/onderzoek/cijfers-en-prognoses/Recidive-monitor/repris/>) van het WODC, waarbij, maar alleen met het oog op incidentie van recidive (strafrechtelijk) het zelfde kan en *Statstat* van Amsterdam (<http://stadstat.osamsterdam.nl/>). Een vierde en hier als laatste te noemen manier is te werken met wiki's en blogs, waar bij het vrijwillig poolen en vrijwillig corrigeren van gepoolde informatie centraal staat. Sunstein's (2006) studie naar de kracht van wat hij noemt *infotopia* laat succesvolle voorbeelden van leren-van-elkaar zien, waarbij ook meer traditionele evaluaties een zekere rol spelen. Het is denkbaar dat deelnemers aan wiki's van in institutioneel verband gepubliceerde evaluaties, dan wel van hun data gebruik maken, maar evenzeer bestaat de mogelijkheid dat officiële evaluaties vanuit wiki's en blogs gekritiseerd worden door deelnemers en contributors.

Figure 1: Outline of guidelines and handbooks from evidence-producing organizations with reference to the basis for the assessments of primary studies.

Organization	Title of guidelines Date of publishing Number of pages	Based on a hierarchy of evidence	Criteria for assessment of the specific design	Comments
The Cochrane Collaboration	Cochrane Handbook for Systematic Reviews of Interventions ver. 4.2.6 August 2006, 257 pp	Yes, RCT has highest priority. Other types of design are reconsidered to be supportive.	Reference to literature on assessing the quality of RCT	About assessment of RCTs: Reference to Moher et al. 1995.
The Nordic Cochrane Centre	-	-	-	Refer to the Cochrane Collaboration
Centre for Reviews and Dissemination (CRD)	Undertaking systematic reviews of research on effectiveness, March 2001, 164 pp	Yes, RCT has highest priority. Other types of design are considered to be supportive.	Criteria for assessing effectivity studies, accuracy studies, qualitative studies and economic evaluations	Discussion on external validity and assessment scales – both numeric and non-numeric scales.
The Campbell Collaboration (C2)	Research Design Policy Brief, November 2004, 25 pp	Discussion on the hierarchy of evidence. Yet, RCT has highest priority. Other types of design are considered to be supportive.	Discussion on assessment of quality of individual design	
	Guidelines for the preparation of review protocols ver. 1.0, January 2001, 9 pp	RCT has highest rank order. Contributions of qualitative studies are made explicit	No	

Organization	Title of guidelines Date of publishing Number of pages	Based on a hierarchy of evidence	Criteria for assessment of the specific design	Comments
	Steps in proposing, preparing, submitting and editing of Camp-bell Collaboration Systematic Reviews, 4 pp	No	No	Flowchart for development of new systematic reviews.
Nordisk Campbell Center (NC2)	How to make a Campbell Collaboration Review: Guide 3 how to make a Review, 135 pp	Yes, RCT has highest priority. Other types of design are considered to be supportive.	-	Refer to C2.
The Evidence for Policy and Practice Information and Coordinating Centre (EPPIC)	EPPIC-Centre methods for conducting systematic reviews ver. 1.0 September 2006, 18 pp	No	Yes	A general guideline which allows multiple methodological approaches.
	Book: Using research for effective health promotion. August 2001, pp 176	-	-	Different contributions of knowledge from research and knowledge from practice are discussed.
	EPPIC-Centre review companion	-	-	Not accessible on the www.

Organization	Title of guidelines Date of publishing Number of pages	Based on a hierarchy of evidence	Criteria for assessment of the specific design	Comments
	Guidelines for the REPORTING of primary empirical research studies in education (The REPOSE Guidelines)8, 2004, 12 pp	No	Yes	How to report empirical research within education.
	EPPI-Centre Keywording strategy for classifying educational research ver. 0.9.79 18 pp	No	No	Guidelines for keywords for classifying research studies.
Social Care Institute for Excellence (SCIE)	The conduct of systematic research reviews for SCIE knowledge reviews. December 2006 85 pp	o	-	Multiple methodological approaches. Criteria of quality appraisal of included studies. Reporting of study quality in the review.
Health Evidence Network (HEN)	-	-	-	HEN has no guidelines.
National Institute for Health and Clinical Excellence (NICE)	The guidelines manual 200610 2006 190 pp	EBM-hierarchy	Assessment form for each study design	Internal and external validity and forms of bias.

Organization	Title of guidelines Date of publishing Number of pages	Based on a hierarchy of evidence	Criteria for assessment of the specific design	Comments
IMS – Institute for Evidence-Based Social Work Practice, Sweden (IMS)	–	–	–	IMS has a guideline in progress.
What Works Clear- inghouse (WWC)	Evidence standards for reviewing studies ¹¹ September 2006 12 pp	Yes	Yes	

Bijlage 4

Geraadpleegde personen

Dhr. Joost Andriessen
Dhr. Peter Arnoldus
Mw. Laetitia van Asch

Dhr. Cees Bansema
Dhr. Philip Bastiaenen
Dhr. Wiebe de Boer
Dhr. Jaime de Bourbon de Parme
Dhr. Maarten Brouwer
Mw. Reina Buijs

Mw. Yolanda van Cooten

Mw. Ineke Duijvestein

Dhr. Reneko Elema
Mw. Wilma van Esch

Dhr. Jan Henneman
Mw. Betty van Hooidonk

Dhr. Wepke Kingma
Dhr. Pieter Jan Kleiweg de Zwaan
Dhr. Jan Klugkist
Mw. Margriet Leemhuis
Mw. Sandra Louiszoon

Dhr. Bernard Meijerman
Mw. Sophie Middelhoff
Dhr. Henk Molenaar
Dhr. Jan-Peter Mout

Dhr. Bram van Ojik

Dhr. Han Peters
Mw. Marijke Roes
Mw. Marianne Sinke
Dhr. Arjan Schuthof

Dhr. Sjeng Smeets
Dhr. Fred Smiet
Mw. Natascha Soons
Dhr. Herman Specker

Mw. Birgitta Tazelaar
Mw. Rita Tesselaar

Dhr. Arjen Uijterlinde

Dhr. Roel van der Veen
Dhr. Hidde van der Veer
Dhr. Rob Visser

Dhr. Frits van der Wal
Dhr. Ewout de Wit
Dhr. Ronald Wormgoor

Dhr. Jaap van der Zeeuw
Dhr. Ton van der Zon

Bijlage 5

Geraadpleegde bronnen

Publicaties

Algemene Rekenkamer. 2007. *Leren van parlementaire onderzoeken*. <http://www.rekenkamer.nl/Actueel/Onderzoeksrapporten/Introducties/2007/10/Lerenvanparlementaironderzoek>.

Argyris, Chris and Donald Schön. 1974. *Theory in Practice: Increasing professional effectiveness*, San Francisco: Jossey - Bass.

Argyris, Chris and Donald Schön. 1978. *Organizational learning: A theory of action perspective*. Reading, Mass: Addison Wesley.

Argyris, Chris and Donald Schön. 1996. *Organizational learning II: Theory, method and practice*. Reading, Mass: Addison Wesley.

Brandt, T. 2007. Sozialer Kontext der Evaluation. In: *Handbuch zur Evaluation: Eine praktische Handlungsanleitung*. Stockmann, R. (Hrsg.): 164 - 194. Münster: Waxmann Verlag.

Clements, Paul, Thomaz Chianca and Ryoh Sasaki. 2008. Reducing World Poverty by Improving Evaluation of Development Aid. In: *American Journal of Evaluation*, 29: 195-214.

Chelimsky, E. 1995. Where we stand today in the practice of evaluation. In: *Knowledge and Policy*, 8: 8 - 19.

Dahler-Larsen, Peter. 2000. The Administrative Use of Evaluations in Danish Municipalities: Surviving the Routinization of Evaluation. In: *Administration & Society*, 32: 70 - 92.

Eijk, L. van en J. P. Hartkamp. 1992. Van zwarte schapen en vreemde eenden: Een verhandeling over de positie van het WODC. In: *Justitiële verkenningen*, 18(3), 90 - 108.

Fabian, Carlo. 2009. Nutzung der Evaluationsfunktion im Vergleich. In: *Evaluation; ein systematisches Handbuch*, T. Widmer et al (Hrsg.): 501 - 506. Wiesbaden: VS Verlag fuer Sozialwissenschaften.

Farrington, D.P. en Welsh, B.C. 2005. Het belang van experimentele evaluaties in de criminologie. *Justitiële Verkenningen*, (8), 11 - 41.

Frey, B.S. and M. Osterloh. 2006. *Evaluations: hidden costs, questionable benefits, and superior alternatives*. Working Paper No. 302. University of Zurich: Institute for Empirical Research in Economics.

Furubo, Jan-Eric, Rist, Ray and Rolf Sandahl. ed. 2002. *The International Atlas of Evaluation*. London: Transaction.

Gaay Fortman, Bas de, en Arie de Ruijter. 2008. *Van Overkant Tot Overkant, Onderzoek naar aard en aanpak van de IOB-evaluatie van het Nederlands Afrikabeleid Bilateraal 1998-2006*. Den Haag: Ministerie van Buitenlandse Zaken.

Hellstern, G. M. 1985. Assessing evaluation research. In: *Guidance, control and evaluation in the public sector*, ed. F.X. Kaufmann, G. Majone and V. Ostrom. Berlin: de Gruyter.

Henry, Gary T. 2003. Influential Evaluations. In: *American Journal of Evaluation*, 24: 515 - 524.

Henry, Gary T. and Melvin M. Mark. 2003. Beyond use: Understanding evaluation's influence on attitudes and actions. In: *American Journal of Evaluation*, 24: 293 - 314.

Horton, D., Alexaki, A., Bennett-Lartey, S., Brice, K.N., Campilan, D., Carden, F., de Souza Silva, J., Duong, L.T., Khadar, I., Maestrey Boza, A., Kayes Muniruzzam, I., Perez, J., Somarriba Chang, M., Vernooy, R. and J. Watts. 2003. *Evaluating capacity development: Experiences from research and development organizations around the world*. The Hague: International Service for National Agricultural Research (ISNAR) and Technical Centre for Agricultural and Rural Cooperation (ACPEU) (CTA) and Ottawa, Canada: International Development Research Centre (IDRC).

Johnson, K., Greenesid, L.O., Toal, S.A., King, J.A., Lawrenz, F. and B. Volkov. 2009. Research on Evaluation Use. A Review of the Empirical Literature. From 1986 to 2005. In: *American Journal of Evaluation*, 30 (3): 377 - 410.

- Kleemans, E., C. Klein Haarhuis, F. L. Leeuw and M. van Ooyen-Houben. 2007. Law enforcement interventions in the Netherlands: mapping interventions and 'browsing' for evidence. In: *Evidence & Policy*. 4: 487 - 504.
- Kogel, C.H. 2008. *Hersenen in beeld*. Amsterdam: Boom.
- Leeuw, Frans L. 1998. Doelmatigheidsonderzoek van de Rekenkamer als regelgeleide organisatiekunde met een rechtssociologisch tintje? In: *Recht der Werkelijkheid*, 14: 35 - 71.
- Leeuw, Frans L. 2009. Evaluation policy in the Netherlands. In W.M.K. Trochim, M. M. Mark, & L. J. Cooksy, ed. *Evaluation policy and evaluation practice. New Directions for Evaluation*, 123, 87 - 102.
- Leeuw Frans L. and Leslie Cooksy. 2004. Evaluating the performance of development agencies: the role of meta-evaluations. In: *Evaluating development effectiveness*, ed. George Keith Pitman, Osvaldo N. Feinstein and Gregory K. Ingram: 95 - 108. New Brunswick: Transaction Publishers.
- Leeuw, Frans L. en M. Crijns. 2005. Normnaleiving door organisaties. Sociologie toegepast op een actueel vraagstuk. In: *Bestuurskunde*, 14: 40 - 49.
- Leeuw, Frans L. and Jan Eric Furubo. 2008. Evaluation systems: what are they and why study them? In: *Evaluation*, 14 (2): 157 - 169.
- Leeuw, Frans L., Rist, Ray C. and Richard C. Sonnichsen. ed. 1994. *Can governments learn: comparative perspectives on evaluation and organizational learning*. New Brunswick, NJ: Transaction.
- Leeuw Frans L. and Jos Vaessen. 2009. *Impact evaluations and development*. NONIE Guidance on Impact Evaluation. Washington DC: NONIE (The Network of Networks on Impact Evaluation).
- Leviton, L. and E. Hughes. 1981. Research on the utilizations of evaluations: A review and synthesis. *Evaluation*, 5, No. 4, 525 - 548.

March, J. G. and Johan P. Olsen. 1995. The Uncertainty of the past: Organizational Learning under Ambiguity. In: *European Journal of Political Research*. 3: 147 - 171.

Mises, L. von. 1976. *A Critique of Interventionism: Inquiries into Present Day Economic Policy and Ideology*. Jena: Gustav Fischer Verlag.

Nutley, S., Walter, I. and H. T.O. Davies. 2003. From knowing to doing: a framework for understanding the evidence-into-practice agenda. In: *Evaluation*, 9 (2): 125 - 149.

Pawson, R. and Tilley, N. 1997. *Realistic Evaluation*. London: Sage Publications.

Pollitt, C. 1999. Stunted by stakeholders? Limits to collaborative evaluation. In: *Public Policy and Administration*. 14 (2): 77 - 90.

| 114 |

RPE, Rijksbegrotingsvoorschriften. 2006. http://rbv.minfin.nl/rbv1/departementale-begroting/begroting_begroting.

Schön, Donald. 1983. *The Reflective Practitioner. How professionals think in action*. London: Temple Smith.

Shulha, L.M. and J. Bradley Cousins. 1997. Evaluation Use: Theory, Research, and Practice Since 1986. In: *American Journal of Evaluation*, 18(1): 195 - 208.

Stamm, Margrit. 2003. *Evaluation und Ihre Folgen für die Bildung – eine unterschätzte pädagogische herausforderung*. Münster: Waxmann.

Sunstein, Cass. 2003. *Why Societies Need Dissent*. Cambridge: Harvard University Press.

Sunstein, Cass. 2006. *Infotopia: how many minds produce knowledge*. New York: Oxford University Press.

Swanborn, Peter G. 1999. *Evalueren, het ontwerpen, begeleiden en evalueren van interventies: een methodische basis voor evaluatie-onderzoek*. Amsterdam: Boom.

Walters, R. 2008. Government manipulation of criminological knowledge and policies of deceit. In: *Critical thinking about the uses of research*, ed. T. Hope and R. Walters: 7 - 26. Centre for Crime and Justice studies.

Weiss, C., Murphy-Graham, E., Petrosino, A. And A.G. Gandhi. 2008. The Fairy Godmother—and her warts: making the dream of evidence-based policy come true. In: *American Journal of Evaluation*. vol. 29, no. 1: 29 - 47.

World Bank. 2004. *Influential evaluations. Evaluations that Improved Performance and Impacts of Development Programs*. Washington: the World Bank.

Deze publicatie is een uitgave van:
Ministerie van Buitenlandse Zaken
Directie Financieel Economische Zaken
Postbus 20061, 2500 EB Den Haag
www.minbuza.nl

Auteurs:

Ger H.C. van Gils

Frans L. Leeuw

Marie Gildemyn

Liesbeth Inberg

Redactie: David Korthals Altes en Margret Verwijk

Print: vijfkeerblauw.nl Rijswijk

© Buitenlandse Zaken | juli 2010