

# Samenvatting

## Inleiding en vraagstelling

In de borging van wetgevingskwaliteit spelen evaluaties een belangrijke rol. Willen deze daadwerkelijk een bijdrage leveren aan de kwaliteit van wetgeving, dan is het zaak om: (1) te zorgen dat ze worden verricht; (2) ze zó op te zetten dat ze bruikbare, valide en betrouwbare inzichten genereren; en (3) dat ze worden gebruikt door beleidsmakers en wetgevingsjuristen in de terugkoppeling naar hun producten.

Dit is in de praktijk echter niet vanzelfsprekend. Wetten en wetswijzigingen zijn doorgaans complexe evaluatieobjecten: een mix van regels, normenkaders en beleidsinterventies. Behalve empirische en beleidsevaluatie-vaardigheden vereisen wetsevaluaties dus ook een juridische blik. Veel individuele aanvragers zeggen niet goed te weten wat te doen als zij een wetsevaluatie (proces- of ex post) moeten gaan opleveren. Deze factoren kunnen eraan hebben bijgedragen dat wetsevaluaties in de praktijk heel verschillend worden aangepakt: zowel qua inrichting van het aansturingproces als methodologisch-inhoudelijk. Een centraal wetsevaluatiebeleid is er niet. Dit laatste hoeft niet problematisch te zijn, ware het niet dat eerder in meta-onderzoek ook aanwijzingen zijn gevonden voor verschillen in de validiteit en betrouwbaarheid van wetsevaluaties.

De gevarieerde praktijk van wetsevaluatie houdt mogelijk verband met verschillen in *evaluatievermogen*: het vermogen om wetsevaluaties te verrichten en ze bovendien te gebruiken. In hoeverre zijn binnen beleidsdepartementen de voorwaarden aanwezig om het doen en gebruiken van evaluaties mogelijk te maken? Die voorwaarden kunnen zowel het evaluatieproces als de inhoud van evaluaties betreffen.

Doel van deze studie is om een beeld te krijgen van bestaande departementale praktijken rond het verrichten en gebruiken van evaluaties – met name wetsevaluaties. Aan de basis ligt een streven bij de aanvrager van dit onderzoek, de directie Wetgevingskwaliteitsbeleid (WKB) van het Kenniscentrum Wetgeving en Juridische Zaken (KCWJ), naar meer vergelijkbaarheid tussen wetsevaluaties. Dit om te zorgen dat ervaringen *met* en kennis *uit* wetsevaluaties worden geaccumuleerd om bredere lessen te kunnen trekken.

## Onderzoeksvragen en aanpak

- 1 Welke inzichten bieden de internationale (a) literatuur en (b) normen en handreikingen met betrekking tot *evaluatievermogen*? (hoofdstuk 1 en 2)
- 2 Welke praktijkervaringen met evaluatievermogen zijn opgedaan door beleidsmedewerkers en wetgevingsjuristen, betrokken bij (wets)evaluaties? (hoofdstuk 3)
- 3 Welke inhoudelijke en methodologische eigenschappen van *ex post* evaluaties en -wetsevaluaties acht men in de praktijk van belang? (hoofdstuk 3)
- 4 Welke lessen bieden de antwoorden op voorgaande onderzoeksvragen voor verdere gedachtevorming over de ontwikkeling van evaluatievermogen – vooral met betrekking tot wetgeving? (hoofdstuk 4)

Ruggengraat van dit onderzoek vormden de volgende aspecten van evaluatievermogen en hun kenmerken – gebaseerd op literatuur over evaluatievermogen:

- structuur: doelstellingen en organisatieonderdelen voor evaluatie;
- programmering en evaluatieverplichtingen;

- middelen voor evaluatie: financiële en andere;
- evaluatiekader: evaluatieprogramma's en richtlijnen bij de aanpak (proces en inhoud);
- borging van evaluatiekennis en -ervaring;
- bevordering van gebruik.

In deze structuur, waarmee een deel van onderzoeksvraag 1 is beantwoord, zijn eerst een aantal aanpakken en ervaringen van internationale organisaties en grotere landen op een rij gezet. Vervolgens zijn, aan de hand van *desk study* en een twintigtal gesprekken met 35 sleutelinformanten, praktijkgegevens verzameld bij negen Nederlandse beleidsdepartementen.

Evaluaties en wetsevaluaties vinden niet in een vacuüm plaats; ze worden verricht vanuit verschillende motieven. Soms zijn ze verplicht uit oogpunt van verantwoording richting de Tweede Kamer (bijv. zijn begrotingsgelden goed besteed?) of een wettelijke evaluatiebepaling, een andere keer komen ze van binnenuit, bijvoorbeeld om te leren met het oog op verdere beleidsontwikkeling. Ook kunnen evaluaties strategisch worden ingezet. In de literatuur worden twee ideaaltypische evaluatiemotieven onderscheiden: (a) verantwoorden en (b) op 'leren' gerichte motieven zoals kennisvergroting of *enlightenment* en eventuele beleidsaanpassing of *em-betterment*. Bij het eerste motief passen eindevaluaties gericht op de vraag 'heeft het gewerkt', bij het tweede passen tussentijdse evaluaties, gericht op het begrijpen van de uitwerking en eventuele bijsturing. De politieke context bepaalt dus mede de motieven voor evalueren, die kunnen doorwerken in hoe evaluaties worden opgezet, aangestuurd en gebruikt. Uiteindelijk werkt ze ook door in evaluatievermogen.

### **Internationale ervaringen met evaluatievermogen (Onderzoeksvraag 1)**

Uit internationale handboeken van bijvoorbeeld de Verenigde Naties (VN), de Verenigde Staten (VS), het Verenigd Koninkrijk (VK), de Wereldbank en de Europese Commissie (EC) blijkt dat beleidsmedewerkers een steeds belangrijker rol krijgen toebedeeld in het managen van evaluaties. De kwaliteit en het gebruik van evaluatieonderzoek zijn niet alleen afhankelijk van hoe onderzoekers werken en hun resultaten presenteren, maar juist ook van de aanvragers, zo stelt men.

In de eerste plaats nemen die aanvragers bepalende beslissingen: over evaluatie-programmering, budget, reikwijdte en onderzoeksvragen. Ook hebben ze vaak een rol in de verdere aansturing of begeleiding van de evaluatie. Er circuleren uiteenlopende normen en leidraden ten aanzien van opdrachtformulering, aanbesteding, begeleiding, eindbeoordeling en het gebruik van evaluaties, waarvan enkele ter illustratie zijn opgenomen in bijlage 3 van dit rapport.

Niet alleen het evaluatietraject zelf, maar ook de inhoud en opzet van te evalueren beleid of wetgeving en de onderbouwing daarvan zijn medebepalend voor evaluatiekwaliteit. Evalueren wordt steeds vaker beschouwd als onderdeel van de beleids-cyclus, en minder als een activiteit die daar los van staat. Ontbreekt bijvoorbeeld de argumentatie voor beleidskeuzes (beleidstheorie), dan wordt het lastiger om een evaluatie te focussen. Dit zou volgens de EC het evalueren zelfs duurder kunnen maken. Ook in het monitoren van beleid vervult de (latere) evaluatie-aanvrager een cruciale rol. Meerdere instituties vinden het een gemiste kans voor hoogwaardige evaluaties als niet, of niet op tijd, wordt begonnen met nulmetingen en de verdere monitoring van belangrijke proces- en uitkomstindicatoren. Ook doelgroepen, zoals gemeenten, burgers en bedrijven kunnen in de monitoring een rol vervullen.

Ook gidsen voor het design van evaluaties, de inhoudelijke evaluatie-aanpak, zijn vaak niet alleen voor onderzoekers maar ook voor aanvragers bedoeld. Hierin zijn verschillende stromingen te onderscheiden. Illustratief is het onderscheid tussen de 'attributieve' (*black box*) en de 'contributieve' (*clear box*) stroming. In de eerste stroming staat het kunnen toeschrijven van resultaten aan beleid centraal: heeft het beleid gewerkt en in welke mate? Met de designs in deze stroming wordt gepoogd de effecten van het beleid of de wet in kwestie zo goed mogelijk te isoleren van andere, om tot een gerichte, netto-uitspraak over doeltreffendheid te komen. In de praktijk komt het dan vaak neer op experimentele designs. De aanpakken in de tweede stroming zijn gericht op het begrijpen van hoe en waarom – onder welke condities – veranderingen zijn opgetreden en resultaten zijn geboekt. De stromingen sluiten elkaar niet uit, ze kunnen wederzijds aanvullend worden gebruikt. Terugkerend uitgangspunt bij de keuze voor een benadering is *fit for purpose*: keuzes in het evaluatiedesign worden afhankelijk gesteld van de situatie, onder andere soort beleid, beleidsfase, evaluatieonderwerp, hoofdvragen en soort doelgroep. In dit licht ontwikkelde onder meer de Amerikaanse federale Rekenkamer handgrepen voor de evaluatie van complexe federale programma's. Net als Nederlandse wetten zijn ook deze meestal geen 'cleane' interventies waarvan de werking eenvoudig gemeten en bovendien geïsoleerd kan worden van de context, zoals externe invloeden en parallel beleid. Met wat voor soort evaluatiedesigns valt dan toch het maximale te zeggen?

Naast het evalueren als zodanig beschouwen grote instituties en landen ook het evaluatiegebruik steeds meer als onderdeel van het beleidsproces en niet slechts als resultante van een eindrapport. Zo voorziet het *Smart Regulation*-programma van de EC in terugkoppeling van ex post evaluatiebevindingen naar de regelgevingscyclus, via de verplichte ex ante evaluaties waarin die bevindingen een plek dienen te krijgen. Ook bij de VN wordt evaluatiegebruik voorgeschreven en gemonitord; in dat verband stimuleert men ook het aanleggen van metastudies en kennisbanken om het overzicht over evaluatiekennis te behouden. Uit recent internationaal onderzoek blijkt dat op landenniveau het systematisch gebruiken van evaluaties lastig te bewerkstelligen. Wel spelen evaluatiegemeenschappen en parlementen zeer waarschijnlijk een stimulerende rol in de totstandkoming en het gebruik van evaluaties.

Behalve de aanvrager speelt de evaluatieonderzoeker nog steeds een belangrijke rol in de bevordering van gebruik. Dit blijkt bijvoorbeeld uit leidraden om bevindingen toegankelijk en bruikbaar te presenteren. Ook blijkt het uit competentiebeschrijvingen van onderzoekers: niet alleen methodologisch-inhoudelijke vaardigheden worden belangrijk gevonden, maar ook interpersoonlijke en communicatieve vaardigheden, flexibiliteit en een gevoelige antenne voor de (politieke en beleids-) context.

Hoewel internationaal veel is verschenen over het bevorderen van uitvoering en gebruik van evaluaties, is nog niet aangetoond dat dit een positieve uitwerking heeft op de doeltreffendheid van beleid of wetgeving. Ook is gewaarschuwd voor ongewenste effecten van evaluatievermogen. Zo blijkt dat het ontwikkelen en handhaven van standaarden opportunistisch gedrag of ritualisering in de hand kan werken.

## Praktijkervaringen met evaluatievermogen bij beleidsdepartementen (onderzoeksvragen 2 en 3)

Beleidsdepartementen hebben heel verschillend vorm gegeven aan evaluatievermogen, in termen van zowel de betrokken organisatieonderdelen als procesmatige en inhoudelijke evaluatiekaders.

Het verantwoorden van begrotingsuitgaven blijkt een belangrijk doel van de evaluaties bij (rijks)beleidsdepartementen. Veruit het belangrijkste verantwoordingsinstrument is de beleidsdoorlichting op begrotingsartikelniveau, die volgens de Regeling Prestatiegegevens en evaluatieonderzoek (RPE) elke vier tot maximaal zeven jaar dient plaats te vinden en aan de Kamer moet worden gestuurd. Voor de inhoud van de beleidsdoorlichtingen, een synthesesdocument, zijn beleidsdirecties verantwoordelijk, de directie Financieel-economische zaken (FEZ) heeft bij elk departement een coördinerende en bewakende rol. De beleidsdoorlichting werkt door in de ambtelijke evaluatiepraktijken waar beleidsevaluaties plaatsvinden: idealiter stapelt zich hier in de loop van elke vier tot zeven jaar het benodigde evaluatiemateriaal op.

### Organisatieonderdelen gericht op evalueren

Evaluaties worden voor het overgrote deel uitbesteed aan externe partijen, vanwege interne capaciteitsrestricties en vermeende objectiviteit. Binnen beleidsdepartementen vervullen verschillende organisatieonderdelen een rol in het managen van evaluaties. Doorgaans ligt het initiatief voor ex post evaluaties van wetgeving en beleid bij de verantwoordelijke onderzoekscoördinatoren van beleidsdirecties of directeuren-generaal (DG's). Geen van de onderzochte departementen heeft een aparte evaluatieafdeling – afgezien van onderzoeksinstituten die ook evaluaties doen, zoals het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het ministerie van Veiligheid en Justitie (VenJ). Wel zijn er steeds een of meerdere stafafdelingen, units of directies die de evaluatiefunctie ondersteunen en stimuleren. Een gemeenschappelijke noemer is de rol in deze van de directie FEZ. Wat evalueren betreft, vullen directies FEZ hun rol per departement verschillend in – hoewel verantwoording uiteraard wel steeds het hoofddoel is. Zo heeft FEZ bij het ministerie van Sociale Zaken en Werkgelegenheid (SZW) een stimulerende en aanjagende rol in planning, kwaliteitsborging en gebruik van evaluaties en beleidsdoorlichtingen, terwijl andere directies FEZ dichter op hun formele controletaken blijven. Binnen het ministerie van Economische Zaken (EZ) is er sinds 2012-2013 een netwerk van organisatieonderdelen inclusief FEZ, met een Beleidskwaliteit- en Evaluatiecommissie (BEC) en een Regiegroep monitoring & evaluatie (M&E).<sup>1</sup> Dit netwerk richt zich specifiek op de kwaliteit van doeltreffendheidsevaluaties.

Anders dan FEZ dragen kennisorganen zorg voor meer strategisch en toekomstgericht onderzoek. In de praktijk maakt evaluatieonderzoek in zeer beperkte mate deel van uit van het kennisgeoriënteerde werk. Dit onderscheid tussen directies FEZ en kennisdirecties is echter niet zwart-wit. Neem bijvoorbeeld de rol die het Centraal Planbureau (CPB) en een enkele kennisdirectie hebben gespeeld in de bevordering van quasi-experimentele doeltreffendheidsstudies; en omgekeerd de rol die de *Chief Science Officer* (CSO) bij SZW vervult in relatie tot evaluatie- en onderzoeksinitiatieven.

Een aantal van de organisatieonderdelen en -structuren zijn nog vers en vereisen een wenperiode: het is zoeken naar een rolverdeling die niet alleen op papier, maar ook in de praktijk werkt. Een uitzondering is de Commissie Evaluatie Regelgeving (CER) van ZonMw, die al langer actief is voor een deel van de wetsevaluaties in het domein van met name het ministerie van Volksgezondheid, Welzijn en Sport (VWS).

---

<sup>1</sup> De Regiegroep M&E is in 2016 opgegaan in het Beleidsanalistenteam, de BAT.

### **Programmering en middelen**

In elke jaarlijkse departementale begroting wordt een evaluatieplanning opgenomen. Eventuele extra of parallelle onderzoeksprogrammeringen, met name strategische kennisagenda's, bevatten nauwelijks ex post evaluatieonderzoek en zijn vooral toekomstgericht. Programmering vindt ook plaats via (periodieke) evaluatieverplichtingen. 10-20% van de wetten kent een dergelijke verplichting. De resulterende beleids- en wetsevaluaties worden door directies gebruikt als basismateriaal voor beleidsdoorlichtingen.

Een adviespanel bij het ministerie van Buitenlandse Zaken (BZ) oordeelde dat de nadruk in de evaluatieprogrammering sterk op verantwoord en ligt – het afdekken van begrotingsartikelen – en minder op 'leren'. Dit lijkt relevant voor meer bestudeerde departementen. De politieke context werkt door in de evaluatieprogrammering. Sommige respondenten relativiseren het feitelijke belang van programmering: zij stellen dat evaluaties vooral samenhangen met veranderingen in kabinetssteun voor een bepaald beleid of met beleidsaanpassingen en -ontwikkelingen.

De meeste departementen stellen eerst een integraal budget voor onderzoek vast. Uitgevoerde evaluaties moeten binnen dat financiële kader passen. In de praktijk gaat men vaak uit van de bestedingsplannen. Onderzoeksinstituten beheren vaak hun eigen budget.

Wat de niet-financiële middelen betreft, melden respondenten vaak dat beleidsdirecties weinig tijd hebben voor evaluaties en onderzoek, onder meer blijkend uit beperkte capaciteit voor onderzoekscoördinatie. Dit zet ook de programmering onder druk. Bovendien werkt tijdgebrek in de hand dat men te laat begint aan geplande evaluaties, en belangrijke meetkansen 'mist' – zoals het aanleggen van een monitoringsysteem of het vormen van een controlegroep. Het laatste is een veelvoorkomend euvel.

In de context van verantwoording zijn de afgelopen paar jaar door FEZ en sommige andere organisatieonderdelen initiatieven genomen om de evaluatieprogrammering proactiever te maken, vooral met het oog op aankomende beleidsdoorlichtingen en het meer doordacht onderzoeken van doeltreffendheid/doeltreffendheid en doelmatigheid. We hebben niet overkoepelend kunnen vaststellen wat dit heeft opgeleverd.

### **Evaluatiekader**

Met evaluatiekader doelen we voornamelijk op evaluatie(bevorderende) programma's – die in de praktijk vaak zowel het proces als de inhoud van evaluaties betreffen (onderzoeksvragen 2 en 3).

In hoofdstuk 3 zijn praktijkervaringen met drie speciaal op wetgeving gerichte evaluatieprogramma's de revue gepasseerd, te weten het Programma Evaluatie Regeling (PER), de evaluaties van de Algemene wet bestuursrecht (Awb) en het programma Structurele Evaluatie Milieuwetgeving (STEM). Deze bevatten lessen over het evaluatieproces (programmering, afbakening hoofdvragen door de aanvrager), het evaluatiedesign en evaluatiecriteria. Een overkoepelende les die uit de praktijk van elk van deze programma's is getrokken, betreft de meerwaarde van een multidisciplinaire aanpak van wetsevaluaties. Het voldoen aan de juridische randvoorwaarden voor doeltreffendheid betekent nog niet dat de wet in de werkelijkheid ook doel zal treffen – en omgekeerd. In het PER is een gecombineerd juridisch-sociaalwetenschappelijke aanpak de eerste van zeven basisvereisten voor een wets-evaluatie.

Daarnaast zijn een aantal algemene beleidsevaluatieprogramma's besproken, die van toepassing zijn op zowel beleid als wetgeving. Bij het ministerie van Onderwijs, Cultuur en Wetenschappen (OCW) is beschreven, hoe de directie Kennis combinaties van *black box*- en *clear-box*- (experimentele en verklarende) evaluatiebenaderingen poogt te bevorderen. Bij EZ worden evaluatiehandvatten – in lijn met wat inter-

nationaal wordt voorgestaan – geïntegreerd aangeboden in een handreiking voor beleidskwaliteit. Het overkoepelende doel van die handreiking is effectiever beleid bewerkstelligen en één van de middelen is evalueren. De handreiking omvat alle fasen van de beleidscyclus, van probleemanalyse, instrumentkeuze en ex ante gevolgenbeoordeling, tot en met politieke beslissing, uitvoering en – uiteindelijk – ex post-evaluatie. Ze bevat ook aanwijzingen en tips bij het opstellen van (a) vroege beleidsexperimenten en (b) het opzetten van een monitoringsysteem.

In de Nederlandse praktijk verschilt de inhoudelijke evaluatiefocus enigszins tussen departementen, vaak passend bij de rol van de meest actieve organisatieonderdelen. Is FEZ relatief actief, dan weerspiegelt zich dat in een meer summatieve, op verantwoording gerichte aanpak, zoals bij EZ en SZW. Bij EZ staat men aanvragers bijvoorbeeld bij met uitleg en checklists voor de opzet van doeltreffendheidsevaluaties (procesmatig en inhoudelijk) en voor de beoordeling van offertes en rapporten. De ervaring met de Wet Bevordering speur- en ontwikkelingswerk (WBSO) laat zien dat een empirische wetsevaluatie volledig kwantitatief kan zijn; het fiscale instrumentarium keert in meerdere wiskundige formules terug. Nauwe betrokkenheid van een directie Kennis of onderzoeksinstituut uit zich vaak in verschillende soorten voorgestane aanpakken, zowel experimenteel (*black box*) als verklarend. Het onderscheid is dus niet zwart-wit; zo wijzen respondenten van meerdere directies FEZ en ook het ministerie van Financiën ook vaak op toetsing van veronderstellingen in de beleidstheorie.

Het Integraal afwegingskader voor beleid en wetgeving (IAK) wordt nauwelijks genoemd in het verband van ex post evaluatie; bij het ministerie van Infrastructuur en Milieu (IenM) probeert men bij ex post-beleidsdoorlichtingen terug te grijpen op de resultaten van het ex ante te gebruiken IAK voor beleid en wetgeving.

### **Borging van evaluatiekennis en -ervaring**

De aanvrager wordt ook in ons land steeds vaker gezien als een belangrijke schakel in procesverloop en evaluatiekwaliteit. Hoe is bij beleidsdepartementen geborgd dat aanvragers beschikken over de kennis en ervaring om evaluaties in goede banen te leiden? Bij de meeste DG's en directies vervullen de onderzoekskoördinatoren hierin een spilfunctie en gezien hun relevante expertise en ervaring fungeren zij tevens als vraagbaak voor collega's. Verder verschilt de ondersteuning sterk tussen departementen, afhankelijk van de betrokken organisatieonderdelen. Zo heeft bij SZW de *Chief Science Officer* (CSO) inhoudelijk een ondersteunende rol; bij OCW is dat de directie Kennis en bij EZ de vroegere Regiegroep M&E (inmiddels Beleidsanalistenteam (BAT)). Bij VenJ wordt bij een belangrijk deel van de (wets)evaluaties de aanvragersrol voor een deel overgenomen door projectbegeleiders binnen het WODC – de aanvrager maakt wel deel uit van de begeleidingscommissie.

Meerdere departementen en rijksacademies bieden bovendien cursussen of trainingen aan voor beleidsmedewerkers, die deels of geheel over evaluatie gaan.

Bij de selectie van onderzoekers wordt vooral gebruik gemaakt van de competitie tussen onderzoeksgroepen in het aanbestedingsproces. Daarnaast zijn een beperkt aantal initiatieven genoemd, zoals het gericht werken met selectiecriteria en het bijhouden van databases (WODC) met capaciteiten van externe onderzoeksgroepen. Door verloop is het een uitdaging gebleken om dergelijke gegevens up-to-date te houden. Juist bij wetsevaluaties is, zoals aangegeven, een multidisciplinair onderzoeksteam van belang. In de praktijk is die combinatie lang niet altijd waargemaakt – maar er zijn ook tegenvoorbeelden, zoals de vaste procedure bij ZonMw-wetsevaluaties en de latere Awb-evaluatie van versnellingsinstrumenten.

Ook begeleidingscommissies (BC's) worden vaak ingezet ter bewaking van voortgang, inhoud of kwaliteit. Bij WODC en CER-ZonMw is het werken met een BC

standaardpraktijk; in geval van WODC-evaluaties wordt altijd een wetenschapper als voorzitter aangewezen, bij ZonMw is dat een CER-lid. Het uitwisselen van ervaringen op het vlak van evaluatie gebeurt binnen departementen vaak via informele overdracht of in het afdelingsoverleg. Betwijfeld kan worden of dit volstaat in het borgen van evaluatie-ervaring. Interdepartementaal vindt uitwisseling vooral plaats tussen FEZ-directies en Financiën, in de Interdepartementale Begeleidingsgroep Prestatiegegevens en Beleidsevaluatie (IBP), wat positief wordt gewaardeerd.

Initiatieven ter versterking van het evaluatievermogen lijken kwetsbaar want dreigen soms in de knel te komen: beleidsmedewerkers zijn er niet klaar voor, of er zijn reorganisaties (zoals bij inkoop), of snel personeelsverloop. Dit bevestigt een observatie uit de internationale praktijk dat evaluatievermogen een kwestie van een lange adem is.

### **Bruikbaarheid en gebruik**

Evaluatiegebruik geeft zin aan alle eerder behandelde aspecten van evaluatievermogen: zonder gebruik vindt geen terugkoppeling naar de beleids- of wetgevingscyclus plaats en rijst de vraag wat het nut van de evaluatie is geweest.

Contact en overleg tussen aanvrager en onderzoekers gedurende het evaluatietraject – al dan niet in een BC – is volgens de literatuur en ook volgens dit onderzoek een cruciale manier om de kans op gebruik te vergroten. Ook timing en tijdigheid in relatie tot de beleids- en besluitvorming zijn dat.

Over het algemeen hanteert elk onderzocht departement een vergelijkbare formele wijze van communicatie en (korte-termijn)kennisname en gebruik van evaluatieonderzoek. In de regel worden rapporten openbaar gemaakt – en vaak ook naar de Tweede Kamer gestuurd – met een beleidsreactie. Door onderzoekers vindt daarnaast vaak nog verspreiding plaats in de vorm van academische en vakpublicaties en via toelichtingen op de departementale werkvloer. Bij BZ is voor 'leren' een adviespanel ingesteld om de bruikbaarheid en terugkoppeling van evaluaties naar beleid te vergroten.

Vervolgens is het aan beleidsmedewerkers en politiek wat er met bevindingen gebeurt. Er zijn bij departementen geen 'verplichte' constructies gevonden voor terugkoppeling naar de beleidscyclus, los van de horizonbepaling in financiële regelingen die stelt dat de regeling na vijf jaar vervalt. Eventuele voorzetting dient te worden gemotiveerd, bijvoorbeeld met onderzoek naar doeltreffendheid en doelmatigheid. Niettemin percipiëren enkele respondenten dat de Tweede Kamer toenemende belangstelling laat zien in ex post evaluaties en in wat de minister daarmee wil gaan doen. Het betreft wel met name de verantwoordingsfunctie en financiële lasten.

Om de leerfunctie van evaluatie een belangrijker plaats te geven, zouden ex post-evaluaties volgens sommige gebruikers meer actie- en toekomstgericht moeten zijn. Ook worden evaluaties leerzaam gevonden als ze de impliciete aannames waarop wetgeving of beleid rust, boven water halen en toetsen aan de empirie. Een aanvullende weg is om – in lijn met wat bijvoorbeeld in de VN en bij BZ wordt gestimuleerd – meer werk te maken van de systematische stapeling van kennis uit evaluaties, bijvoorbeeld in metastudies en kennisbanken. Zo verzet de aanvrager van dit onderzoek (WKB) op dit punt al geruime tijd werk.

### **Slotbeschouwing (Onderzoeksvraag 4)**

Veel van de beschreven constructies, procedures en aanpakken zijn nog vers, waardoor niet kan worden nagegaan – dit was ook niet het doel – in hoeverre ze struc-

turele verbeteringen hebben opgeleverd in termen van leren en verantwoorden of zelfs tot effectiever beleid hebben geleid. Uit ervaringen elders is opgemaakt dat evaluatievermogen ook een kwestie van doorzettingsvermogen is. Dit maakt het interessant om de verdere ontwikkelingen te volgen.

Een randvoorwaarde voor 'lerend evaluatievermogen' lijkt dat initiatieven ter vermeerdering van beleidsrelevante kennis niet worden afgestraft, maar ondersteund door de politiek en door de instrumenten in het centrale evaluatiestelsel. Sommige respondenten betreuren de nadruk die met het centrale evaluatiestelsel en de beleidsdoorlichtingen op (eind)doeltreffendheid en verantwoording is komen te liggen. Van plan- en procesevaluatie – en uiteraard vele andere soorten beleidsonderzoek – kan bijvoorbeeld veel geleerd worden, ondanks dat de 'hamvraag' van doeltreffendheid en doelmatigheid niet of niet volledig wordt beantwoord.

*Wat betekenen bovenstaande bevindingen nu voor wetgevingsdirecties en de aanvrager, WKB?* Directies Wetgeving en/of juridische zaken lijken op grond van dit onderzoek nauwelijks aangehaakt bij ex post beleidsevaluaties, hoewel zij soms wel een wetsevaluatie begeleiden, zeker als er meerdere directies (bijv. Spoorwetgeving) of departementen (bijv. Awb) bij betrokken zijn. Over het algemeen zijn beleidsdirecties leidend, terwijl de wetgevingsjurist bijna los staat van het centrale evaluatiestelsel met de beleidsdoorlichtingen. Aan de voorkant zijn wetgevingsdirecties nadrukkelijker betrokken, in het kader van ex ante onderzoek en het IAK. Wetgevingsjuristen en andere specialisten kunnen echter een waardevolle rol spelen in ex post wetsevaluaties; zuiver sociaalwetenschappelijke aanpakken zijn immers te arm gebleken. Zij zouden vooralsnog eerst meer betrokken kunnen raken bij het evalueren, voordat wordt ingezet op een aparte lijn van 'wetsevaluatievermogen'. Ook kunnen wetgevingsjuristen een aantal andere dingen betekenen. Zo vallen evaluatiebepalingen in wetten zelf meer in de richting van 'lerend evalueren' te formuleren; standaard vragen om onderzoek naar 'doeltreffendheid en doelmatigheid' is daarvoor niet altijd genoeg. Er kan bijvoorbeeld tevens om inzicht gevraagd worden in de uitvoeringspraktijk en in het hoe en waarom van de werking. Daarbij zijn realistische evaluatietermijnen in de bepaling nodig om recht te doen aan de ontwikkeling van het beleid dan wel het uitkristalliseren van doeltreffendheid en neveneffecten. Ten slotte draagt een adequate toelichting op de wet in de memorie van toelichting (MvT) bij aan de werking van de evaluatiebepaling: wat is de bedoeling van de wetgever; wat zijn de aannames, welke resultaten van de wet worden op basis hiervan verwacht? Op welke punten kan een wetsevaluatie hier over x jaar aan bijdragen? WKB heeft inmiddels een schrijfwijzer voor MvT's voltooid, die kan bijdragen aan de onderbouwing en verheldering van wetgeving en beleid (beleids-theorie) en zo aan beter gefocuste evaluaties.

Het conceptrapport heeft de inbreng gevormd voor een expertmeeting met als doel de discussie en gedachtenvorming over de (eventuele) verdere ontwikkeling van evaluatievermogen – in het bijzonder rond wetgeving – aan te jagen.