

Zicht op evaluatiecapaciteit

**Een onderzoek naar het vermogen van de Nederlandse ministeries
om (wets)evaluaties te verrichten**

Masterthesis Bestuurskunde
Specialisatie: Organisatie & Management
Andreea Parapuf (s4263723)

Begeleider: Dr. Ellen Mastenbroek
Begeleider WODC: Dr. Carolien Klein Haarhuis

Faculteit der Managementwetenschappen
Radboud Universiteit Nijmegen

Inhoudsopgave

Hoofdstuk 1: Inleiding	6
1.1 Probleemschetsing.....	6
1.2 Probleemstelling.....	10
1.2.1 Doelstelling.....	10
1.2.2 Centrale vraag	11
1.2.3 Deelvragen	12
1.3 Relevantie van het onderzoek	13
1.3.1 Relevantie voor de aanvrager van het onderzoek	13
1.3.2 Maatschappelijke relevantie.....	14
1.3.3 Wetenschappelijke relevantie	15
1.4 Voorbeschouwing theoretisch kader	16
1.5 Voorbeschouwing methodologisch kader	18
1.6 Leeswijzer	19
Hoofdstuk 2: Beleidskader.....	21
2.1 Inleiding	21
2.2 Ontwikkelingen in het Nederlandse overkoepelende evaluatiebeleid.....	21
2.2.1 Evaluatie in de jaren 1970.....	21
2.2.2 Evaluatie in de jaren 1980.....	22
2.2.3 Evaluatie in de jaren 1990.....	23
2.2.4 Evaluatie in het eerste decennium van de jaren 2000	24
2.3 Het huidige evaluatiebeleid.....	24
2.4 De drie functies van evaluatie	26
2.5 Vormen van evaluatie.....	28
2.6 Wetgevingskwaliteitsbeleid	31

2.7 Evaluatie van wet- en regelgeving	33
2.8 Samenvatting.....	38
Hoofdstuk 3: Theoretisch kader.....	39
3.1 Inleiding	39
3.2 Modellen voor de diagnose van evaluatiecapaciteit.....	39
3.2.1 Het vermogen om evaluaties te doen en het vermogen om evaluaties te gebruiken ...	39
3.2.2 Vraagzijde en aanbodzijde van evaluatie.....	41
3.2.3 Micro-, meso- en macroniveau	43
3.2.4 Ontwerpen van evaluatiecapaciteit	46
3.3 Conceptueel model.....	49
3.3.1 Institutionele inrichting.....	50
3.3.2 Individuele competenties	52
3.3.3 Omgevingsfactoren.....	54
3.4. Samenvatting.....	54
Hoofdstuk 4: Methodologisch kader.....	56
4.1 Inleiding	56
4.2 Onderzoeksstrategieën	56
4.3 Kwalitatieve methoden van dataverzameling	59
4.4 Keuze van methoden en databronnen.....	60
4.5 Operationalisatie van variabelen	62
4.6 Kwaliteit van empirisch onderzoek.....	64
4.6.1 Validiteit	64
4.6.2 Betrouwbaarheid.....	65
4.7 Samenvatting.....	65
Hoofdstuk 5: Resultaten.....	67

5.1. Inleiding	67
5.2 De institutionele aspecten van evaluatiecapaciteit	67
5.2.1 Structuur van organisatie	67
5.2.2 Evaluatieplanning en -programmering	77
5.2.3 Evaluatiekader.....	83
5.2.4 Financiële en niet-financiële middelen	91
5.2.5 Communicatie	93
5.2.6 Gebruik van evaluatiebevindingen	95
5.2.7 Leren van evaluaties	98
5.3 Individuele competenties.....	100
5.3.1 Kennis over evaluatie.....	100
5.3.2 Benodigde vaardigheden voor evaluatieonderzoek	103
5.3.3 Ervaring van medewerkers met evaluatieonderzoek	105
5.4 Omgevingsfactoren	107
5.4.1 Politieke betrokkenheid	108
5.4.2 Stakeholdersbetrokkenheid.....	109
5.5 Samenvatting.....	110
Hoofdstuk 6: Conclusie en aanbevelingen.....	113
6.1 Beantwoording van de onderzoeksvraag.....	115
6.2 Kritische terugblik op het onderzoek	117
6.3 Aanbevelingen naar aanleiding van het onderzoek.....	118
6.4 Suggesties voor vervolgonderzoek.....	120
Bronnen.....	122
Bijlage A: Topiclijst (gebruikt bij het interviewen van respondenten).....	127
Bijlage B: Lijst van interviews	134

Hoofdstuk 1: Inleiding

1.1 Probleemschetsing

De huidige samenleving stelt de overheid voor grote uitdagingen. Ten eerste zijn de burgers mondiger geworden. Ze komen voor zich zelf op en eisen hun rechten op als ze de overheidsinterventie als te streng of te laks ervaren. Burgers kunnen zich tegen de overheid verzetten als ze denken dat hun rechten worden geschonden of dat ze ongelijk worden behandeld in gelijke omstandigheden. Dit kan de acceptatie en de legitimatie van de overheid schenden. Goede wetgeving vormt de garantie dat de burgers op een gelijke manier in gelijke omstandigheden worden behandeld. Daardoor speelt goede wetgeving een belangrijke rol in de acceptatie en de legitimatie van de overheidsinterventie in de samenleving.

Ten tweede zijn economische competitiviteit en innovatie belangrijke kenmerken van de Westerse samenleving. Om innovatie en competitiviteit te bevorderen, probeert de overheid bedrijven aan te trekken en gunstige regelingen op te stellen die bedrijven aanmoedigen zich in Nederland te vestigen. Daarvoor zijn wet- en regelgeving onmisbaar. Te veel regelgeving kan echter bedrijven weggagen en moet daarom worden vermeden. De overheid staat dus voor bijzondere uitdagingen: burgers gelijk behandelen, maar ook tevreden stellen, het economische welzijn veilig stellen, maar ook het publiek belang vrijwaren (Ministerie van Justitie, 2000).

Om de samenleving goed te laten functioneren en adequaat om te gaan met deze uiteenlopende eisen, zijn goede wet- en regelgeving onontbeerlijk. Wet- en regelgeving reguleren het gedrag van burgers, bedrijven en publieke organisaties en vormen het kader waarbinnen de maatschappij goed kan functioneren. De ervaring van de wetgevingsjuristen is dat de wet- en regelgeving in de afgelopen jaren steeds vaker moet worden aangepast om de consequenties van de snelle technologische veranderingen en de ontwikkelingen en innovaties die daaruit voortvloeien, op te kunnen vangen (Ministerie van Justitie, 2000). Om goede wetgeving te maken is het dus van belang dat wetgevingsjuristen¹ zich zoveel mogelijk door de praktijk laten

¹ Wetgevingsjuristen houden zich bezig met het opstellen of beoordelen van voorgenomen wet- en regelgeving met als doel de juridische houdbaarheid van de regels te garanderen. Onder hun werkzaamheden tellen het doen van onderzoek en schrijven van teksten van wetten. Wetten zijn onderdeel van breder beleid, dat ook andere interventies bevat. Omdat wetten en beleid gemaakt worden om maatschappelijke problemen op te lossen, werken wetgevingsjuristen samen met beleidsambtenaren om de juiste instrumenten te ontwikkelen en de verwachten effecten hiervan te voorspellen (Academie voor Wetgeving, 2015).

informereren en inspireren over hoe een bepaalde wet of regel in de praktijk uitpakt. Om deze informatie te vergaren, worden wetsevaluaties opgesteld en uitgevoerd (KWJZ, 2015).

Een wet is een juridisch instrument gericht op gedragsverandering van doelgroepen (burgers, bedrijven, lokale instanties, uitvoerders etc.) of op het aanleggen van de infrastructuur van het besturen (*governance*). Enerzijds moeten wetten rechtsbeginselen zoals gelijkheid, proportionaliteit, subsidiariteit of rechtszekerheid vrijwaren (Klein Haarhuis & Niemeijer, 2009: 404). Anderzijds wil men met wetten bepaalde beleidsdoelstelling bereiken en maatschappelijke interventies teweeg brengen. De sociale context waarin interventies worden gedaan, kan het succes of het falen daarvan bepalen. Wetten zijn in het bijzonder lastig te evalueren omdat ze zowel juridische als sociale aspecten bevatten. Sociaalwetenschappelijke methodologie is toepasbaar bij de evaluatie van wetten, maar is niet voldoende, omdat de wetsevaluatie rekening moet houden met de juridische houdbaarheid van de wet. Uitspraken over de effectiviteit van de wet moeten zich daarom niet alleen baseren op de uitleg van hoe de wet juridisch in elkaar zit, maar in even grotere mate op het empirisch onderzoek van hoe de wet heeft gewerkt, in welke context en onder welke omstandigheden de wet wel of niet heeft gewerkt. Zo kan dezelfde wet wel effectief zijn in gunstige omstandigheden en niet-effectief in hinderlijke omstandigheden.

De verhouding tussen wetten en maatschappelijke interventies is ook complex. Enerzijds kan een wet het “voertuig” van beleid zijn (Veerman, 2004: 15; Klein Haarhuis & Niemeijer, 2009: 403). Dat wil zeggen dat één of meerdere interventies uit het beleid wettelijk worden vastgelegd. Anderzijds bevatten wetten vaak meerdere interventies en interventies kunnen op verschillende doelgroepen doelen (Klein Haarhuis & Niemeijer, 2009: 405).

Wetsevaluaties streven te beoordelen of de wet en de bedachte reeks interventies de beoogde situatie daadwerkelijk teweeg hebben gebracht zoals bedacht in theorie. Dit doet een evaluatieonderzoek door gegevens over de precieze werking van de wet, over de doelstellingen en de (direct, indirecte, gewenste en ongewenste) effecten van de wet te verzamelen en door deze systematisch te analyseren. Elke evaluatie bevat een *waardebepaling* ten aanzien van de uitkomsten en de implementatie van de wet of het beleid (Swanborn, 2007). De combinatie van het juridische en het sociale aspect van wetten maakt wetsevaluaties bijzonder complex. Het is moeilijk om strakke conclusies te trekken over hoe en in hoeverre een wet aan de beleidsdoelstellingen bijdraagt, aangezien de wet maar één van de beleidsinstrumenten is (Klein Haarhuis & Niemeijer, 2009: 404). Wetten zijn ook vaak domein specifiek. De methodologie van

wetsevaluaties is daarom vaak context- en domeinafhankelijk. Het is vaak moeilijk om algemene lessen uit eerdere wetsevaluaties te trekken over hoe je een wetsevaluatie moet inrichten en uitvoeren. Meta-evaluaties leggen verbanden tussen de onderliggende theorieën van verschillende wetten en interventies om gemeenschappelijke draden te identificeren die zich in de praktijk herhaaldelijk als succesfactoren hebben bewezen bij een diversiteit van wetten, interventies en domeinen (Klein Haarhuis & Niemeijer, 2009: 405).

Het domein van evaluatieonderzoek is in de afgelopen decennia zowel op landelijk als op internationaal niveau sterk gegroeid (Leeuw, 2008: 157; Dahler-Larsen, 2013). Kwaliteitszorg inzake wetgeving en wetsevaluaties is een groot aandachtspunt voor politici, beleidsmakers en wetgevingsjuristen, zowel in het internationaal wetenschappelijke debat als in het beleid van de Nederlandse overheid. In Nederland bestaat er binnen de Directie Wetgeving en Juridische Zaken (DWJZ) van het ministerie van Veiligheid en Justitie (VenJ) de sector Wetgevingskwaliteitsbeleid. Deze sector is verantwoordelijk voor “het bevorderen van de kwaliteit van de Nederlandse wetgeving” (Ministerie van Justitie, 2015). Dit doet de sector Wetgevingskwaliteitsbeleid concreet door ondersteuning aan andere ministeries te bieden en zorg te dragen voor de “inhoud, vormgeving, procedures en organisatie van wetgeving” van alle ministeries (Ministerie van Justitie, 2015).

De laatste paar jaren lijkt in enkele landen het besef door te dringen dat de kwaliteit van evaluatie niet alleen afhangt van criteria en standaarden voor evaluatie, maar ook van de inrichting van de evaluatiefunctie binnen de organisatie en de capaciteit die de organisatie heeft om evaluatie te verrichten of uit te zetten. Bij steeds meer publieke organisatie dringt het bewustzijn door dat de inrichting van de evaluatiefunctie een belangrijke voorwaarde voor de kwaliteit van evaluaties is. Grote internationale instituties zoals EU, OESO, UN en Wereldbank zijn aanjagers van deze ontwikkeling. In de literatuur zijn er lijsten opgesteld met vaardigheden voor evaluatieonderzoekers (Cooksy, 2012; Stevahn et al., 2005) en modellen voor het meten en verbeteren van de evaluatiefunctie.

Uit recente empirische onderzoeken (Cooksy, 2012; Bourgeois & Cousins, 2013; Jacobs et al., 2015) blijkt dat de kwaliteit van het evaluatieproces en de evaluatierapporten deels beïnvloed wordt door de evaluatiecapaciteit en de leercapaciteit binnen een organisatie en de evaluatiecultuur in het land. Het onderzoek naar evaluatiecapaciteit staat in Nederland echter nog in de kinderschoenen. De enige bestaande studies bespreken de casus van Nederland in internationaal

vergelijkend perspectief (Jacob et al., 2015; Furubo et al., 2002). Nederland scoort goed in internationaal perspectief. In een recent vergelijkend onderzoek naar de mate van institutionalisering van evaluatieonderzoek onder 19 OESO-landen eindigt Nederland op de vijfde plaats met een score van 15,3 uit een maximum van 18 punten (Jacobs et al., 2015).

Terwijl Nederland op landelijk niveau goed scoort op evaluatiecapaciteit, zijn er tussen de ministeries grote verschillen in de manier waarop de evaluatiefunctie wordt ingericht. Uit een rapport van de onafhankelijke Visitatiecommissie Wetgeving uit 2000 blijkt dat er verschillen zijn tussen ministeries in de manier en de mate waarin de wetgevingsfunctie is georganiseerd (Ministerie van Justitie, 2000: 6). De Nederlandse rijksoverheid heeft een gedecentraliseerde karakter, waarbij elk ministerie zijn eigen beleidstraditie en -cultuur heeft. Een gevolg hiervan is dat elk ministerie zijn eigen evaluatieprogrammering opstelt en verantwoordelijk is voor de evaluatie van het beleid dat op de terreinen van het ministerie speelt. Dit kan een reden zijn voor het verschil tussen ministeries in de kwaliteit van de uitgevoerde wetsevaluaties. Uit de meta-evaluaties die uitgevoerd zijn in de afgelopen jaren blijkt ook dat de kwaliteit van evaluatierapporten sterk verschilt (Klein Haarhuis & Niemeijer, 2009; Veerman et al., 2013). De vraag is dan of deze verschillen in de kwaliteit van evaluatierapporten, die door eerder onderzoek is geconstateerd, ook de diversiteit in de inrichting van evaluatiecapaciteit tussen ministeries weerspiegelen.

Voor zover bekend bestaan er geen vergelijkende studies over de verschillen in de evaluatiecapaciteit van de Nederlandse ministeries en de invloed ervan op de kwaliteit van wetsevaluaties en evaluatieonderzoek breder gezien. Deze scriptie beoogt dit kennistekort te ondervangen. Door inzicht te geven in de verschillen in evaluatiecapaciteit van de Nederlandse ministeries en door *best practices* enerzijds en tekorten en gebreken anderzijds in kaart te brengen, draagt dit onderzoek bij aan de bewustwording van het belang van evaluatiecapaciteit voor de borging van de kwaliteit van evaluatieonderzoek. Bewustwording mondt uit in het verbeteren van de voorwaarden en voorzieningen voor evaluatieonderzoek zowel in termen van benodigde kennis bij de uitvoerders als in termen van het opstellen van maatregelen, instrumenten arrangementen die de kwaliteit van evaluatieonderzoek bevorderen.

Dit onderzoek wordt gedaan in opdracht van de sector Wetgevingskwaliteitsbeleid van de Directie Wetgeving en juridische zaken van VenJ. Deze sector monitort de kwaliteit van de wetgeving van de andere ministeries en is verantwoordelijk voor de verspreiding van *best practices*

van deze ministeries. Een van de aandachtspunten voor de Directie Wetgeving is het vergroten van de leercapaciteit en het gebruik van de evaluatieresultaten voor het ontwikkelen van toekomstig wetgevingskwaliteitsbeleid.

1.2 Probleemstelling

1.2.1 Doelstelling

Dit onderzoek beoogt de inrichting van de evaluatiecapaciteit bij de Nederlandse ministeries te evalueren aan de hand van een lijst van criteria voor evaluatiecapaciteit. Het doel van het onderzoek is inzicht verschaffen in het vermogen van de Nederlandse ministeries (en verschillen hiertussen) om (wets)evaluaties van goede (wetenschappelijke) kwaliteit te verrichten en de bevindingen uit deze evaluaties te gebruiken in besluitvorming.

Allereerst zal dit onderzoek inzicht verschaffen in wat *evaluatiecapaciteit* omvat en hoe deze theoretisch uiteengeerafeld kan worden. Om deze doelstelling te bereiken, zullen recente inzichten uit de internationale wetenschappelijke literatuur worden uiteengezet. Op basis hiervan zal een conceptueel model worden gepresenteerd waarin een omvattend beeld wordt geschetst van alle belangrijke *aspecten van de evaluatiecapaciteit* van een organisatie. De evaluatiecapaciteit van de Nederlandse ministeries wordt geëvalueerd tegen dit model.

Evaluatiecapaciteit wordt breed opgevat in dit onderzoek. Zij omvat niet alleen het moment van evalueren zelf, met daaronder het verrichten van onderzoek en het schrijven van het onderzoeksrapport. Het omvat ook de andere fases van het evaluatietraject: de voorbereidingsfase van evaluatie en de verspreiding en het gebruik van evaluatiebevindingen bij het opstellen van toekomstig beleid en wetgeving. Dit onderzoek zal dus analyseren in hoeverre de ministeries voorbereid zijn om startnotities voor evaluaties op te stellen, de vraag- en probleemstelling voor evaluaties te formuleren, programma's en handreiking voor evaluaties te ontwikkelen en in de praktijk te gebruiken, de methodologische criteria en standaarden voor evaluatie te gebruiken, extern uitgevoerde evaluaties te begeleiden en de bevindingen uit evaluatierapporten terug te koppelen naar de wetgever en de beleidsmakers voor toekomstige besluitvorming. Het vermogen om evaluatiebevindingen te gebruiken in toekomstige besluitvorming beschouwen wij als onderdeel van de evaluatiecapaciteit van een organisatie in aansluiting op de definitie van evaluatiecapaciteit van Bourgeois en Cousins (2013). Zo zal de studie ook onderzoeken wat met de evaluatierapporten gebeurt na de afronding van evaluaties, in hoeverre de opdrachtgevende

ministeries (directies of afdelingen) ook in staat zijn om evaluatiebevindingen binnen en buiten het ministerie te verspreiden en invloed uit te oefenen op toekomstige besluitvorming met behulp van de evaluatiebevindingen.

De aspecten van het conceptueel model dienen als criteria voor het evalueren van het vermogen van ministeries om evaluaties uit te zetten, te verrichten, te begeleiden en in besluitvorming te gebruiken. Het empirische onderzoek zal inzicht geven in de verschillen tussen ministeries in de evaluatiecapaciteit. Deze studie richt zich op zowel *best practices* – ministeries die zich positief op een of meerdere aspecten van de evaluatiecapaciteit onderscheiden – als op de knelpunten in de inrichting van evaluatiecapaciteit bij ministeries. Dit onderzoek zal een oordeel vellen over het vermogen van de Nederlandse ministeries om goede (wets)evaluaties te verrichten en uit te zetten. Op basis van het empirische onderzoek zullen een aantal aanbevelingen en aanknopingspunten worden geformuleerd, die WKB kunnen helpen een plan van aanpak op te stellen om de kwaliteit van het proces, inhoud en kwaliteit van evaluaties bij de Nederlandse ministeries te borgen en indien nodig te verbeteren.

1.2.2 Centrale vraag

In hoeverre voldoen de Nederlandse ministeries aan de voorwaarden voor evaluatiecapaciteit die gesteld worden in theoretische modellen, ten einde de kwaliteit van de (wets)evaluaties die ministeries verrichten, te borgen?

Dit onderzoek is evaluatief van aard. Het evalueert in hoeverre de voorwaarden voor evaluatiecapaciteit aanwezig zijn bij de Nederlandse ministeries. De voorwaarden worden afgeleid uit modellen die in internationale wetenschappelijke literatuur zijn gevonden. De modellen zijn constructen die op ervaringen steunen van andere landen met de ontwikkeling van evaluatiecapaciteit bij publieke organisaties. Aan de hand van theoretische inzichten uit de wetenschappelijke nationale en internationale literatuur, documentenanalyse en interviews met wetgevingsjuristen en ambtenaren die betrokken zijn bij wetsevaluaties en evaluatietrajecten in het algemeen, zal dit onderzoek de verschillen in de evaluatiecapaciteit bij verschillende ministeries verkennen en evalueren.

1.2.3 Deelvragen

Om de centrale vraag te beantwoorden zal dit onderzoek worden opgesplitst in vier onderdelen. Ieder onderdeel wordt in een apart hoofdstuk behandeld. Bij ieder hoofdstuk horen één of meer deelvragen zoals hieronder opgesteld.

1. Hoe heeft het Nederlandse evaluatie- en wetsevaluatiebeleid zich vanaf jaren 1970 tot heden ontwikkeld?
2. Wat zijn de functies en de vormen van evaluatie?
3. Welke rol speelt de evaluatiecapaciteit voor het borgen van evaluatiekwaliteit?
4. Hoe wordt evaluatiecapaciteit in de wetenschappelijke literatuur gedefinieerd en opgevat?
5. Aan welke voorwaarden moeten publieke instellingen voldoen volgens wetenschappelijke literatuur om over goede evaluatiecapaciteit te beschikken?
6. Op welke manier kunnen de inzichten uit de internationale wetenschappelijke literatuur worden geoperationaliseerd om de mate van evaluatiecapaciteit van de Nederlandse ministeries te evalueren?
7. Welke voorwaarden voor evaluatiecapaciteit zijn terug te vinden bij de onderzochte ministeries en welke ontbreken?
8. Welke *best practices* rondom evaluatieonderzoek van wetten kunnen uit het voorgaande worden gehaald?
9. Op welke dimensies van evaluatiecapaciteit verschillen de Nederlandse ministeries in de praktijk van wetsevaluatie?
10. Welke aanbevelingen kunnen gedaan worden op basis van de resultaten van deze empirische studie?

Om de centrale vraag over de mate van evaluatiecapaciteit van de Nederlandse ministeries te beantwoorden, zal gebruik worden gemaakt van *literatuurstudie* (wetenschappelijke theorieën over de voorwaarden van evaluatiecapaciteit, de borging van de kwaliteit van evaluatieonderzoek en de verhouding tussen evaluatiecapaciteit en evaluatiekwaliteit), aangevuld met een kwalitatieve onderzoeksstrategie waarbij twee methoden van dataverzameling worden gebruikt: *inhoudsanalyse van documenten* en *interviews*.

Het beleidskader (hoofdstuk 2) geeft antwoord op de eerste twee deelvragen. Deelvragen drie, vier en vijf worden in het theoretisch kader (hoofdstuk 3) beantwoord. Het methodologisch kader (hoofdstuk 4) beantwoordt deelvraag zes. Tot slot worden deelvragen zeven, acht en negen in het analyse-hoofdstuk (5) beantwoord. Hoofdstuk 6 grijpt naar de probleemstelling terug, geeft antwoord op de centrale vraag, trekt conclusies uit het uitgevoerde onderzoek en formuleert aanbevelingen voor de opdrachtgever en voor het verbeteren van de evaluatiecapaciteit bij de Nederlandse ministeries.

1.3 Relevantie van het onderzoek

Dit paragraaf licht de relevantie van dit onderzoek op drie vlakken toe: ten eerste, de relevantie voor de aanvrager, de sector Wetgevingskwaliteitsbeleid (WKB), onderdeel van de Directie Wetgeving en Juridische zaken van het ministerie van Veiligheid en Justitie, ten tweede de maatschappelijke relevantie, en tot slot de wetenschappelijke relevantie.

1.3.1 Relevantie voor de aanvrager van het onderzoek

Dit onderzoek wordt uitgevoerd in opdracht van de sector Wetgevingskwaliteitsbeleid van het Ministerie van Veiligheid en Justitie. Het ministerie is verantwoordelijk voor de kwaliteit van het wet- en regelgevingsbeleid van alle ministeries (Kenniscentrum Wetgeving en Juridische zaken, 2015). WKB beoogt de kwaliteit van wetsevaluaties te bevorderen en indien nodig te vergroten. Om dit te doen is het echter allereerst nodig om te weten of en in hoeverre de ministeries bereid en voorzien zijn om goede evaluatie te verrichten. De aanname afkomstig uit de wetenschappelijke literatuur en uit de internationale praktijkervaringen is dat hoe beter een organisatie de evaluatiefunctie inricht – in termen van voldoende middelen voor evaluatie, goed opgeleide medewerkers die ervaring hebben met het verrichten en uitzetten van evaluaties, goede communicatie tussen aanvragers en onderzoekers, gebruik van methodologische criteria voor evaluatie – des te beter is de kwaliteit van evaluatieonderzoek en des te groter is de kans dat de evaluatiebevindingen in de besluitvorming door werken. Zo beweert Veerman dat “met een institutionele inbedding van kwaliteitszorg en een actief beleid de kans op een goed product wordt vergroot” (Veerman, 2014: 210). Met een goed product doel hij hier op de kwaliteit van evaluatierapporten.

Door inzicht te bieden in mate van evaluatiecapaciteit bij de rijksoverheid, de verschillen in evaluatiecapaciteit tussen ministeries en de mogelijke knelpunten die ministerie in het evaluatieproces ervaren, draagt dit onderzoek bij aan de doelstelling van WKB om de kwaliteit van wetsevaluaties te borgen en indien nodig, te verbeteren.

Het analyseren van de verschillen in de evaluatiecapaciteit van ministeries is relevant omdat dit een van de verklaring is voor de fluctuerende kwaliteit van evaluaties die verricht zijn in opdracht van verschillende ministeries. Gebrekkige evaluatiecapaciteit is een reden voor het produceren van minder goed kwalitatieve (methodologisch goed onderbouwde en bruikbare) evaluatierapporten. Dit onderzoek zal identificeren waar de problemen in het evaluatietraject zich voordoen en welke ministeries hulp nodig hebben om de evaluatiefunctie beter in te richten en zo de kans op betere evaluaties te vergroten. Dit onderzoek biedt WKB een overzicht van de aspecten van evaluatiecapaciteit die relevant zijn voor het verrichten van goede evaluaties en van de aspecten van de evaluatiecapaciteit waarop ministeries minder goed scoren. Naast het identificeren van *best practices* en knelpunten in het evaluatietraject bij ministeries, zal dit onderzoek ook handvatten aan WKB bieden om maatregelen te nemen voor het verbeteren van de inrichting van de evaluatiefunctie bij ministerie die het minder goed doen. Dit onderzoek zal WKB in staat stellen om een plan van aanpak op te stellen om de individuele departementen te ondersteunen bij het voorbereiden van evaluaties, alsook bij het verspreiden van nuttige kennis en *best practices* tussen departementen. Tot slot helpt het onderzoek de overheid om haar instrumenten (middelen, bronnen) gericht in te zetten en zo de randvoorwaarden voor een goed verloop van de beleidscyclus. Idealiter zou de besluitvorming (in zowel beleid als wetgeving) gebruik maken van bevindingen uit evaluaties.

1.3.2 Maatschappelijke relevantie

Naast de directe relevantie voor de opdrachtgever, heeft dit onderzoek ook een bredere maatschappelijke relevantie. De vraag naar een bepaalde minimale kwaliteit van wetsevaluaties is een permanente zorgreden, maar ook een actueel onderwerp. De hedendaagse maatschappij is sterk gericht op de verantwoording op een rationele, wetenschappelijke, bewijsbare manier van alle keuzes die gemaakt worden in de publieke sector. Deze druk komt van zowel de burger die wil weten hoe de overheid het uitgaven spendeert, als door de overheid zelf die wil laten zien dat haar keuzes worden gebaseerd op wetenschappelijke en bewezen inzichten. Onder deze

omstandigheden is evaluatie beschouwd als een op praktijk gerichte *wetenschappelijke* activiteit. Evaluatieonderzoekers worden verwacht om verantwoording af te leggen over alle gemaakte keuzes: de onderzoeksdesign, de gebruikte methoden, het trekken van conclusies en het rapporteren hiervan op toegankelijke wijze, wat het gebruik van evaluatiebevindingen door beleidsmakers, politici en derden zou bevorderen. Dit alles met oog op het verbeteren van de wet- en regelgeving.

Via wet- en regelgeving wordt de infrastructuur aangelegd die het gedrag van burgers beïnvloedt en de interactie tussen burgers, bedrijven en publieke instellingen reguleert. Volgens de visie van de overheid roept de kwaliteit van de wetgeving veel aandacht op, omdat iedereen op de ene of de andere manier te maken heeft met wetgeving (Ministerie van Justitie, 2000). Slechte wetgeving kan door de burgers, bedrijven en administratieve instellingen als onterecht of lastig worden ervaren. Dit kan leiden tot verzet tegen de overheid. Evaluaties spelen een grote rol in het bijsturen van inadequate, ongeschikte of ingewikkelde wetgeving. Hierdoor dragen evaluaties bij aan het versoepelen van de relatie tussen burgers, bedrijven en overheid, en tussen verschillen overheidslagen. Deze studie is maatschappelijk relevant omdat deze bijdraagt aan het ontwikkelen van *evidence-based* wetgeving en beleid.

1.3.3 Wetenschappelijke relevantie

Dit onderzoek draagt bij aan de groeiende literatuur over evaluatiecapaciteit bij publieke instellingen. Dit onderwerp heeft in de afgelopen vijf jaren vooral in de internationale omgeving veel aandacht ontvangen. In landen zoals Canada, Australië, Groot-Brittannië en Verenigde Staten zijn al modellen ontwikkeld met factoren die van invloed zouden zijn op de evaluatiecapaciteit. Deze modellen zijn vervolgens getoetst in de praktijk door de verschillende voorwaarden voor evaluatiecapaciteit bij publieke instellingen te onderzoeken. Hoewel Nederland een vooraanstaande positie bezet als het om het aantal uitgevoerde evaluatieonderzoeken gaat, is er echter weinig bekend over de mate van evaluatiecapaciteit van de Nederlandse ministeries. Om dit tekort te ondervangen, stelt deze studie een conceptueel model van evaluatiecapaciteit op – gebaseerd op modellen die in de internationale wetenschappelijke literatuur zijn ontwikkeld – op basis waarvan het vermogen van de Nederlandse ministeries om (wets)evaluaties te verrichten wordt geëvalueerd.

1.4 Voorbeschouwing theoretisch kader

Om het theoretische gedeelte van de vraagstelling – deelvragen twee, drie en vier – te beantwoorden, worden wetenschappelijke inzichten in de relatie tussen evaluatiecapaciteit en de kwaliteit van evaluatieonderzoek besproken. Vervolgens wordt met behulp van nationale en internationale wetenschappelijke literatuur het begrip “evaluatiecapaciteit” gedefinieerd en afgebakend. Daarnaast zullen we een overzicht geven van dimensies of aspecten die gebruikt worden in de wetenschappelijke literatuur om de mate van evaluatiecapaciteit bij publieke organisaties in te schatten. Men kan daarbij denken aan de aan- of afwezigheid van een aparte evaluatieafdeling, zoals kennisdirecties of strategiedirecties die verantwoordelijk zijn voor het evaluatietraject vanaf de aanvraag en begeleiding van evaluaties tot de beoordeling, het gebruik en de verspreiding van de evaluatiebevindingen. Men kan verder denken aan de aan- of afwezigheid van protocollen, handleidingen, procedures, leidraden of toetsingskaders die het evaluatieonderzoek helpen standaardiseren en uniformiseren om tot vergelijkbare bevindingen te komen die vervolgens in een systematisch aan elkaar kunnen worden gekoppeld om conclusies met een bredere toepassing te trekken. Naast deze structurele voorzieningen zullen we ook informatie verzamelen over incidentele aspecten van evaluatiecapaciteit, zoals de verhouding tussen opdrachtgever (evaluatieaanvrager), eventuele intermediaire partijen (onderzoeksinstituten van de overheid) en opdrachtnemers (uitvoerders van evaluatieonderzoeken zoals universiteiten, onderzoeksbureaus, enz.). Hieronder geven wij een voorbeschouwing in de wetenschappelijke theorieën die in hoofdstuk drie worden gebruikt om de evaluatiecapaciteit te definiëren, de voorwaarden hiervoor te identificeren en de verhouding tussen evaluatiecapaciteit en de evaluatiekwaliteit toe te lichten.

Van Aeken (2011: 63) betoogt dat de institutionalisering en formalisering van evaluatie een aantal voordelen heeft voor de borging van de kwaliteit van evaluatie. Een organisatie met een vastgestelde evaluatieplanning en -programmering kan ten eerste voorkomen dat de evaluatie strategisch door partijen in eigenbelang wordt gebruikt. In de afwezigheid van vastgestelde evaluatieplanning en -programma kunnen partijen de doelstellingen van evaluatie, het tijdstip van de evaluatieaanvraag en het gebruik van resultaten makkelijker manipuleren. Landen met een meerpartijenstelsel zoals Nederland zouden het meest gevoelig zijn voor het gebruik van deze strategie, genoemd “submarine evaluation” (Van Aeken, 2011). In een geïnstitutionaliseerde organisatie staat al vastgesteld bij wie het initiatief voor wetsevaluaties ligt, wie de evaluatie mag

uitvoeren en hoe de resultaten worden gevalideerd. Ten tweede zou een duidelijk plan voor evaluatie, een goede infrastructuur en uitgebreide voorzieningen de kans vergroten dat de kwaliteit van evaluatie toeneemt. Zo kan ook aandacht worden besteed aan het opleiden van onderzoekers, de verspreiding van kennis en de ontwikkeling van leidraden, handvatten voor evaluatiepraktijk, standaarden en *best practices*. Een goed ontwikkelde infrastructuur voor evaluatie vergroot de kans dat de evaluaties op een adequate manier worden uitgevoerd en dat men waardevolle kennis ontwikkelt, deze in de besluitvorming gebruikt en met andere publieke instellingen deelt, die hier iets van kunnen leren.

Volgens Cooksy (2012) wordt de kwaliteit van evaluaties voor een deel bepaald door drie met elkaar samenhangende factoren: de *competentie* van de evaluatieonderzoekers, de *context* en *omgeving* waarin de evaluatie plaatsvindt en de mate en de aard van de ondersteuning die de *evaluatiegemeenschap* geeft.

In een studie naar de verschillen tussen instellingen binnen de Canadese overheid, ontwikkelen en valideren Bourgeois en Cousins (2013) een construct voor het meten van de evaluatiecapaciteit. Dit construct houdt zes dimensies in, opgesplitst in twee onderdelen – het *vermogen om evaluatie te doen* en het *vermogen om evaluatie te gebruiken*. Het construct geeft volgens de onderzoekers een compleet beeld van de mate waarin een overheidsinstelling in staat is om evaluatieonderzoek uit te voeren en de resultaten hiervan te gebruiken en te verspreiden.

De eerste dimensie van evaluatiecapaciteit verwijst naar het vermogen om evaluaties te plannen en uit te voeren. Daaronder vallen volgens de onderzoekers drie aspecten: individuele middelen, organisatorische middelen, evaluatieplanning en activiteiten.

De tweede dimensie van evaluatiecapaciteit geeft de mate aan waarin de organisatie in staat is uitkomsten uit de evaluatieonderzoeken te gebruiken, terug te koppelen naar de besluitvorming en te integreren in het proces van het maken van beleid en wetgeving. De betrokkenheid van de stakeholders en het management, de integratie van evaluatie in het proces van besluitvorming en de leerwaarde zijn de drie dimensies die volgens Bourgeois en Cousins (2013) een compleet beeld geven van de leercapaciteit van overheidsinstellingen en dus over het vermogen om de kennis en inzichten uit wetsevaluaties voor beter beleid en wetgeving in te zetten. De evaluatiecapaciteit van de organisatie hangt af van de mate waarin de twee dimensies – het vermogen om evaluatie te doen en het vermogen om evaluatie te gebruiken, met daaronder de zes onderscheiden aspecten – ontwikkeld zijn.

Om de evaluatiecapaciteit te kunnen inschatten, zijn een aantal criteria opgesteld in het theoretisch kader (hoofdstuk 3). Op basis van de inzichten uit de wetenschappelijke literatuur. Deze criteria aan de hand waarvan de evaluatiecapaciteit in dit onderzoek wordt geëvalueerd zijn voor een groot deel al in de praktijk gebruikt door verschillende auteurs voor het meten van de evaluatiecapaciteit bij publieke instellingen (Bourgeois & Cousins, 2013; Cooksy, 2012; Nielsen et al., 2011; Preskill & Boyle, 2008). De bijdrage van dit onderzoek is dat de selectie van de criteria rekening houdt met omstandigheden in de Nederlandse context. Zo zijn bijvoorbeeld enkele criteria weggelaten, omdat ze de interne groei- en leercapaciteit van een organisatie meten, terwijl ons focus helemaal niet het groeien van organisatie is, maar de kwaliteit van evaluaties. Verder is er al bekend dat veel evaluatie die in opdracht van ministeries worden gedaan, extern worden gericht door onderzoeksbureaus en academische teams. Andere criteria die apart in de wetenschappelijke modellen stand hebben wij samengevoegd omdat ze betrekking hadden op dezelfde dimensie en omdat niet nodig was, anders dan in onderzoeken naar leerorganisaties, om criteria voor elk onderdeel van dimensies te gebruiken.

1.5 Voorbeschouwing methodologisch kader

Voor het beantwoorden van deelvragen zes, zeven en acht zal de huidige toestand van de evaluatiecapaciteit bij een aantal geselecteerde ministeries en de recente ontwikkelingen hiervan worden onderzocht. Om een beeld te krijgen van de evaluatiecapaciteit zal een kwalitatief onderzoek worden verricht waarbij twee methoden van dataverzameling worden gebruikt. Ten eerste zal een breed scala aan documenten worden geraadpleegd en geanalyseerd: websites van de ministeries, organogrammen, jaarverslagen, agenda's, bestaande evaluatieprogramma's, protocollen en richtlijnen voor de uitvoering van evaluatie. Het is te verwachten dat deze documenten rijke informatie bevatten over hoe bereid ministeries zijn om (wets)evaluaties te verrichten. Om zicht te krijgen op de verschillen tussen ministeries op het gebied van evaluatie zullen we semigestructureerde interviews houden met beleidsmakers en opstellers van evaluatieaanvragen. Het doel hiervan is om extra informatie te verzamelen die niet beschikbaar is via (openbare) documenten. Het is te verwachten dat niet alle departementen met gestandaardiseerde documenten en uitgebreide programmateksten werken. Om toch zoveel mogelijk erachter te komen hoe ze met evaluatieverplichtingen omgaan, zijn de interviews van groot belang.

Het voordeel van de kwalitatieve strategie is dat zij inzicht geeft in complexe verschijnselen en concurrerende verklaringen, waardoor de “objectieve” verklaringen die verzameld worden via de methode van documentenanalyse, aangevuld en verdiept worden door semigestructureerde gespreken met sleutelinformanten. We zullen de respondenten naar hun concrete ervaringen en belevingen van het evaluatieproces doorvragen en deze terugkoppelen naar de verklaringen die besproken worden in de nationale en internationale wetenschappelijke literatuur over de randvoorwaarden voor evaluatiecapaciteit van organisaties die betrokken zijn bij de evaluatie van wetten en beleid. De kwalitatieve onderzoeksstrategie waarbij gebruik wordt gemaakt van triangulatie zal een meer genuanceerd beeld geven dan het toekennen van scores aan indicatoren via de kwantitatieve methode.

Het doel is de evaluatiecapaciteit bij de ministeries in kaart te brengen en de verschillen hiertussen te verkennen en te evalueren aan de hand van een ideaal beeld van hoe de inrichting van de evaluatiecapaciteit eruit zou moeten zien. De selectie van de respondenten zal vervolgens rekening houden met deze verschillen tussen ministeries.

Mogelijke respondenten zijn deskundigen die werkzaam zijn bij Directies Wetgeving van het ministerie van Veiligheid en Justitie, bij de kennis- of strategiedirecties van verschillende ministeries, onderzoekers bij onderzoeksinstituten (bijv. ZonMw, RIVM, WODC, Alterra, SCP), een beperkt aantal private onderzoeksbureaus met ruime ervaring in wetsevaluatie (bijv. Regioplan, NIVEL, TNO, Panteia) en universiteiten/onderzoekscholen (bijv. binnen de RUG, UU, UL, VU, RU). In de selectie van respondenten zullen we zoveel mogelijke rekening houden met verschillen in de evaluatiecapaciteit tussen departementen zoals gebleken uit de documentenanalyse.

1.6 Leeswijzer

Het beleidskader (hoofdstuk 2) beantwoordt de eerste deelvraag door het huidige beleid op het gebied van evaluatie, wetsevaluatie en wetgevingskwaliteit evenals de ontwikkelingen in de rol de en doelen van evaluatieonderzoek te schetsen. Om deelvragen twee, drie en vier te beantwoorden, bakent het theoretisch kader (hoofdstuk 3) het begrip “evaluatiecapaciteit” met behulp van inzichten uit de wetenschappelijke literatuur af. Hetzelfde hoofdstuk onderscheidt uit wetenschappelijke modellen de meest relevante voorwaarden voor evaluatiecapaciteit bij publieke organisaties. Deze worden gebruikt in het empirisch deel van het onderzoek als standaard voor bepalen van de mate van evaluatiecapaciteit bij de Nederlandse ministeries. Daarna licht

methodologisch kader (hoofdstuk 4) de onderzoeksstrategie, de methoden van dataverzameling, de operationalisatie van variabelen en methodologische aspecten van validiteit en betrouwbaarheid toe. In hoofdstuk vijf worden de data geanalyseerd. Hier wordt antwoord gegeven op deelvragen zes tot en met acht. Hoofdstuk zes vat de bevindingen samen en beantwoordt de centrale vraag van het onderzoek. Daarnaast reflecteert hoofdstuk zes terug en formuleert aanbevelingen voor de opdrachtgever en voor vervolgonderzoek.

Hoofdstuk 2: Beleidskader

2.1 Inleiding

Dit hoofdstuk bespreekt de ontwikkelingen in het Nederlandse evaluatie- en wetsevaluatiebeleid vanaf 1970 tot heden en beantwoordt de eerste deelvraag. Achtereenvolgens zullen het huidige beleid op het gebied van evaluatie, wetsevaluatie en wetgevingskwaliteit alsmede de ontwikkelingen in de rol en de doelen van evaluatieonderzoek de revue passeren. Paragraaf 2.2 schetst de ontwikkelingen in de visie van de Nederlandse overheid op evaluatie vanaf de jaren zeventig van de vorige eeuw. Vervolgens komen de hoofdlijnen van het huidige beleid aan de orde (paragraaf 2.3). Aansluitend licht paragraaf 2.4 de functies van evaluatie. Paragraaf 2.5 bespreekt de tegenwoordig meest gebruikte vormen van evaluatie bij de rijksoverheid kort toe. Vervolgens komt in paragraaf 2.6 het beleid over de wetgevingskwaliteit aan bod. Daarna gaat paragraaf 2.7 specifiek op het beleid rondom wetsevaluaties in. Het hoofdstuk sluit met een samenvattend paragraaf (2.8).

2.2 Ontwikkelingen in het Nederlandse overkoepelende evaluatiebeleid

Leeuw (2009), Bemelmans-Videc (2002) en Bemelmans-Videc et al. (1990) onderscheiden vier fasen in de ontwikkeling van het evaluatiebeleid in Nederland vanaf 1970 tot jaren 2000, die we schetsen in de volgende vier deelparagrafen.

2.2.1 Evaluatie in de jaren 1970

De eerste fase in de ontwikkeling van het evaluatiebeleid was in de jaren zeventig. Toen was er voor de eerste keer aandacht voor het belang van beleidsanalyse in het besturen van het land. Het ministerie van Financiën richtte de interdepartementale Commissie voor de Ontwikkeling van Beleidsanalyse (COBA) op, waarbij alle departementen samen met Centraal Planbureau (CPB) en Centraal Bureau voor de Statistiek zitting hadden. COBA had als taak het geven van advies en het rapporteren aan de Rijksoverheid inzake beleidsanalyse, het bevorderen van de toepassing van beleidsanalyse bij departementen en het ontwikkelen van methoden om de beleidsanalyse vooruit te helpen (Bemelmans-Videc et al., 1990). De oprichting van COBA spiegelde de behoefte weer

om de overheidsprogramma's onder de loep te nemen. Dit had vooral te maken met de wens van het Kabinet om op beleid te bezuinigingen.

Terwijl er steun kwam voor beleidsonderzoek en -analyse, was de aandacht voor evaluatie beperkt. Informele evaluaties hielpen slecht bij het beheren van de overheidsuitgaven en bij de herverdeling van het budget (Bemelmans-Videc, 2002). Een duidelijk vastgesteld evaluatiebeleid ontbrak nog in de jaren 1970 (Leeuw, 2009; Bemelmans-Videc et al., 1990). De overheid had toen nog geen beleid of visie op evaluatie. Het onderwerp evaluatie ontbrak volledig in de Comptabiliteitswet van 1976. Slechts twee ministeries, het ministerie van Onderwijs en Wetenschap en het ministerie van Beleidswetenschap, werkten op dat moment met informele regels voor ex ante studies in de voorbereidingsfase van het maken van beleid. Er kwam vaak veel weerstand vanuit ministeries bij de roep van COBA om programmadoestellingen vast te stellen en prestatie-indicatoren te ontwikkelen. De roep om efficiëntie en effectiviteit van het beleid aan te tonen had nog geen prioriteit. Een methodologisch probleem bij de meting van de effecten van het beleid was de kloof tussen de brede en vage beleidsdoelstelling (die tot stand kwamen via consensus in de politieke arena) en het meetbaar maken en operationaliseren van deze beleidsdoelstellingen.

2.2.2 Evaluatie in de jaren 1980

De tweede fase in de ontwikkeling van het evaluatiebeleid dekt de jaren 1980. Hoewel een formeel beleidskader voor evaluatie nog steeds ontbrak, zagen de jaren 1980 een groei in de informele manier van evaluatieonderzoek. Er waren volgens Leeuw (2009) vier aanleidingen voor de groei van evaluatieonderzoek in jaren 1980 en de roep om evaluatieonderzoek een formele en systematische plek te geven in het vaststellen van beleidsdoelstellingen.

Ten eerste uitte het Parlement zich kritisch over het kennisgebrek bij ministeries over de kosten en de effecten van het beleid. Ten tweede speelde de Algemene Rekenkamer een vooruitstrevende rol in het stimuleren van het evaluatieonderzoek. Zo voerde de Algemene Rekenkamer enkele evaluatieonderzoeken waarbij ministeries met elkaar werden vergeleken op het gebruik van evaluatie en hun kennis over de efficiëntie en de effectiviteit van het gevoerde beleid in kaart werd gebracht (Leeuw, 2009).

Ten derde leidde de groei van New Public Management tot de *framing* van het beleid in termen van efficiëntie, legitimatie en transparantie over de overheidsuitgaven inzake beleid. Als

gevolg hiervan werd een nieuwe evaluatievorm, de ‘heroverwegingsoperaties’ vanaf 1981 ingevoerd. Deze heroverwegingsoperaties waren studies bedoeld om de efficiëntie van het beleid te beoordelen. Waren de doelstellingen bereikt en hoe efficiënt was de doelbereiking? Elke heroverwegingsstudie moest tevens aangeven hoe het beleid eruit zou zien bij een bezuiniging van 20 procent op financiering (Bemelmans-Videc et al., 1990; Leeuw, 2009). Deze vraag is sinds 2014/2015 opgenomen in beleidsdoorlichtingen, die vanaf 2006 voor alle ministeries verplicht zijn. In paragraaf 2.3 die de huidige inrichting van evaluatie in Nederland bespreekt, wordt hierop verder ingegaan.

De vierde aanleiding voor de groei van evaluatieonderzoek is volgens Leeuw (2009) de Nederlandse administratieve cultuur. Kenmerkend voor deze is de openheid tot innovatie, data-gerichte besluitvorming en de opvatting dat evaluatie bedoeld is om lessen uit praktijk te trekken en uit fouten te leren.

2.2.3 Evaluatie in de jaren 1990

De derde fase van het evaluatieonderzoek zag de ontwikkeling van het *formele* evaluatiebeleid in Nederland. Evaluatieonderzoek werd een activiteit bedreven door professionals. De relevantie van evaluatieonderzoek drong zowel tot de overheid als bij de maatschappij door (Bemelmans-Videc, 2002). In de jaren 1990 werd de infrastructuur voor evaluatie op overheidsniveau aangelegd. De Algemene Rekenkamer inventariseerde voor de eerste keer de institutionalisering van evaluatie bij rijksdepartementen. Het rapport gaf een beeld over de structuren en de processen van evaluatie zoals de inrichting van *evaluation units*, de werkwijzen hiervan, de kosten van evaluatie, het gebruik van evaluatiebevindingen binnen departementen, de communicatie van bevindingen naar beleidsmakers en politici (Leeuw, 2009; Bemelmans-Videc, 2002).

Als gevolg van de actieve inzet van de Algemene Rekenkamer kwam een visie tot stand over hoe evaluatie opgesteld en uitgevoerd moest worden. De belangstelling voor evaluatie groeide en de prestatie-indicatoren bleken noodzakelijk te zijn voor het opstellen van goed beleid. Evaluatie werd door de overheid beschouwd als onderdeel van de beleidscyclus en in dienst van het beleidsproces. Evaluatie was onder anderen bedoeld als informatie op basis waarvan meerdere alternatieven in het besluitvormingsproces tegen elkaar werden afgewogen. Het evaluatiesysteem wordt op dat moment ‘gedecentraliseerd’, met evaluatieonderzoek verdeeld over verschillende ministeries, beleidsdomeinen en thema’s. Elk ministerie was verantwoordelijk zowel voor de

ontwikkeling van benodigde methoden, strategieën en standaarden voor evaluatie als voor het gebruik van evaluatiebevindingen in het opstellen van toekomstig beleid. Hierdoor verschilden ministeries onderling in aanpak. Er moest tevens meer aandacht voor evaluatie van wet- en regelgeving komen (Leeuw, 2009). Evaluatie heeft echter in de jaren 1990 nooit de centrale plek gekregen die in deze visie werd geschetst.

2.2.4 Evaluatie in het eerste decennium van de jaren 2000

Het evaluatiebeleid werd geformaliseerd en gestandaardiseerd in het begin van de jaren 2000. In 2001 werd een nieuwe *Comptabiliteitswet* opgesteld, waarin evaluatie voor de eerste keer plaats inneemt. Het belang van evaluatie is daardoor formeel erkend. In aanvulling werden de hoofdpunten over evaluatie in de Regeling prestatiegegevens en evaluatieonderzoek Rijksoverheid uit 2001 uitgewerkt. Deze zijn vervolgens in 2004, 2006 en 2012 aangepast. Het huidige document *Regeling periodiek evaluatieonderzoek* is in 2014 voor het laatst aangepast.

Het oorspronkelijke document uit 2001 legde de nadruk op het belang van het gebruik van verschillende evaluatievormen. De frequentie van de evaluatieactiviteiten moest explicieter worden en er moest meer aandacht worden geschonken aan zowel de methodologische aspecten van evaluatie als aan het gebruik en de verspreiding van de evaluatiebevindingen. Verder moest het evaluatieonderzoek transparanter worden georganiseerd.

Het aantal programma-evaluaties en gerelateerde activiteiten zoals inspectie, toezicht en *ad hoc* evaluatie naar aanleiding van crises en rampen nam spoedig toe. Met de oprichting van nieuwe instellingen zoals gemeentelijke en provinciale rekenkamers en de Beroepsvereniging VIDE in 2002 en de druk gezet door de Europese Unie, ging het evaluatieonderzoek uitdijen, waaronder ook de evaluatie van de naleving van de regels, voorschriften en regelingen. Evaluatie werd langzamerhand onderdeel van bureaucratische processen. Richtlijnen, protocollen en indicatoren werden ontwikkeld over hoe evaluaties op te stellen en uit te voeren (zie RPE, die herhaaldelijk in 2004, 2006, 2012 en 2014 werd herzien; de Comptabiliteitswet, 2001; richtlijnen voor het opstellen van beleidsdoorlichting).

2.3 Het huidige evaluatiebeleid

Tegenwoordig is evaluatie een belangrijk instrument bij het bereiken van “good governance” en voor de controle van de efficiëntie en de effectiviteit van het gevoerde beleid. De

Comptabiliteitswet (artikel 20, lid 2) schrijft voor dat elk ministerie verantwoordelijk is voor het onderzoeken van de effectiviteit van zijn beleid (als daar uitgaven mee gemoeid zijn).

Er bestaat geen overkoepelend evaluatiebeleid dat alle vormen van evaluatie afdekt zoals ex ante evaluatie, ex post effectenevaluaties, beleidsdoorlichtingen, monitoring en inspecties. Elk ministerie geeft in zijn jaarlijkse begroting de planning van het evaluatieonderzoek voor het komende jaar aan. Om financiering te krijgen zou in principe ieder programma moeten steunen op een ex ante evaluatie die inzicht geeft in de plausibiliteit van de geplande interventies. Ministeries moeten de evaluatiebevindingen over het gevoerde beleid in het jaarverslag rapporteren. Verdere financiering van programma's hangt van de evaluatiebevindingen af. Daarnaast komt er steeds meer aandacht voor het gebruik van adequate methodologische criteria in het uitvoeren van evaluaties (Davidson, 2005; Leeuw, 2009).

Het evaluatieonderzoek in Nederland is chronologisch gezien steeds systematischer geworden. Het aantal evaluaties groeide voortdurend, evaluatie is tegenwoordig geaccepteerd als een moment van de beleidscyclus. Daarnaast groeide ook de vraag naar het ontwikkelen van standaarden, criteria en toetsingskaders voor evaluaties. Bovendien gaf men meer aandacht voor de methodologische vereisten van evaluatieonderzoek en organisatorische inrichting van evaluatiefunctie bij de overheid (Leeuw, 2009).

Leeuw (2009) concludeerde dat de overheid tegenwoordig transparanter en meer toerekenbaar handelt inzake evaluatie in vergelijking met de vorige decennia, maar dat er nog veel ruimte voor verbetering is, vooral wat betreft de ontwikkeling van een overkoepelend evaluatiebeleid en de inzet van experts bij evaluatieonderzoek. Leeuw (2009) noemt enkele sterke punten en verbeterpunten van de Nederlandse inrichting van evaluatieonderzoek. De doelen van evaluaties zijn vaak te breed en open voor interpretatie. Er zijn vaak weinig consequenties bij het niet naleven van evaluatieverplichtingen. Meta-evaluaties, synthese-onderzoeken gebaseerd op losse evaluaties, hebben in de afgelopen jaren meer aandacht gekregen.² Er worden echter nog steeds weinig meta-evaluaties gedaan. Verder stelt de overheid de noodzakelijke middelen voor evaluatieactiviteiten beschikbaar en bestaan er een aantal onderzoeksinstituten, zoals WODC van het ministerie van VenJ en ZonMw van het ministerie van VWS, die evaluatie namens ministeries opstellen en aanvragen of zelf uitvoeren. Verbeterpunten zijn het management, de methodologische aspecten en het gebruik van evaluaties. Evaluatiebevindingen zouden meer

² De verplichte beleidsdoorlichtingen zijn in feite óók metastudies, d.w.z. syntheses van eerder evaluatiewerk.

impact kunnen hebben bij het vaststellen van toekomstig beleid en wetgeving. Ministeries zouden meer rekening moeten houden met evaluatiebevindingen bij hun beslissingen ten aanzien van welke beleidsprogramma's financiële steun zouden krijgen.

2.4 De drie functies van evaluatie

De vraag hoe evaluaties het beste kunnen worden uitgevoerd en wat een goede evaluatie is hangt nauw samen met de visie op evaluatie: welke doeleinden evaluatie moet dienen volgens de organisatie en welke definitie van evaluatie wordt binnen de organisatie gehanteerd. De visie op evaluatie - hoe naar evaluatie binnen de organisatie wordt gekeken – bepaalt voor een groot deel ook het evaluatiestelsel, d.w.z. wie verantwoordelijk is voor het uitvoeren van evaluaties, welke voorzieningen voor evaluaties worden ontwikkeld en wat wordt gedaan met de evaluatiebevindingen.

Vedung (1997), een klassieker van programmaevaluaties, onderscheidt drie doeleinden of functies van evaluatie: (1) het verantwoorden en legitimeren van beleidsinterventies, (2) het verbeteren van interventies (3) het vergaren van kennis en informatie. Evaluaties kunnen worden opgesteld met oog op minstens één van deze doeleinden. De keuze voor een van deze doeleinden of een combinatie daarvan is bepalend voor hoe het evaluatiestelsel in elkaar zit, hoe evaluaties concreet worden uitgevoerd en hoe evaluatiebevindingen naar andere fases in de beleidscyclus doorwerken.

Allereerst worden evaluatie vaak uitgevoerd om te toetsen of het beleid, de wet of de interventie heeft gewerkt zoals gepland en bedoeld. De achterliggende gedachte hiervan is dat de overheid verantwoording moet afleggen voor de uitgaven voor (maatschappelijke) interventies. De vraag die de evaluaties moet beantwoorden is of de (financiële) middelen op een adequate en zuinige manier zijn besteed en of de interventies de beoogde doelen hebben bereikt. Een mogelijk probleem met de eerste doelstelling van evaluatie, het verantwoorden, is dat het ministerie dat het beleid of de wet heeft ontwikkeld ook de opdrachtgever is van de evaluatie en vaak ook verantwoordelijk is voor de besluitvorming en de interventie. Om de objectiviteit en transparantie van het evaluatieonderzoek te waarborgen, worden evaluaties in opdracht van ministeries vaak uitbesteed en extern uitgevoerd door onafhankelijke teams zoals academische onderzoekers of onderzoeksbureaus. Als een organisatie de evaluatie veelal als verantwoordingsinstrument ziet, wordt de nadruk gelegd op effectiviteitevaluaties, zoals evaluaties naar doelmatigheid,

doeltreffendheid en naar de (gewenste en ongewenste) effecten van de wet of het beleid. In deze visie zijn andere evaluatievormen zoals procesevaluaties of evaluaties naar onderdelen (interventies) van de wet of beleid onvoldoende om de verantwoordingsfunctie in te vullen. De voorkeur voor effectiviteitsevaluaties is een gevolg van de assumptie dat de evaluaties de vraag moeten beantwoorden of de wet of het beleid hebben gewerkt en of de geconstateerde effecten aan de interventies kunnen worden toegeschreven.

De tweede en de derde functie van evaluatie zijn volgens Vedung (1997) het verbeteren van de gevoerde interventie en het vergaren van kennis en informatie. Andere onderzoekers zien deze twee aspecten als onderdelen van de leerfunctie van evaluatie (Leeuw et al., 1994). In deze visie worden evaluaties veelal gedaan om na te gaan of er een kloof is tussen de interventie zoals in theorie bedacht en de interventie zoals deze daadwerkelijk heeft gewerkt. Omdat maatschappelijke problemen vaak heel complex zijn en de interventie in een dynamische omgeving plaatsvindt, is de werking van de interventie volgens het plan niet vanzelfsprekend. Het doel van evaluaties is in dit perspectief begrijpen en verklaren waarom de interventie anders heeft gepakt. Evaluatiebevindingen dienen als aanleiding voor de aanpassing of de opheffing van de interventie. De hoofdvraag die een evaluatie in dit geval stelt is: hoe kan het beleid of de wet worden verbeterd? (Vedung 1997: 108). Vooral beleidsambtenaren en juristen die de interventie hebben bedacht, hebben veel baat aan evaluaties die gericht zijn op het leerproces. Deze evaluatie is ook bekend onder de naam “formative evaluation” (Vedung 1997: 109). Organisaties die veel waarde aan de leerfunctie van evaluaties hechten, zijn open tot een diversiteit aan evaluatievormen. Het is de verwachting dat een organisatie die evaluaties ziet als een leerinstrument, meer bereid zijn te experimenteren. Een ander verschil is dat het perspectief dat nadruk legt op het *leren* uit evaluaties beter rekening kan houden met de tijdscomponent van evaluatie. De werking van een beleidsinterventie of wet heeft een dynamisch karakter. Dezelfde interventie kan op verschillende wijzen uitpakken afhankelijk van externe, vaak veranderende factoren. Dezelfde factoren kunnen in verschillende omstandigheden remmende of bevorderende invloeden op de werking van de interventie hebben. Het succes en het falen van interventies kan tijdelijk zijn en kan het gevolg zijn van bevorderende of remmende externe factoren. Daarom is een lange termijn perspectief van belang.

Evaluatie is in principe het moment waarop kritisch wordt gereflecteerd op de onderliggende theorie van het beleid of van de wet. Evaluaties bieden inzicht in de knelpunten in

de implementatie van de wet en discrepanties tussen de werking van de interventies in theorie en de werking van de interventie in de praktijk. Bovendien spreken evaluaties de ambitie uit om lessen te trekken over wat wel of niet werkt en waarom. Deze uitkomsten van het evaluatieonderzoek bieden handvatten aan juristen en beleidsmakers in het formuleren van toekomstige wetgeving en beleid. Beleid verantwoorden, het verbeteren van het beleid en de wet- en regelgeving, en kennisvergaring zijn de drie doelen van evaluatieonderzoek (Ministerie van Financiën, 2015; Veerman, 2007; Vedung, 1997).

Het domein van evaluatieonderzoek is zeer divers en rijk (Klein Haarhuis & Niemeijer, 2009; Bovens, 2008). Naast de verscheidenheid aan doeleinden waarvoor een evaluatie kan worden gebruikt, neemt ook de diversiteit aan evaluatievormen toe. Men kan daarbij denken aan *ex ante*, *ex durante* en *ex post* evaluaties maar ook aan syntheseonderzoek van evaluatie, de zogenaamde meta-evaluaties (Leeuw, 2009: 95). Daarnaast groeit het domein van het evaluatieonderzoek door nieuwe technieken en methoden te ontwikkelen en toe te passen. Tevens neemt de kennis en ervaring van experts voortdurend toe mede als gevolg van de multidisciplinariteit van het evaluatieonderzoek (Cooksy, 2012: 83). De diversiteit in evaluatievormen, de groei van expertise en kennis van professionals alsook de verscheidenheid aan evaluatiemethoden maakt het echter moeilijk om het evaluatieonderzoek te standaardiseren en evaluaties met elkaar te vergelijken. Standaardisering en uniformisering zijn echter heel hard nodig om lessen te trekken die breder inzetbaar zijn, om evaluatieresultaten te kunnen vergelijken en om verbanden te leggen tussen assumpties en onderliggende theorieën van diverse wetten. Vergelijkbaarheid kan de leercapaciteit en de kennisoverdracht en -verspreiding vergroten.

2.5 Vormen van evaluatie

De overheid kan op drie momenten evalueren: voorafgaand aan het opstellen van de wet of beleid (*ex ante* evaluatie), tussentijds – deze evaluatie is bedoeld om het implementatieproces te monitoren (*ex durante* evaluatie) en achteraf – bedoeld om te beoordelen of het doel van de interventies is bereikt en over welke bedoelde en onbedoelde effecten de wet of het beleid heeft veroorzaakt (*ex post* evaluatie) (Ministerie van Financiën, 2015b).

Volgens het ministerie van Financiën zouden departementen zich op dit moment op drie vormen van evaluatie richten: beleidsdoorlichtingen, effectevaluaties en maatschappelijke kosten-batenanalyses (Ministerie van Financiën, 2015a). Deze volgorde is niet chronologisch, maar laat

de voorkeuren van de overheid inzake evaluatieonderzoek zien. Beleidsdoorlichtingen moeten volgens het ministerie van Financiën de primaire focus van evalueren zijn. Evaluatie dient dus vooral het gevoerde beleid te verantwoorden.

Beleidsdoorlichting is een instrument ontwikkeld door het ministerie van Financiën. Zij geven een overzicht van de mate waarin de doelstellingen van het voorgestelde beleid op een groot terrein daadwerkelijk zijn bereikt. Beleidsdoorlichtingen kijken “naar de samenhang tussen verschillende vormen van beleid” (Ministerie van Financiën, 2015a). Zij integreren de kennis die losse evaluatie over de effectiviteit en doelmatigheid van verschillende interventies hebben verzameld. De lossen evaluatie fungeren als onderbouwing voor beleidsdoorlichtingen.

De beleidsdoorlichting is een relatief nieuw instrument. Het wordt sinds 2006 ingezet als controle-instrument om meer transparantie en verantwoording over de begrotingsuitgaven te verkrijgen³ (Regeling Periodiek Evaluatieonderzoek, 2006). Het beleid dat elk ministerie uitvoert moet minstens eens in de zeven jaar worden geëvalueerd. De beleidsdoorlichtingen rapporteren de bevindingen hiervan en koppelen het beleid aan één of meerdere beleidsartikelen uit de rijksbegroting.

Ieder ministerie wordt geacht om de bevindingen uit verschillende effectiviteitsevaluaties samen te vatten in dit syntheseonderzoek. Als syntheseonderzoek steunt de beleidsdoorlichting zoveel mogelijk op al bestaand ex post evaluaties naar de effectiviteit van het beleid of wet. Beleidsdoorlichting heeft een synthesesekarakter. Volgens artikel 3 van de *Regeling Periodiek Evaluatieonderzoek* (RPE) uit 2014 hoort een beleidsdoorlichting informatie over de volgende aspecten te bevatten:

- Een afbakening van de onderzochte beleidsterreinen;
- De motivering voor het beleid en de beoogde doelen die de overheid met het beleid wil bereiken;
- Een beschrijving van het beleidsterrein en de financiële middelen die zijn ingezet;
- Een overzicht van het al bestaand onderzoek naar doeltreffendheid en doelmatigheid en de onderbouwing van de evaluatieprogrammering;
- De effecten van het gevoerde beleid, de doeltreffendheid en doelmatigheid hiervan en de effecten van het beleid op economische groei en regeldruk.

³ De rapporten zijn openbaar op de website van de Rijksbegroting:
www.rijksbegroting.nl/beleidsevaluaties/evaluaties-en-beleidsdoorlichtingen/archief.

- Suggesties voor de verhoging van de doeltreffendheid en doelmatigheid van het gevoerde beleid.
- Een beschrijving van het beleidsalternatief indien een daling van middelen met 20 procent zich voordoet.

Departementen verschillen in de kwaliteit van beleidsdoorlichting. Beleidsdoorlichting is een instrument dat rijksbreed is ingevoerd en door het ministerie van Financiën wordt bewaakt. Het rapport *Tussenevaluatie Beleidsdoorlichting* (Meyenfeldt, Schrijvershof & Wilms, 2008) heeft geconcludeerd dat “op een aantal ministeries het instrument beleidsdoorlichting goed blijkt te zijn opgepakt. Het gaat dan veelal om ministeries waar een cultuur bestaat, waarin betrokkenen alert zijn op de effectiviteit van het ontwikkelde beleid en open staan voor aanpassingen om dat beleid verder te verbeteren. In andere ministeries is die cultuur minder duidelijk aanwezig, met als gevolg dat beleidsdoorlichtingen daar veel meer als een verplicht nummer gezien worden, als een extra taak die niets oplevert.” (Von Meyenfeldt, Schrijvershof & Wilms, 2008: 74).

Effectevaluaties, de tweede vorm van evaluatieonderzoek waarop de overheid nadruk legt, is de evaluatie van de doelmatigheid of de doeltreffendheid van het beleid of wet. De evaluatie van de doeltreffendheid gaat na of en in hoeverre de beleidsdoelstellingen zijn bereikt dankzij de ingezette instrumenten (RPE, 2014). Dat wil zeggen in hoeverre de geconstateerde effecten aan de ingezette instrumenten kunnen worden toegeschreven. Dit fenomeen is bekend in de literatuur als het attributie-probleem (Coglianese, 2012). De evaluatie van de doelmatigheid beoordeelt “de relatie tussen de effecten van het beleid en de kosten van het beleid” (RPE, 2014). De vraag die deze evaluatie moet beantwoorden is of de ingezette middelen opwegen tegen de geconstateerde effecten die met de interventie zijn bereikt.

De derde vorm van evaluatie die genoemd wordt door het ministerie van Financiën op de website van Rijksbegroting is de maatschappelijke kosten-batenanalyse (MKBA). Deze is vooral als ex ante onderzoek ingezet. 16 % van alle ex ante evaluatie die uitgevoerd zijn tussen 2005 en 2013, waren MKBA-studies (Klein Haarhuis et al., 2014). MKBA's toetsen de efficiency van het beleidsvoorstel in termen van de kosten en de baten die vereist zijn om de gewenste effecten te bereiken. Andere vormen van ex ante evaluatie zijn planevaluaties, risicoanalyse, kosten-effectiviteitsanalyses (KEA) en verkenningen (Klein Haarhuis et al., 2014: 4-6). Alleen MKBA is

echter genoemd en kort toegelicht op pagina van de Rijksbegroting. Dit geeft aan dat de overheid voorkeur heeft voor de drie bovengenoemde evaluatievormen.

De Algemene Rekenkamer heeft in 2012 onderzocht hoeveel van het beleid dat tussen 2006 en 2010 door departementen is uitgevoerd om maatschappelijke doelen te bereiken, op effectiviteit is geëvalueerd. In ongeveer 350 uit rond de 1200 verrichte evaluatieonderzoeken is het beleid op effectiviteit beoordeeld. Deze onderzoeken hadden betrekking op 46% van de uitgaven voor beleid. Dit komt neer op € 51 miljard uit € 111 miljard beleidsuitgaven in 2010 (Algemene Rekenkamer, 2012). De andere helft van de beleidsuitgaven is ten dele geëvalueerd op het proces van implementatie en niet op effectiviteit, en deels helemaal niet geëvalueerd. De Algemene Rekenkamer concludeert dat het risico van het niet evalueren van beleid is dat ineffectief beleid blijft doorlopen en publieke middelen verkeerd worden besteed (Algemene Rekenkamer, 2012). Bij dit onderzoek hebben vele ministeries kritische kanttekeningen geplaatst. De Algemene Rekenkamer heeft een aantal (effect)evaluaties niet meegeteld omdat deze niet direct aan de doelstellingen van beleidsartikelen konden worden gekoppeld. Na verdere gesprekken met de ministeries heeft de Algemene Rekenkamer in 2013 een nieuw onderzoek uitgebracht waarin zeven verklaringen worden gegeven voor het niet evalueren van het beleid op effectiviteit. (1) Het is niet duidelijk hoe het beleid op effectiviteit geëvalueerd kan worden. (2) Het is nog te vroeg om de effectiviteit te kunnen vaststellen. (3) Het beleid wordt al door anderen op effectiviteit geëvalueerd. (4) Effectiviteitsonderzoek voegt weinig toe. (5) Het beleid is inmiddels stopgezet. (6) De kosten van effectiviteit zijn te hoog (Algemene Rekenkamer, 2013).

2.6 Wetgevingskwaliteitsbeleid

De overheid heeft niet duidelijke keuzes gemaakt over de functie en de vormen van evaluatie in het breed, maar heeft ook een apart kwaliteitsbeleid ontwikkeld, dat toegespitst is op wetgeving. De zorg voor de kwaliteit van de wetgeving is sinds de jaren tachtig expliciet in het regeringsbeleid genoemd. De wortels van het huidige wetgevingskwaliteitsbeleid liggen in de nota *Zicht op wetgeving* uit 1991 (Ministerie van Justitie, 1991). Deze gaf de visie van de overheid inzake kwaliteit van wetgeving weer. De nota bevatte de hoofdlijnen voor de bevordering van goede juridische en bestuurlijke kwaliteit van wetgeving. Zij schetste de kwaliteitseisen die aan de grondslag liggen van het huidige wetgevingskwaliteitsbeleid: rechtmatigheid, doeltreffendheid en

doelmatigheid, subsidiariteit en evenredigheid, uitvoerbaarheid en handhaafbaarheid, onderlinge afstemming, eenvoud, duidelijkheid en toegankelijkheid (Ministerie van Justitie, 1991).

Tien jaar later verscheen de nota *Wetgevingskwaliteitsbeleid en wetgevingsvisitatie* (2000) als aanscherping en aanvulling op de vorige nota. Deze liet zien dat het toenmalige kabinet waarde hechtte aan dezelfde kwaliteitseisen als in de nota daarvoor. De kwaliteitszorg voor wetgeving bleef ook in de jaren 2000 een aandachtspunt en werd door de overheid gezien als een permanente taak (Ministerie van Justitie, 2000: 4). Vanwege voortdurende veranderingen in maatschappelijke omstandigheden en het brede kwaliteitsbegrip moet het kabinet regelmatig zijn visie over de kwaliteit herzien om aangesloten te blijven bij de praktijk en de problemen die in de maatschappij op ieder moment spelen. Twee maatschappelijke veranderingen, die effect hebben op wat goede wetgeving is volgens de overheid, zijn ten eerste de eisen die vanuit Europa worden gesteld aan wetgeving en ten tweede de razendsnelle ontwikkelingen op het gebied van informatie- en communicatietechnologie (Ministerie van Justitie, 2000: 4). Wetten moeten wegens deze factoren regelmatig worden herzien.

Een ander belangrijk punt in de nota *Wetgevingskwaliteitsbeleid en wetgevingsvisitatie* (2000) was de zorg voor de kwaliteit van wetgeving op alle fasen, op alle niveaus en voor alle betrokken partijen bij het wetgevingsproces. In vergelijking met de nota *Zicht op wetgeving* uit 1991 signaleert deze verschuiving een uitbreiding van het belang van de kwaliteitszorg voor wetgeving naar alle vier fasen van het wetgevingsproces: de voorbereidende fase, de bestuurlijke vaststelling van een wetsvoorstel, de parlementaire behandeling en de terugkoppeling (Ministerie van Justitie, 2000: 14). Daarnaast benadrukte het kabinet in de nota uit 2000 dat wetgeving slechts één van de beleidsinstrumenten was waarover de politiek en het ambtelijke apparaat beschikten om hun plannen te verwezenlijken en veranderingen in de maatschappij te laten doordringen. In het opstellen van het wetgevingskwaliteitsbeleid ging het Kabinet uit van de visie dat de nieuwe wetgeving binnen het samenstel van de al bestaande wettelijke regelingen moet passen. Verder moesten wetgevingsjuristen bij het opstellen van nieuwe wetgeving rekening houden met alle betrokken actoren die een rol spelen in het wetgevingsproces: internationale actoren, actoren bij de centrale overheid, de decentrale overheden zoals provincies, gemeenten, lokale rekenkamers, uitvoerders, handhavers en wetenschappers. Het kabinet benadrukt het belang van samenwerken en het benutten van de kennis die vanuit alle niveaus wordt aangereikt. Op deze manier kan de kwaliteit van de wetgeving worden geborgd.

De jaren 2000 zagen een accentverschuiving van het belang van een evenwichtig zicht op alle kwaliteitscriteria naar het belang van de uitvoerbaarheid van wetgeving (Ministerie van Justitie, 2000: 5). De zorg voor uitvoerbaarheid houdt in dat de nieuwe regelgeving geen grote lasten voor de uitvoering en handhaving mag veroorzaken.

Als gevolg van de ontwikkeling van een stevige visie over het wetgevingskwaliteitsbeleid werden steeds meer aanwijzingen, checklists en toetsen op interdepartementaal niveau opgesteld om wetgevingsjuristen te helpen de juiste afwegingen te maken. Enerzijds ligt de verantwoordelijkheid voor het wetgevingskwaliteitsbeleid bij het ministerie van VenJ, anderzijds is elk ministerie zelf verantwoordelijk voor de organisatie en inrichting van de wetgevingsfunctie. Uit een rapport van de onafhankelijke visitatiecommissie wetgeving uit 2000 blijkt dat er verschillen zijn tussen departementen in de manier en de mate waarin de wetgevingsfunctie is georganiseerd (Ministerie van Justitie, 2000: 6). Dit kan leiden tot een verschil in de kwaliteit van de wetgeving tussen aparte beleidsdomeinen. Om de kwaliteit van wetgeving op peil te houden moet informatie worden verzameld over welke wetten werken, onder welke omstandigheden en waarom. Hieraan dragen wetsevaluaties direct bij (Ministerie van Justitie, 1991).

2.7 Evaluatie van wet- en regelgeving

Deze paragraaf geeft inzicht in de evaluatieaspecten die specifiek voor wet- en regelgeving gelden, zoals beleid dat specifiek op wetgeving is gericht en de institutionele randvoorwaarden die opgesteld zijn om de kwaliteit van wetsevaluaties te waarborgen.

De inrichting van wetsevaluatie komt op grote lijnen overeen met de inrichting van evaluatie in algemene zin. De arrangementen voor wetsevaluaties verschillen niet structureel van degene voor evaluatie breder gezien. De taak om wetten te evalueren is belegd bij dezelfde instanties binnen departementen die zich ook met beleidsevaluaties bezig houden. Departementen hebben geen aparte afdelingen of inrichtingen voor wetsevaluaties en dus ook geen apart beleid voor wetsevaluatie.

Er zijn weinig auteurs (Bussmann, 2010; Veerman, 2007) die speciaal aandacht schenken aan evaluatie van wetgeving afgezet tegen beleidsevaluaties. Degene die dat doen, merken snel op dat de uitspraken over wetsevaluatie in dezelfde mate ook voor beleids- en programmaevaluatie gelden. Zo wijdt Veerman in zijn boek *Over wetgeving* een apart hoofdstuk aan het onderwerp wetsevaluatie, waarin de mogelijke vraagstellingen van wetsevaluatie, de onderzoeksmethoden,

onderzoeksproblemen en het gebruik van de resultaten van het evaluatieonderzoek aan bod komen. Al deze aspecten gelden echter in zekere mate niet alleen voor wetsevaluatie, maar ook voor beleidsevaluaties. Veerman stelt dat “het verschil tussen een wetsevaluatie en een beleidsevaluatie op voorhand niet duidelijk (is). Een wet bestaat uit recht, maar is een voertuig van beleid (...). Zij maakt deel uit van een breed beleid. Wet en beleid zijn niet goed meer te scheiden.” (Veerman, 2007: 228).

Bussmann (2010) merkt op dat de wetenschappelijke literatuur over evaluatiemethodologie vaak het evaluatieonderzoek als een geheel domein ziet zonder aandacht te schenken aan de verschillen tussen evaluatie van beleid, programma's of wetten. De weinige studies die op de specifieke methodologische vereisten van wetsevaluaties ingaan, zijn afkomstig van juristen. Deze laten vaak de sociaalwetenschappelijke aspecten van wetsevaluaties onderbelicht (Bussmann, 2010: 280). Bussmann bespreekt enkele beperkingen of uitdaging voor wetsevaluaties in verschil tot andere vormen van evaluatieonderzoek.

Ten eerste zijn beleidsinterventies beperkt tot een bepaalde periode, terwijl wetten in principe geen eindtermijn bevatten. Ten tweede is de doelgroep van beleidsinterventies afgebakend, terwijl wetten voor de hele bevolking van een land of voor iedereen in een bepaalde categorie gelden. Dit stelt de wetsevaluaties voor de eerste methodologische uitdaging: een experiment met controlegroep is uitgesloten. Onderzoeken waarbij het werken met controlegroepen is uitgesloten, maakt het moeilijk de geconstateerde effecten aan de wet of de interventie toe te schrijven (het zogenoemde attributieprobleem). Het verband tussen de wet en de geconstateerde effecten moet dus op andere manieren worden aangetoond.

Een alternatief hiervoor is het expliciteren en reconstrueren van de onderliggende theorie van de wet. Deze theorie beschrijft hoe de wet dient te werken en via welke mechanismen de gewenste effecten bereikt zullen worden. Wetsevaluatie moeten de assumpties uit de theorie reconstrueren en vervolgens deze vergelijken met wat in de praktijk is waargenomen om vragen te beantwoorden zoals: zijn de gewenste effecten bereikt? Zijn deze bereikt via dezelfde mechanismen als de theorie had aangenomen? Zijn er ook ongewenste effecten na de invoering van de wet? Een ander alternatief om het attributieprobleem op te lossen is door het *counterfactual* verhaal te bedenken over hoe de huidige situatie eruit zou zien, had de wet niet bestaan.

Naast de complexiteit van wetten ziet Delahais (2014) een andere uitdaging voor wetsevaluaties. Beleidsmakers hechten minder waarde aan wetsevaluaties dan aan

beleidsevaluaties en hebben daarom minder belang bij het evalueren van wetten dan van beleid en programma's.

Wetten bevatten vaak een evaluatiebepaling. De evaluatiebepaling is een artikel van de wet dat vaststelt wanneer de wet zal worden geëvalueerd om inzicht te geven in de werking van de wet. Dat gebeurt in de regel vijf jaar na de inwerkingtreding van de wet, hoewel een tussentijdse evaluatie ook mogelijk is in het geval dat de implementatie van de wet anders verloopt dan verwacht of de Tweede Kamer om informatie vraagt op een eerder tijdstip. Een wet kan eenmalig of periodiek worden geëvalueerd (conform aanwijzing 164 voor de regelgeving). Voor wetten die geen evaluatiebepaling bevatten, kan het Parlement, indien noodzakelijk, *ad hoc* opdracht geven voor evaluatie van de wet (Bemelmans-Videc, 2002). Dit laat zien dat de wetsevaluaties zich in een politieke context bevinden en soms ook direct uit de politieke realiteit kunnen ontstaan.

Een evaluatie naar de “werking” van de wet kan op heel verschillende manieren worden aangepakt. Veerman (2007: 229) onderscheidt zes mogelijke onderzoeksvragen naar de werking van de wet, die tevens ook voor beleids- en programmaevaluaties gelden:

- onderzoek naar de doelbereiking (zijn de doelstelling van de wet bereikt?);
- onderzoek naar de effecten van de wet (bedoelde en onbedoelde, gewenste en ongewenste effecten);
- onderzoek naar de doelmatigheid van de wet (ook bekend als kosteneffectiviteitsonderzoek);
- onderzoek naar het uitvoeringsproces;
- onderzoek naar de knelpunten in de toepassing en invoering van de wet;
- onderzoek naar de gelding van de wet (gericht op de juridische aspecten van de wet).

De verplichting tot het evalueren van belangrijke wetten en het rapporteren van evaluatiebevindingen heeft soms geleid tot het doen van symbolische, routineachtige wetsevaluaties, waarvan weinig wordt geleerd over de effectiviteit van de wet. Voor wetten die een evaluatieverplichting bevatten, kan het proces van evalueren een ritueel worden. Zo concludeert Dahler-Larsen (2013) dat de druk om te evalueren in veel maatschappijen veel sterker is dan de inspanning om goede evaluaties te verrichten.

Schouten (2013) concludeert in een rapport over evaluatie uit 2013, geschreven in opdracht van de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) dat de kwaliteit van

wetsevaluaties geborgd kan worden door richtlijnen, evaluatieplanning en -programmering van evaluatie op te stellen. Verder kunnen mechanismen in gang worden gezet om het gebruik en de verspreiding van evaluatiebevindingen te verhogen. Het *Integraal Afwegingskader voor beleid en regelgeving* (IAK), dat sinds 2010 in Nederland gebruikt wordt, is zo'n set criteria, bedoeld om de kwaliteitsborging van wetgeving te bevorderen.⁴ Deze omvat zeven vragen en achttien kwaliteitseisen, die beleidsmakers en wetgevingsjuristen moeten toegepassen bij de voorbereiding van nieuw beleid en regelgeving. Ieder nieuw voorstel van beleid en regelgeving moet worden begeleid door een toelichting die onder anderen antwoord geeft op de zeven IAK-vragen. IAK is dus een toets dat bij de voorbereiding van zowel beleid als van regelgeving wordt toegepast. Voor ex post evaluatie (van beleid en regelgeving) bestaat er geen vergelijkbaar toetsingskader.

Het opstellen van wetgeving vindt in een gedecentraliseerd systeem plaats. Dat geldt niet alleen voor de wetgeving, maar al sinds de jaren 1990 voor het evaluatiebeleid in zijn geheel (zie paragraaf 2.2.3). Dat wil zeggen dat ieder ministerie verantwoordelijk is voor de wetgeving op zijn eigen terreinen. Het ministerie van VenJ draagt zorg voor het rijksbrede evaluatiebeleid van wet- en regelgeving. Het ministerie van VenJ ondersteunt via het Kenniscentrum Wetgeving en Juridische zaken (KWJZ) de directies Wetgeving van de afzonderlijke departementen bij het opstellen van wetgeving op hun eigen terreinen (KWJZ, 2015a).

Het KWJZ heeft twee hoofddoelen. Enerzijds verzamelt het KWJZ alle wetsevaluaties die de afgelopen jaren in opdracht van alle ministeries worden verricht. Het KWJZ vergelijkt vervolgens de verzamelde wetsevaluaties om conclusies te trekken over de effectiviteit van de wetgeving. Anderzijds streeft het KWJZ om de kennis, ervaringen en *best practices* op het gebied van wetsevaluaties over departementen te verspreiden (KWJZ, 2015a). Zo draagt het KWJZ zorg ervoor dat de leerervaringen uit evaluaties gedeeld en verspreid worden en dat wetsevaluaties aan de verbetering van de kwaliteit van wetgeving bijdragen. KWJZ is vanaf 2013 begonnen met het verzamelen van wetsevaluaties van alle departementen.⁵ *Het KWJZ schenkt veelal aandacht aan*

⁴ Het IAK wordt vooral in een vroeg stadium van het beleidsproces toegepast. Voordat voorstellen van beleid en regelgeving aan het Parlement worden voorgelegd, moet men deze toetsen op zeven vragen: 1. Wat is de aanleiding van het voorstel? 2. Wie zijn betrokken? 3. Wat is het probleem? 4. Wat is het doel? 5. Wat rechtvaardigt de overheidsinterventie? 6. Wat is het beste instrument hiervoor? 7. Wat zijn de gevolgen van het beleid of wet (KWJZ, 2015b). IAK is opgesteld naar het voorbeeld van het Europese Impact Assessment.

⁵ Een overzicht van de verzamelde wetten van jaren 2013 en 2014 is te vinden op de internetpagina van KWJZ: <https://www.kcwj.nl/wetgevingsbeleid/evaluatiebeleid-wetgeving/ex-post-evaluaties/overzicht-wetsevaluaties>.

de vraag of de verrichte evaluaties de effectiviteit van de wet hebben onderzocht, of slechts het proces van de invoering van de wet of de doelmatigheid ervan.

De problemen die zich voordoen bij het onderzoek naar de werking van de wet zijn de problemen van evaluatief onderzoek en van sociaalwetenschappelijk onderzoek in het algemeen. Om conclusies te trekken over de effectiviteit van de wetgeving, moet een evaluatie kunnen vaststellen of de geconstateerde effecten aan de ingezette middelen te danken zijn of dat er minstens een verband ligt tussen de ingezette middelen en de geconstateerde effecten. Daarvoor zou men een quasi-experimentele onderzoeksstrategie moeten ontwikkelen met een aselect getrokken steekproef en een controlegroep, alsook voor- en nametingen. Deze strategie is echter vaak onuitvoerbaar met sociaalwetenschappelijk onderzoek. Enkele redenen hiervoor: de situatie is vaak heel complex met meerdere factoren die invloed erop kunnen uitoefenen, wat onmogelijk maakt het effect van de interventie te isoleren; de doelen van de wet blijven vaak impliciet, vaag of worden in algemene termen omschreven. Bij gebrek van duidelijke indicatoren kan ook niet worden vastgesteld of de situatie naar de interventie aan de indicatoren voldoet (Veerman, 2007: 229-232).

Zoals al eerder aangegeven, wordt het Nederlandse evaluatiesysteem gekenmerkt door decentralisatie op departementen en terreinen. Elk ministerie is verantwoordelijk voor de evaluatie van de wet- en regelgeving op zijn eigen terrein. Een wetsevaluatie kan intern of extern worden uitgevoerd. De keuze of de wetsevaluatie intern of extern wordt uitgevoerd heeft in principe met verschillende factoren te maken hebben. Zo kunnen de omvang van de wet, de politieke gevoeligheid rondom de wet, de behoefte naar grote of mindere mate van onafhankelijkheid in evaluaties en de aanwezige expertise binnen het ministerie bepalend zijn voor wie de evaluatie gaat uitvoeren. Dit onderzoek gaat na welke van deze factoren de keuze voor interne versus externe wetsevaluaties bij de Nederlandse ministeries verklaart.

Het evaluatieonderzoek wordt idealiter door een begeleidingscommissie gemonitord, die op de wetenschappelijke kwaliteit en onafhankelijkheid van het onderzoek toeziet. Vaak zitten in de begeleidingscommissie één of meer wetenschappelijke experts op het terrein waarop de wet betrekking heeft, een jurist en iemand van de directie die de wetsevaluatie heeft aangevraagd (vaak directies Wetgeving, Kennis of Strategie).

Verder hebben enkele ministeries aparte semionafhankelijke wetenschappelijke instituten. Deze zijn WODC bij VenJ, KiM bij IenM, ZonMw bij VWS. Bij deze ministeries loopt de

aanvraag van wetsevaluaties via deze instituten. De wetsevaluaties kunnen dan intern worden uitgevoerd of aan externe partijen uitbesteed. Bij extern uitbestede evaluaties speelt het onderzoeksinstituut een intermediërende rol. Zo wordt de wetsevaluatie door het ministerie aangevraagd. Vervolgens stelt het onderzoeksinstituut de vraag in een startnotitie op en begeleidt het hele traject van wetsevaluatie tot aan het indienen van het rapport aan de directie. De uitbestede wetsevaluaties worden meestal door adviesbureaus of universiteitsonderzoekers uitgevoerd.

2.8 Samenvatting

Dit hoofdstuk heeft het beleid en de inrichting van evaluatie in het algemeen en wetsevaluatie in het bijzonder geschetst. De boodschap van de overheidsvisie op evaluatie komt erop neer dat evaluatie een belangrijk instrument in de beleidscyclus, dat bijdraagt aan het proces van besluitvorming en het opstellen van beleid en regelgeving. Het doel van evalueren is door de tijd heen ontwikkeld van een economisch instrument in dienst van budgettaire herverdeling tot een instrument tot het verantwoorden van het gevoerde beleid en een instrument om lessen uit de praktijkervaring te trekken over wat werkt en waarom. Door hun vaak verplicht karakter hebben wetsevaluaties een bijzondere positie in vergelijking met beleidsevaluaties. De kenmerken van de institutionele inrichting van evaluatie gelden echter in dezelfde mate voor wetsevaluatie als voor beleidsevaluaties.

Hoofdstuk 3: Theoretisch kader

3.1 Inleiding

Dit hoofdstuk beantwoordt deelvragen twee, drie en vier. In de volgende paragrafen wordt het begrip “evaluatiecapaciteit” concreet gemaakt door middel van de inzichten uit vier wetenschappelijke modellen die allemaal verschillende aspecten van evaluatiecapaciteit prioriteren (deelvraag drie). In deze modellen en de wetenschappelijke theorieën daarachter komt de aanname naar voren, dat de kwaliteit van evaluatieonderzoek en -rapporten beïnvloedt wordt door de evaluatiecapaciteit in den brede zin, met andere woorden door de manier waarop de aanvragers van evaluaties zorg dragen voor het proces van evalueren en voor het gebruik van evaluatieresultaten (deelvraag twee). Aansluitend hierop worden de randvoorwaarden voor een hoge mate voor evaluatiecapaciteit van publieke instellingen geïdentificeerd (deelvraag vier).

Om te komen tot de voor het empirische onderzoeksdeel relevante variabelen is gebruik gemaakt van vier modellen die de evaluatiecapaciteit definiëren en de dimensies van evaluatiecapaciteit uiteenzetten (3.2). Op basis van deze modellen wordt in paragraaf 3.3 een conceptueel model ontworpen. De factoren die de evaluatiecapaciteit bepalen zijn in drie categorieën ingedeeld. Deze zijn in drie sub-paragrafen toegelicht: individuele factoren (3.3.1), organisatorische factoren (3.3.2) en omgevingsfactoren (3.3.3). Het hoofdstuk sluit af met een conceptueel model (3.4) waarin de relevante variabelen geoperationaliseerd worden. Dit model vormt de basis voor het empirische deel van het onderzoek. Dit conceptueel model van evaluatiecapaciteit is opgesteld op basis van de selectie van de meest geschikte randvoorwaarden die in de wetenschappelijke modellen naar voren komen. Dit ideaal en breed beeld van wat evaluatiecapaciteit allemaal moet omvatten om een goed evaluatieonderzoek te verrichten, wordt vervolgens gebruikt als richtlijn voor het empirisch onderzoek waarbij de evaluatiecapaciteit van verschillende departementen zal worden geëvalueerd.

3.2 Modellen voor de diagnose van evaluatiecapaciteit

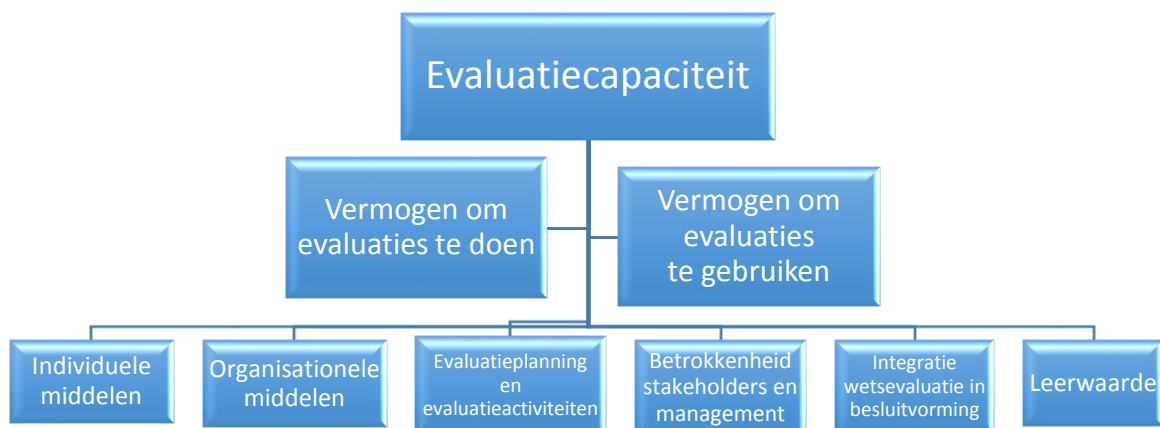
3.2.1 Het vermogen om evaluaties te doen en het vermogen om evaluaties te gebruiken

In een studie naar de verschillen in evaluatiecapaciteit tussen instellingen van de Canadese overheid ontwikkelen en valideren Bourgeois en Cousins (2013) een construct voor het meten van

evaluatiecapaciteit. Evaluatiecapaciteit houdt volgens hen twee dimensies in: het *vermogen om evaluaties te doen* en het *vermogen om evaluatiebevindingen te gebruiken*. Het construct geeft volgens de onderzoekers een compleet beeld van de mate waarin een overheidsinstelling in staat is om evaluatieonderzoek uit te voeren en de resultaten hiervan te gebruiken en te verspreiden.

De eerste dimensie van evaluatiecapaciteit verwijst naar het vermogen om evaluaties te doen. Dit betreft het hele evaluatietraject vanaf het aanvragen, opstellen van startnotitie, uitvoeren en begeleiden van evaluatieonderzoek. Daaronder zijn volgens Bourgeois en Cousins (2013) drie aspecten van belang: *individuele middelen*, *organisatorische middelen*, *evaluatieplanning en evaluatieactiviteiten*.

De tweede dimensie van evaluatiecapaciteit is het *vermogen om evaluatiebevindingen te gebruiken*. Deze geeft de mate aan waarin de organisatie in staat is kennis en inzichten uit (wets)evaluaties terug te koppelen naar het proces van beleid en wetgeving. Deze wordt gemeten aan de hand van drie aspecten: de *betrokkenheid van de stakeholders en het management*, de *integratie van evaluatie* in het proces van besluitvorming en de *leerwaarde*.



Figuur 1. Model van Bourgeois en Cousins (2013)

Dit model stelt een brede opvatting van evaluatiecapaciteit, die ook het huidige onderzoek naar de evaluatiecapaciteit bij de Nederlandse ministeries hanteert. In deze brede opvatting houdt het evaluatietraject het proces van evalueren in, vanaf het opstellen en aanvragen van (wets)evaluaties via het uitvoeren en begeleiden van het evaluatieonderzoek tot het terugkoppelen van de

evaluatiebevindingen naar het beleid, het leren uit onderzoek en het communiceren en verspreiden van de bevindingen naar andere vergelijkbare organisaties.

Dit brede perspectief op evaluatiecapaciteit is het meest geschikt voor ons onderzoek aangezien de opdrachtgever – Directie Wetgeving van het ministerie van VenJ – inzicht wil krijgen niet alleen in de mate van evaluatiecapaciteit bij verschillende Nederlandse ministeries, maar ook in de mogelijkheid om de mechanismen te identificeren waardoor ministeries uit evaluatiebevindingen daadwerkelijk leren om de kwaliteit van evaluatie, beleid en wetgeving te op peil te houden en te verbeteren. Door rekening te houden niet alleen met het vermogen om evaluaties te doen, maar ook met het vermogen om evaluatiebevindingen te gebruiken, kan dit onderzoek conclusies trekken over het leervermogen van de Nederlandse ministeries inzake (wets)evaluaties en aanbevelingen doen over mogelijke manieren om de communicatie van bevindingen te vergroten.

3.2.2 Vraagzijde en aanbodzijde van evaluatie

Volgens Nielsen et al. (2011) heeft evaluatiecapaciteit twee onderdelen: de vraagzijde van evaluatie (in hoeverre heeft de organisatie evaluatie nodig en welke waarde wordt hieraan gehecht) en de aanbodzijde van evaluatie (de competenties van onderzoekers). Nielsen et al. (2011) definiëren evaluatiecapaciteit als het vermogen van een organisatie om haar *doelstellingen, structuren en processen, menselijk kapitaal en technologie* te ontwerpen en voort te bouwen om evaluatieve kennis te produceren. Daarmee zijn ook de vier meest belangrijke randvoorwaarden voor evaluatiecapaciteit volgens Nielsen et al. (2011) genoemd.

Nielsen et al. (2011) onderscheiden deze randvoorwaarden in twee categorieën: de vraagzijde en de aanbodzijde. Onder de eerste categorie vallen twee randvoorwaarden: *doelstellingen* en *structuren en processen*. Ten eerste moet een organisatie die naar een hoge mate van evaluatiecapaciteit streeft duidelijke en expliciete *doelstellingen* hebben voor het doen van evaluaties. Het moet duidelijk vastgesteld worden wat het einddoel van het verrichten van evaluatieonderzoek is. Nielsen et al. (2011) onderscheiden drie grote doelstellingen van evaluatieonderzoek: ten eerste, verantwoording afleggen voor het gevoerde beleid of wetgeving, ten tweede, begrijpen en leren welke interventies in de maatschappij werken en hoe deze precies werken (mechanismen van interventies) en ten derde, het managen van organisaties. Duidelijkheid krijgen binnen de instelling over de doelstellingen die het evaluatieonderzoek zou moeten dienen

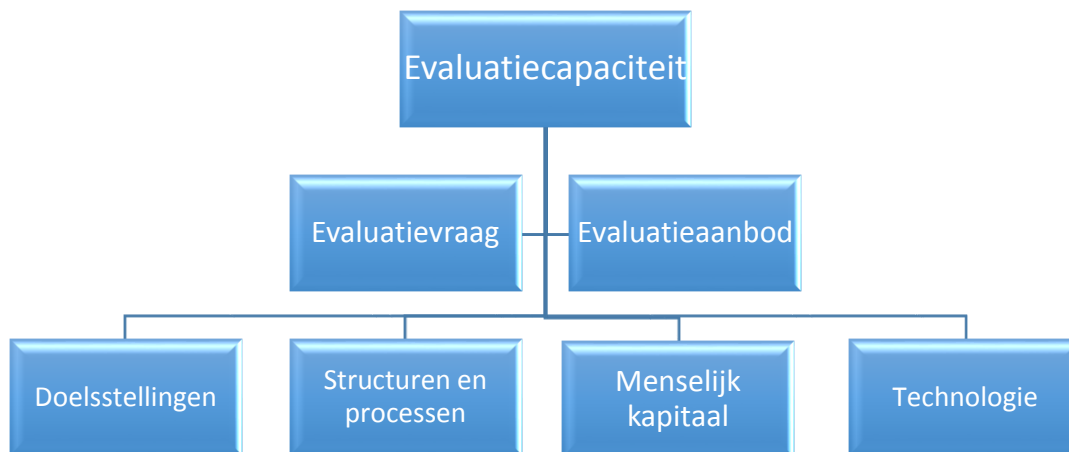
is volgens Nielsen et al. (2011) de eerste indicator voor een hoge evaluatiecapaciteit van een instelling. *Doelstellingen* kunnen eventueel worden geformaliseerd in *richtlijnen* of *handreikingen* voor evaluatieonderzoek of een meerjarige *evaluatieplanning* en *-programmering*. Omdat evaluatieonderzoek door Nielsen et al. (2011) als één van de vele vormen van kennisvergaring wordt gezien – naast bijvoorbeeld monitoren en onderzoeken – die vervolgens de besluitvorming kunnen beïnvloeden, moeten de doelstellingen van evaluatieonderzoek ook aansluiten bij een brede kennisstrategie van de instelling waarin verschillende vormen van kennisvergaring op elkaar afgestemd dienen te zijn. Een andere indicator voor hoe serieus een organisatie met evaluatieonderzoek omgaat en wat zij ermee wil bereiken, is volgens Nielsen et al. (2011) de mate en de manier waarop bevindingen uit evaluatieonderzoek gebruikt worden in het proces van besluitvorming en in de beleidscyclus. De aansluiting van evaluatiebevindingen bij andere fases van de beleidscyclus is een goede indicator dat het evaluatieonderzoek meer is dan alleen maar een ritueel proces. Evaluatie heeft dan meer dan meer dan een symbolische waarde.

Het tweede aspect van de vraagzijde van evaluatiecapaciteit zijn de *structuren en processen* waarover de organisatie beschikt bij het verrichten of ondersteunen van evaluatieonderzoek. Daarbij kan men denken aan de aanwezigheid van een *evaluatieafdeling* (*evaluation unit*) binnen de organisatie en de plek die deze afdeling bezet binnen de organisatie. De afdeling kan bijvoorbeeld een stafdirectie zijn, maar ook een financiële directie óf een beleidsdirectie. Andere mogelijkheden die Nielsen et al. (2011) niet noemen, maar die wel denkbaar zijn in de Nederlandse context zijn kennisdirectie, strategiedirectie of onderzoeksafdeling die zich bezig houden met evaluatiecoördinatie of met een onderdeel van het evaluatietraject. Verder spelen de omvang van de evaluatieactiviteit en -onderzoek, de tijd en de financiële middelen die organisatie vrij maakt voor evaluaties ook een rol in het bepalen van de evaluatiecapaciteit van de organisatie.

Het twee onderdeel van het model van Nielsen et al. (2011) verwijst naar de aanbodzijde van evaluatie. Er zijn volgens Nielsen et al. (2011) twee aspecten van de aanbodzijde waar een organisatie rekening mee moet houden, wil zij een hoge evaluatiecapaciteit hebben: het *menselijk kapitaal* en *technologie*. Menselijk kapitaal omvat zowel de competenties van (beleids)medewerkers om evaluatieaanvragen op te stellen als de competenties van onderzoekers om evaluaties uit te voeren. Onder *menselijk kapitaal* vallen bijvoorbeeld de vaardigheden van mensen die zich met evaluatie bezig houden, zoals hun *opleiding*, *praktijkervaring* met het opstellen, uitvoeren en begeleiden van evaluaties. De praktijkervaring met evaluatieonderzoek is

een volgens Nielsen et al. (2011) sterk bepalend voor hoe geavanceerd en genuanceerd een evaluatieonderzoek wordt. *Technologie* verwijst naar de middelen en bronnen die de organisatie beschikbaar stelt of ontwikkeld specifiek met als doel goed evaluatieonderzoek te kunnen verrichten: ICT-infrastructuur, betrouwbare evaluatiemodellen en methoden voor dataverzameling waarmee veel ervaring is gedaan, en systemen voor monitoring.

De vraagzijde en de aanbodzijde zijn afhankelijk van elkaar (Cooksy, 2012; Nielsen et al., 2011). Als een organisatie weinig belang aan evaluatie hecht, weinig middelen (tijd en geld) ervoor vrij maakt of de bevindingen niet terugkoppelt naar andere fases van de beleidscyclus, daalt ook de kwaliteit van de individuele competenties. Waar het belang, de middelen en de motivatie voor goede evaluatie ontbreekt, wordt ook de uitvoering van evaluatie minder professioneel en kwalitatief minder goed. Er is in dat geval niet genoeg geld om de meest capabele onderzoekers aan te stellen, of de onderzoekers doen niet hun best omdat ze al weten dat de evaluatie niet meer dan een ritueel is met als doel het beleid of de wet te legitimeren. Mensen zetten zich minder voor goede evaluaties als ze al van tevoren weten dat het onderzoek slechts een symbolische waarde heeft en weinig impact op besluitvorming zal hebben.



Figuur 2. Model van Nielsen et al. (2011)

3.2.3 Micro-, meso- en macroniveau

Volgens Cooksy (2012: 80) wordt de kwaliteit van evaluaties voor een deel bepaald door de samenhang van drie factoren die samen de evaluatiecapaciteit van een organisatie typeren: de

competentie van de evaluatieonderzoekers, de *organisatiecontext* en *-omgeving* waarin evaluatie plaatsvindt en de ondersteuning vanuit de *evaluatiegemeenschap*.⁶ Cooksy benadrukt dat de technische criteria waaraan een evaluatieonderzoek moet voldoen – denk aan het toepassen van geschikte methoden, het op tijd verzamelen van data, nauwkeurigheid in de interpretatie van data, ethische aspecten in het omgaan met deelnemers enz. – noodzakelijk, maar niet voldoende zijn om de kwaliteit van evaluaties te borgen. Evaluatiecapaciteit, afgebakend aan de hand van de drie bovengenoemde aspecten, is ook van groot invloed op de kwaliteit van evaluaties (Cooksy, 2012: 79-80).

Onder *competenties* van evaluatieonderzoekers, de eerste randvoorwaarde voor evaluatiecapaciteit, vallen niet alleen de opleidingen van onderzoekers en beleidsmedewerkers, hun methodologische en inhoudelijke kennis, vaardigheden en praktijkervaring met evaluatie, maar ook interpersoonlijke vaardigheden zoals communicatieve en projectmanagement vaardigheden en het vermogen om te reflecteren.

De tweede randvoorwaarde voor evaluatiecapaciteit is de context waarin evaluatieonderzoek plaatsvindt. De context betreft zowel de voorzieningen die binnen de organisatie aanwezig zijn voor evaluatieonderzoek, als de politieke context. Een eerste indicator voor de voorzieningen voor evaluatie binnen een organisatie is het *evaluatiebeleid* (Cooksy, 2012: 81-82). Het evaluatiebeleid spiegelt de waarde weer die een organisatie aan evaluatie hecht en laat zien hoe bereid en bereidwillig een organisatie is om goed evaluatieonderzoek te verrichten. Evaluatiebeleid kan bijvoorbeeld bepalen welke financiële middelen beschikbaar worden voor evaluatie, welke doelstelling het evaluatieonderzoek moet dient, in hoeverre de evaluatiebevindingen worden gebruikt, welk publiek met evaluatie wordt bereikt. Niet alleen de aan- of afwezigheid van evaluatiebeleid zegt iets over hoe serieus de organisatie zich voor evaluatieonderzoek wil inzetten. Men kan ook op een subtiele en impliciete manier de evaluatiecapaciteit binnen een organisatie beïnvloeden, en daardoor ook de kwaliteit van evaluatiecapaciteit aantasten. Men kan bijvoorbeeld minder of weinig middelen voor evaluatieonderzoek toeschrijven, bepaalde methodologische keuzes prioriteren of afdwingen, die minder geschikt zijn voor het ene programma, beleid of doelstelling. Een tweede indicator voor de

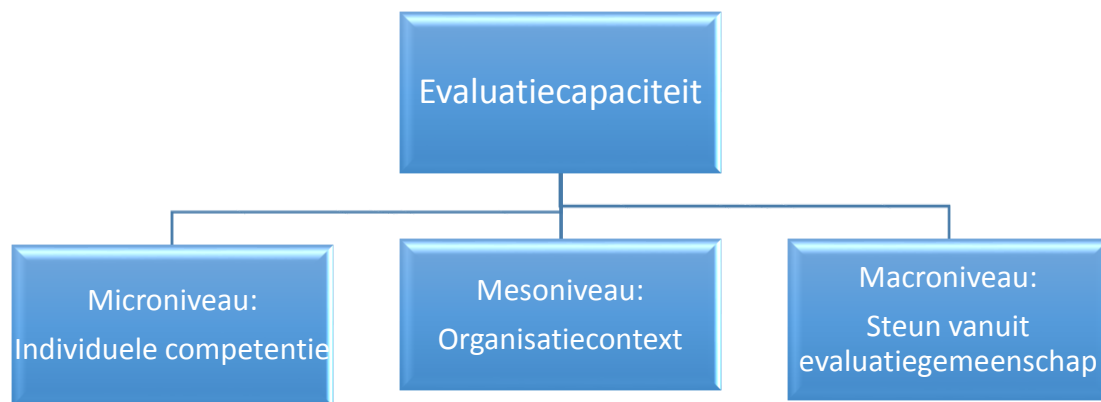
⁶ De steun van de evaluatiegemeenschap als voorwaarde voor evaluaties is contextgebonden. Cooksy bekijkt het onderwerp door de bril van de Verenigde Staten, waar de evaluatiegemeenschap een sterke positie en rol heeft in het evaluatieonderzoek.

organisatiecontext zijn de *interne strategieën voor kwaliteitscontrole*. Deze kunnen checklists, systemen of onafhankelijke onderzoekers zijn die nagaan of de conclusies van evaluatieonderzoeken goed onderbouwd zijn met bronnen zoals documenten, interviewverslagen, statistische toetsen en andere vormen van gegevensverzameling. Naast de organisatiecontext – met daaronder evaluatiebeleid en interne strategieën voor kwaliteitscontrole – speelt ook de *politieke context* een belangrijke invloedfactor op hoe ontwikkeld de evaluatiecapaciteit bij een organisatie is. Zo kan volgens Cooksy (2012: 82) een verandering in de politieke top leiden tot de stopzetting van een programma terwijl het programma nog niet is geëvalueerd. Als politici minder interesse hebben in het gevoerde beleid, kan de evaluatie vroegtijdig worden afgerond voordat het genoeg bewijs wordt verzameld om te concluderen of het beleid of de wet wel of niet werkt en via welke mechanismen dit gebeurt.

Als het om de derde randvoorwaarde, de ondersteuning van de evaluatiegemeenschap, gaat, schrijft Cooksy (2012: 82) er een zeer positieve rol aan toe in de borging van de kwaliteit van evaluatie. Een relatief sterk ontwikkelde en actieve evaluatiegemeenschap zorgt voor een goede omgeving en –sfeer waarin veel gedacht wordt over hoe men beleid en evaluatie doet. Zo stellen meeste evaluatiegemeenschap algemene richtlijnen, handreikingen, checklist en standaarden voor evaluatieonderzoek. Er ontstaan verschillende benaderingen van evaluatie waarover er ook veel samen wordt besproken en van gedachte gewisseld. Evaluatiegemeenschappen kunnen aandacht trekken voor het belang van kwalitatief goede evaluatieonderzoeken, waardoor de beleidsmakers bewuster met evaluatieverplichtingen omgaan. Evaluatiegemeenschappen kunnen ondersteuning bieden door trainingen, cursussen, conferenties, presentaties en lezingen te organiseren. Dit bevordert de kennisuitwisseling, de reflectie op de huidige instrumenten met zicht op toekomstige instrumenten en manieren om evaluaties aan te pakken.

De drie randvoorwaarden voor evaluatiecapaciteit en evaluatiekwaliteit – competenties, organisatiecontext en ondersteuning vanuit de evaluatiegemeenschap – hangen met elkaar samen. Zo kunnen bijvoorbeeld ondersteunende activiteiten, die georganiseerd worden door de evaluatiegemeenschap, de vaardigheden van onderzoekers doen toenemen. Door cursussen en trainingen te volgen, leren onderzoekers maar ook beleidsmakers van elkaar en wisselen ze voorbeelden van goede evaluatiepraktijken en kunnen ze beter inschatten welke instrumenten bij welke evaluatie en op welk terreinen het meest geschikt zijn. De drie randvoorwaarden liggen op

drie verschillende niveaus: micro, meso en macro. De competenties van onderzoekers hebben vooral betrekking op de afzonderlijke individuen of op kleine groepen en liggen dus op het micro niveau. De organisatiecontext bevindt zich op het mesoniveau en zegt iets over hoe in de organisatie over evaluatie wordt gedacht en welke voorzieningen ervoor bestaan. De toestand van de evaluatiegemeenschap bevindt zich op het macroniveau en geeft een beeld van de mogelijkheden voor evaluatieonderzoek bekeken van boven de geïsoleerde instituties.

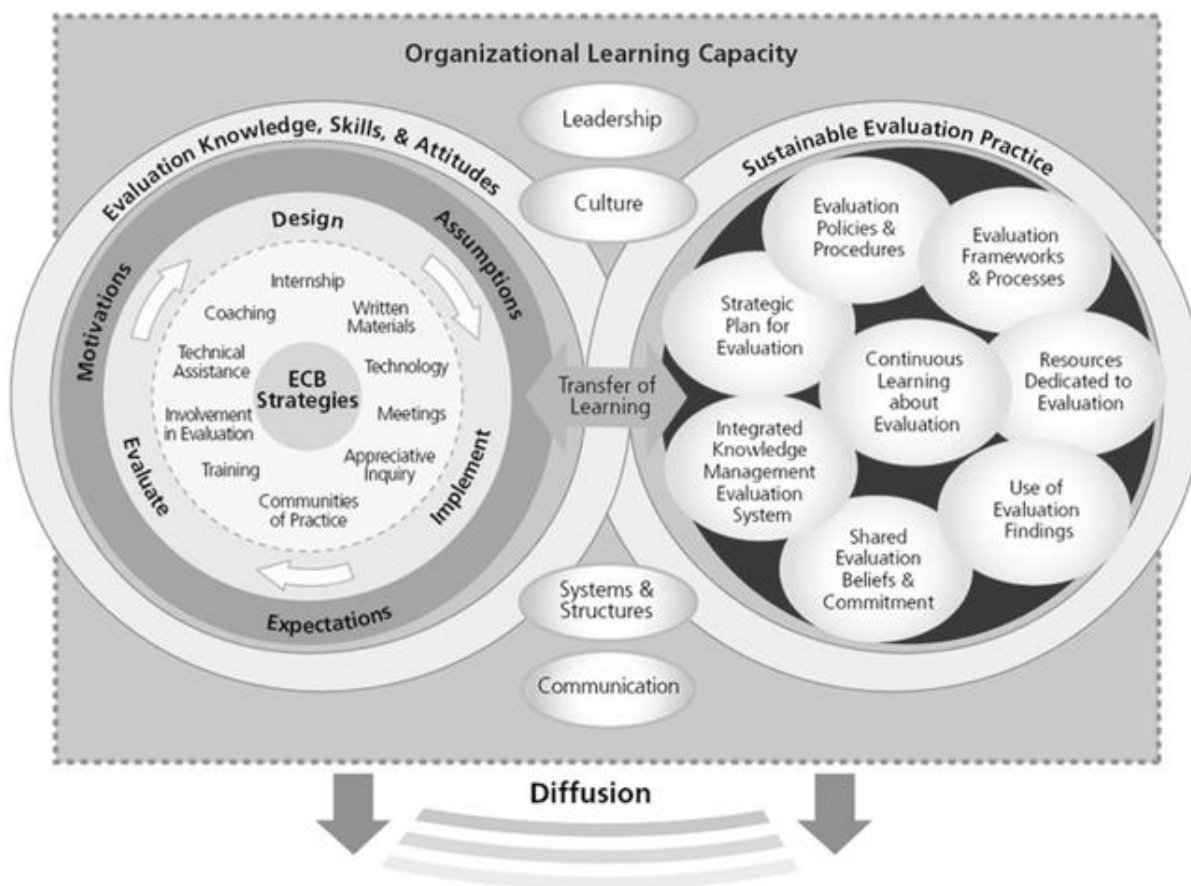


Figuur 3. Model van Cooksy (2012)

3.2.4 Ontwerpen van evaluatiecapaciteit

Preskill en Boyle (2008: 445) hebben een model ontwikkeld dat gericht is op de verhoging van evaluatiecapaciteit op organisatieniveau. Volgens dit model vindt een wisselwerking plaats tussen de linker zijde – individuele factoren zoals de kennis, vaardigheden en gedragingen van onderzoekers en de rechter zijde – evaluatiepraktijken zoals de aanwezigheid van evaluatiebeleid en -procedures, evaluatiekader en -processen, middelen voor evaluatie, een strategische evaluatieplanning et cetera. Zij beïnvloeden elkaar onderling. Evaluatiekader (‘evaluation framework’) is een breed begrip dat verwijst naar de aanwezigheid en het gebruik binnen organisaties van evaluatiehandboeken en -richtlijnen die de praktische stappen en processen beschrijft die de medewerkers kunnen volgen bij het opstellen van startnotities voor evaluatie en bij begeleiden van evaluaties. (Preskill & Boyle, 2008: 454). De achtergrondkenmerken van de organisatie (leiderschap, cultuur van de organisatie, systemen en structuren en communicatie

binnen de organisatie) bepalen de mate waarin de individuele factoren en de evaluatiepraktijken op elkaar reageren.



Figuur 4. Model van het Ontwerpen van Evaluatiecapaciteit volgens Preskill en Boyle (2008)

Preskill en Boyle (2008) stellen dat organisaties die adequate evaluaties willen produceren, een integrale visie over evaluatiepraktijken moeten ontwikkelen met daarin acht elementen: (1) een *evaluatiebeleid* en *-procedures*. Deze geven de mate aan waarin de evaluatie is geïstitutionaliseerd. Mogelijke indicatoren voor de institutionalisering van evaluatie binnen een organisatie zijn de aanwezigheid van een *evaluation unit*, vastgestelde taken en verantwoordelijkheden voor degene die betrokken zijn bij evaluatieaanvraag of –onderzoek, en interne processen die evaluatieonderzoek een plek in de dagelijkse praktijk geven; (2) een *evaluatiekader, processen en richtlijnen*. Deze omschrijven de praktische stappen en de methoden die volgens de organisatie in het evaluatieonderzoek moet worden gezet; (3) *financiële en non-financiële middelen* zijn noodzakelijk om goede onderzoekers aan te stellen en om de

technologische infrastructuur voor dataverzameling en -verwerking te verschaffen; (4) *evaluatiebevindingen* moeten niet alleen worden *gebruikt* in het proces van besluitvorming, maar ook moeten ook *aangekondigd* en *gevierd* worden. Taylor-Ritzler et al. (2013) vermoeden ook dat incentives en beloningen als ondersteuning vanuit de organisatie de kwaliteit van evaluatie zouden kunnen beïnvloeden; (5) organisatielieden moeten *gemeenschappelijke ideeën en visie over evaluatie* delen. Hiervoor is het noodzakelijk dat deelnemers met elkaar in gesprek gaan over de doelstellingen en standaarden van evaluatie. Professionals moeten in de gelegenheid worden gesteld om herhaaldelijk en voortdurend te reflecteren en het debat aan te gaan over de assumpties, theorieën en methoden die de praktijk van de evaluatie voeden. Hoe worden dingen gedaan hier binnen de organisatie? Moeten zij het anders gaan doen? Is er een betere manier om evaluatie uit te voeren? (Baizerman et al, 2002: 110). Het is interessant om na te gaan hoe verschillende actoren binnen dezelfde organisatie over het belang van de evaluatie en over de praktijken en processen denken. Denken leiders anders dan medewerkers die de evaluatie moeten opstellen, uitvoeren of de resultaten ervan verwerken en verspreiden (Taylor-Ritzler et al., 2013)? (6) een *integraal systeem van kennismanagement* kan helpen om lessen afkomstig uit evaluatie beschikbaar te maken voor toekomstige evaluaties, om data en bevindingen te verspreiden, en al bestaande informatie zo goed mogelijk te gebruiken in nieuwe evaluaties. Dit voorkomt dat evaluaties onnodig worden herhaald of parallel naast elkaar lopen; (7) een *strategisch evaluatieplan* beschrijft welk programma en/of wet moet worden geëvalueerd, hoe, wanneer en door wie de evaluaties wordt gedaan; (8) doorlopende gelegenheden voor de organisatielieden om van en over evaluatiepraktijken te *leren* en hun visie over evaluatie verder op te bouwen en aan te scherpen. Wetenschappelijke literatuur en rapporten over case studies en lessen uit internationale ervaringen helpen om *good practices* uit verschillende landen te ontdekken en verzamelen (Mackay, 2006).

De wisselwerking tussen individuele factoren en de acht evaluatiepraktijken zoals bedacht door Preskill en Boyle geldt vooral voor organisaties die evaluatie intern uitvoeren en waarin de kennis, vaardigheden en gedragingen ten opzichte van evaluatieonderzoek direct invloed uitoefenen op evaluatiepraktijken en andersom.⁷ Taylor-Ritzler et al. (2013) zien de rechterkant van het plaatje van Preskill en Boyle, namelijk de acht evaluatiepraktijken als een resultaat van de

⁷ De wisselwerking geldt echter in mindere mate voor ministeries die evaluaties uitbesteden. De externe onderzoekers kunnen slechts een beperkt invloed hebben, via de interactie met de begeleidingscommissie, op de evaluatiepraktijken van de organisatie die de opdracht tot evaluatie verleent.

individuele factoren gemedieerd door de achtergrond kenmerken van de organisatie (leiderschap, cultuur van de organisatie, systemen en structuren en communicatie binnen de organisatie). Dit geeft wellicht een geschikter beeld voor ons onderzoek omdat wetsevaluatie in Nederland veelal aan externe partijen worden uitbesteed.

3.3 Conceptueel model

Op basis van de modellen die in paragraaf 3.2 zijn gepresenteerd, stelt dit onderzoek het onderstaande conceptueel model op. Tabel 1 hieronder geeft een overzicht van de drie dimensies van het conceptueel model – de institutionele inrichting, de individuele competenties en de omgevingsfactoren – en van de variabelen die onder iedere dimensie vallen. De selectie van de variabelen is gebaseerd op twee criteria: ten eerste, de relevantie van variabelen voor de institutionele situatie en de praktijk van evaluatie in Nederland en ten tweede, de frequentie van de variabelen in de vier eerder besproken modellen.

Tabel 1. Conceptueel model Evaluatiecapaciteit

Dimensies	Variabelen	Bronnen	Vermogen om evaluaties te doen en Vermogen om evaluaties te gebruiken
Dimensie 1: Institutionele inrichting/ vraagzijde/ mesoniveau	• Organisatiestructuur	Documenten; Organigram; Jaarverslagen; Interviews;	
	• Evaluatieplanning en		
	• -programmering		
	• Evaluatiekader		
	• Middelen		
	• Communicatie		
	• Gebruik van evaluatiebevindingen		
• Leren van evaluaties			
Dimensie 2: Individuele competenties/ aanbodzijde/ microniveau	• Kennis	Interviews; Websites;	
	• Vaardigheden		
	• Ervaring		
Dimensie 3: Omgevingsfactoren/ macroniveau	• Politieke betrokkenheid	Interviews; Brieven naar de Tweede Kamer;	
	• Stakeholdersbetrokkenheid		

Wetenschappers zijn het met elkaar eens dat de twee meest belangrijke dimensies van evaluatiecapaciteit die invloed hebben op de kwaliteit van evaluaties zijn de institutionele en de

individuele factoren (Taylor-Ritzler et al., 2013: 192; Nielsen et al., 2011; Preskill en Boyle, 2008). Daarnaast houdt dit conceptueel model ook rekening met de omgevingsfactoren zoals de politieke betrokkenheid en de stakeholdersbetrokkenheid.

3.3.1 Institutionele inrichting

Het eerste aspect dat de mate van evaluatiecapaciteit van een organisatie bepaalt, is de institutionele inrichting, d.w.z. de plek en het belang die de organisatie aan evaluatie toeschrijft. Deze bevinden zich volgens Cooksy (2012) op het mesoniveau. Het *mesoniveau* van Coosky (2012) valt samen met de *vraagzijde* van Nielsen et al. (2011) en met de *organisatorische middelen* van Bourgeois en Cousins (2013). De dimensie van evaluatiecapaciteit houdt vijf variabelen in: organisatiestructuur en -inrichting, evaluatieplanning en -programmering, evaluatiekader, middelen en communicatie.

Naast de brede kijk op evaluatiecapaciteit dat dit onderzoek uit het model van Bourgeois en Cousins leent, wordt hun model gedeeltelijk aangepast. Dit onderzoek vat het aspect “institutionele inrichting” breder dan Bourgeois en Cousins (2013) het doen. Zo houden de organisatorische factoren niet alleen de fysieke infrastructuur voor evaluatie in, zoals de aanwezigheid van een *evaluation unit*, maar ook de concrete activiteit, praktijken en planning van evaluatie. Wat Bourgeois en Cousins (2013) als een apart aspect van het vermogen om evaluatie te doen – de evaluatieplanning en – activiteiten – zien, wordt in ons onderzoek beschouwd als een onderdeel van de organisatorische factoren of randvoorwaarden van de evaluatiecapaciteit. Dit ook omdat de praktijken en planning van evaluatie iets zeggen over hoe een bepaalde organisatie over evaluatie denkt en hoe belangrijk zij het evaluatieonderzoek acht.

Wat Preskill en Boyle (2008) als evaluatiepraktijken beschrijven, zijn voor een groot deel voorwaarden voor evaluatiecapaciteit. Aspecten zoals de aanwezigheid van een evaluatiebeleid, het gebruik van processen, procedures en richtlijnen bij het opstellen en uitvoeren van (wets)evaluaties, de middelen die de organisatie investeert om evaluaties uit te voeren en de aanwezigheid van een strategisch evaluatieplan bepalen de mate waarin een organisatie bereid en in staat is om goed evaluatieonderzoek te doen. Deze aspecten die Preskill en Boyle (2008) als evaluatiepraktijken aanduiden, worden in ons conceptueel model als variabelen van de organisatorische-inrichting (de tweede dimensie van evaluatiecapaciteit) opgenomen.

Om de evaluatiekwaliteit te bevorderen, moet een organisatie volgens Nielsen et al. (2011) expliciet voor ogen hebben wat de *doelstellingen* van evalueren in het algemeen is en wat de organisatie met het doen van evaluaties überhaupt wil bereiken. De intenties van de organisatie met evaluaties kunnen verder worden geformaliseerd in de vorm van *evaluatiebeleid, -richtlijnen en -jaarplannen*. Het bestaan van deze documenten geeft aan hoe belangrijk een organisatie het evaluatieonderzoek vindt.

Van Aeken (2011: 63) betoogt dat de formalisering en institutionalisering van evaluatie een aantal voordelen heeft voor de borging van een goede evaluatie. Ten eerste kan een organisatie met een vastgestelde evaluatieplanning en -programmering voorkomen dat de evaluatie strategisch door de oppositiepartijen voor eigen belang en tegen de regerende coalitie wordt gebruikt. In de afwezigheid van vastgestelde evaluatieplanning en -programma kunnen oppositiepartijen de doelstellingen van evaluatie, het tijdstip van de evaluatieaanvraag en het gebruik van resultaten makkelijker manipuleren. Landen met een meerpartijenstelsel zoals Nederland zouden het meest gevoelig zijn voor het gebruik van deze strategie genoemd *submarine evaluation*. In een geïnstitutionaliseerde organisatie staat al vastgesteld bij wie het initiatief voor wetsevaluaties ligt, wie de evaluatie mag uitvoeren en hoe de resultaten worden gevalideerd.

Bekeken vanuit Cooksy (2012) moet een kanttekening worden geplaatst bij deze beredenering van Van Aeken. Evaluatie kan evengoed door de regerende coalitie als door de oppositie partijen worden gemanipuleerd. Dit kan zowel in de fase van het initiëren van de evaluatie (tijdstip, wie doet het, wie zit er in de begeleidingscommissie, welke bronnen en middelen worden ter beschikking gesteld) gebeuren, als ook in de fase van het bekendmaken van de evaluatiebevindingen (in de la stoppen, in gesloten kring bespreken, de bevindingen niet terugkoppelen naar besluitvorming).

Ten tweede zou een duidelijk plan voor evaluatie, een goede infrastructuur en uitgebreide voorzieningen de kans vergroten dat de kwaliteit van evaluatie toeneemt (Van Aeken, 2011). Zo kan ook aandacht worden besteed aan het opleiden van onderzoekers, de verspreiding van kennis en de ontwikkeling van leidraden, handvatten voor evaluatiepraktijk, standaarden en *best practices*. Een goed ontwikkelde infrastructuur voor evaluatie vergroot de kans dat de evaluaties op een adequate manier worden uitgevoerd en dat men waardevolle kennis ontwikkelt, gebruikt en verspreidt. In een onderzoek naar evaluatie in OESO-landen redeneert Coglianesi (2012) dat het opstellen van een integraal kader voor evaluatie de standaardisering van het

evaluatieonderzoek leidt. De standaardisering zorgt voor betere vergelijkbaarheid tussen evaluaties, wat vervolgens de bruikbaarheid van evaluatieonderzoeken in besluitvorming zou doen toenemen.

Verder zijn de *plaats* die deze *unit* binnen de organisatie krijgt en de *financiële middelen* die beschikbaar worden gesteld voor het evaluatieonderzoek ook van grote betekenis voor de kwaliteit van evaluatie. *Middelen (zowel financiële en niet-financieel)* worden in nagenoeg alle modellen van evaluatiecapaciteit opgenomen als een belangrijke factor.

Bourgeois en Cousins (2013) zien het vermogen van een organisatie om evaluatiebevindingen te gebruiken als een aparte dimensie van evaluatiecapaciteit. Een organisatie kan heel goed in staat zijn om evaluaties op te stellen en/of uit te voeren, zonder vervolgens de bevindingen in te zetten om de wet of het beleid te verbeteren. Omdat de verwachting in ons onderzoek is dat sommige departementen goede evaluatieprogrammering en evaluatieopstelling hebben, maar verder weinig gebruik maken van evaluatiebevindingen, wordt in dit onderzoek ervoor gekozen het gebruik van evaluatiebevindingen en het leervermogen van als aparte variabelen van de organisatorische dimensie mee te nemen. Dit onderzoek volgt in dit opzicht het breed perspectief van Bourgeois en Cousins (2013) waarin evaluatiecapaciteit zowel *het vermogen om evaluaties te doen* als *het vermogen om evaluaties te gebruiken* omvat.

3.3.2 Individuele competenties

Het *microniveau* van Cooksy (2012) valt samen met de *aanbodzijde* van Nielsen et al. (2011) en met de *individuele middelen* van Bourgeois en Cousins (2013). In deze paragraaf wordt deze dimensie van evaluatiecapaciteit verder geconcretiseerd naar drie variabelen – kennis, vaardigheden en ervaring.

Het tweede aspect voor een goede evaluatiecapaciteit is volgens Cooksy (2012) zijn de individuele competenties van evaluatieonderzoekers om evaluaties uit te voeren. Om een goede evaluatie te kunnen uitvoeren, is het van belang dat onderzoekers goed *opgeleid* zijn en voldoende *kennis* en *vaardigheden* hebben ontwikkeld (Cooksy, 2012; Taylor-Ritzler et al. 2013). Zij moeten zowel kennis hebben van kwalitatieve en kwantitatieve methodologische aspecten van evalueren, zoals validiteit, generaliseerbaarheid en meta-analyse, als vaardigheden van hoe de bevindingen te rapporteren en over te brengen. Daarnaast is *ervaring* met evaluatieonderzoek een onontbeerlijke voorwaarde.

Door de nieuwe informatie- en communicatiemiddelen is een enorme hoeveelheid informatie beschikbaar over welke middelen – methoden, modellen, programma’s enz. – geschikt zijn voor een bepaalde evaluatieopdracht. Dit kan de onderzoekers helpen om een steeds hoger niveau van professionaliteit te bereiken. Onderzoekers kunnen het echter soms lastig vinden om de relevante informatie uit de fragmentarische en vaak onoverzichtelijke hoeveelheid kennis te halen. Kennis en vaardigheden kunnen de evaluatie alleen dan positief beïnvloeden, als men de meest relevante en nuttige bronnen over een onderwerp goed kan vinden en tot zich kan nemen. Daarom is het volgens Cooksy (2012) tegenwoordig van grote betekenis dat evaluatieonderzoekers de *relevante informatie* uit de evaluatiepraktijk snel kunnen *destilleren* en inzicht krijgen in wat wel dan niet in de praktijk werkt. Cooksy ziet het snel vinden van de juiste weg in een grote hoeveelheid informatie als even belangrijke vaardigheid als methodologische kennis. Geen individu kan echter een volledig overzicht hebben van de meest actuele informatie. Instituten die op de kwaliteit van evaluaties toezien, zoals het KWJZ in Nederland en kennisdirecties van verschillende ministeries, kunnen hierin een belangrijke rol spelen door ondersteunend beleid te ontwikkelen. Zo kunnen zij informatie over de positieve en negatieve ervaringen van andere landen verzamelen, centraliseren en verspreiden. Bovendien zouden de kennisdirecties ook eisen kunnen stellen aan de minimale competenties van onderzoekers. Verder kan de overheid beleid maken om ervoor te zorgen dat de benodigde middelen zoals tijd en geld beschikbaar worden.

De aanbodzijde in het model van Nielsen et al. (2011) (zie Figuur 2) valt voor een groot deel samen met de individuele vaardigheden van Cooksy. De aanbodzijde omvat twee aspecten: het *menselijk kapitaal* en *technologie*. Het menselijk kapitaal wordt geoperationaliseerd aan de hand van indicatoren zoals *kennis*, *opleiding* en *praktijkervaring* van onderzoekers met evaluatie. Kennis over evaluatie zou de effectiviteit van de lopende evaluatiepraktijken en van de besluitvorming in de politieke arena doen verhogen (Nielsen et al., 2011: 15). Andere individuele factoren die voor de kwaliteit van evaluatie relevant zijn, zijn het *beseft* over de baten van evaluatie en de *motivatie* van onderzoekers om goede evaluaties uit te voeren (Taylor-Ritzler et al., 2013). Deze worden gefaciliteerd volgens Cooksy (2012) door de steun vanuit de evaluatiegemeenschap (macroniveau). *Technologie* omvat *monitoringsystemen* en *methoden voor dataverzameling en –analyse*. Deze laatste dimensie is minder relevant voor ons onderzoek omdat de Nederlandse ministeries het evaluatieonderzoek veelal uitbesteden aan externe bureaus of aan academische onderzoeksgroepen, en minder intern uitvoert.

Gezien vanuit het standpunt van de onderzoekers is evaluatie volgens Nielsen et al. (2011) een technische activiteit. Zij leggen de nadruk op de expertise waarover onderzoekers moeten beschikken en die ze voortdurend moeten ontwikkelen door training, praktijkervaring en uitwisseling met andere professionals. Andere onderzoekers zien een gevaar in deze technocratische visie van evaluatie (Mackay, 2006; Kushner, 2005). De aanname die aan de technocratische visie ten grondslag ligt, is dat de wetenschappelijke en methodologische kwaliteit van evaluatie in het bijzondere en van het sociaalwetenschappelijk onderzoek in het algemeen, gewaarborgd kan worden door de strikte toepassing van administratieve instrumenten en richtlijnen, structuren en procedures (Kushner, 2005: 112). Evaluatie is echter meer dan het toepassen van instrumenten en protocollen. Het doel van evaluatie is het leren hoe een wet of een beleidsinterventie in de praktijk uitpakt om vervolgens tot betere interventies te komen die maatschappelijke problemen oplossen (Rossi & Freeman, 1993: 403). Zo zou evaluatie een zeer waardevol instrument moeten zijn om *good governance* tot stand te brengen (Mackay, 2006).

3.3.3 Omgevingsfactoren

Naast de individuele competenties en de organisatorische inrichting, worden de evaluatiecapaciteit en de evaluatiekwaliteit bevorderd door de *politieke betrokkenheid*. Veranderingen in de *politieke betrokkenheid* en het beleid dat daaruit voortvloeit, kunnen de belangstelling in geplande of al lopende evaluaties sterk doen verlagen. Zo kan de kwaliteit van de evaluatie eronder lijden als de evaluaties sneller dan gepland moet worden afgerond of als er minder geld wordt toegekend. Er is dan weinig tijd om voldoende gegevens te verzamelen en conclusies met voldoende bronnen te onderbouwen. Bij weinig belangstelling wordt de evaluatie oppervlakkig uitgevoerd en heeft evaluatie tot doel slechts het gevoerde beleid te legitimeren, zonder veel aandacht aan de evaluatiekwaliteit te besteden (Cooksy, 2012: 82). *Politieke betrokkenheid* is ook volgens Van Aeken (2011) nodig om goede evaluaties uit te voeren. Evaluatie is deels een politieke activiteit, omdat de politieke controle heeft over de doelstelling, de reikwijdte, het tijdstip van de evaluatie, de middelen ervoor en het gebruik van de evaluatiebevindingen.

3.4. Samenvatting

Dit hoofdstuk heeft het begrip “evaluatiecapaciteit” door middel van de inzichten uit vier wetenschappelijke modellen uiteengezet. Op basis van literatuur is er een conceptueel model

ontwikkeld. Dit model bevat drie aspecten – institutionele inrichting, individuele competenties en omgevingsfactoren – die verder in twaalf variabelen zijn opgesplitst. Het model geeft een ideaal beeld van hoe de evaluatiefunctie bij departementen ingericht zou moeten zijn. Het model vormt de basis voor het empirische deel van het onderzoek. De empirische gegevens en de waarnemingen over hoe de evaluatie in praktijk bij departementen is ingericht en wordt gedaan die we in het empirisch onderzoek gaan verzamelen, zullen met het ideaal beeld (het conceptueel model) worden vergeleken.

Hoofdstuk 4: Methodologisch kader

4.1 Inleiding

Om de onderzoeksvraag naar de aspecten van evaluatiecapaciteit die aanwezig zijn bij Nederlandse departementen te kunnen beantwoorden, is het van belang om de onderzoeksstrategie te bepalen en de keuzes uiteen te zetten die voorafgaand aan het empirisch onderzoek zijn gemaakt. Dit hoofdstuk beschrijft allereerst de twee onderzoeksstrategieën, de kwantitatieve en kwalitatieve onderzoeksstrategie. Op basis van hun voor- en nadelen wordt een afweging gemaakt voor de meest geschikte onderzoeksstrategie (4.2). Vervolgens zet paragraaf 4.3 de methoden voor dataverzameling met hun voor- en nadelen uiteen. Paragraaf 4.4 bespreekt de twee methoden die gebruikt worden in het empirisch deel van het onderzoek: de inhoudsanalyse van documenten en interviews. Paragraaf 4.5 presenteert de operationalisatie van de variabelen. Aan de hand van deze operationalisatie wordt de topiclijst voor het interview opgesteld, die in Bijlage A te vinden is. Paragraaf 4.6 bespreekt twee methodologische kwaliteitsvoorwaarden van het empirische onderzoek: validiteit en betrouwbaarheid. De samenvatting (paragraaf 4.7) werpt een terugblik op wat in dit hoofdstuk is besproken.

4.2 Onderzoeksstrategieën

Dit onderzoek beoogt het vermogen van de Nederlandse departementen te beoordelen om gedegen evaluaties te verrichten en uit te zetten. Dit doel wordt bereikt in een empirisch onderzoek, dat de inrichting van de evaluatiefunctie bij de Nederlandse departementen in kaart brengt en evalueert. Om de evaluatiefunctie te beoordelen, wordt onderzocht in welke mate de aspecten van evaluatiecapaciteit die afgeleid zijn uit de wetenschappelijke literatuur, aanwezig zijn bij de Nederlandse departementen. Een goed ontwikkelde evaluatiecapaciteit wordt in de literatuur gezien als een voorwaarde voor het vermogen van departementen om goede (wets)evaluaties te verrichten en zo de besluitvorming te ondersteunen en het beleidscyclus te stroomlijnen.

Bij het empirisch onderzoek naar bestuurskundige verschijnselen ligt de keuze tussen twee onderzoeksstrategieën: de kwantitatieve en de kwalitatieve onderzoeksstrategie.

Het eerste verschil tussen de kwalitatieve en de kwantitatieve onderzoeksstrategie is de manier waarop de resultaten van het onderzoek worden vastgelegd. Kwantitatief onderzoek

verwijst naar het verzamelen van “cijfermateriaal, zoals reeksen scores in een survey onderzoek, over een sociaal verschijnsel” (Bleijenbergh, 2013; Verschuren & Doorewaard, 2007). Het kwalitatieve onderzoek omvat verschillende vormen van onderzoek die gericht zijn “op het verzamelen en interpreteren van talig materiaal om op basis daarvan uitspraken te doen over een (sociaal) verschijnsel in de werkelijkheid.” (Bleijenbergh, 2013). Het sociale verschijnsel wordt dan in zijn natuurlijke situatie onderzocht (Van Thiel, 2013: 99). Het empirische materiaal bij kwalitatief onderzoek kan onder andere bestaan uit verslagen of transcripten van interviews, verslagen van observatie en informatie uit schriftelijke documenten.

Ten tweede verschillen de twee onderzoeksstrategieën op het aantal onderzoekseenheden of waarnemingseenheden. Bij een enquête (een vorm van kwantitatief onderzoek) moet men een groot aantal respondenten benaderen om vervolgens het verzamelde cijfermateriaal met statistische technieken te analyseren. Kwalitatief onderzoek (bijvoorbeeld de gevalstudie) beperkt zich tot een gering aantal onderzoeks- of waarnemingseenheden. Bij het kwantitatieve onderzoek is het aantal onderzoekseenheden afhankelijk van het aantal variabelen dat in het onderzoek meegenomen wordt. Hoe meer variabelen en verbanden hiertussen men wil toetsen, hoe meer onderzoekseenheden ervoor nodig zijn. Als vuistregel geldt een aantal van 100 onderzoekseenheden als groot genoeg voor statistische analyses.

Ten derde verschillen het kwantitatieve en het kwalitatieve onderzoek van elkaar in arbeidsintensiviteit. Het invullen van een gesloten vragenlijst (via een enquête) en het verwerken van de antwoorden hierop in statistische programma's kost minder tijd dan het houden van interviews door middel van een semigestructureerde vragenlijst. Bovendien wordt met een enquête een groter aantal mensen bereikt (Verschuren & Doorewaard, 2007: 242). Het kwantitatieve onderzoek is daarom arbeidsextensief, terwijl het kwalitatieve onderzoek arbeidsintensief is. Het laatste kost meer tijd, maar levert ook rijker materiaal op (Bleijenbergh, 2013). Het voordeel van het face-to-face interview is dat je het onderwerp dieper kunt onderzoeken door mensen door te vragen naar meer details en naar verklaringen voor hun opvattingen en meningen (Verschuren & Doorewaard, 2007: 242).

Bij de keuze van de geschikte onderzoekstrategie doen er ook andere factoren toe zoals de probleemstelling (beschrijven, toetsend, verklaren of evaluatief) en de al bestaande kennis over het onderzochte verschijnsel (Van Thiel, 2013: 69). Naarmate meer informatie voorhanden is over

een onderwerp, wordt het makkelijker om statistische methoden en technieken te gebruiken (Van Thiel, 2013: 68).

In dit onderzoek wordt gebruik gemaakt van de kwalitatieve onderzoeksstrategie. De keuze hiervoor komt voort uit de probleemstelling, het aantal onderzoekseenheden van dit verschijnsel en de stand van zaken inzake evaluatiecapaciteit. Hieronder worden deze drie aanleidingen voor de keuze van kwalitatieve onderzoeksstrategie uiteengezet.

De probleemstelling in deze studie is *evaluatief* van aard. Bij evaluatief onderzoek stelt de onderzoeker zich ten doel een bepaald verschijnsel te beoordelen op basis van een beoordelingskader. In dit onderzoek bestaat het beoordelingskader uit het conceptuele lijst van voorwaarden voor evaluatiecapaciteit, zoals opgesteld in hoofdstuk 3.

De populatie bestaat uit de elf Nederlandse departementen omdat het doel van dit onderzoek is de evaluatiecapaciteit op rijksniveau te beoordelen en bijzondere arrangementen en voorzieningen voor het evaluatieproces te identificeren. Twee ministeries – Algemene Zaken en Defensie – worden buiten beschouwing gelaten omdat al bekend is dat zij nauwelijks evaluaties uitvoeren. Aangezien het aantal onderzoekseenheden (n=9) gering is, is de kwalitatieve onderzoeksstrategie meer geschikt dan de toepassing van statistische methoden.

Het verschijnsel evaluatiecapaciteit zoals onderzocht in deze studie is heel breed en complex en omvat het hele proces van evalueren beginnend bij het opstellen en aanvragen van evaluaties via het uitvoeren en rapporten van evaluatiebevindingen en het terugkoppelen van bevindingen naar besluitvorming en wetgeving. Om al deze aspecten te onderzoeken is het van belang om zo breed mogelijk te kijken naar hoe de evaluatiefunctie in Nederland is ingericht. Niet alle aspecten van het evaluatieproces zijn schriftelijk vastgesteld. Veel van de benodigde informatie zou zowel via enquête als via interviews kunnen worden verzameld. De interviews hebben echter het voordeel dat je de respondenten kunt doorvragen en op elkaars ideeën laten reageren. Er wordt verder gekeken hoe verschillende partijen binnen hetzelfde ministerie de evaluatiefunctie ervaren en beoordelen, welke positieve en minder positieve aspecten van de inrichting van evaluatiefunctie zien. Een ander voordeel van de kwalitatieve onderzoeksstrategie is dat je als onderzoeker open staat tot suggesties vanuit de praktijk en rekening kan houden met factoren die buiten het theoretische beoordelingskader vallen, maar die toch zeer relevant kunnen zijn voor het begrijpen van het onderzochte sociale verschijnsel. Het geringe aantal onderzoekseenheden, het voordeel van doorvragen en jezelf door praktijk laten informeren zijn

aanleidingen om voor de kwalitatieve onderzoeksstrategie te kiezen over de kwantitatieve onderzoeksstrategie.

4.3 Kwalitatieve methoden van dataverzameling

Bij kwalitatief onderzoek kan men verschillende methoden van dataverzameling gebruiken: het interview, de participerende observatie, de inhoudsanalyse en de meta-analyse of een combinatie daarvan (Van Thiel, 2013: 69). Hieronder bespreekt deze studie de voor- en nadelen van deze vier methoden van dataverzameling om vervolgens een afweging te maken over welke methoden het meest geschikt zijn voor het empirische onderzoek.

“Een interview is een gesprek waarin de onderzoeker, door vragen te stellen aan één of meer personen (respondenten), informatie verzamelt over het onderzoeksonderwerp” (Van Thiel, 2013: 108). Een interview kan open, semigestructureerd of gestructureerd zijn, naarmate de gestelde vragen open of gesloten zijn. Het volledig gestructureerde interview wordt in feite gebruikt bij enquête en daarmee worden vooral kwantitatieve data verzameld. Bij het semigestructureerde interview gebruikt de onderzoeker een vragenlijst (met open vragen), een interviewhandleiding of topiclijst, waarin de onderwerpen voor het gesprek worden vastgelegd en eventueel aangevuld met open vragen. Bij het open interview, ook genoemd het vrije-attitude-interview, wordt nog meer ruimte aan de respondent gegeven om zijn mening over het onderwerp en zijn ervaringen te presenteren en te duiden. De onderzoeker stuurt weinig aan en de vragen die hij of zij stelt ontstaan uit het gesprek zelf.

De participerende observatie, ten tweede, verwijst naar de activiteit waarin de onderzoeker een sociaal verschijnsel of situatie op een heel bewuste manier waarneemt. De activiteit vindt plaats in de natuurlijke omgeving van het sociale verschijnsel (Bleijenbergh, 2013). Het voordeel van deze methode van dataverzameling is dat de onderzoeker zelf deelneemt aan activiteiten die met het verschijnsel te maken hebben. Bovendien gebruikt de onderzoeker alle vijf de zintuigen om de gebeurtenissen te duiden. Het nadeel is dat de aanwezigheid van de onderzoeker de natuurlijke situatie kan verstoren en dat mensen zich op een sociaal gewenste manier kunnen gedragen (Bleijenbergh, 2013).

De inhoudsanalyse van documenten, ten derde, bestaat uit het verzamelen van documenten die met het onderwerp te maken hebben en het selecteren, ordenen en interpreteren van relevante fragmenten hiervan. Met inhoudsanalyse van documenten tracht de onderzoeker het sociale

verschijnsel te reconstrueren door middel van teksten. Documenten kunnen schriftelijke bronnen zijn (raadpleegbaar openbaar of op verzoek), persoonlijke documenten zoals dagboeken, correspondenties, biografieën of interviews in kranten, of niet schriftelijke bronnen zoals film, foto's, posters en anderen. (Van Thiel, 2013: 117-119).

Bij een meta-analyse maakt de onderzoeker een overzicht van al bestaande onderzoeksbevindingen en brengt hij of zij structuur hierin (Van Thiel, 2013: 70). Veel onderzoeken maken gebruik van meerdere methoden van dataverzameling. Dat heet triangulatie van methoden en is een manier om de kwaliteit van het kwalitatief onderzoek te versterken (Van Thiel, 2013).

4.4 Keuze van methoden en databronnen

De studie maakt gebruik van twee methoden van dataverzameling: inhoudsanalyse van schriftelijke documenten en semigestructureerde interviews. De analyse van schriftelijke documenten is noodzakelijk om informatie te verzamelen die al via de websites van de bestudeerde organisatie openbaar toegankelijk zijn. De semigestructureerde interviews worden gehouden om interne informatie te verzamelen over hoe evaluaties daadwerkelijk worden gedaan en om beleidsmakers die betrokken bij evaluaties te laten reageren op onze bevindingen (zie Bijlage B).

Om informatie te vergaren over de stand van zaken inzake evaluatiecapaciteit bij Nederlandse departementen, zullen allereerst *bestaande schriftelijke documenten* worden geraadpleegd. Om documenten te verzamelen, zal op verschillende manieren online worden gezocht. Documenten die informatie kunnen leveren over de voorwaarden voor evaluatiecapaciteit zoals het bestaan van evaluatieafdelingen, evaluatieplanning- en programmering, handreikingen en richtlijnen voor evaluatie etc., worden opgezocht via de websites van departementen en andere overheidsinstellingen die zich met evaluatie bezig houden. Relevante documenten kunnen zijn: organogrammen van departementen, jaarverslagen, documenten over de evaluatieplanning en -programmering, schriftelijk materiaal zoals presentaties en flyers over evaluatieactiviteit bij departementen, openbaar toegankelijke handreikingen en richtlijnen voor het evaluatieproces.

Dit onderzoek zal de informatie die verkregen wordt uit documenten, op twee manieren gebruiken. Allereerst geeft de informatie inzicht in welke aspecten van de evaluatiecapaciteit goed ontwikkeld zijn bij welke departementen en welke aspecten onvoldoende ontwikkeld zijn. Op basis van deze empirische informatie kan dit onderzoek conclusies trekken over de mate waarin

departementen aan het ideaal beeld van evaluatiecapaciteit voldoen, dat in de wetenschappelijke literatuur is geschetst. De informatie uit documenten is ook een hulpmiddel bij de selectie van mogelijke respondenten voor de interviews. Documenten zoals organogrammen laten zien waar de evaluatiefunctie binnen het ministerie wordt geplaatst en wie erbij betrokken is. Daarnaast zullen respondenten worden gevraagd om aan te geven hoe relevant en bruikbaar zij de door ons gevonden documenten achten in hun werkzaamheden rondom evaluatie.

In navolging van de inhoudsanalyse van documenten zullen interviews worden gehouden met respondenten die betrokken zijn bij het evaluatieproces (zie Bijlage B). De selectie van respondenten zal rekening houden met de spreiding van respondenten over verschillende functies – beleidsmakers, wetgevingsjuristen, onderzoekers in dienst bij de overheid en projectleiders van evaluatietrajecten. Het is belangrijk dat geïnterviewde respondenten verschillende functies bij departementen vervullen om zoveel mogelijke opvattingen en standpunten binnen hetzelfde ministerie te horen en de respondenten op elkaars meningen te laten reageren. Deze werkwijze verhoogt de validiteit van de bevindingen (zie voor meer details hierover paragraaf 4.6).

De *semigestructureerde interviews* geven ruimte aan respondenten om over hun ervaring te vertellen, over hoe het evaluatietraject concreet eruit ziet, wat voor hulpmiddelen voor evaluatie ze gebruiken en welke voorzieningen ze missen (Van Thiel, 2013). Verder is het doel van het interviewen van respondenten ook de feiten die verzameld zijn uit documentenanalyse op juistheid te controleren, en aanvullende informatie te verzamelen over hoe met evaluatie bij verschillende departementen concreet wordt omgegaan en welke arrangementen, activiteiten en praktijken prioriteit hebben bij aparte ministeries. Via interviews wordt in dit onderzoek gepolst welke belemmeringen (individuele, organisatorische of externe) de respondenten ervaren in de inrichting, planning en uitvoering van evaluatieonderzoek.

De interviews zullen worden opgenomen indien de respondenten hiermee akkoord gaan. Het opnemen van interviews verhoogt de betrouwbaarheid van het onderzoek doordat de feiten en mening van respondenten altijd kunnen worden gecontroleerd.

Bij de selectie van de respondenten is op een aantal aspecten gelet. Personen kunnen fungeren als informant of als respondent. Een informant is een persoon die de onderzoekers naar respondenten doorverwijzen (Verschuren & Doorewaard, 2007). Bij de meeste interviews zijn meerdere respondenten uit verschillende invalshoeken aanwezig: beleidsmakers, medewerkers van ondersteunende stafdirecties, juristen van directies Wetgeving. Zij zijn zelf bij

evaluatietrajecten betrokken, vooral als opdrachtgevers of projectbegeleiders. Het voordeel van interviews met meerdere respondenten is dat deze op elkaars meningen en opvattingen kunnen reageren. Dat verschaft inzicht in de overeenkomsten en verschillen tussen respondenten in hun interesses en motieven voor evaluaties. Een nadeel van interviews met meerdere respondenten is dat mensen misschien niet snel bereid zijn om hun meningen te uiten. Respondenten kunnen echter zelf aangeven of ze apart willen worden geïnterviewd. In dit geval zullen er meerdere interviews bij ieder ministerie worden gehouden om toegang te krijgen tot een diversiteit van informatie en perspectieven.

De sleutelinformanten zijn geselecteerd via de contacten die de onderzoekers bij elk ministerie hebben. In sommige gevallen worden deze contactpersonen zelf geïnterviewd indien ze direct betrokken zijn bij evaluatietrajecten. Anders verwijzen ze door naar deskundigen (sleutelinformanten) die (veel) kennis en ervaring met het evaluatiewerk hebben. De tweede route bij de selectie van respondenten was door contact te nemen met sleutelinformanten voorgesteld door Directie Wetgeving van het ministerie van VenJ, de aanvrager van het huidige onderzoek naar evaluatiecapaciteit. In een vroege fase van het onderzoek is een exploratief interview gehouden met een sleutelinformant die werkzaam is bij een semionafhankelijk onderzoeksinstituut en betrokken is bij het opstellen, aanvragen en begeleiden van evaluaties op verzoek van het ministerie van VWS.

De topiclijst (zie Bijlage A) voor interviews is afgeleid uit de onderstaande operationalisatie van variabelen en sluit aan op de praktijk van evaluatie. De variabelen zijn in het theoretisch kader (hoofdstuk 3) geïdentificeerd als de voorwaarden van evaluatiecapaciteit. Iedere variabele is op meerdere manieren geoperationaliseerd om het begrip zo goed mogelijk te dekken (zie tabel 2). De topiclijst is vervolgens in een interview getoetst. De respondenten (projectbegeleiders van uitbestede evaluaties) hebben feedback en nuttige suggesties op de topiclijst gegeven. De suggesties zijn verwerkt in de topiclijst.⁸

4.5 Operationalisatie van variabelen

Om een brug te slaan van theorie naar het empirische deel van het onderzoek, zijn de geselecteerde variabelen geoperationaliseerd. Operationalisatie houdt in dat theoretische begrippen

⁸ De eindversie van de topiclijst is in Bijlage A opgenomen.

waarneembaar of meetbaar worden gemaakt (Van Thiel, 2007: 52). Bij het invullen en afbakenen van variabelen is gekeken naar hoe dezelfde theoretische begrippen in verschillende wetenschappelijke artikelen zijn afgebakend. Verder hebben wij ons laten informeren door een exploratief interview met een sleutelinformant die werkzaam is bij een publieke organisatie op het gebied van evalueren. De kolom “Operationalisatie” geeft de elementen weer met behulp waarvan de variabelen worden “gemeten”. De operationalisatie dient als uitgangspunt voor het opstellen van de topiclijst die gebruikt wordt bij de semigestructureerde interviews (zie Bijlage A).

Tabel 2. Operationalisatie variabelen Evaluatiecapaciteit

Dimensies	Variabelen	Operationalisatie	Bronnen	Vermogen om evaluaties te doen en Vermogen om evaluaties te gebruiken
Dimensie 1: Organisatie-inrichting/ vraagzijde/ mesoniveau	Organisatie-structuur	Aanwezigheid <i>evaluation unit</i> ; Aantal medewerkers; Percentage wetsevaluaties onder evaluaties; Percentage uitbestede wetsevaluaties vs. intern uitgevoerde wetsevaluatie; Doelstellingen van evaluatieonderzoek;	Documenten; Organigram; Jaarverslagen; Interviews;	
	Evaluatie-planning en -programmering	Meerjarig programma evaluatie; Flyers; Jaarlijkse bijstelling evaluatieprogramma;		
	Evaluatiekader	Aanwezigheid en inhoud richtlijnen, handreikingen en protocollen voor evaluatieaanvraag; Rol begeleidingscommissie;		
	Middelen	Financiële en niet-financiële		
	Communicatie	Communicatie tussen aanvragers en onderzoekers, en tussen onderzoekers en begeleidingscommissie;		
	Gebruik van evaluatie-bevindingen	Typen gebruik: direct gebruik (in besluitvorming) vs. Indirect gebruik (conceptueel gebruik);		
	Leren van evaluaties	Manier van rapporteren over resultaten en van communicatie van aanbevelingen; Verspreiding van <i>best practices</i> over departementen; Verwerken van aanbevelingen in werkwijze wetgevingsjuristen en standaarden voor evaluatieonderzoek; Verwerken aanbevelingen uit de evaluatie van programmering in nieuwe programmering; Contact afdeling/ <i>evaluation unit</i> met KWJZ;		
Dimensie 2: Individuele competenties/ aanbodzijde/ microniveau	Kennis	Opleiding aanvragers en onderzoekers; Kennis inzake kwalitatieve en kwantitatieve methoden; Voorkeur voor bepaald onderzoeksdesign en -methoden; Manier waarop de onderzoeker de wetenschappelijke kwaliteit van wetsevaluaties waarborgen;	Interviews; Websites;	

	• Vaardigheden	Multidisciplinariteit van team aanvragers en team onderzoekers; Gevolgd trainingen en vakken over evalueren in de afgelopen jaren;		
	• Ervaring	Aantal jaren werkzaam in functies gerelateerd aan evalueren (van wetten);		
Dimensie 3: Omgevings- factoren/ macroniveau	• Politieke betrokkenheid	<i>Ad hoc</i> veranderingen in het verloop van evalueren van wetten;	Interviews; Brieven naar de Tweede Kamer;	
	• Stakeholder-betrokkenheid	Betrokkene partijen bij de bepaling van programmering voor evaluatie, aanvragen en uitvoering van wetsevaluaties; Betrokkenheid doelgroepen bij evalueren;		

4.6 Kwaliteit van empirisch onderzoek

Nu het duidelijk is geworden hoe het onderzoek zal worden uitgevoerd, kunnen enkele conclusies worden getrokken over de kwaliteit van het empirisch onderzoek. Hieronder komen validiteit en betrouwbaarheid aan bod.

4.6.1 Validiteit

Validiteit geeft de mate aan waarin men daadwerkelijk meet wat hij wil meten (Van Thiel, 2013: 58). Een aspect van validiteit is dat de variabelen en de indicatoren (door middel waarvan de variabelen zijn geoperationaliseerd) een goede weergave van het conceptuele model moeten vormen. De operationalisatie in tabel 2 is voor een groot deel door de onderzoeker zelf opgesteld. De indicatoren zijn echter geïnspireerd door de checklists die de auteurs van de theoretische modellen (zie paragraaf 3.2) gebruiken om hun deze modellen te meten en valideren. Meeste indicatoren komen in meer dan één checklist terug, wat bewijst dat ze een goede operationalisatie zijn van de variabelen van evaluatiecapaciteit. Voorts is de tabel met de geoperationaliseerde variabelen bij de stagebegeleidster voor feedback voorgelegd.

Dit onderzoek hanteert twee methoden van dataverzameling, de inhoudsanalyse van schriftelijke documenten en interviews. Zoals al eerder aangegeven, zullen er in dit onderzoek meerdere respondenten worden geïnterviewd. Om de validiteit van de bevindingen en conclusies te waarborgen, zal bij de selectie van de respondenten rekening worden gehouden met het criterium van de verspreiding over verschillende functies. Hierdoor streeft deze studie een zo compleet mogelijk beeld van de evaluatiepraktijken in Nederland te geven. Daarnaast wordt informatie

verkregen uit openbaar toegankelijke schriftelijke documenten aangevuld met intern verkregen documenten die de respondenten aan ons beschikbaar zullen stellen. Het gebruik van meerdere methoden van dataverzameling en de selectie van respondenten uit verschillende invalshoeken verhoogt de kans dat de verzamelde informatie valide is en dat komt ten goede van de kwaliteit van het onderzoek (Van Thiel, 2013).

4.6.2 Betrouwbaarheid

Betrouwbaarheid bij kwalitatief onderzoek is de mate waarin de onderzoeksbevindingen herhaalbaar en navolgbaar zijn (Van Thiel, 2013: 167). Omdat de bevindingen gebaseerd zijn op interpretaties van onderzoekers zelf, worden in de literatuur verschillende maatregelen aanbevolen om de kans te minimaliseren dat waarnemingen op toevallige omstandigheden rusten.

Om de betrouwbaarheid te verhogen past dit onderzoek de volgende maatregelen toe: de triangulatie van methoden van dataverzameling, de keuzes worden vastgelegd in een logboek, de operationalisatie van variabelen en de topiclijst zijn aan twee andere onderzoekers opgelegd. Dit vergroot de interbeoordelaars-betrouwbaarheid (Van Thiel, 2013: 168).

De respondenten zijn zorgvuldig uitgekozen na e-mailcontact met informanten bij departementen en met de opdrachtgever die ook contacten heeft onder de juristen en beleidsmedewerkers die zich met evaluatie bezighouden. Om het gevaar van sociaal wenselijke antwoorden van de respondenten te voorkomen (Verschuren & Doorewaard, 2007: 243) gebruikt deze studie, naast de semigestructureerde vragenlijst, open vragen. De open vragen geven de respondenten de kans om commentaar te geven op hoe het op dit moment de evaluatiefunctie binnen het departementen is ingericht en suggesties te uiten over hoe de evaluatiepraktijk in de toekomst zou kunnen worden verbeterd. Bovendien zullen de gegevens uit de interviews met de gegevens uit de documentenanalyse worden vergeleken. Daarnaast zal bij alle interviews zoveel mogelijk steeds dezelfde topiclijst worden gebruikt. Deze topiclijst sluit aan bij de operationalisatie van de variabelen. Bij semigestructureerde interviews moeten de onderzoekers echter ook aanvullende vragen stellen, rekening houdend met de context van elk ministerie.

4.7 Samenvatting

Dit hoofdstuk heeft de keuze voor de kwalitatieve onderzoekstrategie, voor de methoden en de databronnen verantwoord. Daarnaast zijn in dit hoofdstuk de theoretische begrippen

geoperationaliseerd. Tot slot heeft dit hoofdstuk uitgelegd welke stappen in dit onderzoek worden gezet om aan de twee belangrijke kwaliteitscriteria voor het sociaalwetenschappelijke onderzoek – validiteit en betrouwbaarheid – te voldoen.

Hoofdstuk 5: Resultaten

5.1. Inleiding

Dit hoofdstuk geeft een overzicht van de belangrijkste overeenkomsten en verschillen tussen ministeries met betrekking tot de drie aspecten van evaluatiecapaciteit zoals deze uit de wetenschappelijke literatuur in hoofdstuk 3 zijn afgeleid. De resultaten van de analyse zijn ingedeeld in drie paragrafen. Paragraaf 5.2 evalueert de institutionele inrichting van de evaluatiecapaciteit binnen acht ministerie. Vervolgens analyseert paragraaf 5.3 de manieren waarop ministeries de individuele competenties voor (wets)evaluatie borgen. Paragraaf 5.4 evalueert de invloed van omgevingsfactoren zoals de politieke betrokkenheid en de stakeholdersbetrokkenheid op de keuzes die worden gemaakt door ministeries in het evaluatieproces.

5.2 De institutionele aspecten van evaluatiecapaciteit

De institutionele inrichting is volgens de wetenschappelijke literatuur in hoofdstuk 3 een belangrijke voorwaarde voor een goed ontwikkelde evaluatiecapaciteit en voor het verrichten van goede evaluaties. De institutionele inrichting omvat meerdere aspecten. Deze paragraaf geeft inzicht in de institutionele inrichting van de Nederlandse departementen op basis van de zeven aspecten die uit de wetenschappelijke literatuur zijn afgeleid en die als belangrijke voorwaarden voor evaluatiecapaciteit worden gezien: (1) structuur van de organisatie, (2) evaluatieplanning en -programmering, (3) evaluatiekader, (4) middelen voor evaluatieonderzoek, (5) communicatie binnen en tussen de departementen rondom evaluaties, (6) gebruik van evaluaties en ten slotte, (7) leren van evaluaties. In wat nu volgt, worden de resultaten van de analyse per variabele uiteengezet.

5.2.1 Structuur van organisatie

Om vast te stellen in hoeverre de Nederlandse departementen in staat zijn goede (wets)evaluatie te doen en de kwaliteit van wetsevaluaties te borgen, is het belangrijk om zicht te krijgen in de structuur van departementen. Dit onderzoek heeft de structuur van de organisatie door middel van vier aspecten bestudeerd.

De aanwezigheid en de plaats van evaluatieafdelingen binnen departementen

Volgens de literatuur (Nielsen et al, 2011) zegt de *aan- of afwezigheid van evaluatieafdeling* binnen de organisatie evenals de *plek* die deze afdeling binnen de organisatie krijgt en de *taken* die zij vervult al heel veel over de visie van de organisatie op evaluatie en over het belang dat de organisatie aan evaluatie hecht. De evaluatieafdeling kan ofwel binnen het ministerie, of op afstand van het ministerie worden geplaatst. Binnen het ministerie, kan de evaluatieafdeling verschillende functies bekleden: een overwegend financiële functie, een ondersteunende functie (zowel van het evaluatieproces als van meer inhoudelijke keuzes gerelateerd aan evaluatie), een onderzoek- en kennisfunctie, of een combinatie van deze functies.

Het algemene beeld uit de bevindingen van deze studie is dat de inrichting van de evaluatiefunctie sterk verschilt per ministerie. Bij de meeste departementen is er *geen aparte evaluatieafdeling*, en al helemaal geen afdeling die zich uitsluitend bezighoudt met wetsevaluaties. De afdelingen die de wetsevaluaties (zowel procesmatig als inhoudelijk) ondersteunen en stimuleren, houden zich naast wetsevaluatie ook met andere soorten onderzoek bezig, waaronder beleidsdoorlichtingen, monitoring, experimenten en onderzoek in het algemeen (alle interviews).

Elk ministerie heeft een Directie Wetgeving en een Directie Financiële en Economische Zaken (FEZ). Een overeenkomst tussen departementen is dat de directies Wetgeving verassend genoeg weinig betrokken zijn bij ex post evaluaties van wetten en in ieder geval geen centrale rol hierin spelen (alle interviews). Directies Wetgeving houden zich bezig met het begeleiden van ex ante onderzoek en met de toetsing van nieuwe wetsvoorstellen aan de IAK-richtlijnen (Interviews 9 juli 2015 en 16 juli 2015). Slechts bij grote sectoroverstijgende wetsevaluaties is de Directie Wetgeving de formele opdrachtgever, zoals bij de evaluaties van Awb waarbij Directie Wetgeving van VenJ de aanvrager was. Grote, sectoroverstijgende wetsevaluaties raken meerdere terreinen en dus ook meerdere beleidsdirecties waardoor er niet één directie is waarbij de evaluatie belegd kan worden. In dat geval is Directie Wetgeving de aanvrager en de trekker van de evaluaties (Interview, 1 juli 2015). Volgens een meerderheid van de geïnterviewde respondenten zijn wetsevaluaties in meeste gevallen moeilijk te scheiden van beleidsevaluaties. Wetten worden vaak ingezet als beleidsinstrument naast andere beleidsinstrumenten. Een gevolg hiervan is dat departementen weinig aparte voorzieningen voor wetsevaluaties hebben ontwikkeld.

Ieder ministerie heeft een Directie Financiële en Economische Zaken (FEZ), maar deze directie vervult verschillende rollen en taken bij verschillende departementen. Zo heeft FEZ bij SZW een stimulerende en aanjagende rol in zake kwaliteitsborging van het evaluatieonderzoek, terwijl de Directie FEZ van VenJ en van OCenW zich tot financiële taken beperkt (Interview 1 juli 2015). In dat geval is de ondersteunende en stimulerende rol voor evaluatieonderzoek bij andere afdelingen belegd zoals bij de Directie Kennis (OCenW) of het onderzoeksinstituut WODC (VenJ). Bij alle departementen ligt het initiatief voor ex post evaluaties (wetsevaluaties inclusief) bij de beleidsdirecties, die ook verantwoordelijkheid hiervoor dragen. Ieder ministerie heeft één of meerdere stafafdelingen, -units of -directies die de evaluatiefunctie bij beleidsdirecties ondersteunen en stimuleren.

Tabel 3 presenteert een overzicht van de ondersteunende afdelingen die belast zijn met de kwaliteitsborging van evaluaties en ondersteuning van de evaluatiefunctie.

Tabel 3. Evaluatieafdelingen bij de Nederlandse departementen

Ministerie	Afdelingen/directies betrokken bij evaluatieonderzoek
BZ	• IOB - Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie (doet geen wetsevaluaties)
EZ	• Directie AFP - Algemene Financiële Politiek • DG B&I – Directoraat-Generaal Bedrijfsleven & Innovatie (vooral betrokken bij monitoring en effectmeting) ➤ Regiegroep monitoring en effectmeting binnen DG B&I (fungeert als linking pin naar de directies van DG B&I) • FEZ - Financiële en Economische Zaken • BEC - Beleidsevaluatiecommissie (gezamenlijk secretariaat van AEP en FEZ; draagt zorg voor kwaliteit evaluatie EZ-breed) • BIT - Behavioural Insights Team • DWJZ - Directie Wetgeving en Juridische Zaken (vooral betrokken bij ex ante evaluatie, toetsen wetsvoorstellen aan IAK-richtlijnen)
Financiën	• IRF – Inspectie der Rijksfinanciën
IenM	• FMC - Financiën, Management en Control (met daaronder FEZ) • BIT - Behavioural Insights Team • HBJZ - Hoofddirectie Bestuurlijke en Juridische Zaken • KiM – Kennisinstituut voor Mobiliteitsbeleid

OCenW	• Directie Kennis • Onderzoekscoördinatoren van ieder DG
SZW	• FEZ – Financiële en Economische Zaken • Onderzoekscoördinatoren van beleidsdirecties onder leiding van <i>Chief Science Officer (CSO)</i>
VenJ	• Onderzoeksinstituut WODC met daaronder afdeling EWB (betrokken bij uitbesteding evaluaties) • Directie Wetgeving
VWS	• Onderzoeksinstituut ZonMw ➤ CER - Commissie Evaluatie Regelgeving

In wat volgt geeft deze studie een uitgebreider beeld van de *verschillen tussen departementen* in de manier waarop zij de evaluatiefunctie inrichten.

Uit de empirische studie blijkt dat departementen sterk van elkaar verschillen in de plek die zij aan evaluatie geven binnen de organisatie (zie Box 5.1). De diversiteit in de inrichting van de evaluatie en de plek die zij krijgt binnen het ministerie is een gevolg van het gedecentraliseerde karakter van de Nederlandse rijksoverheid. Departementen hebben veel ruimte om zelf de evaluatie in te richten. De keuze die departementen hebben gemaakt over de inrichting van de evaluatiefunctie zegt iets over hun visie op evaluatie en het belang dat zij aan evaluatie hechten.

Twee departementen werken nauw samen met semi-onafhankelijke onderzoeksinstituten: VenJ met WODC en VWS met ZonMw (zie Box 5.1). Allebei de instituten hebben een semi-onafhankelijke positie ten opzichte van het ministerie. Deze instituten staan op het raakvlak tussen onderzoek, beleid en politiek.⁹ WODC doet drie soorten onderzoek: probleemverkenkend en -verklarend onderzoek, monitoring en evaluatie en R&D (WODC, 2015). ZonMw stimuleert en financiert het onderzoek op het gebied van gezondheid en het gebruik van kennis om de zorg en de gezondheid te stimuleren (ZonMW, 2015). Het verrichten van beleidsevaluaties en wetsevaluaties wordt door deze instituten als onderdeel gezien van een breder doel: het ontwikkelen en vergaren van kennis en het gebruik daarvan om maatschappelijke problemen aan te pakken. Dit laat zien dat wetsevaluatie en evaluaties in de bredere context van onderzoek voor beleid worden geplaatst. Beide instituten besteden een groot deel van hun onderzoek aan externe

⁹ WODC doet veel onderzoek en evaluaties in opdracht van het ministerie van VenJ en in uitzonderlijke gevallen van het ministerie van BZK ook. Verder is WODC ook betrokken bij de onderzoekprogrammering van het ministerie van VenJ door onderwerpen en thema's op agenda te zetten (WODC, 2015; Interview 8 juni 2015).

onderzoeksteams uit¹⁰, maar besteden veel aandacht aan kwaliteitsstandaarden en -criteria voor zowel het proces van evalueren als de inhoud van evaluaties.

Het ministerie van IenM heeft ook een onderzoeksinstituut: het Kennisinstituut voor Mobiliteitsbeleid (KiM). Omdat KiM echter geen ex post evaluaties verricht, rekenen we het ministerie van IenM als een departement waarbij de evaluatiefunctie intern is ingericht. De andere ministeries hebben geen aparte onderzoeksinstituten en regelen de evaluatiefunctie intern. Dat betekent concreet dat veel evaluatie die in de (meerjarige) onderzoeksprogrammering worden opgenomen, uitbesteed worden (zie hieronder het paragraaf over intern versus extern uitgevoerde evaluaties). Uit tabel 3 blijkt dat vaak niet één, maar meerdere directies betrokken zijn bij het verrichten en uitbesteden van evaluaties.

Bij het ministerie van OCenW zijn de onderzoekscoördinatoren van de drie DG's – DG Primair en Voortgezet Onderwijs, DG Hoger Onderwijs, Beroepsonderwijs, Wetenschap en Emancipatie en DG Cultuur en Media – verantwoordelijk voor het uitbesteden van evaluaties. De onderzoekscoördinatoren zijn tevens voor een deel aangesteld als beleidsmedewerkers (OCenW, 2015a). Directie Kennis ondersteunt de onderzoekscoördinatoren bij het opstellen van startnotities voor evaluaties en bij de keuze voor de geschikte methodologische aanpak van evaluaties. Iedere evaluatie is belegd bij een van de beleidsdirecties waarbij een van de beleidsmedewerkers als dossierhouder fungeert (Interviews 1 juli 2015, 26 augustus 2015). De dossierhouder is eindverantwoordelijk voor het evaluatieproces en het eindrapport. Omdat het evaluatieonderzoek een klein aandeel vormt van het onderzoek dat in opdracht van het ministerie van OCenW wordt gedaan, komt het vaak voor dat de dossierhouder voor de eerste keer een dergelijk evaluatie-dossier of wetsevaluatie-dossier moet begeleiden (Interview 26 augustus 2015). De onderzoekscoördinatoren van ieder DG, die beter zicht hebben op evaluatieonderzoek dat in de afgelopen jaren gedaan is, helpen de dossierhouder bij aspecten zoals de formulering van de vraag- en probleemstelling, de opzet van de startnotitie en de keuze van de begeleidingscommissieleden (Interview 26 augustus 2015). Er is hier sprake van informele borging van de kwaliteit van evaluaties.

¹⁰ WODC verricht ongeveer 30-40% van het onderzoek (waaronder ook wetsevaluaties) intern, terwijl ZonMw subsidies verstrekt voor de uitvoering van wetsevaluaties door externe onderzoeksteams, (Interviews 9 april 2015 en 8 juni 2015). Het merendeel van het onderzoek wordt bij WODC via de afdeling Externe Wetschappelijke Betrekkingen uitbesteed. WODC verricht soms ook evaluaties in opdracht van het ministerie van BZK (Interview 8 juni 2015).

Directie Kennis heeft een goed beeld van wat ze met evaluaties willen bereiken en aan welke minimale voorwaarden evaluaties moeten voldoen. De directie probeert het effectiviteitsonderzoek en vooral quasi-experimenteel onderzoek te stimuleren en daarbij ook de evaluatieopdrachten, indien mogelijk, op een experimentele manier te laten uitvoeren. Effectiviteitsonderzoek betreft echter slechts een klein deel van het hele pakket evaluatieonderzoek bij OCenW. Er zijn tot nu toe ongeveer tien afgeronde en tien lopende onderzoeken naar de effectiviteit van het beleid en wetgeving op allebei beleidsterreinen van onderwijs en onderzoek (Interview, 1 juli 2015). Directie Kennis heeft geconstateerd dat ongeveer twee derde van alle onderzoeken die in opdracht van het ministerie worden uitbesteed, effectiviteitsonderzoeken zouden kunnen zijn (OCenW, 2015b). De Directie laat een onderzoeksdesign door CPB en SCP opstellen als richtlijn voor de startnotitie van ieder nieuw onderzoek. De keuze om deze onderzoeksdesign wel of niet te gebruiken ligt bij de directie (Interviews 1 juli 2015 en 26 augustus 2015). Dit gebeurt in de praktijk echter nog heel weinig (Interview 26 augustus 2015). Directie Kennis ondersteunt de dossierhouders en de beleidsdirectie zowel op eigen initiatief als op het initiatief van de dossierhouders. Het ministerie van OCenW werkt met raamcontracten. Voor ieder thema zijn van tevoren acht onderzoeksbureaus geselecteerd die offertes kunnen indienen als een nieuw onderzoek door het ministerie wordt uitbesteed. De selectie is gebaseerd op de ervaring en de capaciteit die de onderzoeksbureaus in huis hebben. Een overzicht van alle onderzoeksinstituten waarmee het ministerie raamcontracten heeft of subsidies voor onderzoek verleent, staat in de Kennisagenda 2015 (OCenW, 2015c).

Het ministerie van SZW belegt de ondersteunende functie op het gebied van evaluaties bij FEZ en bij de onderzoekscoördinatoren van ieder DG. De onderzoekscoördinatoren staan onder leiding van een *Chief Science Officer*. De Directie Financiële en Economische Zaken (FEZ) speelt een belangrijke rol in de inrichting van de evaluatiefunctie bij het ministerie van SZW. FEZ borgt de kwaliteit van de beleidsinformatie en -evaluaties. FEZ is verantwoordelijk voor het systematisch evalueren binnen SZW en ziet toe op de kwaliteit en de benutting van beleidsevaluaties en -doorlichtingen binnen SZW” (SZW, 2012). FEZ draagt zorg voor de procesmatige en inhoudelijke kant van het evaluatieonderzoek waaronder kennisverspreiding over het aanpakken van evaluatieonderzoek en beleidsdoorlichtingen, over de selectie van onderzoeksmethoden en over het proces van begeleiden van evaluatieonderzoek. De beleidsdirecties worden daarnaast ondersteund door onderzoekscoördinatoren onder leiding van

een *Chief Science Officer* (CSO). Sinds enkele jaren is de uitbestedingsprocedure bij het Bureau Inkoop ondergebracht. Ter bevordering van de integriteit van de procedure krijgen alle bureaus dezelfde informatie en de offerteverzoeken worden centraal en vertrouwelijk behandeld. Het Bureau Inkoop speelt geen rol bij de vraagstelling en probleemstelling van het onderzoek, maar begeleidt slechts procesmatig het uitbestedingstraject (Interview, 15 juli 2015).

Het ministerie van EZ heeft een aantal afdelingen dat zich richten op de kwaliteitszorg voor onderzoek, waaronder ook wetsevaluaties: de Beleidsevaluatiecommissie (BEC), een Regiegroep binnen het DG Bedrijfsleven & Innovatie en het Behavioural Insights Team (BIT). Deze houden zich vooral bezig met het ontwikkelen en toepassen van instrumenten om te kwaliteit van effectevaluaties te stimuleren. BEC is sinds 2013 een gezamenlijk secretariaat van de Directie Algemene Financiële Politiek (AEP) en FEZ en draagt zorg voor de kwaliteit van evaluatie EZ-breed. BEC is ontstaan na de publicatie van het rapport *Durf te meten* van Commissie Theeuwes waarin meer aandacht wordt gevraagd voor de effectiviteit van de beleidsmaatregelen en voor het gebruik van kwantitatieve methoden (Commissie Theeuwes, 2012; Interview 9 juli 2015). “Bij BEC gaat het voornamelijk om de ex post evaluaties en de probleemstellende notitie, voordat je met een offertetraject begint en een bureau betreft, moet je de startnotitie (en het logisch model van je interventie) goed hebben opgesteld” (Interview 9 juli 2015). “BEC is één structuur op basis waarvan je iedereen weet te bereiken, aan elkaar blootstellen, de do’s en don’ts delen” Interview, 9 juli 2015). Naast BEC heeft het ministerie van EZ ook een Regiegroep die zich over het uitgevoerde beleid buigt vanuit de vraag: wat levert het beleid de maatschappij op? (Interview 9 juli 2015). De Regiegroep functioneert tegenwoordig alleen voor het DG Bedrijfsleven & Innovatie en komt gemiddeld eens in de zes weken bij elkaar. Als een lid van de Directie van DG B&I met een nieuw instrument bezig is, dan biedt de Regiegroep begeleiding door een monitoring- en effectmetingsplan op te stellen en door de dossierhouder aan te moedigen om te denken wat hij nodig heeft om de evaluatie straks in goede banen te leiden. Het BIT van EZ wordt ook gezien als een innovatief team dat ernaar streeft de gedragswetenschappen een plek in evaluatieonderzoek bij overheid te geven.

Bij het ministerie van IenM worden de wetsevaluaties vooral via Directie Financiën, Management en Control (FMC) en Hoofddirectie Bestuurlijke en Juridische Zaken (HBJZ) uitbesteed. IenM wordt daarnaast bijgestaan door twee kennisplatformen, het Behavioural Insights Team (BIT) en het Kennisinstituut voor Mobiliteitsbeleid (KiM). BIT is een netwerk van

kennisinstellingen op het gebied van gedrag. Het ondersteunt het ministerie van IenM met het toepassen van gedragsinzichten bij het maken en implementeren van beleid en wetgeving (Kennisplatform Crow, 2014). Beleidsevaluatie is een van de zes kennislijnen van KiM (KiM, 2015). Omdat het instituut echter exclusief ex ante evaluatie verricht, is het hier buiten beschouwing gelaten.

Beleidsevaluaties worden bij het ministerie van Buitenlandse Zaken door de Inspectie Ontwikkelingssamenwerking en Beleidsevaluaties (IOB) uitgevoerd. Onder de taken van IOB vallen de programmering en uitvoering (intern en extern) van onafhankelijke beleidsevaluatie en de verzameling van benodigde kennis om de nieuwe interventies doelgericht voor te bereiden (Ministerie van Buitenlandse Zaken, 2009; Ministerie van Buitenlandse Zaken, 2015; Interview 2 september 2015). De aanpak en methodologie die IOB bij beleidsevaluaties gebruikt, is vastgesteld in de handreiking *Evaluatiebeleid en richtlijnen voor evaluaties* uit 2009.

Box 5.1.

- ***Twee departementen (VenJ en VWS) verrichten en besteden (wets)evaluatie met behulp van semionafhankelijke onderzoeksinstituten (WODC en ZonMw).***
- ***Zes uit acht departementen hebben de evaluatiefunctie intern ingericht.***
- ***Er is een diversiteit aan combinaties van directies die de evaluatiefunctie ondersteunen. Ieder ministeries heeft zijn eigen combinatie van directies.***

De positionering van wetsevaluatie binnen het veld van evaluatieonderzoek

Bij alle departementen is het aantal wetsevaluaties zeer klein in vergelijking met het aantal beleidsevaluaties en het evaluatieonderzoek die door departementen worden verricht of uitbesteed. Ieder ministerie maakt jaarlijks tenminste één planning van alle soorten onderzoek bekend als bijlage in het document Rijksbegroting, dat ieder ministerie jaarlijks bij het ministerie van Financiën indient. Het onderzoek is onderverdeeld in twee categorieën: (1) beleidsdoorlichtingen en (2) ander onderzoek naar doeltreffendheid en doelmatigheid. Er is geen apart kopje voor wetsevaluaties. Deze kunnen zowel in de eerste als in de tweede categorie vallen. Dit laat zien dat wetsevaluaties als onderdeel van evaluatieonderzoek worden gezien en dat er geen aparte structuur en programmering hiervoor bestaat (meer over de evaluatieprogrammering in paragraaf 5.2.2). Wetsevaluaties worden gezien als onderdeel van ex post evaluatieonderzoek in het algemeen en

het laatste wordt beschouwd als een beleidsaangelegenheid (Interview 9 juli 2015). Daardoor ligt de verantwoordelijkheid voor wetsevaluaties bij de beleidsdirecties en niet bij de directies Wetgeving. De enige instelling die een apart programma voor de evaluatie van wet- en regelgeving heeft ontwikkeld is ZonMw, dat wetsevaluaties uitbesteedt in opdracht van het ministerie van VWS.¹¹ De hoofdlijnen van het programma worden in de paragraaf over het evaluatiekader (5.2.3) uitgelegd.

Box 5.2

- ***Slechts één ministerie (VWS) uit de acht onderzochte departementen werkt met een speciaal programma voor de evaluatie van wet- en regelgeving (Programma Evaluatie Regelgeving van ZonMw).***
- ***De andere departementen beschouwen wetsevaluaties als een vorm van evaluatieonderzoek in het breed. Daar hebben ze ook nauwelijks programma voor.***

Intern uitgevoerde wetsevaluaties versus uitbestede wetsevaluaties

Zoals aan het begin van dit onderzoek verwacht, is uit de interviews gebleken dat wetsevaluaties grotendeels *uitbesteed* worden aan externe partijen zoals academische onderzoekteams en onderzoeksbureaus. Een verklaring hiervoor is het feit dat het verrichten van evaluaties slechts een van de taken van beleidsmedewerkers is, naast het maken van beleid, het beantwoorden van vragen uit de Tweede Kamer, het verzamelen van gegevens over een bepaald maatschappelijk probleem en het monitoren van de invoering van het beleid (Interview 26 augustus 2015). Het doen van wetsevaluaties wordt vaak aan externe onderzoekers overgelaten, ook omdat ze vaak specialistische kennis vereisen. Een andere reden voor het uitbesteden van (wets)evaluaties is het gebrek aan tijd en capaciteit bij beleidsdirecties (Interview 26 augustus 2015b).

Bij de ministeries van VenJ, BZK en VWS worden soms wetsevaluaties ook intern uitgevoerd door de semionafhankelijke onderzoeksinstituten WODC en ZonMw. Het merendeel van wetsevaluatie wordt echter bij alle departementen uitbesteed. Bij OCenW wordt bijna al het onderzoek uitbesteed, grotendeels via raamcontract. Een klein deel van dit evaluatieonderzoek wordt via het Nationaal Regieorgaan Onderwijsonderzoek (NRO) verricht. Bij uitbestede

¹¹ Niet al het onderzoek dat het ministerie van VWS uitbesteedt, gaat echter via ZonMw. Een reden hiervoor is dat de uitbestedingsprocedure via ZonMw aan veel standaarden moet voldoen waardoor het evaluatietraject langduriger is in vergelijking met (wets)evaluaties dat direct door het ministerie worden uitbesteed (Interview 9 april 2015).

wetsevaluaties doet OCenW soms intern interviewrondes, vooral voor interne kennisdoelen (Interview 1 juli 2015).

Bij SZW verschilt de ratio intern versus extern onderzoek sterk per beleidsdirectie. De beslissing hierop hangt af van de capaciteit binnen de directie, de tijd en de expertise van interne medewerkers. Zo is de beleidsdoorlichtingen van artikel 6 van de begroting zelf door het ministerie gedaan, terwijl bij artikel 7 de synthese zelf wordt gedaan, en de onderliggende studies zijn uitbesteed. Anderzijds is de beleidsdoorlichting van het artikel 11 over het uitvoeringsbeleid van het ministerie helemaal uitbesteed (Interview 15 juli 2015).

Box 5.3

- *Het grootste deel van evaluaties worden door departementen uitbesteed. Departementen kunnen evaluaties meestal niet intern uitvoeren omdat:*
- *Evaluaties vereisen zowel inhoudelijke als methodologische specifieke kennis, waardoor onderzoekers het meest geschikt zijn om evaluaties uit te voeren;*
- *Zorg dragen voor evaluaties is slechts één van de taken van beleidsmedewerkers, naast het maken van beleid, het beantwoorden van de Kamervragen etc. Daardoor concentreren departementen zich op het verbeteren van het proces van het opstellen van startnotities en op het uitbestedingsproces van evaluaties.*
- *Beleid en wetten die door het ministerie worden gemaakt en (deels) uitgevoerd, worden door externe partijen (onderzoeksbureaus of academische onderzoeksteams) uitgevoerd. Dit verhoogt in theorie de objectiviteit en neutraliteit van evaluaties.*

Doelstellingen van wetsevaluatie

Er bestaan geen documenten – rijksbreed of bij departementen – waarin de doelstellingen van wetsevaluaties expliciet worden geformuleerd (met uitzondering van ZonMw, zie paragraaf 5.2.3 onder). De richtlijnen voor het doen van wetsevaluaties vallen samen onder de richtlijnen voor beleidsevaluaties en voor andere soorten onderzoek. Evaluaties en onderzoek moeten de vraag beantwoorden of en in hoeverre het beleid of de wet doelmatig en doeltreffend is geweest en welke gewenste en ongewenste effecten het gevoerde beleid of wet teweeg heeft gebracht (RPE, 2006). Alle geplande ex post onderzoeken naar effectiviteit – zowel evaluatief van aard als anderszins – moet aan één van de beleidsartikelen worden gekoppeld en moet de invulling van beleidsdoorlichtingen dienen (RPE, 2006). Uit alle onderzoeken zou idealiter conclusies worden

getrokken over de doeltreffendheid en doelmatigheid van het gevoerde beleid. De nadruk op doeltreffendheid en doelmatigheid laat zien dat de hoofddoelstelling van evaluatieonderzoek het *verantwoorden* van het gevoerde beleid is (RPE, 2006; Financiën, 2003).

De meerderheid van de respondenten afkomstig uit ondersteunende stafdirecties (zoals, Kennisdirecties, FEZ), maar ook uit beleidsdirecties gaf echter aan dat niet al het onderzoek, dat in de evaluatie- en onderzoeksplanning van de Rijksbegroting staat, ook daadwerkelijk bruikbaar is bij de invulling van de beleidsdoorlichtingen. Bij sommige departementen is het de afgelopen jaren een worsteling geweest om beleidsdoorlichtingen te vullen. De resultaten uit procesevaluaties, waarin bijvoorbeeld de invoering van een wet wordt beschreven, zijn volgens de richtlijnen voor beleidsdoorlichting van het ministerie van Financiën ontoereikend materiaal voor beleidsdoorlichtingen. (Interviews 1 juli 2015 en 16 juli 2015). Anderzijds stellen respondenten dat veel succes is geboekt op het gebied van evaluatieonderzoek gelet op hun perceptie dat “10 jaar geleden de evaluatie naar de effectiviteit van een interventie bijna ondenkbaar was” (Interview, 1 juli 2015).

Box 5.4

- *In het huidige evaluatiestelsel dient evaluatie vooral de verantwoordingsfunctie. Als goede evaluaties worden in dit stelsel veelal de evaluaties gezien die de doelmatigheid en doeltreffendheid van de onderzochte interventies kunnen aantonen.*
- *Er wordt nadruk gelegd op de verantwoordingsfunctie ten schade van de leerfunctie van evalueren.*
- *De meerderheid van departementen doet evaluaties om materiaal te verzamelen waarmee beleidsdoorlichtingen kunnen ingevuld worden.*
- *De nadruk op beleidsdoorlichtingen leidt tot een voorkeur voor effectiviteitsevaluaties. Procesevaluaties en evaluaties die geen causaal verband tussen de interventie en de geconstateerde effecten kunnen aantonen, worden vaak als weinig interessante evaluaties beschouwd.*

5.2.2 Evaluatieplanning en -programmering

Een belangrijke voorwaarde voor de evaluatiecapaciteit is volgens de wetenschappelijke literatuur de structurele planning van evaluaties. Planning moet van tevoren, op een gedegen en uitgebreide manier worden gedaan, als men evaluaties van goede kwaliteit wil verrichten (Bourgeois & Cousins, 2013; Cooksy, 2012; Nielsen et al, 2011; Van Aeken, 2011). De veronderstelling in deze

theorieën is dat door de evaluaties te programmeren men eerder met het verzamelen van gegevens kan beginnen, wat de kans op effectenstudies vergroot.

Volgens het artikel 9 uit het RPE 2006 moet ieder ministerie in het Rijksbegrotingsdocument dat jaarlijks bij het Ministerie van Financiën wordt ingediend, een bijlage toevoegen met de programmering van alle effectenonderzoeken ex post en beleidsdoorlichtingen voor de volgende paar jaren. Dit programma is integraal, d.w.z. bevat alle soorten onderzoek, is afgestemd op beleidsterreinen en niet op de vorm van onderzoek (RPE, 2006). Alle ex post evaluaties en onderzoek worden in twee categorieën ingedeeld: beleidsdoorlichtingen en overig effectenonderzoek ex post. De evaluatieplanning wordt jaarlijks bijgesteld. De bijstelling houdt rekening met veranderingen in de steun van het Kabinet voor een bepaald beleid of onderwerp en de grote maatschappelijke veranderingen waardoor de focus op andere thema's komt te staan. Als gevolg van deze veranderingen kan een onderwerp minder relevant of actueel worden, waardoor het onderzoek of evaluatie wordt opgegeven (Interview 2 september 2015).

Ministeries maken geen aparte planning voor evaluaties en dus ook niet voor wetsevaluaties in het bijzonder. Alleen die evaluaties worden gepland die aan een van de beleidsartikelen worden gekoppeld en voor zover de evaluaties materiaal voor de invulling van beleidsdoorlichting kunnen opleveren. Evaluaties die niet aan deze twee voorwaarden voldoen, worden niet gepland, waardoor het moeilijk is om een overzicht van alle evaluaties te verkrijgen. Omdat wetten en beleid vaak met elkaar worden gecombineerd om een bepaald effect in de maatschappij teweeg te brengen, maken departementen ook geen aparte planning voor wetsevaluaties (Interview 15 juni 2015). Wetsevaluaties betreffen bovendien slechts een klein deel van alle evaluaties (zie de Rijksbegrotingsdocumenten op de website van ieder ministerie). Bij sommige departementen heeft de Directie Wetgeving een wetgevingskalender met daarbij de evaluatietermijnen van wetten. Terwijl deze lijst gebaseerd is op de evaluatiebepalingen in de wet, heeft het KWJZ van VenJ al eerder geconstateerd dat veel evaluaties van wetten met een evaluatiebepaling helemaal niet of later worden uitgevoerd dan in de planning staat (Intern gesprek met KWJZ). De respondenten hebben bevestigd dat evaluatieonderzoek niet automatisch de evaluatieplanning volgt (Interview 1 juli 2015).

Alle respondenten gaven aan dat een omvattende en tijdelijke programmering van (wets)evaluaties veel voordelen heeft, maar dat er op dit gebied nog een wereld te winnen is. Er is

een kloof tussen het aantal wetsevaluaties dat gepland staat en de wetsevaluaties die daadwerkelijk worden verricht en op tijd afgerond. Respondenten stelden dat vele dossierhouders pas over de vraag- en probleemstelling van een evaluatieonderzoek beginnen te denken op het moment dat de evaluatiebepaling nadert (Alle interviews). Op dat moment is het vaak te laat om een goede effectevaluatie te laten uitvoeren. Om de effecten van de interventie te meten moet men beginnen met de gegevensregistratie en -verzameling al voor de invoering van de wet of maatregel of meteen daarna. Bij evaluatieonderzoeken waar te laat is begonnen met het denken over de probleemstelling en de doelen die de wet of de maatregel beoogt te bereiken, is de kwaliteit vaak onder de maat of wordt het rapport niet op tijd ingeleverd (Interview 26 augustus 2015).

Bij het ministerie van SZW voert FEZ gesprekken met iedere beleidsdirectie over de evaluatieagenda en evaluatieplanning. Dit gebeurt vooral in verband met beleidsdoorlichtingen en met als doel gedegen effectenevaluaties te laten verrichten die relevant zijn voor de toelichting van de beleidsartikelen. Ieder onderzoek en evaluaties moeten dus goed aansluiten bij één van de beleidsartikelen. Verder proberen de onderzoekcoördinatoren bij SZW de dossierhouders te stimuleren om pro-actiever te denken over evaluatieplanning: niet alleen over verplicht onderzoek, maar ook over niet verplichte nummers (Interview 15 juli 2015).¹²

Bij het ministerie van VenJ worden jaarlijks twee programmeringsrondes van onderzoek door het WODC geïnitieerd (WODC, 2015). Voor iedere programmeringsronde worden de onderzoekswensen zowel binnen WODC als binnen beleidsdirecties geïnventariseerd. De lijst met onderzoeksvoorstellen wordt in het interne Management Team besproken. Het MT overlegt samen met de onderzoekscoördinatoren van DG's welke onderzoeken in de volgende periode mogen worden verricht. Vervolgens wordt de keuze gemaakt of het onderzoek door de interne onderzoekers van WODC wordt uitgevoerd of via EWB (Externe Wetenschappelijke Betrekkingen) wordt uitbesteed.

¹² Het ministerie van SZW heeft ook een programmering voor het onderzoek dat door SZW is geïnitieerd (evaluatie vallen meestal hierbuiten). De programmering wordt opgenomen in de *Hoofddlijnennotitie Kennisvragen* (vóór 2014 bekend onder de naam *Strategische Kennisagenda*). De programmering van 2015 bevat 10 thema's. Elk onderzoek dat aangevraagd is (ongeacht of dit een verplichtnummer van de Kamer is of niet) moet aan de 10 thema's worden gekoppeld. De *Hoofddlijnennotitie* "is een kennisinstrument om te bundelen en te focussen" (Interview 15 juli 2015). Deze thema's ontstaan in het interne overlegproces van de *Regiegroep Kennis*, waarin de directeurs van alle kolommen samen met FEZ zitting nemen. Daarnaast heeft het ministerie van SZW ook een *KennisPartnerberaad*, een overleggroep met externe partijen. Deze partijen krijgen geen uitbesteed onderzoek, maar worden wel gesubsidieerd om met SZW over de onderzoeksprogrammering te overleggen (Interview 15 juli 2015).

Het ministerie van VWS heeft zelf geen evaluatieplanning, afgezien van de begroting. Dit komt vooral omdat evaluatie gezien wordt als een beleidsinstrument en beleidsaangelegenheid (Interview, 9 april 2015). Dit maakt evaluatie onderdeel van de beleidscyclus. Evaluaties worden uitgevoerd indien het Kabinet of de minister hierom vraagt. Wetsevaluaties worden óf via de Commissie Evaluatie Regelgeving (CER) van ZonMw uitgezet óf direct door de betrokkene directies uitbesteed. Het uitbesteden van wetsevaluaties via ZonMw is een wat langduriger en intensiever proces omdat CER nadrukkelijk aandacht besteedt aan de kwaliteit van de uitbestedingprocedure alsook aan de kwaliteitsstandaarden van het rapport (meer hierover in paragraaf 5.2.3). CER stuurt op de kwaliteit van de inhoud van het rapport door het onderzoek aan het beste onderzoeksteam te gunnen en door zitting te nemen in de begeleidingscommissie. De begeleidingscommissie is volgens ZonMw één van de belangrijkste factoren in de borging van de kwaliteit van evaluaties. Alle relevante disciplines – afhankelijk van het domein van de wet – worden vertegenwoordigd, zoals rechten, gezondheidszorg, psychologie etc.

Bij het ministerie van OCenW maken beleidsdirecties hun eigen onderzoeksprogrammering en zijn zelf hiervoor verantwoordelijk, maar de indruk bij de directie Kennis is dat het programma *Zicht op effectiviteit*, bedoeld voor de stimulering van effectevaluaties echter “in een omgeving landt waar niet gepland wordt nagedacht over het onderzoek dat gedaan moet worden” (Interview 1 juli 2015; meer over dit programma in paragraaf 5.2.3.1). Als het om onderzoek in het brede gaat, maakt Directie Kennis wel ieder jaar een onderzoeksprogrammering. Dit wordt echter niet vanuit de gedachte gedaan welk onderzoek nodig is om de beleidsdoorlichting te laten invullen. Er is meer behoefte bij het ministerie van OCenW aan een overzicht over welk onderzoek is gedaan om de beleidsdoorlichtingen te laten vullen (Interview 1 juli 2015).

Het ministerie van EZ heeft een intern document *Handreikingen evaluatieonderzoek EZ* laten opstellen. Daar staan enkele minimale richtlijnen voor evaluatieprogrammering. “Reeds bij het opzetten van een instrument wordt idealiter nagedacht over het evalueren van het instrument. Voor ieder instrument zou duidelijk moeten zijn wanneer de evaluatie plaatsvindt.” (Bonte, *Handreikingen evaluatieonderzoek EZ*: 9). Een goede evaluatieprogrammering is vereist om de effecten van subsidies en beleid door middel van evaluatie te kunnen aantonen. De programmering van het evaluatieonderzoek wordt afgestemd op de beleidsdoorlichtingen zowel qua timing als qua dekkendheid. Omdat ex post evaluaties bij EZ primair als een beleidsaangelegenheid en onderdeel

van de beleidscyclus worden gezien, speelt Directie Wetgeving geen opvallende rol bij de programmering hiervan. De beleidsdirecties zijn verantwoordelijk voor de evaluatieprogrammering, in overleg met FEZ (Bonte, *Handreikingen evaluatieonderzoek EZ*: 9).

Het ministerie van BZ heeft een programmering voor vijf jaar, die jaarlijks wordt bijgesteld. De programmering betreft geen wetsevaluaties. Daarbij wordt rekening gehouden met de timing voor beleidsdoorlichtingen. De programmering betreft beleid en subsidies. Er worden nauwelijks wetten door het ministerie van BZ gemaakt en dus ook niet geëvalueerd. FEZ en IOB werken nauw samen bij de programmering van evaluatieonderzoek met input vanuit de vier DG's van het ministerie: DG Europese Samenwerking, DG Internationale Samenwerking, DG Politieke Zaken en DG Buitenlands Economisch Beleid. IOB is verantwoordelijk voor de programmering van het DG Internationale Samenwerking en heeft een adviserende rol bij de evaluatieprogrammering van de anderen drie DG's (IOB, 2009: 12). De meest voorkomende oorzaken van de bijstelling van evaluatieprogrammering zijn de veranderingen in de steun van het Kabinet voor een bepaald beleid en de grote maatschappelijke veranderingen zoals oorlog in een ontwikkelingsland (Interview 2 september 2015). De evaluatieprogrammering bevat alleen beleidsdoorlichtingen en andere evaluaties naar doelmatigheid en doeltreffendheid. De respondent van IOB gaf echter aan dat veel evaluaties *de facto* procesevaluaties zijn, omdat de effecten van een bepaald instrument of interventie moeilijk isoleerbaar zijn. Soms kunnen evaluaties wel effecten constateren en meten, maar ze kan de effecten niet direct en exclusief aan het beleid toeschrijven (het attributie-probleem). Veel evaluaties beperken zich daarom tot de reconstructie van de implementatie van het beleid en de activiteiten en instrumenten die ontwikkeld zijn om de bepaalde effecten teweeg te brengen. (Interview, 2 september 2015). De respondent gelooft niettemin dat men leerzame conclusies uit procesevaluaties kan trekken. Het feit dat het causaal verband tussen beleid en de geconstateerde effecten niet hard kan worden gemaakt, kan vraagtekens opleveren over de beleidstheorie waarop het beleid steunt. De onmogelijkheid om de geconstateerde effecten aan het beleid (instrument, interventie) toe te schrijven, kan betekenen dat de beleidstheorie niet klopt en dat de effecten ook via andere mechanismen bereikt worden dan in de beleidstheorie wordt aangenomen.

Samenvattend is een op tijd gemaakte evaluatieprogrammering een noodzakelijke voorwaarde, maar geen garantie voor een goed verloop en bruikbaarheid van evaluaties. Enerzijds kunnen externe omstandigheden zoals veranderingen in de politieke context de geplande evaluatie

irrelevant maken en de planning op haar kop zetten. Anderzijds kunnen evaluatie volgens de evaluatieprogrammering worden verricht en toch onbruikbaar blijken te zijn bij de invulling van beleidsdoorlichtingen omdat de evaluaties de vraag naar de doelmatigheid en doeltreffendheid van het beleid niet kunnen beantwoorden.

Box 5.5

- *Alle departementen maken een standaard (meerjarige) onderzoek- en evaluatieprogrammering. Alleen ex post evaluaties en onderzoek naar effectiviteit wordt hierin vermeld. De programmering focust sterk op beleidsdoorlichtingen en al het geplande onderzoek moet aan een van de beleidsartikelen in de begroting worden gekoppeld.*
- *Politieke belangen kunnen met programmeringen conflicteren. De programmering staat niet vast, maar wordt jaarlijks bijgesteld op basis van politieke belangen en maatschappelijke veranderingen.*
- *In theorie moet de planning van evaluaties tegenwicht bieden aan de politieke belangen en druk. In de praktijk hebben politieke belangen een grote invloed op de uitvoering van de geplande evaluaties. Minder politieke interesse in een thema of probleem kan op twee manieren tot wijzigingen in de evaluatieprogrammering leiden: (1) geplande evaluaties worden helemaal niet uitgevoerd; (2) geplande evaluaties worden wel uitgevoerd, maar door het verminderende belang hiervoor, zijn de bevindingen als niet meer relevant beschouwd waardoor de evaluatie geen/lage doorwerking op besluitvorming heeft.*
- *Er wordt geen aparte programmering van evaluaties of wetsevaluaties gemaakt.*
- *Evaluatieonderzoek loopt niet gelijktijdig met de evaluatieprogrammering. De programmering van evaluaties biedt in de praktijk geen garantie dat evaluaties daadwerkelijk worden uitgevoerd.*
- *Een probleem dat bij alle departementen speelt is dat ze te laat beginnen met evalueren (gegevensverzameling, vraagstelling, uitbesteding evaluaties).*
- *Alle departementen hebben een betere planning van evaluaties en wetsevaluaties nodig.*
- *Alle departementen zouden baat hebben bij eerder beginnen te denken over evaluaties die in de nabije toekomst moeten worden verricht en over welke instrumenten ingezet moeten worden om de vraagstelling van de evaluaties te beantwoorden.*

5.2.3 Evaluatiekader

Om informatie over het evaluatiekader te verzamelen, heeft deze studie onderzocht welke programma's voor (wets)evaluaties bij departementen bestaan, welke doelen deze programma's nastreven en hoe deze programma's in de praktijk van evaluatie vorm worden gegeven. Daarnaast is ook gekeken welke handreikingsboeken, richtlijnen of leidraden ministeries hebben ontwikkeld voor het uitvoeren, aanbesteden en begeleiden van evaluatieonderzoek, en hoe bekend beleidsmedewerkers hiermee zijn. Deze handreikingen omvatten concrete indicaties voor de stappen die dossierhouders moeten zetten bij de verschillende fases van het evaluatietraject. De assumptie afgeleid uit de wetenschappelijke literatuur is dat hoe meer instrumenten, arrangementen en voorzieningen een ministerie heeft voor de begeleiding van het proces van evalueren, hoe beter gestroomlijnd de evaluatiefunctie is en hoe hoger de standaarden voor evaluaties (Jacob et al., 2015; Van Aeken, 2011; Furubo et al., 2002).

Tabel 4 presenteert een overzicht van de belangrijkste programma's, handreikingen en richtlijnen die gebruikt worden bij verschillende departementen om de evaluatie vorm te geven. Tabel 4 geeft ook aan of het document openbaar toegankelijk is (dus ook buiten de overheid), voor intern gebruik is bedoeld (binnen het ministerie) of verspreid is over departementen voor rijksbreed gebruik. Interne documenten zijn bedoeld als hulpmiddel voor beleidsmedewerkers die evaluatieaanvragen opstellen. De *Handreiking Beleidsdoorlichting* van het ministerie van Financiën is bedoeld als richtlijn en hulpmiddel voor alle departementen bij het opstellen van evaluaties, effectiviteitsonderzoek en beleidsdoorlichtingen. Openbaar toegankelijke documenten zijn bedoeld voor zowel beleidsmedewerkers als voor onderzoekers die offertes voor evaluaties bij departementen willen indienen.

Tabel 4 laat zien dat een diversiteit aan handreikingen en richtlijnen voor evaluatieonderzoek bij ministeries bestaat. Het nadeel hiervan is dat de evaluatiefunctie niet centraal op het rijksniveau is geborgd en dat de verschillende handboeken deels overlappen en deels verschillen. Een ander aspect dat opvalt, is dat deze handreikingen en richtlijnen slechts door de directie worden gebruikt en/of gedragen die hen heeft ontwikkeld. Vaak weten andere directies niet van het bestaan van deze handreiking en leidraden.

Tabel 4. Evaluatiekader bij de Nederlandse departementen

Ministerie	Programma's, handreiking en richtlijnen voor evaluatieonderzoek	Openbaar toegankelijk vs. voor intern gebruik
BZ	• <i>Evaluatiebeleid en richtlijnen voor evaluaties (2009)</i> van IOB	Openbaar toegankelijk
EZ	• <i>Handreiking evaluatieonderzoek EZ</i> (door FEZ bij EZ geschreven) • <i>Handreikingenboekje beleidskwaliteit en beleidsevaluatie</i> (door BEC geschreven) • <i>IBO procedures en richtlijnen (2013)</i> (procedurele richtlijnen voor de interdepartementale beleidsonderzoeken) • <i>Durf te meten (2012)</i> (Eindrapport Expertwerkgroep Effectmeting)	Intern gebruik Intern gebruik Verspreid over departementen Openbaar toegankelijk
Financiën	• <i>Handreiking Evaluatieonderzoek ex post (2003)</i> • <i>Handreiking Beleidsdoorlichting</i> (bedoeld voor rijksbreed gebruik)	Intern gebruik Verspreid over departementen
IenM	• <i>Nota Governance programmering en uitvoering beleidsdoorlichting/evaluatie (2013)</i>	Intern gebruik
OCenW	• <i>Notitie Zicht op effectiviteit (2010)</i> van Directie Kennis (bedoeld om de quasi-experimentele design te stimuleren door onderzoeksdesigns voor effectevaluaties door CPB te laten ontwerpen. Vier jaar na de invoering van het programma heeft Directie Kennis geconcludeerd dat 2/3e van de onderzoeken wel effectiviteitsstudies kunnen zijn) • <i>Zicht op Effectiviteit van Beleid (2011)</i>. (CPB Achtergronddocument. Studie naar evaluatieontwerpen voor onderwijs en wetenschapsmaatregelen) • <i>Verklarende evaluatie (2012)</i> van het Kohnstamm Instituut (handboek bruikbaar bij het opzetten van verklarend onderzoek)	Intern gebruik Openbaar toegankelijk Openbaar toegankelijk
SZW	• <i>Handreikingen voor beleidsdoorlichtingen</i> (door FEZ ontwikkeld met daarin formats voor de uitbesteding van het onderzoek, het opstellen van de startnotitie, planning, aandachtspunten bij onderzoeksvragen, de hoogte van vergoeding voor onafhankelijk onderzoek, de afspraken met onafhankelijke deskundigen en het opstellen van de nota aan minister en de kabinetsreactie).	Intern gebruik

VenJ	• Handboek EWB voor uitbesteed onderzoek (2008) • Wegwijzer EWB (2012) (bij contractonderzoek in opdracht van het WODC van VenJ (2012)) • RIPi's (Recently Introduced Policy Instruments and Interventions Theories, 2005) artikel door P. Kautto & J. Similä • Evaluatie van justitiële (beleids)interventies (2010). WODC-notitie door M. van Ooyen-Houben en F. Leeuw	Intern gebruik Openbaar toegankelijk Openbaar toegankelijk Openbaar toegankelijk
VWS	• Programma Evaluatie Regelgeving (PER, 1997; door ZonMw ontwikkeld). In 2003 is PER door een externe onderzoeker geëvalueerd en de aanbevelingen zijn in de nieuwe versie verwerkt. • Brochure Programma Evaluatie Regelgeving (2013) (een verkorte en voor iedereen toegankelijk versie van de oorspronkelijke programmatekst uit 1997).	Openbaar toegankelijk Openbaar toegankelijk

Programma's voor evaluatie: aanwezigheid, doelen en gebruik

In het volgende beschrijft deze studie in het kort de doelstellingen van twee programma's die bij het ministerie van OCenW en bij ZonMw (bij VWS) zijn ontwikkeld met als doelstelling het evaluatieonderzoek en de wetsevaluatie in het bijzonder te stimuleren. Deze zijn bedoeld vooral om beleidsmedewerkers handvat te bieden bij verschillende fases van het (uitbesteden van het) evaluatietraject.

De Directie Kennis van het ministerie van OCenW heeft vanaf oktober 2010 in samenwerking met CPB het programma *Zicht op effectiviteit* ontwikkeld. Het doel van dit programma is “de kwaliteit van evaluatieonderzoek, dat in opdracht van het ministerie OCenW wordt verricht, te verbeteren en de methodische opzet van dit type onderzoek voor zover mogelijk te harmoniseren.” (OCenW, 2015b). Directie Kennis zet zich in voor het stimuleren van een combinatie van experimenteel onderzoek en verklarende evaluaties. Dit doet zij door voor elke effectevaluatie een onderzoeksdesign te laten opstellen door CPB of SCP. Beleidsdirecties kunnen gebruik maken van de desbetreffende onderzoeksdesign bij het opstellen van offerteaanvragen voor evaluaties, maar zijn hierom niet verplicht. (OCenW, *Zicht op effectiviteit*, 2015; Reininga, 2014). Dat doen de beleidsdirecties echter heel weinig omdat volgens de

respondenten niet al het onderzoek de vraag naar effectiviteit toelaat (Interviews 1 juli 2015 en 26 augustus 2015). Sinds 2013 heeft Directie Kennis besloten onderzoeks-ontwerp te laten opstellen voor beleid dat aan drie voorwaarde voldoet: “(1) beleid moet een goede kans bieden op een robuuste effectevaluatie gekoppeld aan een goede verklarende evaluatie; (2) beleid moet tot de kern van het beleid van OCenW horen; (3) beleidsdirecties moeten serieuze belangstelling hebben voor deze aanpak” (OCenW, *Zicht op Effectiviteit*, 2015). Verder is Directie Kennis altijd betrokken bij quasi-experimenteel onderzoek. Bij andere soorten onderzoek bepaalt de verantwoordelijke beleidsdirectie of Directie Kennis bij wordt betrokken. In dat geval kan Directie Kennis gevraagd worden om mee te denken over aanbesteding van offertes en over de onderzoeksdesign of om zitting te nemen in de begeleidingscommissie.

Box 5.6

- *Het ministerie van OCenW heeft via Directie Kennis het programma “Zicht op effectiviteit” ontwikkeld. Het programma heeft als doel “de kwaliteit van evaluatieonderzoek, dat in opdracht van het ministerie OCenW wordt verricht, te verbeteren en de methodische opzet van dit type onderzoek voor zover mogelijk te harmoniseren.” (OCenW, 2015b). Directie Kennis zet zich in op het stimuleren van een combinatie van experimenteel onderzoek en verklarende evaluaties. Dit doet zij door voor elke effectevaluaties een onderzoeksdesign te laten opstellen door CPB of SCP.*
- *Beleidsdirecties kunnen gebruik maken van de desbetreffende onderzoeksdesign bij het opstellen van offerteaanvragen voor evaluaties, maar zijn hierom niet verplicht. Dat doen de beleidsdirecties echter heel weinig. Redenen hiervoor:*
 - *Dossierhouders beslissen zelf of ze de onderzoeksdesign van Directie Kennis wel of niet gebruiken.*
 - *Directie Kennis kan de dossierhouders proberen te overtuigen de onderzoeksdesign te gebruiken. Directie Kennis heeft echter geen stok achter de deur om de dossierhouders te dwingen de onderzoeksdesign te gebruiken.*

ZonMw, het onderzoeksinstituut dat wetsevaluaties in opdracht van het ministerie van VWS uitvoert en uitbestedt, heeft in 1997 een evaluatieprogramma voor wet- en regelgeving ontwikkeld. Het programma is in 2013 samengevat in de brochure *Evaluatie Regelgeving* die openbaar toegankelijk is op de website van ZonMw. Daarin staan de criteria vermeld die de Commissie Evaluatie Regelgeving hanteert bij de selectie van de evaluatie-uitvoerders. De brochure is de verkorte versie van de programmatekst uit 1997 en is bedoeld als richtlijn voor de

potentiële aanvragers en uitvoerders van evaluaties. Het *Programma Evaluatie Regelgeving* (PER) beschouwt wetsevaluaties als een combinatie van juridisch en sociaalwetenschappelijk onderzoek. Dit is het enige programma dat dit punt expliciet noemt. De combinatie van juridische en sociaalwetenschappelijke expertise is uniek voor wetsevaluaties. Respondenten van andere departementen zijn bewust van deze aparte status van wetsevaluaties en de uitdagingen die hiermee gepaard gaan bij het verrichten van wetsevaluaties. Omdat departementen (met uitzondering van ZonMw) geen afzonderlijke programma's en handreikingen voor wetsevaluaties hebben, besteden ze ook geen expliciete aandacht aan de combinatie van juridische en sociaalwetenschappelijke onderzoek en maken geen suggesties voor de aanpak van dit aspect.

PER is in 2003 geëvalueerd na vijf jaar bestaan (Olsthoorn-Heim, 2003). De aanbevelingen zijn in het nieuwe programma verwerkt.

Box 5.7

De zeven criteria voor evaluatie van Programma Evaluatie Regelgeving (ZonMw, 2013: 4-5):

- 1) wetsevaluaties moeten door een multidisciplinaire groep worden uitgevoerd met aandacht voor zowel het empirische als voor het juridische deel;***
- 2) kennismaken van de beginsituatie (voor zover mogelijk een nulmeting) moet onderdeel zijn van de evaluatie;***
- 3) wetsevaluatie beperkt zich niet tot de te evalueren wet, maar houdt ook rekening met de internationale en Europese context;***
- 4) wetsevaluaties kijken zowel naar gewenste als ook ongewenste effecten;***
- 5) wetsevaluaties moeten aandacht schenken aan de relatie tussen wetgeving en zelfregulering (criterium verbonden aan het subsidiariteitsbeginsel);***
- 6) elke nieuwe evaluatie moet naast haar eigen onderzoek ook gebruik maken van bestaande kennis, documenten en resultaten uit eerder onderzoek. Dit voorkomt verspilling van tijd en geld;***
- 7) en tot slot wetsevaluaties besteden aandacht aan diversiteit en aan de toegankelijkheid van alle bevolkingsgroepen tot de wet.***

Verder omvat de brochure een korte toelichting over de relatie tussen de opdrachtgever (ZonMw) en de opdrachtnemer (onderzoekers), de rol van CER, de rol van de begeleidingscommissie, de toekenningsprocedure van wetsevaluatie vanaf het moment van de aankondiging van toekomstige wetsevaluatie via de procedure van beoordeling van aanvragen tot de eindfase van de communicatie en implementatie van evaluatiebevindingen. Daarnaast bevat de brochure een lijst

van de wetsevaluaties die in de afgelopen jaren (tot en met 2014) zijn afgerond. Het voordeel van een dergelijke lijst is dat zij de vergelijking vergemakkelijkt tussen de geplande evaluaties en de afgeronde evaluaties (ZonMw, 2013: 11).

Daarnaast gebruikt CER bij de selectie van het onderzoeksteam een aantal kwaliteitscriteria voor evaluaties. Daaronder vallen de helderheid, originaliteit en haalbaarheid van de vraagstelling, de aansluiting van het plan van aanpak aan de vraagstelling, de theoretische en empirische onderbouwing, de kennis en ervaring van de aanvragers, de haalbaarheid van het project in het breed en de begroting. Deze lijst met kwaliteitscriteria voor wetsevaluaties is ook voor aanvragers toegankelijk en dienen als richtlijnen bij het schrijven van offertes (ZonMw, 2013; Interview 9 april 2015).

Box 5.8.

Criteria voor het toekennen van wetsevaluaties (ZonMw, 2013):

- ***helderheid, originaliteit en haalbaarheid van de vraagstelling;***
- ***de aansluiting van het plan van aanpak aan de vraagstelling;***
- ***de theoretische en empirische onderbouwing;***
- ***de kennis en ervaring van de aanvragers;***
- ***de haalbaarheid van het project in het breed;***
- ***de begroting.***

Naar aanleiding van het rapport *Effectiviteitsonderzoek bij de rijksoverheid* van de Algemene Rekenkamer, waarin is geconstateerd dat er weinig ex post evaluatie naar doelmatigheid en doeltreffendheid werd gedaan, is Directie Financiën, Management en Control (FMC) van het ministerie van IenM met de ontwikkeling van een programma begonnen. De ontwikkeling is nu nog gaande. Het programma beoogt aandacht te trekken voor de noodzaak om begrippen zoals doelmatigheid en doeltreffendheid al in de fase van het maken van wetten te expliciteren. Als de doelen van de wet niet duidelijk expliciet worden omschreven in de Memorie van Toelichting is het moeilijk te evalueren of de wet effectief is geweest. De evaluatie moet de assumpties achter de invoering van de wet reconstrueren, maar deze kunnen vaak niet uit de oorspronkelijke wet worden achterhaald. FMC bij IenM wordt bij de ontwikkeling van het programma door de interdepartementale beleidsdoorlichtingsgroep (IBP) bijgestaan.

De andere ministeries hebben geen programma's voor evaluatieonderzoek, maar werken met handreikingen, richtlijnen en leidraden.

5.2.3.2 Handreiking en richtlijnen voor (wets)evaluatie

Omdat er maar weinig wetsevaluaties zijn en deze worden beschouwd als een vorm van evaluatieonderzoek in het breed, hebben ministerie ook geen handreiking en richtlijnen specifiek voor wetsevaluaties ontwikkeld. De interne handreikingboekjes en richtlijnen die bij departementen bestaan, zijn bedoeld om een houvast te bieden aan medewerkers bij het opstellen van startnotities, de formulering van de vraag- en probleemstelling, de samenstelling van begeleidingscommissies, de begeleiding van het evaluatietraject en de verspreiding van evaluatiebevindingen.

Box 5.9

Handreikingen en richtlijnen voor (wets)evaluaties houden informatie en toelichting, bedoeld voor de dossierhouders over de volgende aspecten:

- ***hoe startnotities op te stellen;***
- ***hoe de vraag- en probleemstelling te formuleren;***
- ***hoe de begeleidingscommissie samen te stellen;***
- ***hoe het evaluatietraject te begeleiden;***
- ***hoe de evaluatiebevindingen te verspreiden.***

Over de behoefte en de noodzaak voor speciale handreiking en richtlijnen voor wetsevaluaties bestaat er geen overeenstemming tussen respondenten. De respondenten van VenJ die betrokken zijn bij het uitbesteden van evaluatieonderzoek, waaronder ook wetsevaluaties, zijn het niet met elkaar eens over de vraag of de algemene handreikingen en modellen die bij beleidsevaluaties worden gebruikt voldoende basis bieden voor wetsevaluatie. Als voorbeeld wordt de studie van Campbell over de experimentele design genoemd. Volgens een van de respondenten zijn de vijf niveaus in de benadering van Campbell niet geschikt om goede wetsevaluaties te verrichten, simpelweg omdat er bij wetsevaluaties meestal niet met controlegroep kan worden gewerkt. Het specifieke karakter van wetten is dat zij voor de hele bevolking of de hele doelgroep gelden (Bussmann, 2010). Het onttrekken van een deel van de doelgroep aan de werking van de wet zou het gelijkheidsbeginsel schenden. Een andere respondent deelt dit gezichtspunt over de

ongeschiktheid van algemene handreikingen en modellen voor wetsevaluaties niet. Klassieke standaarden voor evaluaties en onderzoek in het breed zouden evenzogoed toepasbaar zijn op wetsevaluaties (Interview 8 juni 2015).

Naast de notitie *Zicht op effectiviteit* afkomstig van Directie Kennis zijn er geen andere handreikingen of richtlijnen bekend bij de beleidsdirecties van OCenW. OCenW heeft ook geen aparte voorzieningen voor wetsevaluaties. Vaak is degene die het evaluatie-dossier beheert ook voor de eerste keer betrokken bij een dergelijk taak. Voor concrete vragen over alle fases van het evaluatietraject kunnen dossierhouders bij Directie Kennis aankloppen of bij één van de onderzoekscoördinatoren. Deze hebben een goed overzicht over de verschillende soorten onderzoeken en de evaluaties die binnen het ministerie van OCenW de afgelopen jaren zijn afgerond en kunnen daardoor goede tips geven in de beginfase van het opstellen van startnotities voor evaluaties. Het gebrek aan handreikingen en richtlijnen over hoe men met evaluaties moet omgaan, wordt ondervangen door informele consultaties en overlegmomenten tussen dossierhouders en ervaren senior medewerkers, zoals de onderzoekscoördinatoren van ieder DG. De gesprekken zijn altijd informeel, *ad hoc*, er zijn geen richtlijnen die men volgt of toepast, maar men leert van elkaar via het delen van informatie over hoe een gelijksoortig onderzoek eerder is aangepakt.

Het ministerie van EZ heeft twee handreikingsboekjes voor het doen van evaluatieonderzoek opgesteld. Het ene is geschreven door BEC, dekt de gehele beleidscyclus en is bedoeld om ministerie-breed te worden gebruikt. Het is een zelfstandig leesbare handleiding met basiskennis, tips en tricks voor de opstellers van startnotities voor evaluaties om deze in staat te maken het gesprek met onderzoeker aan te gaan en kritische vragen te stellen. Deze handreiking bevat concrete suggesties voor evaluatieopstellers bij alle fases van het evaluatieproces vanuit de programmering van evaluaties, via het opstellen van een evaluatieopzet, het offertetraject, de uitvoering en de rapportage van evaluatiebevindingen (meer hierover in 5.2.3.2) (Bonte, *Handreiking Evaluatieonderzoek EZ*). De andere handreiking evaluatieonderzoek is door FEZ geschreven, is meer algemeen en beschrijvend en biedt meer verdieping in het proces van evalueren.

Het ministerie van IenM heeft geen interne handreikingen of richtlijnen voor het opstellen van evaluaties, maar leert op een informele manier uit interdepartementale overleggen en uit de debatten in de Tweede Kamer over hoe andere departementen het evalueren van wetten aanpakken

en wat voor onderzoeksvragen ze stellen om een wet goed te evalueren (Interview, 16 juli 2015). Als het om evalueren van wetten gaat, heeft Directie Wetgeving van IenM in een enkel geval een design voor wetsevaluaties afkomstig van het *Clearing House*¹³ gebruikt bij een uitbestede evaluatie. Het onderzoeksbureau dat de wetsevaluatie heeft uitgevoerd, stemde geheel in met het gebruik van de design en gebruikte deze bij de uitvoering van de wetsevaluatie. Ondanks goede ervaring met en tevredenheid over de design van het *Clearing House* is dit verder niet gebruikt bij andere wetsevaluaties. De respondenten konden zich niet herinneren wat de redenen daarvoor waren (Interview, 16 juli 2015).

Box 5.10

- *Alle departementen hebben één of meer boekjes met handreikingen of richtlijnen voor het opstellen, aanvragen en begeleiden van evaluatieonderzoek.*
- *Alle boekjes (behalve die van IRF bij Financiën) zijn bedoeld voor intern gebruik en zijn geschreven als hulpmiddelen voor dossierhouders/beleidsmedewerkers.*
- *De meeste handreikingsboekjes en richtlijnen worden door stafdirecties (FEZ, BEC, Kennis) opgesteld en geschreven.*
- *De handreikingsboekjes en richtlijnen zijn lang niet altijd bekend bij de onderzoekscoördinatoren en bij beleidsmakers.*
- *Handreikingsboekjes en richtlijnen die wel bekend zijn dossierhouders/beleidsmedewerkers, worden vaak niet gebruikt.*
- *Om de kwaliteit van het evaluatieproces te borgen, hanteren de meeste dossierhouders/beleidsmakers informele werkwijzen: informele gesprekken met meer ervaren collega's, met onderzoekscoördinatoren en met regiegroepen;*

5.2.4 Financiële en niet-financiële middelen

Om goed evaluatieonderzoek uit te voeren zijn naast de aanwezigheid van een evaluatieafdeling en handreiking, richtlijnen en leidraden voor evaluatieonderzoek ook middelen nodig. Deze betreffen zowel financiële middelen als niet-financiële middelen waaronder budget, mensen en tijd.

¹³ Het *Clearing House* (van VenJ) heeft aandacht besteed aan het ontwikkelen van wetgevingskwaliteitsbeleid. Het project is inmiddels stopgezet.

Bij alle departementen wordt eerst een integraal budget voor onderzoek vastgesteld. Alle uitgevoerde evaluaties moeten binnen dat financiële kader passen. FEZ zorgt dat de voorgestelde projecten binnen het budget blijven (RPE, 2006). Bij het uitzetten van evaluaties voor extern onderzoek bepaalt het ministerie het budget waarbinnen de offertes moeten blijven.

Wat de niet-financiële middelen betreft, wordt vaak door respondenten gemeld dat beleidsdirecties weinig tijd hebben voor evaluaties en onderzoek. De meeste tijd wordt besteed aan het maken van beleid. Onderzoek en beleid staan vaak op gespannen voet met elkaar (Interviews 1 juli 2015 en 2 september 2015). Volgens de respondenten bestaat niet altijd een continuïteit tussen onderzoek en beleid. Zo landen enerzijds methodologisch gedegen onderzoeken soms in de la omdat de onderwerpen niet meer relevant zijn voor het voorgenomen beleid en de beleidsmakers er dus niets aan hebben. Anderzijds worden onderzoeken die kwalitatief onder de maat zijn wel eens aangehaald om beleid die al op tafel ligt, te onderschrijven.

Een andere consequentie van het tijdgebrek is dat beleidsdirecties vaak te laat beginnen te denken over evaluaties die op de agenda staan. Het gevolg van te laat beginnen met het opstellen van een evaluatieopdracht is dat de informatie, benodigd om conclusies over de effecten van de interventie (wet of beleid) te trekken, niet voorhanden is. Dit punt is door alle departementen erkend en bevestigd. Dit probleem zou ten dele kunnen worden vermeden door meer tijd in te plannen voor aankomende evaluatie in de agenda van de beleidsmedewerker en door het systeem van dataverzameling en -registratie te verbeteren opdat genoeg informatie voorhanden is om met controlegroep, nulmeting en statistische methodes te kunnen werken. Ondersteunende afdelingen (zoals FEZ bij SZW, Directie Kennis bij OCenW, de onderzoekscoördinatoren bij zowel SZW als OCenW, BEC en de Regiegroep bij EZ en IRF bij Financiën) hebben zich als expliciete taak voorgenomen de beleidsmedewerkers te stimuleren om eerder te beginnen met het denken over wat nodig is als evaluaties (of beleidsdoorlichtingen) komen om deze op gedegen manier te laten uitvoeren.

Naast tijd vereisen evaluaties ook specifieke kennis over het evaluatieproces en de technieken en methoden voor evaluaties. Beleidsmedewerkers werken in een dynamische omgeving en onder snel wisselende omstandigheden. Daarom zijn beleidsmakers met breed inzetbare kennis meer gewaardeerd dan beleidsmakers met specifieke kennis. Het komt niet zelden voor dat de dossierhouder voor de eerste keer een evaluatie moet uitzetten en belast is met het

opstellen van de startnotitie (Interview 26 augustus 2015). Evaluaties die specifieke methodologische en inhoudelijke kennis vereisen, worden ook om deze reden vaak uitbesteed.

Box 5.11

- *De financiële middelen voor evaluaties worden altijd vrij gemaakt uit het onderzoeksbudget.*
- *Offertes moeten binnen het door FEZ bepaalde budget voor evaluatie blijven.*
- *Gebrek aan tijd is de vaakst voorkomende beperking voor het goed opstellen en begeleiden van evaluaties door departementen.*
- *Te laat beginnen met opstellen van evaluatieaanvragen blijkt uit de empirische studie een belangrijke reden te zijn voor de slechte kwaliteit van evaluaties. Om de kwaliteit van evaluaties te verbeteren, zou men meer ruimte en tijd voor evaluatieonderzoek in de agenda van beleidsmakers boeken.*
- *Gebrek aan data en gegevens is vaak een van de redenen dat evaluaties de vraag naar doeltreffendheid van de wet / het beleid niet kunnen beantwoorden.*
- *Om de vraag naar de doeltreffendheid van de wet / het beleid te kunnen beantwoorden, zou men aan evaluatie moeten denken voordat de wet / het beleid wordt ingevoerd.*
- *Behalve WODC en ZonMw waarbij dezelfde groep projectleiders vaak evaluaties begeleiden, zijn dossierhouders bij de andere departementen in de regel slechts een keer in hun baan verantwoordelijk voor een evaluatie. Dit komt door het kleine aantal evaluaties ten opzichte van andere onderzoeksvormen en door de doorstroom van beleidsmedewerkers binnen en tussen departementen.*
- *Gebrek aan tijd, mensen en specifieke kennis in verband met de methodologische en inhoudelijke vereisten van evaluaties binnen departementen zijn redenen om te concluderen dat departementen niet voldoen aan de voorwaarde van niet-financiële middelen voor evaluaties.*

5.2.5 Communicatie

Communicatie is een belangrijke voorwaarde voor een goed ontwikkelde evaluatiecapaciteit van een organisatie. Een goede communicatie tussen aanvragers en uitvoerders van evaluaties is noodzakelijk om ervoor te zorgen dat het onderzoek de informatie verschaft die de aanvrager nodig heeft om de volgende stappen in de besluitvorming te zetten. Verder is communicatie tussen verschillende directie (juridisch, beleid, financiën) belangrijk omdat de meeste evaluaties om een

interdisciplinaire aanpak van onderzoek vragen. Daarom is het belangrijk dat alle relevante directies betrokken worden bij de keuzes rondom evaluaties. Een goede communicatie tussen directies draagt ook bij aan de verspreiding van zowel hulpmiddelen zoals handreiking en richtlijnen voor evaluatie, als van kennis en ervaring met evaluaties aan. Tot slot is communicatie tussen ministeries belangrijk voor de verspreiding van *best practices* en het leren van elkaars ervaringen met evaluatieonderzoek.

Aangezien de mensen die betrokken zijn bij (wets)evaluaties diverse functies vervullen in het evaluatietraject – medewerkers van beleidsdirecties, wetgevingsdirecties en ondersteunende afdelingen, de begeleidingscommissie en uitvoerende onderzoekers – verschilt de vorm van communicatie zowel per groep als per fase van het evaluatietraject.

De meest voorkomende vorm van communicatie binnen departementen is de informele vorm. Beleidsmedewerkers leren van elkaar hoe je bijvoorbeeld de vraag- en probleemstelling in startnotities moet formuleren. Handreiking en richtlijnen (vaak geschreven door ondersteunende stafdirectie) worden over beleidsdirecties verspreid. Medewerkers van stafafdelingen stellen zich open voor vragen en proberen de dossierhouders (beleidsmedewerkers) te stimuleren wetenschappelijke kwaliteitscriteria toe te passen bij het opstellen van startnotities. Stafdirecties die voor de kwaliteit van evaluaties zorgen, hebben echter geen stok achter de deur om te dossierhouder af te dwingen gebruik te maken van de bestaande handreikingen, richtlijnen of formaat. Het initiatief en de eindbeslissing liggen bij de beleidsmedewerker die als dossierhouder fungeert.

Bij twee ministeries (Financiën en SZW) is gebleken dat de interne handreikingen die wel bestaan – met daarin voorbeelden van aanpakken voor evaluaties, vaak voorkomende aandachtspunten bij evaluatie, suggesties voor het formuleren van vraagstelling, en richtlijnen voor de samenstelling van begeleidingscommissies – niet bekend waren bij minstens één van de andere afdelingen die direct betrokken zijn bij evaluaties: de onderzoekscoördinatoren en/of de Wetgevingsdirectie. Een betere communicatie tussen afdelingen en de verspreiding van handreikingen over beleidsdirecties zou de kans verhogen dat dossierhouders de handreikingen daadwerkelijk gebruiken bij het opstellen van startnotitie en bij het schetsen van de vragen die evaluaties moeten beantwoorden.

Dit onderzoek beperkt zich tot de beoordeling van de inrichting van de evaluatiefunctie binnen departementen. De bijdrage van externe onderzoekers die wetsevaluaties in opdracht van

departementen uitvoeren, is in verband met de beperkte tijd en de complexiteit van het onderzoek buiten beschouwing gelaten. De topic van de communicatie tussen opdrachtgever en onderzoekers is slechts vanuit het standpunt van de opdrachtgever gepolst. Adequate communicatie tussen opdrachtgever en onderzoeker is naast de deskundigheid van het onderzoeksteam als een cruciale factor gezien door sommigen van de geïnterviewde opdrachtgevers (Interview, 9 april 2015).

Box 5.12

- *De communicatie tussen beleidsmedewerkers en stafdirecties loopt niet altijd soepel. Stafdirecties hebben richtlijnen en handreikingen ontwikkeld, maar deze zijn vaak niet bij de beleidsmakers bekend.*
- *Beleidsmedewerkers leren van elkaar via informele communicatiekanalen hoe de inhoudelijke en procesmatige kwaliteit van evaluaties te borgen.*

5.2.6 Gebruik van evaluatiebevindingen

Zoals afgeleid uit Bourgeois en Cousins (2013) is het vermogen om evaluatiebevindingen te gebruiken – naast het vermogen om evaluaties te verrichten – gevolg van de evaluatiecapaciteit van organisaties. De vraag is op welke manieren en in welke vormen de evaluatiebevindingen hun weg weten te vinden terug naar de besluitvorming, de wetgever en de politieke arena.

Om inzicht te krijgen in het vermogen om evaluatiebevindingen te gebruiken, heeft deze studie onderzocht welke procedures departementen hebben om bevindingen uit wetsevaluaties naar de wetgever en beleidsmakers terug te koppelen. Het onderzoek heeft ook rekening gehouden met het verschil tussen enerzijds direct gebruik van evaluatieresultaten en -aanbevelingen in de besluitvorming en anderzijds indirect of conceptueel gebruik in de vorm van kennisname van het rapport en de evaluatiebevindingen. Daarnaast zijn de geïnterviewden om concrete voorbeelden van wetten gevraagd die veranderd of opnieuw opgesteld zijn naar aanleiding van evaluaties.

Alle departementen volgen over het algemeen dezelfde formele procedure na de afronding van wetsevaluaties. De opdrachtgever zoals de betrokken beleidsdirectie of wetgevingsdirectie moet binnen drie maanden na de oplevering van het evaluatierapport een beleidsreactie formuleren en naar de Tweede Kamer sturen. Daarin moet worden gemeld wat de minister van plan is met het onderzoek te gaan doen. Als de beleidsreactie is gepubliceerd, is de betrokkenheid van de ondersteunende afdeling (Kennis, FEZ, EWB etc.) voorbij en kunnen ze de ontwikkelingen slechts

van afstand zien. Het rapport en de kabinetsreacties zijn openbaar of kunnen op aanvraag worden geraadpleegd.

Respondenten van alle departementen betwijfelen of de evaluatiebevindingen daadwerkelijk tot aanpassing van het beleid of regelgeving leiden. Bij de meeste evaluaties is de enige vorm van gebruik het *conceptuele gebruik*, waarbij de belangstellende – minister, Tweede Kamer, beleidsambtenaren en doelgroepen – kennis nemen van de evaluatiebevindingen en eventueel aanbevelingen. Kennisname van evaluatiebevindingen is een indirecte manier van doorwerking omdat de evaluatiebevindingen niet direct aan toekomstige besluitvorming worden gekoppeld.

Bij sommige departementen is het gemakkelijker dan bij andere om te controleren of de evaluatiebevindingen tot verandering in het beleid, subsidie en regelgeving hebben geleid. Het ministerie van EZ heeft bijvoorbeeld een systeem waardoor de verstrekking van subsidies afhankelijk is van de evaluatie van het succes van eerdere subsidies voor hetzelfde beleid. Omdat de financiële regelingen in de regel om de vijf jaar vervallen, moeten ze opnieuw naar monitoringcommissie worden gestuurd om goedgekeurd te worden voordat de regeling afloopt (Interview 9 juli 2015). De beslissing over het wel of niet verstrekken van een nieuwe subsidiecyclus is gebaseerd op informatie over de efficiëntie en effectiviteit van de afgeronde subsidiecyclus. De subsidieregeling voor innovatie in de scheepbouw is bijvoorbeeld ingetrokken als gevolg van de evaluatie van de subsidieregeling (Interview 9 juli 2015). De evaluatie heeft aangetoond dat de subsidie niet effectief en efficiënt werkte. Andere voorbeelden van goede evaluaties en doorlichtingen waarvan veel is geleerd en die direct of indirect gevolgen hebben gehad voor wetgeving en beleid zijn het bedrijvenbeleid, de agro-regeling en het landbouwbeleid (Interview 9 juli 2015).

De procedure voor de borging van het gebruik van evaluatiebevindingen bij het ministerie van EZ omvat enkele aspecten: na de afronding van evaluatie komt het evaluatierapport naar de Regiegroep en BEC terug. BEC en de Regiegroep zitten aan tafel met deskundigen en uitvoerders om de conclusies van het rapport te bespreken en eventuele onduidelijkheden te verhelderen. Vaak wordt het evaluatierapport naar de Tweede Kamer gestuurd samen met een brief van de minister. Een andere manier om de evaluatiebevindingen bekend te maken en te verspreiden is door korte presentatie en toelichtingen op de werkvloer te geven. Onderzoeksbureaus die evaluaties

verrichten schrijven zelf soms brieven en notities over hun eigen evaluatieonderzoek (Interview, 9 juli 2015).

Een voorbeeld van een wetsevaluatie die grote consequenties heeft gehad is de Evaluatie Politiewet, uitgevoerd in opdracht van het ministerie van VenJ (via de uitbestedingsprocedure van EWB). Uit de evaluatie van de Politiewet bleek dat de wet heel ingewikkeld was en dat de wet in praktijk moeilijk uit te voeren was, wat tot fouten heeft geleid in de uitvoering. Op basis van de evaluatieresultaten en -aanbeveling heeft de wetgever een nieuwe wet gemaakt (Interview, 8 juni 2015).

Box 5.13

- *Men kan in theorie onderscheid maken tussen enerzijds het directe gebruik en anderzijds het indirecte of conceptuele gebruik van evaluatiebevindingen. Het ene houdt in dat de wet of het beleid wordt verbeterd aan de hand van de evaluatiebevindingen en -aanbevelingen en dat deze hun weg weten te vinden naar de besluitvorming. Het andere houdt in dat verschillende partijen kennis nemen van de resultaten van de evaluaties.*
- *Bij de meeste evaluaties is de enige vorm van gebruik het indirecte of conceptuele gebruik, waarbij de belangstellende – minister, Tweede Kamer, beleidsambtenaren en doelgroepen – kennis nemen van de evaluatiebevindingen en eventueel aanbevelingen.*
- *Alle departementen twijfelen dat evaluatiebevindingen daadwerkelijk tot aanpassing van het beleid of regelgeving leiden.*
- *Slechts 3 departementen konden minstens één voorbeeld geven van een wetsevaluatie die tot aanpassingen in de wet of tot nieuwe besluitvorming heeft geleid. Evaluaties en wetsevaluaties hebben dus weinig gevolgen voor de besluitvorming.*
- *Wat het conceptuele gebruik betreft, volgen alle departementen over het algemeen dezelfde formele procedure na de afronding van (wets)evaluaties. De opdrachtgever zoals de betrokken beleidsdirectie of wetgevingsdirectie moet binnen drie maanden na de oplevering van het evaluatierapport een beleidsreactie formuleren en naar de Tweede Kamer sturen. Daarin moet worden gemeld wat de minister van plan is met het onderzoek te doen. Het rapport en de kabinetsreacties zijn in de meeste gevallen openbaar.*
- *Omdat de overheid in het huidige evaluatiestelsel een sterke en bijna exclusieve nadruk op beleidsdoorlichtingen legt, worden evaluaties die geen direct verband hebben met de doelstellingen van beleidsartikelen als minder relevant beschouwd. Een nadeel hiervan is dat de verantwoordingsfunctie de leerfunctie overschaduw.*

5.2.7 Leren van evaluaties

Departementen hebben geen ingebouwd systeem om van evaluaties te leren of leerervaring uit evaluaties over andere departementen te verspreiden. Het leerproces vindt op een informele manier plaats. Dossierhouders vragen om hulp, suggesties en feedback bijvoorbeeld op startnotities bij ondersteunende afdelingen (zoals FEZ, Kennisdirectie, BEC of Regiegroep), bij de onderzoekscoördinatoren van DG's of bij collega's die ooit evaluaties hebben uitgezet of begeleid.

Het interdepartementale overlegplatform IBP, onder de coördinatie van het ministerie van Financiën, is een forum waarin de vertegenwoordigers van alle departementen hun vragen kunnen stellen en om voorbeelden van goede aanpakken van evaluaties kunnen vragen. Daarnaast geven enkele departementen aan dat ze veel van Kamerdebatten leren. De vragen die aan de minister gesteld worden in de Tweede Kamer zijn goede indicaties van hoe ze in de toekomst grote wetsevaluaties moeten aanpakken en welke onderzoeksvragen aan bod moeten komen om de Tweede Kamer tevreden te stellen (Interview 16 juli 2015).

Leren is naast verantwoorden de belangrijkste functie van evalueren (RPE, 2006). Sommige respondenten noemen ook agenderen als de derde functie van evaluatie (Interview 15 juli 2015). Uit de deskstudie en interviews blijkt dat evaluaties gedaan worden met als primaire focus de verantwoording van de manier waarop het geld uitgegeven wordt aan een bepaald beleid, subsidie, wet of een ander instrument. In de evaluatieprogrammering die ieder ministerie bij het jaarlijkse Rijksbegrotingsdocument moet bijvoegen, ligt de focus op beleidsdoorlichtingen en ander onderzoek naar doelmatigheid en doeltreffendheid. Alleen die evaluaties tellen mee die een bijdrage leveren aan het beantwoorden van de vraag of het beleid succesvol is geweest, met andere woorden of de interventie van de overheid de gewenste effecten teweeg heeft gebracht. Sommige respondenten klagen over deze aanpak van evaluatie waarin de focus exclusief op de verantwoordingsfunctie van evaluatie valt. De leerfunctie zou net zo belangrijk zijn als de verantwoordingsfunctie en zou vaker naar voren moeten worden gehaald (Interviews 15 juli 2015 en 2 september 2015). In hun mening kan men even veel leren van procesevaluatie als van effectevaluatie hoewel de twee vormen van evaluatie twee verschillende vragen stellen en twee verschillende aanpakken vereisen.

Om de leerfunctie van evaluatie een belangrijker plaats te geven, zouden evaluaties volgens sommige respondenten meer toekomstgericht moeten zijn. In plaats van de *status quo* te beschrijven en de vraag te beantwoorden of het beleid of de wet de gewenste effecten wist te

bereiken, zouden evaluaties de vraag moeten durven stellen of het beleid of wet nog zinvol is binnen het huidige stelsel (Interviews 15 juli 2015 en 2 september 2015). Respondenten stelden dat zij vooral veel leren van evaluaties die in staat zijn nieuwe discussies te starten over hoe een bepaald maatschappelijk probleem zou kunnen worden aangepakt. Als voorbeelden van evaluaties waarvan veel is geleerd, noemden de respondenten de Pensioenwet, de Algemene nabestaandenwet en de Politiewet. Een goede evaluatie waarvan veel wordt geleerd is een evaluatie die goed aansluiten bij de huidige maatschappelijke discussie (Interview 15 juli 2015). Een tweede leerfunctie van (wets)evaluaties is dat ze de impliciete aannames waarop de wet of beleid rust, expliciet maken. Wetsevaluatie dragen bij aan de verduidelijking van de beleidstheorie en aannames waarop de wet steunt en die vaak impliciet door de wetgever zijn gehanteerd. Een goede wetsevaluatie zou dus de vraag moeten stellen of de aannames van de wetgever kloppen en via welke mechanismes de wet bepaalde effecten heeft bereikt. Het kan namelijk blijken dat de gewenste effecten wel zijn bereikt, maar via andere mechanismen dan het in de beleidstheorie wordt gedacht (Interviews 8 juni 2015 en 2 september 2015).

Box 5.14

- *Departementen hebben geen ingebed system om van evaluaties te leren of leerervaring uit evaluaties over andere departementen te verspreiden.*
- *Leren over het opstellen en aanvragen van evaluaties vindt binnen departementen op een informele manier plaats via gesprekken met ervaren collega's.*
- *Departementen klagen over de aanpak van evaluaties in het huidig evaluatiestelsel waarbij de overheid exclusief op de verantwoordingsfunctie focust.*
- *De leerfunctie van evalueren is even belangrijk als de verantwoordingsfunctie en zou vaker naar voren moeten worden gehaald.*
- *Evaluaties zouden vaker toekomstgericht moeten zijn (niet alleen de vraag beantwoorden of het beleid of de wet de gewenste effecten heeft bereikt, maar ook de vraag stellen of het beleid of wet nog zinvol is binnen de huidige maatschappij).*
- *Bekeken vanuit de leerfunctie is een goede evaluatie een evaluatie die goed aansluit bij de maatschappelijke discussies en die de impliciete aannames waarop de wet of het beleid rust, expliciet maken.*

5.3 Individuele competenties

De tweede voorwaarde voor evaluatiecapaciteit zijn de individuele competenties. Uit de literatuur hebben wij drie soorten individuele competenties geïdentificeerd: kennis (zowel over de inhoud en methoden van evalueren, als over het proces van evaluatie), vaardigheden en ervaring met evaluatie. Op basis van de empirische studie zijn de volgende observaties gedaan over de kennis, vaardigheden en ervaring met (wets)evaluaties binnen departementen.

5.3.1 Kennis over evaluatie

Departementen beschouwen kennis als een belangrijke voorwaarde voor het verrichten van goede evaluaties. Kennis omvat in deze studie de volgende aspecten: kennis over het evaluatieproces (startnotitie opstellen; vraag- en probleemstelling van evaluatieopdracht formuleren, begeleidingscommissie samenstellen) en kennis over de methoden en technieken voor evaluatie. Doordat de borging van de kennisbehoefte niet op het centrale niveau, maar binnen departementen ligt, verschillen departementen sterk in de manier waarop zij in de kennisbehoefte van dossierhouders en beleidsmedewerkers voor evaluatieonderzoek voorzien. In het volgende beschrijft deze studie op welke manier departementen de benodigde kennis voor evaluatie borgen.

De stafdirectie FEZ van SZW, die de ondersteunende taak heeft om zorg te dragen voor de kwaliteit van evaluatie, bestaat uit vijf medewerkers. De FEZ-medewerkers doen geen inhoudelijk onderzoek, maar spelen een aanjagende rol als het om evaluatie gaat. Zij houden zich verder bezig met kennisverspreiding over de aanpak van bijvoorbeeld beleidsdoorlichtingen, over methoden en over welke gegevens voorhanden zijn. Zij wonen, samen met vertegenwoordigers van alle departementen, de vergaderingen van het interdepartementale overlegplatform Interdepartementale Begeleidingsgroep Prestatiegegevens en Beleidsevaluatie (IBP) bij. Het doel hiervan is uit andere departementen te leren over hoe evaluatieonderzoek methodologisch en procesmatig beter kan. Een uitdaging voor FEZ is de coördinatie van beleidsdirecties en de afstemming tussen onderzoeken dat bij verschillende directie belegd zijn. Op deze manier wordt voorkomen dat de kennis binnen directies verkokert en dat de onderzoeken bij directies langs elkaar lopen.

Elke DG van SZW heeft daarnaast een eigen onderzoekcoördinator. De onderzoekscoördinatoren staan onder de leiding van CSO. Deze is actief op departementsniveau en ondersteunt alle aanvragers met kennis en feedback over vraag- en probleemstelling,

onderzoeksdesign en -methoden (Interview 15 juli 2015). De afgelopen tijd zijn er door FEZ gesprekken met beleidsdirecties gevoerd. Sommige directies vonden dat de onderzoekscoördinatie op departementsniveau moet worden gecentraliseerd in een aparte stafunit, anderen vonden dat één unit voor het hele ministerie onvoldoende zou zijn. Uiteindelijk is besloten dat de onderzoekcoördinatoren bij elke directie blijven. De rol van onderzoekcoördinatoren is de beleidsdirectie te stimuleren om pro-actiever te denken over bijvoorbeeld evaluatieplanning (beschreven in paragraaf 5.2.2), meer initiatief te nemen en niet alleen verplicht onderzoek te verrichten (zoals beleidsdoorlichtingen), maar ook niet verplichte nummers.

Bij het ministerie van OCenW heeft elk van de drie Directoraten-Generaal een onderzoekscoördinator die de beleidsmedewerkers ondersteunt bij zaken zoals de formulering van de vraag- en probleemstelling van onderzoeken, de opstelling van startnotities, de keuze voor onderzoeksdesign en -methoden. De onderzoekscoördinatoren doen dit vaak parttime als aanvulling op hun beleidsverantwoordelijkheden. Sommige directies hebben de functie van onderzoekcoördinator verdeeld over diverse beleidsmedewerkers. Dit is bij SZW ook herkenbaar: de meeste mensen hebben meerdere petten op, zijn breed inzetbaar en moeten de ene keer een beleidsdoorlichting doen, de andere keer een wetsevaluatie aanvragen. De functie onderzoekscoördinator bedraagt opgeteld wellicht 1fte bij ieder DG, maar meestal ligt ze onder de 0,5 fte. Er is weinig zich op verbeteringen hiervan aangezien dat de financiële middelen hiervoor teruglopen (Interview 1 juli 2015). Het *Top Institute for Evidence Based Education Research* (TIER) biedt een Masterprogramma voor beleidsmedewerkers van het ministerie van OCenW aan. Het programma bestaat uit vier modules, waarvan de eerste over *evidence based policy* gaat. Dit is vooral bedoeld voor de onderzoekscoördinatoren. Beleidsmedewerkers nemen vooral deel aan tweedaagse cursustrainingen. Directie Kennis speelt ook een rol bij de opleiding van beleidsmedewerkers door collega's te stimuleren aan het masterprogramma en cursustrainingen mee te doen (Interview 1 juli 2015).

Het ministerie van EZ heeft een extern onderzoeker (hoogleraar van VU) een training voor beleidsmedewerkers laten ontwikkelen. Die is tot op heden drie keer verzorgd op het ministerie voordat de cursus naar de Rijksacademie is overgeheveld. Het is de bedoeling dat de cursus rijksbreed wordt verzorgd. De cursus biedt aan de beleidsmedewerkers een eerste kennismaking met de kwantitatieve methodes, omdat “voor negen van de tien gewone beleidsmedewerkers de evaluatiemethoden toch lastig zijn” (Interview 9 juli 2015). Aangezien de meeste wetsevaluaties

worden uitbesteed, hoeven de beleidsmedewerkers geen specialistische kennis te hebben, wel moeten zij als aanvrager weten wat je van kwantitatief onderzoek kunt verwachten. Binnen EZ zijn er nog meer andere manieren om de kennis van beleidsambtenaren op peil te houden. Zo bestaat er de Regiegroep die prikkels inzet zoals een “*quick response team*” of een nieuwsbrief waarmee mensen geattendeerd worden op nieuwe tools en publicaties.

Verschillende ministeries (SZW, OCenW) gaven aan geen voorkeur te hebben voor een bepaald design of methode voor wetsevaluaties. Ministeries zijn eigenlijk niet voldoende in staat om evaluatieopdrachten uit te zetten en te begeleiden. De verantwoordelijke beleidsambtenaren laten de keuzes voor methoden en aanpak aan de uitvoerende onderzoeksteams. Een ander reden hiervoor, die vaak door respondenten wordt genoemd, is om de onafhankelijkheid en neutraliteit van het onderzoek te vrijwaren. De beleidsambtenaren sturen via de formulering van de vraag- en probleemstelling in de startnotitie die ze meestal samen met de ondersteunende afdeling van het ministerie (zij het Directie Kennis, de onderzoekcoördinator, of FEZ) opstellen.

Een andere manier om de kwaliteit van wetsevaluaties te borgen bij Nederlandse ministeries is via de selectie van onderzoeksteams. Alle departementen werken graag vooral met onderzoeksteams van universiteiten en onderzoeksbureau. De reden hiervoor is dat wetenschappers goede methodologische kennis hebben.

Departementen zoals SZW (via de onderzoekscoördinatoren) en VenJ (via EWB) hebben een database opgezet met de specifieke kennis en contactpersonen van alle onderzoeksbureaus en onderzoeksgroepen die ooit een offerte hebben ingediend of een evaluatie hebben gedaan. Dit helpt de departementen om overzicht te krijgen over welke kennis waar beschikbaar is en welke teams op welk onderzoeksterrein expertise hebben. Een uitdaging voor deze werkwijze is zowel volgens SZW als VenJ het feit dat de dossierhouders veel meer weten over het hele evaluatietraject en de praktische en concrete uitdagingen in het uitbestedingsproces dan wat geregistreerd wordt. De kennis zit vaak in het hoofd van een persoon (onderzoeker of beleidsambtenaar). Dit zijn mensen met veel ervaring. “Het in kaart brengen van deze kennis kost veel energie, want de kennis zit soms niet bij een bureau, maar bij een persoon” (Interviews 8 juni 2015 en 15 juli 2015). Bovendien op het moment dat de persoon bij het ministerie of onderzoeksbureau vertrekt, is de kennis weg. De consequenties zijn tweeledig: enerzijds kost het tijd om die kennis en ervaring te vervangen en anderzijds is het moeizaam om de database op peil te houden door deze steeds tijdig bij te werken.

Een andere manier om de kwaliteit van (wets)evaluaties te borgen is om te zorgen dat het onderzoek wordt begeleid door een begeleidingscommissie. De begeleidingscommissie is bedoeld voor de “wetenschappelijke toets” en zorgt voor de onafhankelijkheid van het onderzoek (Interview 8 juni 2015). Naast de wetenschappelijke toets zorgt de begeleidingscommissie ook voor een praktijktoets, omdat daar ook een veldvertegenwoordiger bij dient te zijn. De voorzitter is meestal hoogleraar en heeft diepgaande kennis op het gebied waaronder de geëvalueerde wet of beleid valt. De functie van de begeleidingscommissie is “inhoudelijke bijdrage te leveren aan de uitvoering en de kwaliteit van het onderzoek, het ‘betrekken van veldorganisaties’ of het ‘creëren van draagvlak’.” (WODC, 2012), het stellen van kritische vragen, het tijdig ingrijpen en bijsturen als het onderzoek niet goed verloopt (Interview 16 juli 2015). Bij EWB betreuren de projectbegeleiders dat het vaak lastig is om mensen te vinden die zowel methodologische als domein-specifieke kennis hebben. Bij specifieke onderwerpen, bijvoorbeeld onderwerpen die juridische aspecten inhouden, is de groep die aan deze twee voorwaarden voldoet en die dus zitting zou kunnen nemen in de begeleidingscommissie, vaak heel klein. Bovendien hebben veel wetenschappers met kennis van sociaalwetenschappelijke methoden weinig interesse in het begeleiden van wetsevaluaties. Er zijn over het algemeen geen strakke regels of vaste formats over wie voor de begeleidingscommissie wordt uitgenodigd. De samenstelling van de begeleidingscommissie is “maatwerk”, verschilt per domein, terrein en vorm van evaluaties (WODC, 2012). De enige richtlijn is het gebruiken van het gezonde verstand om de juiste disciplines in de begeleidingscommissie te verenigen. Aan tafel zit er meestal iemand met uitvoeringspraktijk samen met experts van de betrokkene disciplines (Interview 16 juli 2015).

5.3.2 Benodigde vaardigheden voor evaluatieonderzoek

Beleidsdirecties stellen in principe geen strenge eisen aan de uitvoerende onderzoekers/onderzoeksteams wat hun opleiding en kennis van evaluatie betreft. De verantwoordelijkheid ligt veelal bij (externe) uitvoerders van evaluaties. Het ministerie van SZW maakt wel gebruik van formulieren waar informatie moet worden ingevuld over de opleiding en ervaring van de winnende onderzoeksgroep. De voorkeur gaat naar onderzoeksgroepen waarbij ook hoogleraren en/of ervaren onderzoekers betrokken zijn. Wat de opleiding van de betrokken beleidsambtenaren betreft, is kennis met de methoden en technieken van evaluatie geen voorwaarden bij aanstelling. Deze wordt echter sterk gewaardeerd (Interview 15 juli 2015).

Alle departementen gaven aan de *multidisciplinariteit* van het onderzoeksteam heel sterk te waarderen. Dit is bij wetsevaluaties nog belangrijker dan bij andere vormen van evaluatie, omdat wetten een combinatie van juridische en sociale elementen bevatten (toegelicht in hoofdstuk 1). Om recht te doen aan de complexiteit van wetten en eenzijdigheid te voorkomen, zou de multidisciplinariteit van de uitvoerende onderzoeksgroepen en de begeleidingscommissie aan de kwaliteit van wetsevaluaties moeten bijdragen. Bij WODC zijn interne, informele afspraken gemaakt over de samenstelling van onderzoeksteams. Vooral bij wetsevaluaties is het zo dat je zowel juristen als sociaalwetenschappers nodig hebt “omdat de spanning tussen deze twee kanten permanent aanwezig is; hoewel de spanning wel per sector kan verschillen” (Interview 8 juni 2015).

Het bewust zijn van het belang van multidisciplinariteit heerst echter niet overal. Medewerkers werkzaam bij FEZ (SZW) en onderzoekers van EWB merken op dat wetsevaluaties waarbij Directies Wetgeving de aanvrager is vaak eenzijdig op de juridische kant focussen. Dit is ook door medewerkers van Directie Wetgeving bevestigd met als voorbeeld de evaluatie van de Awb (Interview 11 augustus 2015).

Onderzoekers van EWB bij VenJ zien een duidelijk verschil tussen wetsevaluaties waarbij grotendeels of uitsluitend juristen zijn betrokken en wetsevaluaties die door een multidisciplinaire groep worden gedaan. In het eerste geval beperkt zich de wetsevaluatie tot de beschrijving en de uitleg van hoe de wet heeft gewerkt ten opzichte van hoe de wet zou moeten werken. Deze laatste vraag wordt beantwoord door informatie te halen uit de Memorie van Toelichting die de wet begeleidt – voor zover die dit vermeldt. Wat juristen vaak niet doen in tegenstelling tot (een deel van de) wetenschappers met een sociaalwetenschappelijke of economische achtergrond, is het expliciteren van de assumpties waarop de wet impliciet rust en het reconstrueren van beleidstheorie. Dit is volgens opstellers van wetsevaluaties in opdracht van het ministerie van VenJ een noodzakelijke fase in het traject van wetsevaluatie. Een andere uitdaging voor juristen vormen methoden en technieken van sociaalwetenschappelijk onderzoek, met als voorbeeld een steekproef trekken. Vervolgens zijn de juristen ook minder bekend met het sociaalwetenschappelijk jargon. Zo staat bijvoorbeeld in evaluatiebepalingen dat de evaluatie moet concluderen of de interventie de beoogde effecten heeft bereikt, maar volgens de respondenten van EWB zijn juristen niet bekend met termen zoals “effecten van wet”. De hoofdvraag die juristen en sociaalwetenschappers stellen bij wetsevaluaties is ook verschillend: “Juristen stellen de vraag of de wet geldt, de

beleidsdirecties en sociaalwetenschappelijk onderzoek willen weten of de wet werkt en onder welke omstandigheden” (Interview 8 juni 2015). Bij gebrek aan multidisciplinariteit wordt enerzijds niet genoeg aandacht aan gedragsverklarende mechanismen geschonken en anderzijds blijft het juridische kader onvoldoende uitgewerkt.

Hoewel alle departementen het ermee eens zijn dat zowel juridische kennis als de sociaalwetenschappelijke onderzoeksmethoden noodzakelijk zijn voor goede wetsevaluaties, wordt toch soms voorkeur gegeven aan de juridische benaderingswijze. De reden daarvoor is volgens een jurist van de Directie Wetgeving van VenJ, die betrokken is geweest bij de drie evaluaties van de Algemene wet bestuursrecht (Awb), heel eenvoudig. Hij verklaarde de keuze voor een onderzoeksteam met een overwegend juridische expertise door het feit dat de Directie Wetgeving veel beter bekend was met het netwerk van juristen in de academische wereld dan met andere groepen wetenschappers. Juristen en beleidsmedewerkers van de Directie Wetgeving hadden nauwelijks contacten onder wetenschappers uit andere disciplines, die eventueel de benodigde expertise konden inbrengen. Om deze reden is ervoor gekozen een uitsluitend juridisch onderzoeksteam als evaluatoren te selecteren. Een andere reden hiervoor was dat de evaluaties op korte termijn afgerond moesten worden. Dit maakte het volgens de juridische respondent onmogelijk om een complexe evaluatie met aandacht voor zowel de juridische als de sociaalwetenschappelijke aspecten te laten uitvoeren. Een combinatie van juridische en sociaalwetenschappelijke wetsevaluatie vraagt dat de monitoring en dataverzameling veel eerder begint. Dit was toen niet haalbaar aangezien men volgens de respondent (te) laat was begonnen met het voorbereidende werk op de evaluaties van AWB.¹⁴

5.3.3 Ervaring van medewerkers met evaluatieonderzoek

Het derde aspect op het vlak van individuele competenties die wij hebben onderzocht is de ervaring van opstellers en aanvragers en van (wets)evaluaties. Volgens Nielsen et al. (2011) faciliteert de praktijkervaring van medewerkers met evaluatieonderzoek de kwaliteit van evaluaties zowel in termen van hoe evaluaties worden gedaan (evaluatieproces) als in termen van de methodologische en inhoudelijke keuzes tijdens het onderzoek. Een rijke ervaring met evaluatie is volgens Cooksy

¹⁴ Dit zegt iets over de evaluatieplanning, -programmering en over hoe bereid departementen zijn voor het verrichten van verplichte wetsevaluatie (voor meer informatie hierover, zie paragraaf 5.2.2).

(2012) noodzakelijk om de relevante informatie te kunnen destilleren en om inzicht te krijgen in welke instrumenten in welke omstandigheden werken.

Ministeries verschillen van elkaar op dit aspect. Zo is eerdere ervaring met het uitbesteden en begeleiden van onderzoeks- en evaluatietrajecten een criterium in de werving van de projectbegeleiders bij het onderzoeksinstituut van WODC (VenJ). Anderzijds stelt de Directie FEZ van SZW, die vijf medewerkers telt, geen eisen aan nieuwe medewerkers, maar beschouwt de ervaring met wetenschappelijke methoden als een bonus (zoals het behalen van een proefschrift op het gebied van sociale wetenschappen) (Interview 15 juli 2015). Bij de afdeling EWB van WODC krijgt elke projectbegeleider zijn eigen terreinen toebedeeld, wat de ruimte geeft om van veel ervaring met wetsevaluaties op een bepaald terrein op te doen. Deze specialisatie is echter alleen mogelijk als de afdeling veel evaluaties en wetsevaluaties doet. EWB krijgt opdrachten aangereikt door het ministeries van VenJ en in mindere mate door het ministerie van BZK. Er is bij EWB een aantal procedures bedoeld voor opstellers van startnotities en begeleiders van evaluatietrajecten om ervaring met elkaar te delen en van elkaar te leren. De bestaande, grotendeels informele procedures zijn: regelmatig afdelingsoverleg met collega's, feedback vragen bij het afdelingshoofd en het managementteam, inlichting krijgen van de directeur over nieuwe inzichten in de (wetenschappelijke) literatuur. De respondenten van EWB verschillen echter van mening over of deze procedures geschikt en voldoende zijn om te zorgen dat ervaring kan worden gedeeld en niet "verkokert". Zo vindt de ene de bestaande procedures voldoende, terwijl de andere respondent juist denkt dat veel ervaring niet gedeeld kan worden (Interview 8 juni 2015).

Box 5.15

- *Doordat de borging van de kennisbehoefte niet op het centrale niveau, maar binnen departementen ligt, verschillen departementen sterk in de manier waarop zij in de kennisbehoefte van dossierhouders en beleidsmedewerkers voor evaluatieonderzoek voorzien.*
- *De kennis die binnen departementen wordt geborgd, betreft de stappen in het evaluatieproces: de formulering van de vraag- en probleemstelling, de opstelling van startnotities, de keuze voor onderzoeksdesign en –methoden, de samenstelling van de begeleidingscommissie.*
- *Slechts vier departementen hebben onderzoekscoördinatoren bij wie de beleidsmakers met kennisvragen over evaluatie kunnen aankloppen.*

- *Onderzoekscoördinatoren vervullen deze functie parttime naast hun werkzaamheden als beleidsmakers.*
- *Twee departementen (EZ en OCenW) hebben masterprogramma's en cursustrainingen voor beleidsmedewerkers laten ontwikkelen door hoogleraren of onderzoeksinstituten. Het doel van deze cursussen is de beleidsmedewerkers meer bekend te maken met evidence based policy.*
- *Beleidsmedewerkers hebben minder methodologische en inhoudelijke kennis over evaluatie nodig omdat de meeste evaluaties door externe onderzoekers worden uitgevoerd. Als gevolg hiervan focussen departementen zich hoofdzakelijk op de borging van de kennis over het proces van evalueren en in mindere maten op de methoden en de inhoud die vereist zijn bij evalueren.*
- *Veel kennis zit in het hoofd van ervaren beleidsmedewerkers. Slechts één ministerie (EWB bij VenJ) heeft een systeem waarin informatie over de startnotities van evaluaties en over de kwaliteit van offertes zijn vastgelegd en op een later tijdstip kunnen worden geraadpleegd.*
- *De borging van de inhoudelijke en methodologische kennis voor evaluatie wordt aan de uitvoerders (meestal externe onderzoekers of onderzoeksbureaus) gelaten.*
- *Alle departementen waarderen de multidisciplinariteit van zowel het onderzoeksteam als de begeleidingscommissie heel sterk. Multidisciplinariteit is vooral bij wetsevaluatie van belang omdat de evaluatie zowel juridische als empirische kennis vergt.*
- *Bij de twee departementen waar de meeste evaluaties via een onderzoeksinstituut worden verricht (WODC bij VenJ en ZonMw bij VWS) is er veel kennis, vaardigheden en ervaring opgebouwd. Het ministeries van EZ en OCenW bieden ook via de ondersteunende groepen, handleidingen en programma's hulp bij kennisvragen. Bij de andere ministeries wordt er minder aandacht aan individuele competenties besteed. Dossierhouders borgen hun kennis en vaardigheden voor evaluatie door van meer ervaren collega's op een informele manier te leren.*

5.4 Omgevingsfactoren

Het derde aspect voor evaluatiecapaciteit zijn de omgevingsfactoren. Op basis van inzichten uit de wetenschappelijke literatuur hebben wij geconcludeerd dat de twee omgevingsfactoren die een grote rol kunnen spelen bij evaluaties zijn: de politieke betrokkenheid en de stakeholdersbetrokkenheid.

5.4.1 Politieke betrokkenheid

Uit de interviews blijkt dat de politieke context waarin evaluaties plaatsvinden van grote invloed is op zowel het vermogen van overheidsinstellingen om evaluaties te verrichten als op de kwaliteit van evaluaties.

De politieke wil is naast grote maatschappelijke veranderingen de belangrijkste aanleiding voor het maken van wetgeving en voor het evalueren hiervan. Zoals in paragraaf 5.2.2 uitgelegd, is de jaarlijkse bijstelling van de evaluatieplanning bij departementen onderworpen aan de wil en de interesse van de politieke arena in een bepaalde thema. Hoewel de verplichte evaluaties vaak wel afgerond worden, komt het soms voor dat ze door de veranderingen in de samenstelling van de politieke context als minder relevant worden gezien (Interview 2 september 2015). Dan is het gebruik van evaluatiebevindingen beperkt tot de kennisname van het rapport zonder de aanbevelingen terug te koppelen naar de besluitvormingsfase. Soms worden evaluatie uitgesteld of nooit uitgevoerd omdat de belangstelling van de Tweede Kamer voor het betreffende beleid, wet of terrein is weggezakt.

Een belangrijke voorwaarde voor een goede evaluatie is volgens een meerderheid van de respondenten het tijdsverloop tussen het moment van de invoering van de wet of het beleid en het moment van evaluatie. Respondenten gaven voorbeelden van wetsevaluaties waaruit geen conclusies getrokken konden worden over de effecten van de ingevoerde wet of maatregelen, omdat de Tweede Kamer om een evaluatie vroeg pas twee jaar na de invoering van de wet of beleid (Interview 16 juli 2015).

Respondenten meldden dat de Tweede Kamer steeds meer belangstelling toont in effectevaluaties. Kamerleden stellen steeds vaker vragen over de effectiviteit van wetten en beleid en zetten daardoor druk op ministers om verantwoording te leggen voor de gevoerde wetten en beleid, ook in verband met financiële lasten (Interview, 16 juli 2015). Alleen beleid dat is aangetoond te werken en dat op *evidence based* onderzoek steunt, mag worden doorgevoerd. Naar aanleiding van kabinetsreacties op evaluatierapporten stellen Tweede Kamerleden daarnaast meer vragen over wat de minister van plan is te doen met de evaluatiebevindingen en aanbevelingen. In verband met beleidsdoorlichtingen moeten ministers aan de Tweede Kamer op zijn minst onderbouwen waarom niet alle wetten en beleidsinterventie die als één van de hoofddoelen van het ministerie moeten dienen, worden geëvalueerd. De redenen voor het buiten beschouwing laten van een wet of beleidsinstrument – bijvoorbeeld dat je soms niet alle wetten die onder een

beleidsartikel tegelijkertijd kunt evalueren – moet altijd voldoende worden verantwoord in de betreffende beleidsdoorlichting.

Uit deze studie blijkt ten eerste dat de politieke betrokkenheid in de vorm van het politieke belang en de politieke druk een dubbele werking heeft op evaluatieactiviteit van ministeries, zowel bevorderend als remmend. Evaluaties bieden idealiter tegenwicht aan de dynamiek en instabiliteit die de politieke sfeer kenmerken. Zo wordt in de wetenschappelijke literatuur aangetoond dat een vaste planning en programmering van de evaluaties een voorwaarde is voor goede en objectieve evaluaties. Hoewel alle ministeries aan deze voorwaarde voldoen – ze stellen meerjarige programma's voor evaluatie en onderzoek in de rijksbegroting van ieder ministerie op – biedt deze planning geen zekerheid voor het op tijd verrichten van geplande evaluaties en ook niet voor de kwaliteit en neutraliteit van de evaluaties. De vermindering van de politieke interesse in een bepaald onderwerp of thema is naast de maatschappelijke verandering de oorzaak van de aanpassingen die in de evaluatieprogrammering worden gebracht. Als de politiek een bepaald programma of wet minder relevant acht of als problematisch beschouwd, kan de aandacht hiervoor worden afgeleid door de evaluatie uit de agenda te schrappen. Dit laat zien dat evaluaties nog steeds gebruikt worden als argumenten in het politieke spel en dat ministeries de objectiviteit van evaluaties niet kunnen garanderen. Men kan concluderen dat de politieke-bestuurlijke context waarin evaluaties worden gepland en gedaan zeer belangrijk is voor het uitvoeren van goede en objectieve evaluaties.

5.4.2 Stakeholdersbetrokkenheid

Er zijn buiten de rijksoverheid weinig andere stakeholders, doelgroepen en belangenorganisatie, die belang hebben bij evaluaties. Daaronder vallen de uitvoeringsorganisaties zoals gemeenten en provincies. In enkele gevallen, zoals bij infrastructurele projecten van het ministerie van IenM (die grote subsidies verstrekt) hebben vaak meerdere partijen belangen bij de evaluatieuitkomsten. Om de objectiviteit en neutraliteit van het evaluatieproces en de evaluatiebevindingen te waarborgen, besteden departementen (wets)evaluaties uit. Dit verhoogt de kans op draagvlak en op de acceptatie van evaluatiebevindingen door alle partijen (Interview 16 juli 2015).

- *Alle departementen noemen de politieke betrokkenheid (druk van de politieke top of de Tweede Kamer, het belang van de politiek in een bepaalde thema) als de belangrijkste omgevingsfactor die zowel bevorderend als remmend kan op de activiteit van evalueren kan werken.*
- *De politieke betrokkenheid heeft een grote invloed zowel op het vermogen van overheidsinstellingen om evaluaties te verrichten als ook op de kwaliteit van evaluaties. Zo is de jaarlijkse bijstelling van de evaluatieplanning bij departementen onderworpen aan de wil en de interesse van de politieke arena in een bepaalde thema.*
- *Enerzijds worden evaluaties uitgesteld o.a. omdat de politieke aandacht op een ander urgent probleem is gevestigd. Anderzijds is de politieke druk een van de belangrijkste redenen dat evaluaties soms te vroeg (bijvoorbeeld twee jaar na de inwerkingtreding van de wet) worden verricht.*
- *De stakeholdersbetrokkenheid speelt een minder belangrijke rol bij evaluaties. Stakeholders zoals uitvoeringsorganisaties (gemeenten en provincies) worden door departementen uitgenodigd hun inbreng vooral in de fase van de voorbereiding van de wet of het beleid te doen.*
- *Bij grote projecten waarbij veel stakeholders betrokken zijn, is het laten uitvoeren van evaluaties door externe onderzoeksbureaus een manier om de objectiviteit van evaluatie en het dialoog tussen partijen te borgen. Dit verhoogt de kans op draagvlak en op de acceptatie van evaluatiebevindingen door alle partijen.*

5.5 Samenvatting

Dit hoofdstuk heeft de bevindingen van het empirisch onderzoek weergegeven en uitgelegd. Het empirisch onderzoek heeft de mate van de evaluatiecapaciteit van acht ministeries geanalyseerd aan de hand van drie aspecten van evaluatiecapaciteit: de institutionele inrichting van evaluatie, de individuele competenties en de omgevingsfactoren die de invloed van buiten uitoefenen op de evaluatieactiviteit zoals ingericht en uitgevoerd door ministeries.

Het algemene beeld uit de bevindingen van deze studie is dat de inrichting van de evaluatiefunctie sterk verschilt per ministerie. De diversiteit in de inrichting van de evaluatie is het gevolg van het gedecentraliseerde karakter van de Nederlandse rijksoverheid.

Een overeenkomst tussen departementen is dat de directies Wetgeving verassend genoeg weinig betrokken zijn bij ex post evaluaties van wetten en in ieder geval geen centrale rol hierin spelen. Volgens een meerderheid van de geïnterviewde respondenten zijn wetsevaluaties in de

meeste gevallen moeilijk te scheiden van beleidsevaluaties. Wetten worden vaak ingezet als beleidsinstrument naast andere beleidsinstrumenten. Een gevolg hiervan is dat departementen weinig speciale voorzieningen en speciale programma's en handreikingen voor het doen wetsevaluaties hebben ontwikkeld. Slechts één ministerie (VWS) werkt met een speciaal programma voor de evaluatie van wet- en regelgeving (Programma Evaluatie Regelgeving van ZonMw).

Een probleem dat bij alle departementen speelt is dat ze te laat beginnen met evalueren (gegevensverzameling, vraagstelling, uitbesteding evaluaties). Alle departementen hebben een betere planning van evaluaties en wetsevaluaties nodig. In theorie moet de planning van evaluaties tegenwicht bieden aan de politieke belangen en druk. In de praktijk hebben politieke belangen een grote invloed op de uitvoering van de geplande evaluaties.

Bij de meeste evaluaties is de enige vorm van gebruik het indirecte of conceptuele gebruik, waarbij de belangstellende – minister, Tweede Kamer, beleidsambtenaren en doelgroepen – kennis nemen van de evaluatiebevindingen en eventueel aanbevelingen. Alle departementen betwijfelen dat evaluatiebevindingen daadwerkelijk tot aanpassing van het beleid of regelgeving leiden. Slechts drie departementen konden minstens één voorbeeld geven van een wetsevaluatie die tot aanpassingen in de wet of tot nieuwe besluitvorming heeft geleid. Evaluaties en wetsevaluaties hebben dus weinig gevolgen voor de besluitvorming.

Kennis, vaardigheden en ervaring met het verrichten en begeleiden van evaluatieonderzoek is een belangrijke voorwaarde voor een goede kwaliteit van evaluaties. Doordat de borging van de kennisbehoefte niet op het centrale niveau, maar binnen ministeries ligt, verschilt de borging van individuele competenties sterk tussen ministeries. Slechts vier departementen hebben onderzoekscoördinatoren bij wie de beleidsmakers met kennisvragen over evaluatie kunnen aankloppen en de onderzoekscoördinatoren vervullen deze functie parttime naast hun werkzaamheden als beleidsmakers. De borging van de inhoudelijke en methodologische kennis voor evaluatie wordt te vaak aan de uitvoerders (meestal externe onderzoekers of onderzoeksbureaus) gelaten.

Alle departementen noemen de politieke betrokkenheid en druk vanuit de politieke top als de belangrijkste omgevingsfactoren die zowel bevorderend of remmend op de activiteit van evalueren kan werken.

De politieke betrokkenheid heeft een grote invloed zowel op het vermogen van overheidsinstellingen om evaluaties te verrichten als ook op de kwaliteit van evaluaties. Zo is de jaarlijkse bijstelling van de evaluatieplanning bij ministeries onderworpen aan de wil en de interesse van de politieke arena in een bepaalde thema. Enerzijds worden evaluaties uitgesteld o.a. omdat de politieke aandacht op een ander urgent probleem is gevestigd. Anderzijds is de politieke druk een van de belangrijkste redenen dat evaluaties soms te vroeg (bijvoorbeeld twee jaar na de inwerkingtreding van de wet) worden verricht. Dit kan de inhoudelijke kwaliteit van evaluaties, de objectiviteit van het onderzoek en de doorwerking van de evaluatiebevindingen naar besluitvorming belemmeren.

Hoofdstuk 6: Conclusie en aanbevelingen

Evaluatie is een activiteit waarbij de mens waarde aan waarnemingen en dingen toeschrijft (Swanborn, 2007). We leven in een gemeenschap waar een permanente spanning heerst om alles te meten en te controleren (Dahler-Larsen, 2012). Evaluaties dienen dit doel op een uitstekende wijze. Enerzijds worden evaluaties gedaan om verantwoording af te leggen over de keuzes die de overheid maakt en daardoor de acceptatie en de legitimatie van de overheidsinterventie in de samenleving te borgen. Anderzijds dienen evaluaties de leerfunctie. Zij leveren inzichten in de werking van wetten en andere interventies in de praktijk op met als doel de kwaliteit van wetgeving en beleid op peil te houden en zo nodig te verbeteren.

De laatste paar jaren is op internationaal niveau het besef doorgedrongen dat de kwaliteit van evaluatie niet alleen afhangt van het gebruik van criteria en standaarden voor evaluatie, maar ook van de inrichting van de evaluatiefunctie binnen de organisatie en de capaciteit die de organisatie heeft om evaluatie te verrichten of uit te zetten. Kwaliteitszorg inzake wetgeving en beleid is een groot aandachtspunt voor politici, beleidmakers en wetgevingsjuristen van de Nederlandse overheid.

De sector Wetgevingskwaliteitsbeleid (van de Directie Wetgeving en juridische zaken van het ministerie van VenJ), die verantwoordelijk is voor de kwaliteit van de wetgeving op rijksniveau, heeft geconstateerd dat de kwaliteit van verrichte wetsevaluaties fluctueert. Zo wisselt de opbouw van evaluaties heel sterk afhankelijk van uitvoerders. Daarnaast is vaak onmogelijk uit een wetsevaluatie te concluderen of de wet wel of niet heeft gewerkt en of de effecten van de onderzochte wet positief of negatief zijn. De sector Wetgevingskwaliteitsbeleid ervaart dit als een probleem omdat evaluaties veel tijd, geld en inspanning vergen, terwijl de rapporten niet altijd heldere conclusies presenteren. Dit onvermogen om heldere conclusies te formuleren kan aan verschillende factoren worden toegeschreven. Bijvoorbeeld kan dit aan de wet of de interventie liggen. Als het tekst van de wet of het beleid niet duidelijk aangeeft wat de wet/het beleid wil bereiken en via welke mechanismen dit zal worden gedaan, is het lastig om te concluderen of de doelstellingen van de wet/beleid wel of niet zijn bereikt. Een ander reden voor de onvoldoende kwaliteit van evaluaties kunnen de methodologische en inhoudelijke keuzes van zowel aanvragers als uitvoerders van evaluaties zijn voor en tijdens het onderzoek.

De sector Wetgevingskwaliteitsbeleid heeft om handvatten gevraagd die de vergelijkbaarheid en de kwaliteit van wetsevaluaties zullen vergroten. Uit de eerdere internationale onderzoeken (zie hoofdstuk 3) is gebleken dat de kwaliteit van evaluaties afhankelijk is van de evaluatiecapaciteit van de organisatie, met name van de manier waarop de evaluatiefunctie in de organisatie is ingericht en afhankelijk van de kennis over evaluatie waarover medewerkers beschikken.

Om de sector Wetgevingskwaliteitsbeleid handvatten te bieden om de kwaliteit van evaluaties te borgen, heeft deze studie onderzocht *in hoeverre de Nederlandse ministeries aan de voorwaarden voor evaluatiecapaciteit die gesteld worden in theoretische modellen, voldoen?* Zoals hoofdstuk 3 heeft afgeleid uit de wetenschappelijke literatuur, is een sterk ontwikkelde evaluatiecapaciteit van een organisatie een noodzakelijke voorwaarde voor het verrichten van evaluatie van goede kwaliteit. Dit onderzoek evalueert het vermogen van ministeries om evaluaties uit te zetten, te begeleiden en te verrichten, identificeert de uitstekende aspecten waarbij sommige ministeries te boven gaan en signaleert de knelpunten die sommige ministerie ondervinden.

Dit onderzoek heeft de evaluatiecapaciteit in lijn met de definitie van Bourgeois en Cousins (2013) afgebakend als het *vermogen* om evaluaties te *verrichten* en het vermogen om evaluatiebevindingen te *gebruiken*. Door vier theoretische modellen over evaluatie te reconstrueren, heeft dit onderzoek een conceptueel model ontwikkeld met drie aspecten van evaluatiecapaciteit: de institutionele inrichting, de individuele competenties en de omgevingsfactoren (die zowel remmend als bevorderend kunnen werken).

Om de onderzoeksvraag naar de evaluatiecapaciteit van Nederlandse ministeries te beoordelen, hebben wij interviews gehouden met sleutelrespondenten van acht ministeries die direct betrokken zijn bij de voorbereiding van evaluatietrajecten, bij de borging van de evaluatiefunctie binnen ministeries of interne evaluaties hebben verricht of begeleid. Ook hebben wij veel schriftelijk materiaal verzameld en geanalyseerd over hoe de afzonderlijke ministeries evaluaties verrichten, wat hun visies op evaluatie is, hoe ze evaluatiesopdrachten concreet aanpakken, voorbereiden en begeleiden, hoe ze de kwaliteit van evaluaties borgen en wat ze met de bevindingen uit evaluaties doen. Bij het beantwoorden van de deelvragen en de hoofdvraag van het onderzoek is gebruik gemaakt van de triangulatie van deze twee methoden: interviews en analyse van schriftelijke documenten.

6.1 Beantwoording van de onderzoeksvraag

De onderzoeksvraag wordt beantwoord door in te gaan op de drie categorieën voorwaarden voor evaluatiecapaciteit en door te beoordelen in hoeverre de Nederlandse ministeries aan deze voorwaarden voldoen.

Om de voorwaarden te identificeren zijn vier theoretische modellen gebruikt die in hoofdstuk 3 uitgebreid zijn uitgelegd. Daaruit is gebleken dat om de evaluatiecapaciteit van een organisatie afhankelijk is van drie belangrijke voorwaarden: de institutionele inrichting, de individuele competenties en omgevingsfactoren. Met behulp van wetenschappelijke literatuur over *evaluation capacity building* heeft deze studie de drie voorwaarden concreter gemaakt met behulp van de 12 aspecten: structuur van de organisatie, evaluatieplanning- en programmering, evaluatiekader, financiële en niet-financiële middelen, de communicatie, gebruik van evaluatiebevindingen, leren van evaluaties, kennis, vaardigheden en ervaring van betrokkenen bij evaluaties, politieke betrokkenheid en stakeholderbetrokkenheid.

Uit deze studie is gebleken dat de meeste aandacht bij de Nederlandse ministeries wordt gegeven aan de eerste voorwaarde voor evaluatiecapaciteit, namelijk de institutionele inrichting. Niet alle zeven aspecten van de institutionele inrichting krijgen zijn even goed ontwikkeld bij de Nederlandse ministeries. Hieronder volgen de meest belangrijke bevindingen uit het onderzoek.

Alle onderzochte ministeries hebben één of meer afdelingen die zich bezig houden met alle vormen van onderzoek waaronder ook evaluaties. Ministeries hebben echter geen evaluatieafdelingen, d.w.z. geen afdeling die exclusief op *evaluaties* van beleid en wetgeving focussen. Het enige ministerie dat een aparte instantie heeft voor de evaluatie van regelgeving is het ministerie van VWS (via de Commissie Evaluatie Regelgeving van ZonMw). Buiten ZonMw krijgt evaluatie van regelgeving geen speciaal aandacht binnen het veld van evaluatieonderzoek in het brede.

De evaluatieplanning is de afgelopen jaren gestandaardiseerd. Alle ministeries maken een meerjarige evaluatieplanning en stellen ze jaarlijks bij. Door evaluatie op tijd te plannen, worden ook de financiële middelen hiervoor veiliggesteld.

Bijna alle ministeries (met uitzondering van het ministerie van IenM) hebben een programma voor evaluatie of handreikingen, richtlijnen en leidraden ontwikkeld voor het omgaan met evaluatie. Daarbij wordt aandacht besteed aan het evaluatieproces en (in mindere mate) aan de inhoudelijke en methodologische kwaliteitscriteria voor evaluaties. Deze handreikingen zijn

echter slechts in beperkte mate gebruikt door de dossierhouders van evaluatieonderzoeken. Soms is de aanwezigheid van dergelijke handreiking niet of weinig bekend bij de beleidsmedewerkers buiten de evaluatieafdelingen; soms zijn ze wel bekend, maar worden niet gebruikt wegens concurrerende factoren zoals de politieke druk.

De meeste ministeries investeren weinig in niet-financiële middelen zoals menselijk kapitaal en maken weinig ruimte in de agenda van beleidsmedewerkers voor evaluatieonderzoek. Enkele positieve voorbeelden hiervan zijn het ministerie van EZ, het ministerie van OCenW en het ministerie van VenJ. EZ en OCenW bieden de beleidsmedewerkers de kans om cursussen en trainingen te volgen over kwantitatieve en kwalitatieve methoden van onderzoek. Ministeries hebben echter geen stok achter de deur om beleidsmedewerkers te dwingen om hun kennis over het evaluatieproces en de methoden en technieken voor evaluatie te verfijnen. Het onderzoeksinstituut WODC (dat in opdracht van het ministerie van VenJ evaluaties verricht of uitbesteed) heeft een goede basis van menselijk kapitaal voor evaluaties (onderzoekers, opstellers van startnotities voor evaluaties en projectbegeleiders).

Het gebruik van evaluatiebevindingen is niet systematisch geborgd. Er zijn formele procedures voor de presentatie van het evaluatierapport aan de Tweede Kamer samen met de kabinetsreactie. Het komt echter niet heel vaak voor dat evaluaties een directe werking hebben op de besluitvorming en het nieuwe beleid, wet- en regelgeving. De twee meest voorkomende aanleidingen voor het evalueren van beleid en wetten zijn de politieke interesse voor het succes of vallen van een bepaald beleid of een wet en de verplichting om te evalueren (zoals de evaluatiebepaling van sommige wetten).

De leerfunctie van evaluaties is secundair aan de verantwoordingsfunctie. Ministeries evalueren vaak omdat de evaluatie een evaluatiebepaling bevatten of omdat de evaluatie in de planning staan. Evaluatie worden in veel mindere mate uitgevoerd om er lessen uit te trekken. Evaluaties dienen als *evidence based* bewijs voor of tegen het gevoerde beleid of wet. Als het om het verspreiden van *good practices* van evaluatieonderzoek gaat, leren dossierhouders vooral op een informele manier: door feedback te vragen bij meer ervaren collega's binnen ministerie en door vergaderingen van interdepartementale overlegplatformen bij te wonen.

Wat het tweede aspect van de evaluatiecapaciteit betreft, besteden ministeries weinig aandacht aan de individuele competenties voor evaluatieonderzoek. Evaluatieonderzoek is slechts een klein onderdeel van het grote takkenpakket van beleidsmakers. Omdat het merendeel van

evaluaties bij de Nederlandse ministeries wordt uitbesteed, zijn de individuele competenties voor evaluatieonderzoek weinig ontwikkeld bij de Nederlandse ministeries.

Uit deze studie is gebleken dat uit de twee onderzochte omgevingsfactoren – het derde aspect van evaluatiecapaciteit – slechts de politieke betrokkenheid een invloedrijke uitwerking heeft op de evaluatiecapaciteit van ministeries. Enerzijds heeft de politieke betrokkenheid een grote invloed op de evaluatiecapaciteit bij ministeries. De interesse van politieke actoren om te leren hoe efficiënt het beleid of een wet is geweest en de politieke steun voor evaluatieactiviteit bepaalt voor een groot deel zowel de kwaliteit van evaluaties als ook het vermogen van ministeries om evaluaties te verrichten. Anderzijds is de overheid, naast de politieke arena, de enige actoren met grote interesse in de evaluatiebevindingen. Andere stakeholders zoals uitvoeringsorganisaties, privé partijen en doelgroepen hebben nauwelijks invloed op het vermogen van ministeries om evaluaties te verrichten. Om draagvlak te creëren voor het verrichten van evaluaties en om te waarborgen dat de stakeholders aan evaluaties meewerken, worden ze bij evaluaties met veel impact betrokken via consultatierondes en panels.

Al met al zijn de ministeries niet helemaal in staat om (wets)evaluaties van goede kwaliteit te verrichten. Geen ministerie voldoet aan alle voorwaarden voor een goede evaluatiecapaciteit.

6.2 Kritische terugblik op het onderzoek

Deze paragraaf geeft inzicht in de belemmering van het onderzoek. Hierbij wordt ingegaan op de onderzoeksmethode, de onderzochte populatie en de selectie van respondenten.

In het empirisch onderzoek is gebruik gemaakt van de kwalitatieve onderzoeksmethode met daarin twee methoden van dataverzameling: de inhoudsanalyse van documenten en semigestructureerde interviews. Het voordeel hiervan is dat we *in-depth* gesprekken hebben kunnen houden met mensen die vanuit verschillende functies betrokken zijn bij (wets)evaluaties. Het nadeel is dat we slechts een paar mensen per ministerie hebben kunnen benaderen. Om meer mensen te benaderen en een uitgebreider overzicht te krijgen van hoe betrokkenen de evaluatiecapaciteit van ministeries inschatten, zou het huidige onderzoek met een kwantitatieve onderzoeksmethode, zoals het survey kunnen worden aangevuld.

Gezien de beperkte tijd die beschikbaar was voor het onderzoek, heeft dit onderzoek slechts acht uit de elf ministeries kunnen betrekken. Het ministerie van Algemene Zaken en het ministerie van Defensie hebben nauwelijks wetten en doen bijgevolg geen wetsevaluaties. Wij hebben

daarom deze twee ministeries van het onderzoek uitgesloten. Het ministerie van Binnenlandse Zaken is uitgesloten uit de analyse omdat de benaderde respondenten niet binnen de periode van het onderzoek hebben gereageerd.

Verder is de selectie van respondenten beperkt tot de vraagzijde van het evaluatietraject, dat wil zeggen tot medewerkers van ministeries die evaluaties aanvragen, begeleiden of ondersteunen. Er zijn onderzoekers geïnterviewd die uitbestede evaluaties uitvoeren. Ten eerste was het niet haalbaar binnen de korte tijd van het onderzoek ook respondenten van buiten de ministeries te interviewen. Ten tweede zou dit slechts ten dele relevant zijn voor de invalshoek waaruit dit onderzoek naar evaluatie kijkt. Uitvoerders van uitbestede evaluaties zoals academische onderzoekers en onderzoeksbureaus zouden wel een ander blik kunnen werpen op de kwaliteit van evaluaties en over de belemmeringen voor het verrichten van goede evaluaties. Hun ervaring met evaluatie zou echter niet even relevant zijn voor het begrijpen van hoe goed de evaluatiecapaciteit van ministeries is ontwikkeld.

6.3 Aanbevelingen naar aanleiding van het onderzoek

Naar aanleiding van het empirisch onderzoek en de resultaten hiervan formuleren wij de volgende aanbevelingen. Deze kunnen fungeren als hulpmiddel voor de sector Wetgevingskwaliteitsbeleid en voor andere directies die zorg dragen voor de inrichting van de evaluatiefunctie bij de Nederlandse ministeries om een plan van aanpak op te stellen om de kwaliteit van evaluaties en het gebruik daarvan te verbeteren.

Institutionele inrichting van de evaluatie

Bij alle ministeries zouden de *samenwerking* en *communicatie* tussen de beleidsdirecties, ondersteunende directies, de Directie Wetgeving en FEZ kunnen worden versterkt. Alle ministeries hebben minstens één handreikingenboekje voor het uitzetten en uitvoeren van evaluaties. Het probleem is dus niet het bestaan van handreikingen, maar de *beperkte verspreiding* hiervan. Veel medewerkers die betrokken zijn bij het uitzetten, begeleiden of verrichten van evaluaties zijn niet op de hoogte dat hun ministerie handreikingen heeft ontwikkeld. Het is belangrijk dat alle directies elkaar op de hoogte stellen van elkaars plannen, agenda's en hulpmiddelen inzake evaluatie.

Elk ministerie zou de relevante informatie met betrekking tot evaluatie – zoals de betrokken directies, de namen van coördinatoren, hulpmiddelen zoals handreikingen en leidraden voor selectie van uitvoerende onderzoeker en de begeleiding van het evaluatieproces etc. – op een gecentraliseerde website of interne portaal kunnen bundelen.

De sector Wetgevingskwaliteitsbeleid van het ministeries van VenJ zou een netwerkgroep of klankbord kunnen oprichten, vergelijkbaar aan de interdepartementale beleidsdoorlichtingsgroep waarin vertegenwoordigers van alle ministeries zitting nemen en waar ministeries met elkaar *beste practices* delen. De rol van een dergelijk netwerkgroep is kennis en ervaring tussen ministeries uit te wisselen. Als een ministerie voor een uitdagend evaluatieonderzoek staat, kan het suggesties en voorstellen bij deze klankbord inwinnen.

Een andere aanbeveling is dat ministeries eerder met het voorbereiden van evaluaties moeten beginnen. Zodra een wet of een interventie is ingevoerd, zou het verantwoordelijke ministerie al erover moeten nadenken wat voor informatie nodig is om de effecten van de interventie te evalueren en de informatie daadwerkelijk te vergaren.

Tot slot zou de leerfunctie een centralere plek moeten krijgen dan in het huidige evaluatiestelsel.

Individuele competenties

Dit onderzoek heeft geconcludeerd dat veel kennis en expertise over evaluatie binnen de ministeries op een informele manier wordt gedeeld. Het is daarom ook vooral moeilijk te zeggen of de kennis, vaardigheden en ervaring voldoende of onvoldoende zijn ontwikkeld bij ministeries. Wat dit onderzoek wel kan concluderen is dat veel kennis en expertise vaak geconcentreerd worden bij een klein aantal personen. Het begeleiden van evaluaties op een professionele manier is te veel een kwestie van toeval, of de mensen met de juiste kennis beschikbaar zijn voor de begeleiding van het ene of het andere evaluatietraject. De oplossing hiervoor is dat alle beleidsmedewerkers die in aanraking (zouden kunnen) komen met evaluaties een minimale training in onderzoeksmethoden en in het formuleren van vraag- en probleemstelling moeten volgen. Daarvoor kan een korte cursus of een reeks presentaties op het rijksniveau door een groep van onderzoekers (die van methodische en inhoudelijke kennis beschikken) en betrokkene uit het veld worden verzorgd.

Een andere aanbeveling is dat alle onderzoeksteams aan de voorwaarde van multidisciplinariteit moeten voldoen. Vooral bij wetsevaluaties is het van groot belang dat het onderzoeksteam de juridische en de sociologische kennis en expertise bundelen.

Omgevingsfactoren

Evaluaties bieden idealiter tegenwicht aan de dynamiek en instabiliteit die de politieke sfeer kenmerken. Zo zouden evaluaties een objectief en neutraal beeld moeten geven op de effectiviteit en efficiëntie van de interventie of de ingevoerde wet. Zoals het huidige onderzoek laat zien, is de invloed van politiek echter een zeer belangrijke remmende omgevingsfactor. Zo worden soms evaluaties te vroeg verricht omdat de Tweede Kamer erom vraagt. De politieke agenda zet een mogelijke blokkade ook in de doorwerking van evaluatiebevindingen. Soms worden wetten en beleid ingevoerd omdat er al afspraken zijn gemaakt, terwijl eerdere evaluatiebevindingen een bepaalde wet of beleid afraden.

Hoewel de verplichte evaluaties vaak wel afgerond worden, komt het soms voor dat ze door de veranderingen in de samenstelling van de politieke context als minder relevant worden gezien. Dan is het gebruik van evaluatiebevindingen beperkt tot de kennisname van het rapport zonder de aanbevelingen terug te koppelen naar de besluitvormingsfase. Soms worden evaluatie uitgesteld of nooit uitgevoerd omdat de belangstelling van de Tweede Kamer voor het betreffende beleid, wet of terrein is weggezaakt.

Als oplossing hiervoor zouden de ministers in de brief naar de Tweede Kamer concrete maatregelen moeten noemen over hoe de bevindingen daadwerkelijk een plek zullen krijgen in toekomstige wetten en beleid. Zo zou al het nieuw beleid en wetten direct moeten verwijzen naar de resultaten van al afgeronde evaluaties en laten zien dat en hoe de besluitvorming gebaseerd is op conclusies en aanbevelingen uit eerdere evaluaties. Op deze manier kan de beleidscyclus als een geheel functioneren en wordt de continuïteit tussen de fases van de beleidscyclus waarborgt.

6.4 Suggesties voor vervolgonderzoek

In het volgende stellen wij een aantal punten voor vervolgonderzoek voor, waarbij wordt aangegeven waarom het nodig is om deze punten verder te onderzoeken.

Ten eerste zou het huidige onderzoek met een vergelijking tussen ministeries aangevuld kunnen worden. Het huidige onderzoek heeft zich geconcentreerd op het identificeren van de

aspecten die essentieel zijn voor het evalueren van de evaluatiecapaciteit van publieke organisaties en het toepassen van deze aspecten om de evaluatiecapaciteit van de Nederlandse ministeries te beoordelen. Dit onderzoek heeft *best practices* geïdentificeerd. *Best practices* zijn voorbeelden van ministeries die het bijzonder goed doen op het ene of andere aspect in vergelijking met de andere onderzochte ministeries. Geen enkel ministerie doet het uitstekend goed op alle aspecten. Daarnaast heeft het huidige onderzoek ook voorbeelden van ministeries geïdentificeerd, die slecht scoren op het ene of andere aspect. In dit geval is het ene of andere aspect totaal verwaarloosd bij het ene of andere ministerie waardoor de evaluatiecapaciteit niet wordt geborgd. Het was niet de bedoeling van dit onderzoek om een volledig en systematisch beeld te geven van alle aspecten van de evaluatiecapaciteit van alle ministeries. Dit zou in een vervolgonderzoek kunnen worden bewerkstelligd.

Hiervoor zou de kwalitatieve methodologie met een kwantitatieve methodologie aangevuld kunnen worden. Zo zou men een survey kunnen opzetten en onder de medewerkers van ministeries verspreiden om alle aspecten van de evaluatiecapaciteit volledig af te dekken.

Ten tweede is in dit onderzoek de mate van de evaluatiecapaciteit *binnen* ministeries bestudeerd. Gezien de scheiding die Nielsen et al. (2011) tussen de *vraagzijde* en de *aanbodzijde*, zou men ook de aanbodzijde, d.w.z. de externe (uitvoerende) kant van de evaluatiecapaciteit in het onderzoek kunnen betrekken. Het bestuderen van de interactie tussen de vraagzijde en de aanbodzijde, tussen de aanvragers van evaluaties en de uitvoerende onderzoekers (academische onderzoeksteams en onderzoeksbureaus) zou een beter beeld van het evaluatiecapaciteit van ministeries kunnen geven.

Ten derde zou het verband tussen de evaluatiecapaciteit en de kwaliteit van evaluaties kunnen worden bestudeerd. Dit verband is als assumptie genomen in het huidige onderzoek. In een vervolgonderzoek zou men de *black box* moeten openen en toetsen of een goede inrichting van de evaluatiefunctie daadwerkelijk tot een betere kwaliteit van evaluaties leidt.

Bronnen

Academie voor Wetgeving. (2015). *De Wetgevingsjurist*. Geraadpleegd op 23 oktober 2015 op <http://academievoorwetgeving.nl/page/de-overheidsjurist>.

Aeken, K van. (2011). From Vision to Reality: *Ex post* Evaluation of legislation. *Legisprudence*, 5(1), 41-68.

Algemene Rekenkamer. (2013). *Effectiviteitsonderzoek bij de rijksoverheid: vervolgonderzoek*. Den Haag: Algemene Rekenkamer.

Algemene Rekenkamer. (2012). *Effectiviteitsonderzoek bij de rijksoverheid*. Den Haag: Algemene Rekenkamer.

Baizerman, M., Compton, D.W. en Stockdill, S. H. (2002). New Direction for ECB. *New Directions for Evaluation*, 93, 109 -119.

Bemelmans-Videc, M.L. (2002). Evaluation in the Netherlands 1990-2000: Consolidation and Expansion. In J.-E. Furubo, R.C. Rist en R. Sandahl (Eds.). *International Atlas of Evaluation*. New Brunswick: Transaction Publishers.

Bemelmans-Videc, M.L., Elte, R. en Koolhaas, E. (1990). Policy Evaluation in the Netherlands: Institutional Context and State of Affairs. In R.C. Rist. (Ed.). *Program Evaluation and the Management of Government. Patterns and Prospects across Eight Nations*. New Brunswick: Transaction Publishers.

Bleijenbergh, I.L. (2013). *Kwalitatief onderzoek in organisaties*. Den Haag: Boom Lemma Uitgevers. (1^e druk).

Bonte, W. *Handreikingen evaluatieonderzoek EZ*. Intern document van EZ, Jaartal onbekend.

Bourgeois, I. en Cousins, J.B. (2013). Understanding Dimensions of Organizational Evaluation Capacity. *Americical Journal of Evaluation*, 34(3), 299-319. DOI: 10.1177/1098214013477235.

Bussmann, W. (2010). Evaluation of Legislation: Skating on Thin Ice. *Evaluation*, 16(3), 279-293. DOI: 10.1177/1356389010370252.

Coglianesi, C. (2012). Measuring Regulatory Performance. Evaluating the Impact of Regulation and Regulatory Policy. OECD.

Commissie Theeuwes. (2012). *Durf te meten. Eindrapport Expertwerkgroep Effectmeting*. Geraadpleegd op 8 september 2015 op <https://www.rijksoverheid.nl/documenten/rapporten/2012/11/23/durf-te-meten-eindrapport-expertwerkgroep-effectmeting>.

Cooksy, L.J. (2012). Influences on evaluation quality. *American Journal of Evaluation*, 33(1), 79-87. DOI: 10.1177/1098214011426470.

Dahler-Larsen, P. (2013) *The Evaluation Society*. Stanford: Stanford University Press.

Delahais, Th. (2014). Ex post evaluation of regulation and regulatory policies - The case of EU regulation. *Quadrant Conseil*.

Davidson, E.J. (2005). *Evaluation Methodology Basics. The Nuts and Bolts of Sound Evaluation*. Thousand Oaks: Sage Publications.

Jacob, S., Speer, S. en Furubo, J.E. (2015). The institutionalization of evaluation matters: updating the International Atlas of Evaluation 10 years later. *Evaluation* 21(1), 6-31.

Kenniscentrum Wetgeving en Juridische zaken (KWJZ). (2015a). *Evaluatiebeleid wetgeving*. Geraadpleegd op 28 mei 2015 op <https://www.kcwj.nl/wetgevingsbeleid/evaluatiebeleid-wetgeving>.

Kenniscentrum Wetgeving en Juridische zaken (KWJZ). (2015b). *Integraal Afwegingskader beleid en regelgeving*. Geraadpleegd op 15 september 2015 op https://www.kcwj.nl/kennisbank/integraal-afwegingskader-beleid-en-regelgeving?ip_login_no_cache=6d668dfff93b38d6b81672d6721a234f.

Kennisplatform Crow. (2014). *Leren en beleid bijstellen*. Geraadpleegd op 8 september 2015 op <http://www.crow.nl/mobiliteit-en-gedrag/weblog/november-2014-%281%29/leren-en-beleid-bijstellen>.

Klein Haarhuis, C., Smit, M. en Keulemans, S. (2014). Ex ante onderzoek in beeld: over aard, aantal en gebruik van ex ante onderzoek bij beleidsvoorbereiding, *Beleidsonderzoek Online*. DOI: 10.5553/Beleidsonderzoek.00004. Geraadpleegd op 17 juni 2015 op <http://www.beleidsonderzoekonline.nl/downloads/pdf/bso/Beleidsonderzoek-2014-10.pdf>.

Klein Haarhuis, C.M. en Niemeijer, E. (2009). Synthesizing Legislative Evaluations. Putting the Pieces Together. *Evaluation*, 15(4), 403-425. DOI: 10.1177/1356389009341897.

Kushner, S. (2005). Qualitative Control. A Review of the Framework for Assessing Qualitative Evaluation. *Evaluation*, 11(1), 111-122. DOI: 10.1177/1356389005053194.

Leeuw, F.L. (2009). Evaluation Policy in the Netherlands. In W.M.K. Trochim, M.M. Mark en J. Cooksy (Eds.). *Evaluation policy and evaluation practice. New Directions for Evaluation*, 123, 87-102.

Leeuw, F.L. en Furubo, J.-E. (2008). Evaluation Systems. What are They and Why Study Them?. *Evaluation* 14(2), 157-169.

Leeuw, F.L., Rist, R.C en Sonnichsen, R.C. (1994). *Can Governments Learn? Comparative Perspectives on Evaluation & Organizational Learning*. New Brunswick: Transaction Publishers.

Mackay, K. (2006). Evaluation Capacity development. Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management. *ECD Working Paper Series, 15*, Washington D.C.: The World Bank.

Meyenfeldt, L. von, Schrijvershof, C. en Wilms, P. (2008) *Tussenevaluatie beleiddoorlichting*. Den Haag: Aarts De Jong Wilms Goudriaan Public Economics (APE), rapport nr. 570.

Ministerie van Buitenlandse Zaken. (2015). *Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie (IOB)*. Geraadpleegd op 8 september 2015 op <https://www.rijksoverheid.nl/ministeries/ministerie-van-buitenlandse-zaken/inhoud/organisatie/beleidsevaluatie/iob>.

Ministerie van Buitenlandse Zaken. (2009). *Evaluatiebeleid en richtlijnen voor evaluaties*. Den Haag: Inspectie Ontwikkelingssamenwerking en Beleidsevaluaties.

Ministerie van Financiën. (2015a). *Evaluaties en beleidsdoorlichtingen*. Geraadpleegd op 29 mei 2015 op <http://www.rijksbegroting.nl/beleidsevaluaties/evaluaties-en-beleidsdoorlichtingen>.

Ministerie van Financiën. (2015b). *Evaluatie van beleid*. Geraadpleegd op 29 april 2015 op <http://www.rijksoverheid.nl/onderwerpen/evaluaties-van-beleid/evaluaties>.

Ministerie van Financiën. (2014). *Regeling Periodiek Evaluatieonderzoek*. Staatscourant Nr. 27142.

Ministerie van Financiën. (2006). *Regeling Periodiek Evaluatieonderzoek*. Staatscourant Nr. 83.

Ministerie van Financiën. (2003). *Handreiking Evaluatieonderzoek ex post*. Den Haag: Directie Begrotingszaken. Geraadpleegd op 28 augustus 2015 op www.minfin.nl/vbtb.

Ministerie van Infrastructuur en Milieu. (2015). *Werkprogramma 2015*. Den Haag: Kennisinstituut voor Mobiliteitsbeleid.

Ministerie van Veiligheid en Justitie (2015). *Overheidsorganisaties*. Geraadpleegd op 24 oktober 2015 op [https://almanak.overheid.nl/46787/VenJ/Directie_Wetgeving_en_Juridische_Zaken_\(DWJZ\)/Sector_Wetgevingskwaliteitsbeleid/](https://almanak.overheid.nl/46787/VenJ/Directie_Wetgeving_en_Juridische_Zaken_(DWJZ)/Sector_Wetgevingskwaliteitsbeleid/).

Ministerie van Justitie. (2000). *Nota Wetgevingskwaliteitsbeleid en wetgevingsvisitatie*. Tweede Kamer, vergaderjaar 2000–2001, 27475, nr. 2. Den Haag: Sdu Uitgevers.

Ministerie van Justitie. (1991). *Nota Zicht op wetgeving*. Tweede Kamer, vergaderjaar 1990-1991, 22008, nrs. 1-2. Den Haag: Sdu Uitgevers.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2015a). *Organogram*. Geraadpleegd op 7 september 2015 op <https://www.rijksoverheid.nl/ministeries/ministerie-van-onderwijs-cultuur-en-wetenschap/inhoud/organisatie/organogram>.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2015b). *Notitie over Zicht op Effectiviteit*. Intern document van Directie Kennis van het ministerie van Onderwijs, Cultuur en Wetenschap.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2015c). *Kennisagenda OCenW*. Geraadpleegd 30 juni 2015 op <https://www.rijksoverheid.nl/documenten/publicaties/2015/03/13/kennisagenda-ocw-2015>.

Ministerie van Sociale Zaken en Werkgelegenheid. (2012). *Introductiedossier nieuwe bewindspersonen 2012*. Geraadpleegd op 7 september 2015 op [dossier-nieuwe-bwp-szw.pdf](#).

Nielsen, S.B., Lemire, S. en Skov, M. (2011). Measuring Evaluation Capacity – Results and Implications of a Danish Study. *American Journal of Evaluation*, 1-21. DOI: 10.1177/1098214010396075.

Olsthoorn-Heim, E.T.M. (2003). *Vijf jaar evaluatie regelgeving via ZonMw*. Den Haag: ZonMw.

Preskill, H. en Boyle, S. (2008). A Multidisciplinary Model of Evaluation Capacity Building. *American Journal of Evaluation*, 29(4), 443-459. DOI: 10.1177/1098214008324182.

Reininga, T. (2014). *Evalueren en leren op een departement?*. Presentaties over het programma Zicht op effectiviteit. Ministerie van Onderwijs, Cultuur en Wetenschap.

Rossi, P.H. en Freeman, H.E. (1993). *Evaluation: a systematic approach*. 5e ed. Thousand Oaks, CA: SAGE.

Schouten, S. (2013). *Meerwaarde van evaluaties voor democratie en openbaar bestuur. Presentatie van de diagnose en de denkrichting van het WRR rapport over evaluatie*. Den Haag. Geraadpleegd op 17 juni 2015 op <https://www.kcwj.nl/wetgevingsbeleid/evaluatiebeleid-wetgeving/clearing-house-voor-wetsevaluatie/bijeenkomst-over-de?cookie=yes.1434476648920774270597>.

Stevahn, L., King, J.A., Ghore, G. en Minnema, J. (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation* 26 (1), 43-59. DOI: 10.1177/1098214004273180.

Swanborn, P.G. (2007). *Evalueren. Het ontwerpen, begeleiden en evalueren van interventies. Een methodische basis voor evaluatie-onderzoek*. Den Haag: Boom Lemma Uitgevers.

Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D.B. en Balcazar, F.E. (2013). Understanding and Measuring Evaluation Capacity: A Model and Instrument Validation Study. *American Journal of Evaluation* 34(2), 190-206. DOI: 10.1177/1098214012471421.

Thiel, S. van. (2013). *Bestuurskundig onderzoek. Een methodologische inleiding*. Bussum: Uitgeverij Coutinho. Tweede herziene druk.

Vedung, E. (1997). *Public Policy and Program Evalution*. New Brunswick: Transaction Publishers.

Veerman, G.J. (2014). Klarheid over het Clearing House. *RegelMaat*, 29 (4), 199-211.

Veerman, G.J., m.m.v. Mulder, R.J. en Meijsing, E.S.M. (2013). *Een empathische wetgever. Meta-evaluatie van empirisch onderzoek naar de werking van wetten*. Den Haag: Sdu Uitgevers.

Veerman, G.J. m.m.v. Hendriks-de Lange, S.R. (2007). *Over wetgeving. Principes, paradoxen en praktische beschouwingen*. Den Haag: Sdu Uitgevers.

Verschuren, P.J.M. en Doorewaard. H. (2007). *Het ontwerpen van een onderzoek*. Den Haag: Uitgeverij LEMMA. Vierde druk.

WODC. (2015). *Onderzoeksprogramma WODC*. Geraadpleegd op 8 september 2015 op <https://www.wodc.nl/onderzoek/onderzoeksprogramma/index.aspx>.

WODC. (2012). *Wegwijzer bij contractonderzoek in opdracht van het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het ministerie van Veiligheid en Justitie*. Den Haag: WODC, ministerie van Veiligheid en Justitie.

ZonMw. (2015). *Over ZonMw*. Geraadpleegd op 15 september 2015 op <http://www.zonmw.nl/nl/over-zonmw/over-zonmw/>.

ZonMw. (2013). *Programma Evaluatie Regelgeving*. Den Haag: ZonMw.

Bijlage A: Topiclijst (gebruikt bij het interviewen van respondenten)

Dimensie 1: Individuele competenties (kennis, vaardigheden, ervaring)

Achtergrondkenmerken aanvragers van wetsevaluaties

- Leeftijd / jaren ervaring als...
- Aantal jaren in huidige functie
- Eventuele ervaringen met dit onderwerp elders (ander ministerie / directie) --> mogelijk ook resulterend in leerervaringen

Dimensie 2: Institutionele inrichting

Arrangementen (en bijkomende activiteiten) voor wetsevaluatie

Voor leidinggevendenden directies wetgeving (en evt. directies kennis)

- Hoeveel keren per jaar heeft uw [afdeling....] te maken met wetsevaluaties?
- Welke structurele arrangementen [procedures, processen, handreikingen] zijn er binnen uw directie / ministerie met betrekking tot het opstellen, begeleiden en uitvoeren van wetsevaluatie en evaluatie algemeen?
- Voor wie zijn deze bedoeld?
 - Voor u
 - Voor onderzoekers
 - Voor allebei
- Welke doelen worden hiermee nagestreefd?
- Welke 'rijksbrede' of interdepartementale arrangementen zijn u bekend? In hoeverre maakt u (of de onderzoekers die uw aanvraag behandelen) daarvan gebruik?

Voor aanvragers van wetsevaluaties (en evt. ook leidinggevendenden)

- Maakt u bij het aanvragen, uitbesteden, uitvoeren van een (wets)evaluatie gebruik van bestaande arrangementen [procedures, processen handreikingen]?
 - Zo ja, van welke arrangementen maakt u dan gebruik?
 - Binnen uw ministerie of directie
 - Interdepartementaal/ rijksbreed
 - Zo nee, waarom niet?

- Kent u ze niet / zijn ze er niet / vindt u ze niet bruikbaar of niet van toepassing / andere reden?

Voor iedereen

- Hoeveel procent van het onderzoek binnen/voor uw directie/ministerie betreft een evaluatie? Hoeveel procent van de evaluaties betreft een wetsevaluatie? (schatting volstaat)
- Werkt u met een evaluatieprogrammering of evaluatieplanning?
 - Zo ja, hoe gaat dit in zijn werk?
 - Met welk doel is dat?
 - Hebben wetsevaluaties hierin een aparte plek?
 - Indien wettelijk verplicht: komen ze te vroeg; zit hier een ontwikkeling in?
- In hoeverre beschouwt u / uw directie / uw ministerie wetsevaluatie als een bijzondere / afzonderlijke vorm van evaluatie?
 - Waar blijkt dat uit? (bijv. complexiteit van wetten en methodologische implicaties daarvan; infrastructureel; *awareness*; juridische elementen)
 - Zijn er aparte voorzieningen voor wetsevaluaties zoals handreikingen, procedures, standaarden etc.?
 - Departementaal
 - Interdepartementaal
 - Is er een (visie op) *Betere Regelgeving* waar wetsevaluatie wellicht deel van uitmaakt?
 - Welke eigenschappen van wetten maken volgens u dat de aanpak van de evaluatie van wetgeving anders is dan die van bijv. beleidsinterventies of een programma?
- In hoeverre worden evaluaties en wetsevaluaties in het bijzonder binnen uw ministerie/directie *uitbesteed*?
- Hoe gaat het verdere proces van evalueren / uitbesteden binnen uw directie / ministerie in zijn werk?

- Graag ook aangeven als er géén vaste praktijk is (en toelichten waarom u denkt dat dit zo is)
- Kiest u wel eens voor een vóóronderzoek? Licht toe.
- Op welke vaardigheden van de uitvoerders van het evaluatieonderzoek zet u in?
 - Bijv. toegankelijkheid, gevoeligheid voor de evaluatiecontext; of meer technische vaardigheden zoals juridisch, sociaal- wetenschappelijk en methodologisch, inhoudelijke kennis en expertise.
- Worden bij de opzet en/of uitvoering en communicatie van de evaluatie (overige) *stakeholders* betrokken? (of eventueel al eerder in het proces?)
- Heeft u ervaring met interdepartementale evaluatietrajecten?
- In hoeverre weet u ‘wat te doen’ als er een wetsevaluatie op handen is?
 - Licht toe, hoe komt dat?
 - Geldt dat ook voor de gemiddelde aanvrager binnen uw directie/ministerie?
- Waar let u op, als u een wetsevaluatie aanvraagt / uitbesteedt en vervolgens als dit proces eenmaal loopt?
- Op welke fronten mist u zelf als aanvrager belangrijke kennis / vaardigheden, naar uw idee?
 - Licht toe
 - Wat zou hier achter kunnen liggen?
- Heeft u ervaring met wetsevaluaties, met overige evaluaties (bijv. van interventies / beleidsprogramma’s, of met beide?).
 - Indien met beide: waar ligt het zwaartepunt?
- Wat zijn uw ervaringen met (wets)evaluaties (door de tijd) ten aanzien van:
 - Afbakening en formulering van de aanvraag
 - Constructies (bijv. stuurgroep / evaluatiecommissie, begeleidingscommissie
 - En hun samenstelling
 - Budgettair
 - Keuzes in de aanpak
 - Verhouding aanvragers / onderzoekers
 - Waarborgen inhoud

- Waarborgen kwaliteit
- Neerslag van bevindingen
- Communicatie van bevindingen
- Stappen / beslissingen t.a.v. gebruik

Als de respondent meer focus nodig heeft, het volgende uitleggen (als handreiking geven).

Mogelijke arrangementen binnen uw organisatie:

- (A) Units, functies speciaal ingericht voor kwaliteitsvergroting, evaluatietraining of controle op inhoud / kwaliteit /gebruik;
- (B) (online) Procedures, werkprotocollen, richtlijnen, standaarden;
- (C) (online) Tools en voorzieningen, handreikingen, kennisoverzichten;
- (D) Studiemiddagen, (lunch)lezingen, bezoeken van conferenties, etc;

Concrete structurele activiteiten binnen uw organisatie:

- Personeel en administratie voor evaluatie
 - Aantrekken van externe expertise
 - Aanbestedingsprocedures
 - Evaluatieplanning of – programmering
 - Budget voor (wets)evaluatie (structureel)
 - Bevordering van de inhoud en kwaliteit van evaluaties en/of (overig) onderzoek (bijv. door training, hulp, controle)
 - Bevorderen of zelfs waarborgen van (elke denkbare vorm van) gebruik van (wets)evaluaties voor feedback op / ontwikkeling van nieuwe wetgeving/beleid
- *(Controle vraag)* Werkt u als aanvrager met handreikingen / evaluatierichtlijnen die u helpen bij het begeleiden van de evaluatie?
 - Kunt u deze zo precies mogelijk beschrijven? Betreffen deze het proces, de inhoud (zie boven) of de evaluatiemethodologie; of meerdere?
 - Hoe heten de richtlijnen, handreikingen etc.
 - Zijn deze voor ons toegankelijk? Zo nee, kunnen wij deze via u wellicht ontvangen?

- In hoeverre maakt u gebruik van bijv. aanwijzing 164 van de Ar? Of het IAK (hoewel vooral ex ante)? Andere interdepartementale instrumenten: de aanwijzingen in de RPE en o.a. het handboek 'Evaluatie als fundering voor beleid'.
- Welke van voornoemde inhoudelijke en methodologische criteria acht u zelf – als ervaren aanvrager van wetsevaluaties – van belang?
 - Inhoud
 - Aanpak
 - Op welke manieren waarborgt u die criteria in de praktijk /helpt u die te realiseren?
- In hoeverre mist u die aspecten in de praktijk? Licht toe.

Voor iedereen: tijdelijke en overige arrangementen / bronnen [voor zover nog niet aan de orde geweest]

Met tijdelijke voorzieningen doelen we op bijvoorbeeld: een periodieke planning van een specifieke Wet, de inrichting van een tijdelijke Stuurgroep / een Evaluatiecommissie, begeleidingscommissie, klankbordgroep, externe beoordelaars of experts, de organisatie van expertsessies voor maatschappelijke, inhoudelijke en/of methodologische inbreng.

- Welke (min of meer) *tijdelijke* voorzieningen worden bij een wetsevaluatie ingericht – afgaande op uw ervaringen?
- Van welke overige informatie en (ervarings-)kennis maakt u gebruik bij het aanvragen van wetsevaluaties? Denk onder meer aan afdelingsoverleg / andere vergaderingen, overleg of informele gesprekken met collega's enz. --> evt. vragen wat de functie van die collega is of interne hulp-lijstjes met tips en aanwijzingen
 - En/of: bij wie klopt u aan met vragen?
- In hoeverre worden uw en andermans ervaringen met (grotere) wetsevaluatie-trajecten geborgd/ ergens vastgelegd, zodat toekomstige wetsevaluatietrajecten daar gebruik van kunnen maken? Graag toelichten waarom wel/ niet/ deels.

Iedereen: gebruik van evaluatiebevindingen

- Wat is uw inschatting van de mate waarin de resultaten van ex post wetsevaluaties worden gebruikt ten behoeve van de (verdere) wetgevings- of beleidsontwikkeling?
 - Gaarne onderscheid maken tussen de ontwikkelings- en besluitvormingsfase

- Ook graag onderscheid tussen kennisname - standpuntbeïnvloeding - uiteindelijk gebruik
- Wat zijn volgens u (mogelijke) oorzaken van deze mate van gebruik?
- In hoeverre zijn er naar uw weten structurele dan wel incidentele arrangementen die maken dat gebruik plaatsvindt?
 - Hangen die samen met de evaluatieprogrammering ... met andere genoemde arrangementen?
- Indien van toepassing, ziet u verbetermogelijkheden in de relatie tussen evaluatieresultaten en wetgeving – en zo ja, welke?
- Wat is uw inschatting van andere vormen van gebruik van wetsevaluaties (indirect of conceptueel), bijvoorbeeld voor de langere termijn, zoals kennisvergroting, of voor de doorontwikkeling van evaluatiemethoden?

Dimensie 3: Omgevingsfactoren

Welke factoren acht u – gelet op uw ervaring – van invloed op:

- Inhoudelijke keuzes en behandelde topics
 - Juridisch/ niet juridisch
- De aanpak van wetsevaluaties
- De kwaliteit van aanpak en resultaten
- Het gebruik van (de resultaten van) wetsevaluaties
- Hoe verhouden die factoren zich volgens u tot eerder besproken arrangementen: maken ze die extra nodig / overbodig, of doorkruisen / bevorderen ze het nut ervan?
- Welke manieren ziet u om de invloed van (volgens u) positieve factoren te versterken?
- Welke manieren ziet u om tegenwicht te bieden aan factoren die u een negatieve invloed toedicht op inhoud/ kwaliteit / gebruik?
- Politieke betrokkenheid aan evaluatie en wetsevaluatie in het bijzonder.
 - Binnen uw ministerie / uw directie
 - Beschrijvend (lieft door de jaren heen).
 - Ziet u verschil tussen politieke betrokkenheid aan wetsevaluatie?

- Indien u het politieke of betrokkenheid aan (wets)evaluatie beperkt acht, ziet u mogelijkheden voor het vergroten ervan? Zo ja, welke?

Vooruitkijkend (iedereen)

- Mist u bepaalde structurele (of tijdelijke) arrangementen bij het aanvragen en uitbesteden (of anderszins managen) van wetsevaluaties?
 - Zo ja, waarom? Wat gaat er in de huidige praktijk volgens u niet optimaal?
 - Welke arrangementen zou u installeren als u het voor het zeggen had? Wilt u dit toelichten, waarom zou dit een geschikte oplossing voor door u geschetste problematiek?
 - Zo nee, waarom niet?
- Ziet u voordelen in een interdepartementale aanpak ter bevordering van kwaliteit / inhoud / gebruik van wetsevaluaties?
 - Waarom wel/niet?
 - Zo ja, hoe zou zo'n aanpak eruit kunnen zien?

Bijlage B: Lijst van interviews

Ministerie	Aantal respondenten	Functie van respondenten	Datum
BZ	1	1 medewerker IOB	2 september 2015
EZ	5	1 Jurist Directie Wetgeving 1 lid Regiegroep 1 lid BIT 2 leden BEC	9 juli 2015
Financiën	2	2 medewerkers IRF	26 augustus 2015
IenM	3	1 medewerker HBJZ 1 medewerker FMC 1 Directiehoofd	16 juli 2015
OCenW	4	3 medewerkers Directie Kennis 1 Onderzoekscoördinator	1 juli 2015 26 augustus 2015
SZW	2	1 medewerker FEZ; 1 Onderzoekscoördinator.	15 juli 2015
VWS	1	1 lid Commissie Evaluatie Regelgeving	9 april 2015
VenJ	6	4 Projectleiders / aanvragers van evaluaties EWB 1 Onderzoekscoördinator van een Directie 1 Jurist Directie WJZ	8 juni 2015 15 juni 2015 11 augustus 2015