

Opgavegericht evalueren

Beleidsvaluatie voor systeemverandering

| Martijn van der Steen

| Albert Faber

| Andrea Frankowski

| Frank Norbruis

NSOB . .

Nederlandse School voor Openbaar Bestuur

Prof. dr. Martijn van der Steen is co-decaan en adjunct-directeur van de Nederlandse School voor Openbaar Bestuur en hoogleraar Strategie en Toekomst aan de Erasmus Universiteit Rotterdam.

Albert Faber is strateeg bij het ministerie van Economische Zaken en Klimaat (EZK).

Andrea Frankowski is onderzoeker en opleidingsmanager bij de Nederlandse School voor Openbaar Bestuur. Tevens is zij als promovenda verbonden aan Tilburg University.

Frank Norbruis is onderzoeker en opleidingsmanager bij de Nederlandse School voor Openbaar Bestuur.

Opgavegericht evalueren

Beleidsevaluatie voor systeemverandering

Martijn van der Steen

Albert Faber

Andrea Frankowski

Frank Norbruis

2018

ISBN: 978-90-75297-85-0

Inhoudsopgave

	Samenvatting	4
1	Waarom een essay over beleidsevaluatie?	7
	<i>Doelstelling, afbakening en opzet</i>	8
2	Inherente dilemma's bij het evalueren van beleid	9
2.1	Beleidsvaluatie als essentiële functie van beleid en sturing	9
2.2	Inherente dilemma's bij de 'gouden standaard' van beleidsvaluatie	10
	<i>I Principiële vragen bij het werken met een experimenteel design</i>	11
	<i>II Praktische kwesties bij het werken volgens de 'gouden standaard'</i>	12
	<i>III Een veranderend perspectief op beleid en sturing</i>	12
	<i>IV Functiecombinatie werkt door in het evaluatietraject</i>	13
2.3	Omgaan met inherente kwesties rondom beleidsevaluatie: 'trap af' of 'trap over'	16
3	Specifieke dilemma's: evalueren van verweven dynamiek	18
3.1	Sturing in toenemende complexiteit	19
	<i>Vanzelfsprekende systemen</i>	20
	<i>Ingewikkelde systemen</i>	20
	<i>Complexe systemen</i>	20
	<i>Wanordelijke systemen</i>	21
	<i>Uitdagingen voor evaluatie: attributie, contributie en nieuwe betekenis voor klassieke termen</i>	22
3.2	Netwerkende strategieën voor het realiseren van maatschappelijke meerwaarde	23
	<i>Vier perspectieven voor overheidssturing</i>	23
	<i>Publieke waarde realiseren in passende combinaties van perspectieven</i>	25
3.3	Beleidsrealisatie in transitiedomeinen: sturing voor het onbekende	28
3.4	Opgavegericht evalueren: evaluatie die past bij de opgave	31

4	Omgaan met dilemma's: ontwikkelingen in de evaluatiepraktijk	33
4.1	Meervoudig evalueren in de praktijk	33
4.2	Ontwikkelijnen in de evaluatiepraktijk	37
	<i>Functie</i>	37
	<i>Timing</i>	38
	<i>Focus</i>	40
	<i>Rol van de onderzoeker</i>	44
	<i>Motivatie</i>	48
5	Inbedding: opgavegericht evalueren vereist <i>bewust organiseren</i>	51
6	Conclusie: opgavegericht evalueren vereist <i>normalisering</i>	58
	Eindnoten	60

Samenvatting

Waarom dit essay? Evaluatie is een essentieel onderdeel van de beleids-cyclus; het stelt beleidsmakers in staat om de toegevoegde waarde van beleid te bepalen. Daarmee vormt het een cruciale schakel in de democratische verantwoording van beleid, maakt het leren mogelijk en levert het signalen op voor eventuele bijstelling. Om de toegevoegde waarde van beleid vast te stellen zijn systematische methoden beschikbaar waarmee uitspraken kunnen worden gedaan over de doeltreffendheid, de doelmatigheid en rechtmatigheid van beleid. In de praktijk is er een ‘gouden standaard’ ontstaan, die geldt als maat voor goed evaluatie-onderzoek.

Regelmatig komen er echter vraagstukken voorbij die vanwege hun structuur, kenmerken of context slecht passen bij de ontwikkelde klas-sieke methoden van beleidsevaluatie. Het gaat dan om vraagstukken met een hoge mate van verwevenheid en complexiteit, om sturing in en met netwerken of om transitieopgaven gericht op zich op onvoorzien-bare wijze ontwikkelende systeemveranderingen op de lange termijn. Dergelijke vraagstukken worden gekenmerkt door een hoge mate van onzekerheid, onbekendheid en brede maatschappelijke betrokkenheid en lenen zich niet goed voor toepassing van de bij de ‘gouden standaard’ behorende methoden. Toepassing van die methoden leidt namelijk onvermijdelijk altijd tot één van volgende oordelen: ‘het beleid is niet succesvol’, of ‘de resultaten van beleid zijn niet vast te stellen’.

In beide gevallen zegt het oordeel weinig over de werkelijke staat van het beleid, maar des te meer over de mismatch tussen de toegepaste methode van evaluatieonderzoek en de aard van de aanpak van de voorhanden zijnde opgave. Omdat steeds meer opgaven spelen in een complexe context van verweven dynamiek en netwerken en bovendien betrekking hebben op deels nog onvoorzienbare transitie met doelstel-lingen die soms onderweg nog vorm en inhoud moeten krijgen, is het belangrijk om te evalueren op een manier die past bij de aard van de aanpak van de opgave: *opgavegericht evalueren*.

Waarover gaat dit essay? In dit essay verkennen we eerst de dilemma's die bij beleidsevaluatie optreden. We benoemen *inherente dilemma's* die als principiële en praktische vragen opkomen bij klassiek evaluatieonderzoek. Vervolgens gaan we in op enkele *specifieke dilemma's*, die specifiek spelen bij het evalueren van complex beleid met een verweven dynamiek waarvoor sturing en inzet langs netwerken plaatsvindt, waaronder transitieopgaven. We signaleren vervolgens een aantal *ontwikkellijnen* in de praktische omgang met deze dilemma's, aan de hand van enkele voorbeelden die in beeld brengen hoe beleidsevaluatie in een context van complexiteit plaatsvindt. Op basis hiervan formuleren we lessen voor de ontwikkeling van een 'nieuwe standaard' en doen we aanbevelingen voor de manier waarop organisaties deze methoden ook daadwerkelijk kunnen inzetten.

Wat zijn opbrengsten van dit essay? Bij complexe vraagstukken is beleid geen enkelvoudige causale interventie met een afgebakend effect gericht op een aanwijsbare actor. Veel vaker is er sprake van sturingsarrangementen van verschillende actoren, die zijn ingericht om condities voor verandering te scheppen, zodat de kans op een bedoelde beweging wordt vergroot. Dergelijke vraagstukken vragen om bijpassende evaluatiemethoden. Er zijn talrijke methoden beschikbaar voor het evalueren van beleid met een verweven dynamiek. Dit essay verkent een aantal methoden uit de Haagse beleidspraktijk.

In de beleidspraktijk bespeuren we ontwikkellijnen in de toepassing van evaluatieonderzoek die betrekking hebben op de functie, de timing en de focus van evaluatieonderzoek, de rol van de evaluator en de motivatie achter de evaluatie. Zo zijn leren en bijsturen belangrijker geworden, naast verantwoorden (*functie*); wordt er naast *ex post* ook steeds meer *ex ante* en *ex durante* geëvalueerd (*timing*); wordt er behalve aan eindopbrengsten ook waarde toegekend aan tussenopbrengsten en is er behalve voor attributie ook aandacht voor contributie (*focus*); wordt van evaluatoren niet alleen onafhankelijkheid, maar ook betrokkenheid verwacht (*rol*); en is evalueren behalve vanuit politiek en bestuur, ook ingestoken vanuit een ambtelijke logica (*motivatie*).

Het is belangrijk om te onderkennen dat een tekort aan methoden niet het primaire probleem van evalueren in een complexe context van verweven dynamiek is; eerder dat die in de reguliere processen van organisaties (nog) niet tot de geaccepteerde standaard behoren. Ze gelden als incomplete, zwakkere, of zachtere vormen van beleidsevaluatie die als *second best* niet voldoen aan de geldende échte standaard. Daarom is de werkelijke opgave niet het vinden van nieuwe methoden, maar het bijstellen van sommige assumpties over wat goed evalueren is en hoe we de effecten van beleid vaststellen. Dat gaat om onderkenning en normalisering van andere vormen van evaluatie, om het ontwikkelen en accepteren van een ‘andere standaard’ toegepast op complexe, netwerkende aanpakken, kenmerkend voor transitieopgaven, naast de al bestaande ‘gouden standaard’ voor opgaven die op een meer lineaire wijze verlopen. Evalueren is daarmee een opgavegerichte activiteit: de manier van evalueren moet passen bij de gekozen aanpak van de opgave. Opgavegericht evalueren moet uitdrukkelijk niet gezien worden als alternatief voor evaluatiemethoden die breed worden toegepast en nu gelden als ‘gouden standaard’, maar als een aanvulling hierop, geschikt voor specifieke vraagstukken.

Welke conclusies kunnen we aan deze opbrengsten verbinden? Er zijn allerlei vormen van evalueren voorhanden, die én tegemoet komen aan dilemma’s van evalueren, én kwalitatief hoogwaardig zijn. Dit essay is een pleidooi voor *opgavegericht evalueren*, waar de aanpak van de maatschappelijke opgave centraal staat, met rekenschap van de complexiteit en bijkomende onzekerheden van de beleidspraktijk. Hiertoe past een breed evaluatieperspectief, afgestemd op de concrete aanpak van de maatschappelijke opgave. *Opgavegericht evalueren* vereist normalisering. Dat vraagt om een institutionele structuur, een manier van denken over regels (informeel en formeel), taal en routines, die open staat voor een veelheid aan evaluatiemethoden samengebracht in een *toolbox* van keuzeopties die passen bij de diversiteit aan beleidsarrangementen. Dit essay opent de gerichte zoektocht naar *opgavegerichte evaluatie* en de arrangementen van normalisering die dat vereist.

1

Waarom een essay over beleidsevaluatie?

Wat is het effect van een programma gericht op de versnelling van de energietransitie? Wat is de toegevoegde waarde van subsidies voor het stimuleren van elektrisch rijden? Welke meerwaarde heeft het bouwen van netwerken om de bio-based economy aan te jagen? Hoe verhoudt het provinciale beleid zich tot de nationale ambities van het natuurbeleid? Het zijn voor de hand liggende beleidsvragen, die niettemin tot ongemakkelijke gevoelens leiden. Elke beleidsmaker zou er meteen antwoord op moeten weten. Het is niet meer dan logisch dat beleidsmakers hebben nagedacht over indicatoren voor en instrumenten waarmee de toegevoegde waarde van beleid nu of straks is vast te stellen middels een goede evaluatie of verantwoordingsmethodiek. Niet alleen omdat dat een cruciale schakel is in de democratische verantwoording van beleid (in termen van doelmatigheid en rechtmatigheid), maar ook om van ervaringen te kunnen leren en beleid zo nodig bij te stellen.

Evaluatie is een essentieel onderdeel van de beleidscyclus, daar is iedereen het over eens. En toch is het zo helder nog niet. Beleidsmakers weten regelmatig niet of beleid effectief is, of de bedoelde toegevoegde waarde er is en of die waarde te herleiden is naar de gepleegde inspanningen. Soms hangt dat samen met onwetendheid, onkunde of gebrek aan gevoelde urgentie. Soms is het ook onduidelijk hoe verantwoordelijkheden zijn verdeeld, wie waarop kan worden aangesproken, of wat de beoogde resultaten zijn.¹ Maar er zijn ook vraagstukken die zich vanwege hun specifieke structuur lastig lenen voor een 'klassieke' beleidsevaluatie. Het gaat dan om probleemvelden die zich kenmerken door een 'verweven dynamiek'², waarvoor sturing het karakter heeft van het werken in netwerken, langs indirecte weg, met doelstellingen die zich bewust en expliciet onderweg verder zullen moeten vormen.

De evaluatievraag komt steeds nadrukkelijker onder de aandacht te staan, nu vraagstukken op het beleidsterrein van EZK – en andere departementen – steeds meer het karakter hebben van sturen in verweven dynamiek, het sturen in netwerken en het sturen op zich ontwikkelende transitiedomeinen.

Traditionele evaluatiemethoden passen echter niet goed bij de turbulentie van de complexe vraagstukken waar beleidsmakers zich tegenwoordig mee geconfronteerd zien. Dergelijke complexiteit staat vaak in de weg van een adequate evaluatie, maar wellicht kan het ook fungeren als uitgangspunt. Daarom is de vraag die we in dit essay zullen behandelen als volgt: *hoe kunnen we sturing en beleid evalueren op een manier die recht doet aan de complexiteit van vraagstukken van verweven dynamiek en transitie?*

Doelstelling, afbakening en opzet

Dit onderzoek is een samenwerkingsproject van het ministerie van Economische Zaken en Klimaat (EZK) en de NSO, in het kader van het *Interdepartementale Kennisprogramma Intelligent Bestuur*. Het onderzoek is gebaseerd op een theoretische verkenning van bestuurskundige theorieën over beleidsevaluatie, een documentenstudie, interviews en een focusgroepsessie met een tiental experts uit de evaluatiepraktijk. De documentenstudie en individuele gesprekken hebben betrekking op drie specifieke voorbeelden van beleidsevaluatie: de evaluatie van de Topsectorenaanpak (door Dialogic), de Lerende Evaluatie van het Natuurpact (uitgevoerd door het Planbureau voor de Leefomgeving in samenwerking met Wageningen University en Research) en de evaluatiepraktijk van de Directie Internationaal Onderzoek Buitenlandbeleid (IOB, ministerie van BZ). Voor elk van de drie praktijkvoorbeelden hebben we enkele interviews gevoerd met betrokkenen. In de focusgroepsessie is gesproken met beleidsmakers van verschillende departementen (EZK, I&W, LNV, BZ en Financiën) en experts van evaluerende instituties (PBL, IOB, Algemene Rekenkamer en RVO). Daarnaast heeft er een uitgebreide analyse plaatsgevonden van methoden en technieken die gebruikt worden voor evaluatieonderzoek gericht op complexe beleidscontexten. De rapportage van deze analyse is niet integraal opgenomen in dit essay, maar lessen die eruit voortvloeien zijn wel meegenomen.

Doel van het onderzoek is het verkennen van ‘complexiteitserkennende’ manieren van evaluatie van beleid en sturing. Het expliciteren van complexiteit als systeemperspectief vereist dat het vanuit datzelfde systeemperspectief geëvalueerd kan worden, zodat het een volwaardige manier van werken en kijken in de beleidspraktijk kan worden. In de volgende hoofdstukken schetsen we achtereenvolgens inherente (hoofdstuk 2) en specifieke (hoofdstuk 3) dilemma’s van beleidsevaluatie, welke ontwikkelingen we zien in de omgang met deze dilemma’s in de evaluatiepraktijk (hoofdstuk 4) en wat dat betekent bezien vanuit bestuurskundig oogpunt (hoofdstuk 5). We sluiten af met wat er vereist is om voorbij deze dilemma’s te komen (hoofdstuk 6).

2

Inherente dilemma's bij het evalueren van beleid

2.1 *Beleidsevaluatie als essentiële functie van beleid en sturing*

Beleidsevaluatie heeft een vaste plek in de beleidsketen. Evaluatie maakt verantwoording mogelijk en zorgt ervoor dat we kunnen zien of en waarom beleid werkt: worden de bedoelde resultaten behaald? Wat zijn de werkzame bestanddelen? Dat is belangrijk ten behoeve van publieke en politieke *verantwoording*, als aanzet voor mogelijke *bijsturing* en omwille van het *beter leren begrijpen* hoe systemen werken en hoe interventies daarop ingrijpen. Evaluatie is een essentieel onderdeel van beleid, net zoals ontwerp en uitvoering dat zijn.

Een beleidsevaluatie heeft een aantal functies. Het maakt uitspraken mogelijk over de *doeltreffendheid* en de *doelmatigheid* van beleid: zijn er effecten opgetreden en is de verhouding tussen de inzet van middelen en het bereiken van doelen in orde? In het verlengde hiervan kan ook iets worden gezegd over de *rechtmatigheid* van beleid: passen de uitgaven binnen de geldende regels en kaders? Daarnaast maakt evaluatie het mogelijk om *lessen te leren* van het beleid. De implicatie daarvan is dat beleidsevaluatie vaak ook een aanleiding voor *heroverweging en bijstelling* van beleid is. Evaluatie zorgt voor een logisch moment in de beleidsketen om expliciete aandacht te hebben voor mogelijke heroverweging. Een evaluatie is niet primair *bedoeld* om beleid bij te sturen en eventuele implicaties spelen idealiter géén rol bij het uitvoeren van een evaluatie of bijvoorbeeld bij het schrijven van een evaluatierapport. Tegelijkertijd is een inherente implicatie van beleidsevaluatie wel dat het aanleiding tot bijstelling kan geven. Beleidsevaluatie is een analytische activiteit, maar wel één die speelt binnen een beleidscontext die ook met politiek en strijd omgeven is.

In een beleidsevaluatie onderzoekt de evaluator of de empirische praktijk zich heeft ontwikkeld volgens de verwachtingen die vooraf aan het beleid ten grondslag lagen: zijn de verwachte effecten opgetreden en waren deze er zonder het beleid niet (geheel of gedeeltelijk) geweest? In een beleidsevaluatie wordt gekeken ook naar additionele effecten. Dat kan via de weg van het vaststellen van *attributie*: zijn gevonden effecten een direct gevolg

van een aanwijsbare concrete interventie, of zijn effecten mogelijk het gevolg van andere factoren? Een andere benadering gaat uit van *contributie*: heeft het geheel van beleid bijgedragen aan de richting waarin het veld zich ontwikkelt en de effecten die daarbinnen ontstaan? Het principe van contributie legt minder de nadruk op het isoleren van één specifieke actor en/of maatregel die een eveneens te isoleren effect tot gevolg heeft. Als de interventie geloofwaardig heeft bijgedragen aan de *richting* is dat vanuit dit perspectief een relevant resultaat.

Om additionaliteit en attributie of contributie vast te stellen is het werken met een *controlegroep* binnen evaluatie de meest sluitende werkvorm; er is een vergelijkbare groep die de bewuste ‘therapie’ niet krijgt, zodat gevolgen bij de onderzoeksgroep ook daadwerkelijk tot de interventie te herleiden zijn. Idealiter is beleid dan ook te isoleren tot op het niveau van een interventie gericht op een *single-agent*. Hierbij is er één interventie, voor één gebruiker (of één eenduidige gebruikersgroep), die één helder omschreven doel nastreeft, waarbij te controleren is voor andere mogelijke interfererende variabelen. Als het mogelijk is om beleid in een dergelijke relatie te vatten dan kan de evaluatie ervan met wetenschappelijke nauwkeurigheid plaatsvinden.

Op deze manier benadert een beleidsevaluatie de standaard die we kennen uit de *clinical trials* in de medische wetenschap, die ook ten grondslag ligt aan het denken in termen van *evidence-based policy*. Dit wordt ook wel de ‘gouden standaard’ van beleidsevaluatie genoemd, ontwikkeld door een commissie onder leiding van prof. Jules Theeuwes³. Door toepassing van de gouden standaard, een ideale manier van beleidsevaluatie, ontstaat een cyclus die finale kennis genereert over de vraag of beleid wel of niet werkt en onder welke condities. Deze kennis kan vervolgens gebruikt worden in een proces van verantwoording, bijstelling of leren.

2.2 Inherente dilemma's bij de ‘gouden standaard’ van beleidsevaluatie

Een dergelijke manier van beleidsevaluatie stelt specifieke eisen aan de inrichting van beleid en het evaluatieproces. De inzet ervan roept bovendien een reeks inherente dilemma's op. Met ‘inherent’ bedoelen we dat dit dilemma's zijn die niet voortkomen uit het ‘niet goed’ toepassen van de methodiek, maar er per definitie bij horen. Het zijn inherente gevolgen van het evalueren volgens de gouden standaard methodiek.

1. *Principiële vragen bij het werken met een experimenteel design*

Het werken met een controlegroep betekent dat bij het ontwerp van het beleid en in de vroegste fasen van de uitvoering al rekening wordt gehouden met de latere evaluatie. Dat betekent dat het beleid zelf als experiment moet worden ingericht. Dat kan lastig zijn, bijvoorbeeld omdat een controlegroep in de praktijk lang niet altijd beschikbaar is.

Er zijn er ook meer principiële kwesties rond een experimenteel design van beleid; zo moet het praktisch én normatief haalbaar zijn om de controlegroep bewust *niet* 'te behandelen' met het beoogde beleid. Dat is een problematische uitgangspositie voor veel beleid. Stel dat er een beleidsaanpak is voor het voorkomen van problematische schulden onder jongeren; is het dan houdbaar om bewust een controlegroep die interventie *niet* aan te bieden en de groep daarmee – als de beleidstheorie juist is – bewust hulp te onthouden en met vermijdbare problemen op te zadelen? Het werken met een controlegroep betekent dat een groep bewust aan risico wordt blootgesteld, of bewust een verbetering wordt onthouden. Zo ontstaat 'bedoelde ongelijkheid', gericht op kennis over wat werkt en wat niet werkt in beleid. Dicht tegen dit dilemma aan ligt de vraag of het voor de groep die de interventie wel krijgt gepast is om te worden blootgesteld aan een interventie waarvan de werkzaamheid (nog) niet is vastgesteld. Is het gepast om het beleid op hen 'uit te proberen'? Niet voor niets zijn er in de medische wereld ethische commissies die experimenten moeten goedkeuren. In de context van beleid zijn dergelijke commissies er niet en zijn er ook geen gedeelde standaarden of principes die voorschrijven wat er wel of niet kan en mag.

Naast deze normatief-ethische kwesties zijn er ook kwesties rondom *transparantie* van het werken in een experimenteel design; mogen deelnemers weten van welke groep zij deel uitmaken? Waarschijnlijk niet, want dat verstoort het experiment. Dat betekent dat subjecten van beleid informatie wordt ontzegd over beleid dat op hen wordt uitgevoerd, terwijl het beginsel van een transparante en controleerbare overheid voorschrijft dat burgers weten wat de overheid voor of met hen doet en dat er geen 'geheime' of 'verborgen' interventies plaatsvinden. Ook hier liggen principiële belemmeringen: is het wenselijk dat mensen *niet weten* aan welk beleid ze onderhevig zijn en welke beleidstheorie er door de overheid op hen wordt toegepast? In de wereld van de inlichtingen- en veiligheidsdiensten gebeurt dit wel, maar alleen in combinatie met een uitgewerkt stelsel van parlementaire en procedurele controle en borgingsmechanismen. Voor regulier beleid ontbreken dergelijke mechanismen en principes vooralsnog.

II. Praktische kwesties bij het werken volgens de ‘gouden standaard’

Naast deze relatief principiële kwesties rond het werken volgens de gouden standaard is er vanuit de concrete beleidspraktijk nog een aantal meer praktische kwesties. De gouden standaard van beleidsevaluatie suggereert een beleidsproces dat min of meer verloopt volgens de opzet van een evaluatiedesign. Dat wil zeggen: met vooraf gestelde doelen, onder vooraf vastgestelde condities, waarna partijen gedurende een gecontroleerde periode aan het werk gaan met de uitvoering, en er dan gereserveerde momenten zijn waarop meting van effecten plaatsvindt. Die volgordelijkheid is typerend voor het design van de evaluatie, maar past vaak niet bij het daadwerkelijke verloop van beleidsprocessen. Generaties onderzoek naar beleidsprocessen hebben dan ook het ‘fasen-model’ en het model van rationeel beleidsontwerp⁴ gerelativeerd. Daarvoor in de plaats zijn alternatieve vormen en rationaliteiten voor beleid gepresenteerd die een meer grillig en verstrengd beeld van beleidsprocessen schetsen. Voor dat meer grillige perspectief op beleidsprocessen is de gouden standaard minder goed toegerust.

De grilligheid van beleidsprocessen staat haaks op de vermeende stabiliteit én volgtijdelijkheid in het beleidsproces. Grilligheid betekent dat doelen, uitvoering en effecten door elkaar heen lopen. Doelstellingen en kaders van beleid veranderen gaandeweg het proces en onderweg kan er worden bijgestuurd, waarbij hele delen van beleid zelfs kunnen verdwijnen. Volgens de gouden standaard verstoort dit echter de meting. Er is weinig ruimte voor de dynamiek die vanuit onderzoek naar beleidsprocessen als gemeengoed wordt gezien. Vanuit het raamwerk van de methode is dan niet veel anders mogelijk dan een *gemankeerde* of *verwaterde evaluatie*, terwijl er toch sprake kan zijn van een adequaat, ‘gewoon’ of zelfs excellent beleidsproces.

III. Een veranderend perspectief op beleid en sturing

Een derde probleem voor de gouden standaard benadering van beleidsevaluatie vormt de stroming die in de bestuurskunde vaak geduid wordt als een beweging van *government* naar *governance*. Of, in de *policy sciences* literatuur, van *policy design* naar *governance design*. Daarin heeft beleid minder het karakter van een concrete maatregel die als een chirurgische ingreep een genezend effect moet hebben, maar meer het karakter van het scheppen van condities die de kans op genezing doen toenemen. Waar het denken in termen van beleidsevaluatie ‘beleid’ (*policy*) centraal stelt, zien we in de praktijk van de overheid steeds vaker het werken in termen van ‘arrangementen’ (*governance*): geen beleid maken, maar sturing vormgeven.

Bij governance komen opbrengsten tot stand doordat relevante partijen tot een gedeelde beweging komen, die tot de bedoelde publieke waarde(n) moet leiden. Dat betekent dat er minder één heldere ‘therapie’ te onderscheiden is, toegediend door één specifieke betrokkene aan één af te bakenen ontvanger. In plaats daarvan is er sprake van een mix van interventies, van omstandigheden die worden beïnvloed, uitgevoerd in een constellatie van actoren, die samen moeten leiden tot de bedoelde beweging bij een doelgroep.

Een tweede problematiserend element van een governance-benadering voor doelrealisatie is dat het een ruimer perspectief op doelstelling en normstelling hanteert; waar beleid idealiter een beleidstheorie kent waarin doelen, kaders, randvoorwaarden en normen geëxpliciteerd zijn, is bij governance veel minder sprake van een expliciet moment van centrale doelstelling. Dat kan wel, maar dat is eerder uitzondering dan regel. Veel vaker is er in een governance-arrangement expliciet ruimte ingebouwd om onderweg de doelen te ontdekken, bij te stellen, gedeelde normen te ontwikkelen en condities nader in beeld te krijgen. Ze staan niet vast, maar zijn onderwerp van interactie binnen het governance-arrangement. Bij veel governance-arrangementen is het de expliciete bedoeling dat gaandeweg, in interactie tussen betrokkenen, meer zicht en helderheid ontstaat over doelen, kaders, randvoorwaarden, inspanningen en te bereiken resultaten. Termen als adaptiviteit van beleid en sturing proberen dit verschil te vatten. Adaptief beleid betekent: ontwerpen om te veranderen, niet als ‘correctie’ op de ingezette koers, maar als bedoeling van het ontwerp. De elementen die in klassieke beleidsevaluatie gefixeerd zijn, zijn hierbij bedoeld beweeglijk, als sleutel tot doelrealisatie en het oplossen van opgaven.

iv. Functiecombinatie werkt door in het evaluatietraject

Een laatste probleem voor de meer traditionele benadering van beleidsevaluatie schuilt in de interactie van de verschillende functies en implicaties van beleidsevaluatie. Beleidsevaluatie staat niet buiten het beleidsproces, maar maakt er deel van uit. Beleidsevaluatie heeft vaak implicaties. Dat maakt dat beleidsevaluatie meer is dan alleen een objectieve exercitie waarin kennis wordt verzameld, maar ook op enig moment een rol gaat spelen in grillige politieke processen over beleidswijziging. Dat werkt door in het evaluatieproces zelf en is daarmee ook tot op zekere hoogte van invloed op de mogelijkheid van de evaluator om effecten vast te stellen.

Met name verantwoording en bijsturing kunnen majeure consequenties hebben voor betrokkenen, in termen van positie, reputatie, macht, middelen en ruimte. Betrokkenen weten dat verantwoording gevolgen heeft – dat is ook de bedoeling ervan. Dat leidt tot strategisch gedrag ten aanzien van de evaluatie.⁵ Betrokkenen zijn dan bijvoorbeeld terughoudend bij het aanleveren van informatie of ze zijn strategisch in het kiezen van het ijkpunt voor de nulmeting. Zo kan een ‘slecht jaar’ worden gekozen als ijkpunt, zodat een relatief normaal navolgend jaar al snel een verbetering laat zien. Soms wordt gekozen voor indicatoren waarvan duidelijk is dat ze gemakkelijk prestaties kunnen laten zien. Een voorbeeld zien we in de ouderenzorg, waar vanuit het programma “Waardigheid en trots” niet de ‘warmte’ of de ‘waardigheid’ van de zorgverlening wordt gemeten, maar waar wordt gekeken naar het aantal en de duur van ‘contactmomenten’. Dat is informatie die beschikbaar is en voor partijen relatief eenvoudig positief is te beïnvloeden. Daarmee zijn inspanningen eenvoudig te vertalen in een positief evaluatie-oordeel; in dit geval door meer contactmomenten te realiseren, een inspanning die zich direct terugvertaalt in een positief oordeel over toegenomen ‘waardigheid en trots’, hoewel het maar de vraag is of die elementen hetzelfde zijn. Dat beïnvloedt de evaluatie; de bedoelde te meten *outcomes* worden strategisch gereduceerd naar *outputs* of zelfs *inputs*, met alle gevolgen van dien voor strategisch handelen van betrokkenen in het beleidsproces.

Een variatie op ditzelfde thema is de vaak indringend gevoelde spanning tussen verantwoorden en leren.⁶ Beleidsevaluatie kan voor leren en verantwoorden worden gebruikt, maar die beide functies kunnen elkaar ook in de weg zitten. Partijen die het ene moment eerlijk zijn en laten zien dat er in het proces veel mis gaat, en die daarmee belangrijke lessen onder de aandacht brengen, kunnen het volgende moment vanuit de andere doelstelling van de evaluatie ter verantwoording worden geroepen. Dat zorgt soms voor ongemak bij betrokkenen in het beleidsproces. Niemand ontkent het nut van leren van fouten, mislukkingen en *near misses*, maar iedereen herkent ook de reflex dat ‘eerlijkheid’ over fouten ook in een goede evaluatie tot strafing in de verantwoording kan leiden. Betrokkenen stemmen hun strategieën bij deelname aan beleidsevaluaties af op een zo gunstig mogelijk verantwoordingsresultaat, en pas in tweede instantie op het bereiken van een maximaal leereffect.

Samengevat is er dus sprake van meervoudige inherente dilemma’s die zich rondom beleidsevaluatie manifesteren. [Tabel 1](#) biedt een overzicht.

Inherent dilemma van beleidsevaluatie	Assumptie vanuit 'gouden standaard'	Kwestie vanuit ervaring in beleidsprocessen
1. <i>Principieel: ultiem design of bescherming van subjecten</i>	De meting perfectioneren door controlegroep te ontzeggen van beleid, waardoor effect optimaal geïsoleerd kan worden.	Ontzeggen van beleid voor een groep roept ethische bezwaren (hoort het?) op, heeft juridische consequenties (rechtsgelijkheid, willekeur) en zorgt voor politieke problemen (transparant zijn of niet?).
2. <i>Praktisch: gefaseerd volgordelijk of grillig en verstrengd beleidsproces?</i>	Beleid komt in een gefaseerd proces tot stand, waarin doelstelling, ontwerp, uitvoering en evaluatie elkaar opvolgen. Logica van 'vroeg verbinding', bij beleidsontwerp (en doelstelling) rekening houden met ontwerp voor evaluatie.	Beleid komt in grillige processen tot stand, waarin genoemde 'fasen' door elkaar heen lopen, en processen 'fasen' overslaan, terugvallen, of combineren.
3. <i>Enkelvoudige sturing of meervoudige 'governance'?</i>	Beleidsmaatregelen als focus van de evaluatie, een interventie gericht op een aanwijsbare actor, met een afgebakend effect. Beleid als gerichte ingreep die direct effect sorteert. Een incentive, prikkel, maatregel, met een doelgroep, voorwaarden en concrete verwachte effecten.	Sturingsarrangementen scheppen condities bedoeld om de kans op een bepaalde beweging door een constellatie van actoren te vergroten. Hierin is ruimte voor ontwikkeling van doelen, middelen en omstandigheden.
4. <i>Gelijkwaardige functies en implicaties van evaluatie, of hiërarchie en onderlinge beïnvloeding tussen functies?</i>	Functies staan naast elkaar, de evaluatie zelf staat los van het specifieke gebruik. Het latere gebruik is niet van invloed op de evaluatie zelf en speelt daarom ook geen rol in de opzet en uitvoering.	Functies beïnvloeden elkaar sterk. Betrokkenen anticiperen op het gebruik dat ze verwachten en handelen strategisch.

Tabel 1 Inherente dilemma's van beleidsevaluatie.

2.3 Omgaan met inherente kwesties rond beleidsevaluatie: ‘trap af’ of ‘trap over’?

De hier beschreven kwesties zijn *inherente dilemma's*, die er rondom ‘de gouden standaard’ altijd zullen zijn en er ook altijd al waren. Wie volgens de gouden standaard wil werken krijgt altijd met deze vragen te maken. In de beleidspraktijk worden daarom vaak minder strikte werkwijzen toegepast: geen *evidence-based* gouden standaard, maar een meer statistische zilveren standaard. Deze wordt, bijvoorbeeld, gevonden in de toepassing van methodieken uit de econometrie, die kunnen helpen om iets te zeggen over statistisch bepaalde correlaties en causaliteit. Dat kan bijzonder zinvol zijn voor analyses over doelmatigheid van specifieke instrumenten of maatregelen, maar bij grotere meer transitiegerichte beleidsvraagstukken zijn zulke benaderingen lastiger. We lopen dan op tegen de complexiteit van hedendaagse vraagstukken, maar er is ook sprake van een verschuiving in de rol van de overheid naar nieuwe vormen van samenwerking.

Kwantitatieve standaarden van evaluatie verhouden zich niet vanzelfsprekend goed tot de realiteit van beleid en sturing. De complexe praktijk vereist dat ofwel het evaluatie-ontwerp ofwel de realiteit ‘gebogen’ wordt, om te passen binnen het sturingsmodel. De realiteit kan worden gebogen door een beleidsproces dat in de praktijk niet strak gefaseerd is voor de evaluatie wel zo voor te stellen. Ook kan (impliciet of expliciet) worden gekozen om de spanning tussen leren en verantwoorden niet mee te wegen in de evaluatie.

Aan de andere kant kan ook de evaluatie-opzet worden versoepeld of gerelativeerd. Dit wordt vaak voorgesteld als het ‘betalen van een prijs’ in de vorm van een vermeend laagwaardiger evaluatie. Het beeld van de evaluatieladder is hiervoor exemplarisch. Er is een ladder van ‘goed evalueren’, waarbij de evaluator stappen omlaag kan zetten om dichterbij de realiteit ‘buiten’. Het frame van ‘trap op, trap af’ suggereert een hiërarchie in de kwaliteit van evaluaties: er zijn goede evaluaties, volgens de gouden standaard, en daarna is het mogelijk om langzaam af te dalen naar *andere*, ‘lagere’ vormen van evaluatie, naar gelang de context dat vraagt.

Tegemoetkomen aan de inherente dilemma's van de gouden standaard suggereert een ‘concessie’ aan de kwaliteit van evalueren. Dat beeld willen we in dit essay ter discussie stellen en vervangen door een ander beeld: er zijn andere vormen van evalueren, die op een andere manier omgaan met de dilemma's en tegelijkertijd op andere manieren kwalitatief hoogwaardig

zijn. Dat is ook de idee van programma's gericht op realistisch en bruikbaar evalueren: niet primair gericht op de perfectie van de evaluatiemethode, maar op de opgave die aan het beleid ten grondslag ligt. Wat was ook al weer de bedoeling? Dit is een pleidooi voor *opgavegericht evalueren*, waar die 'bedoeling' centraal staat, met rekenschap van de complexiteit en bijkomende onzekerheden van de beleidspraktijk. Hiervoor geldt geen *best practice*, maar inzicht in de gereedschapskist aan beschikbare evaluatiemethoden. Inzicht in specifieke dilemma's van evalueren in een verweven dynamiek is van belang om in deze gereedschapskist de weg te kunnen vinden. Daarover gaat het volgende hoofdstuk.

3

Specifieke dilemma's: evalueren van verweven dynamiek

In het vorige hoofdstuk introduceerden we een reeks inherente dilemma's van beleidsevaluatie: dilemma's die er bij de evaluatie van beleid volgens een gouden standaard aanpak *altijd* zijn, in elke mogelijke context. Daarnaast is er een tweede soort dilemma's: *specifieke dilemma's* die optreden bij de evaluatie van opgaven onder bepaalde specifieke omstandigheden. Dit zijn dilemma's die samenhangen met de doelen die met sturing worden nagestreefd en het soort systeem waarin die sturing speelt.

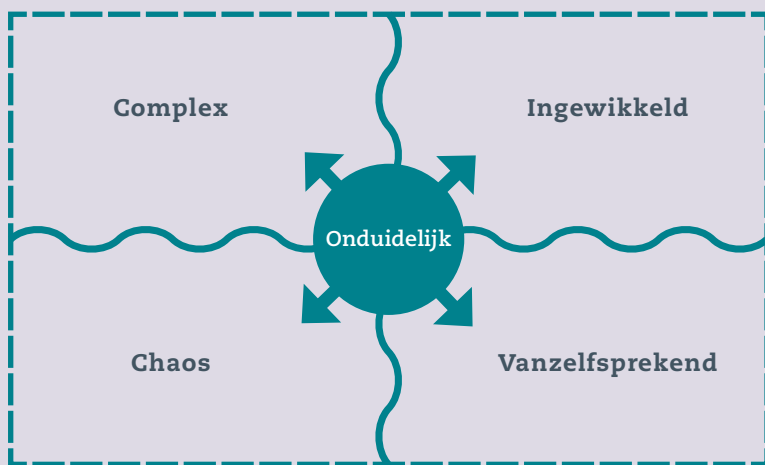
We onderscheiden hier drie soorten omstandigheden die voor specifieke dilemma's bij beleidsevaluatie zorgen. Ten eerste constateren we dat sturing steeds vaker speelt in systemen die *complex* zijn. Complexe systemen worden gekenmerkt door, onder meer, een grote *heterogeniteit aan actoren*, door een evolutionaire dynamiek, met plotselinge kantelpunten van snelle verandering, en door terugkoppelingen die leiden tot vooraf vaak onvoorzijne boemerangeffecten. Ten tweede vindt sturing steeds vaker plaats in een *netwerkend arrangement*. Hiermee bedoelen we dat sturing vaak gebeurt in een aanpak en arrangement dat specifiek is voor de opgave in kwestie. Vaak leidt dat tot netwerkaanpakken waarbij de overheid slechts één van de betrokken partijen is, hetgeen voor specifieke uitdagingen bij evaluatie zorgt. De derde omstandigheid is dat sturing steeds vaker betrekking heeft op domeinen waarin *opgaven* en *transitiedoelen* aan de orde zijn. Dergelijke opgaven zijn doorgaans breed omschreven; meer 'ambities' dan concrete 'beleidsdoelen'. Een transitiedoel is bijvoorbeeld 'klimaatneutraal in 2040', een opgave met vele mogelijke paden er naar toe. Dat geeft vele vrijheidsgraden aan de sturingsmogelijkheden, met gevolgen voor de praktijk van evalueren.

We zullen elk van de drie omstandigheden nader uitwerken, waarbij we eerst kort de problematiek introduceren, daarna aangeven welke implicaties dit heeft voor sturing (wat doet dit met sturing?), en vervolgens

concluderen welke specifieke dilemma's voor beleidsevaluatie hieruit volgen (wat betekent dit voor beleid?). Aan het einde van dit hoofdstuk maken we de balans op en benoemen de gevolgen hiervan voor het denken over beleidsevaluatie.

3.1 Sturing in toenemende complexiteit

Beleidsevaluatie kan niet los gezien worden van het gehanteerde systeembegrip. Veldmodellen en interventiemodellen zijn manieren om een systeem in beeld te brengen, inclusief de verschillende elementen, condities en causale relaties in het beleidsdomein. Het brengt letterlijk het systeem in kaart en maakt expliciet welke causale relaties er tussen de partijen zijn – en welke causale relaties door middel van interventies beïnvloed kunnen worden om bepaalde uitkomsten te genereren. Systeemdenken, het relatieren van causaliteit aan elementen in het systeem, vormt de basis van beleidsevaluatie. Daarom is het van belang om na gaan welk systeem perspectief daarbij aan de orde is. Het Cynefin-framework onderscheidt vier manieren om systemen te bezien (figuur 1)⁷, namelijk: vanzelfsprekend, ingewikkeld, complex en chaotisch.



Figuur 1 Cynefin Framework van David Snowden, met een onderscheid in vier typen systemen (Snowden & Boone, 2007)⁸.

Vanzelfsprekende systemen

Het eerste type dynamiek is die van vanzelfsprekende systemen. In vanzelfsprekende systemen zijn causale relaties tussen componenten stabiel en overzichtelijk. Denk aan een mechanische aandrijving waarin eenvoudige verbindingen bestaan tussen een beperkt aantal elementen; als een radertje wordt gedraaid, komt een volgende component ook in beweging. Dat vereist een set van componenten, die in een bepaalde volgorde en verhouding bij elkaar gebracht moeten worden. Als het raderwerk adequaat op elkaar wordt aangesloten dan ‘werkt’ het. We weten ook *waarom* het zo gaat: we kennen de mechanismen van het raderwerk en begrijpen de werking. Er is een stabiele relatie tussen een beperkt aantal componenten, waardoor het systeem kenbaar, begrijpelijk, te beschrijven en overdraagbaar is in andere contexten. Vanzelfsprekende systemen zijn het domein van *best practices*, waarin we een probleem door waarneming kunnen leren kennen en op basis van routines of blauwdrukken kunnen temmen.

Ingewikkelde systemen

Het tweede soort dynamiek vinden we in ingewikkelde systemen. Ingewikkelde systemen kennen stabiele causale relaties tussen de componenten, maar door de enorme veelheid aan componenten is het vaak lastig om het systeem helemaal te doorzien. Ingewikkelde systemen zijn het domein van experts, die aan de hand van gestructureerde analyses de samenhang en relaties in een systeem kunnen ontrafelen. Dat kan moeilijk zijn, maar ingewikkelde systemen zijn wel kenbaar aan de hand van de afzonderlijke componenten. Hier wordt vaak het beeld van een mechanisch uurwerk gebruikt: als de werking eenmaal ontrafeld is kan een handleiding worden opgesteld, waarmee het systeem voorspelbaar en beheersbaar wordt. Wie het systeem begrijpt kan het bouwen, repliceren, gericht beïnvloeden en voorspellen. Het is kenbaar: niet eenvoudig, maar wel stabiel en voorspelbaar.

Complexe systemen

Complexe systemen zijn niet meer te kennen als de som der afzonderlijke delen, maar als de emergente uitkomst van ontelbare onderlinge interacties tussen componenten en hun omgeving.⁹ Die emergentie is de uitkomst van een veelheid aan terugkoppelingen (*reflexiviteit*), bijvoorbeeld door leereffecten of door veranderende relaties tussen actoren die niet per se te voorzien of te voorspellen zijn. De effecten van de interacties staan bovendien niet los van elkaar: er ontstaan patronen in de interacties, soms stabiliserend en behoudend, soms versnellend en veranderend. Interacties zetten volgende interacties in beweging, waarin het systeem door al die

interacties zelf ook verandert. Een complex systeem is daarom voortdurend in beweging.

Dynamiek in een complex systeem is vanwege de reflexiviteit wezenlijk anders van aard dan in een ingewikkeld systeem. Het schoolvoorbeeld is de zwerm, de school vissen of de spreuwenwolk: aan de hand van een zeer beperkt aantal regels voor de afzonderlijke individuen ontstaan op macroschaal enorm complexe, emergente structuren. Deze structuren zijn niet terug te voeren op de individuen, maar de regels van interactie zijn wel te kennen. Interventie in het systeem is echter bijzonder lastig, omdat uitkomsten altijd anders zijn: kleine interventies kunnen (op termijn) grote effecten hebben, terwijl grote interventies soms nauwelijks invloed hebben. Van tevoren is het effect doorgaans lastig in te schatten. Complexe systemen worden gekenmerkt door interacties en leereffecten, door een evolutionaire dynamiek van verandering en kantelpunten, en door allereerste terugkoppelingen die het systeem een grote veerkracht kunnen geven. Complexe systemen zijn dus het domein van emergentie, te kennen door experiment met aandacht voor proces en interactie. Sturing van complexe systemen volgt zelden een vooropgesteld plan, maar is de uitkomst van adaptatie en bijsturing, op grond van continue waarneming en monitoring, rekening houdend met een veelheid aan lokale interacties. Dit is wezenlijk anders dan het (vanuit een centraal punt) draaien aan een rad, en vervolgens bezien welke schakelingen er in serie plaatsvinden. In werkelijkheid vinden er allerlei dynamieken en tegenbewegingen plaats die alle op elkaar inwerken; sturing dient hier rekening mee te houden.

Wanordelijke systemen

Complexiteit wordt in de literatuur soms gekoppeld aan chaos¹⁰, maar er zijn wezenlijke verschillen: complexe systemen volgen kenbare patronen, terwijl wanordelijke systemen geen ordenende principes meer kennen. Het schoolvoorbeeld van een wanordelijk systeem is een brand, met een op voorhand onkenbare ontwikkeling. Wanordelijke systemen zijn voor bestuurders het domein van directe reactie, waardoor een situatie gestabiliseerd kan worden en orde kan worden hersteld. In de praktijk zijn chaotische systemen bijzonder zeldzaam voor de beleidsmaker, maar dat geldt veel minder voor de beleidsrespons. Vaak worden complexe vraagstukken verward met wanorde of zelfs catastrofes, waardoor sprake is van een snelle reactie in plaats van een meer preventieve of pro-actieve aanpak. De WRR spreekt hier van een 'risico-regel-reflex', gericht op ad hoc herstel van een situatie die zich beter leent voor een meer proactieve aanpak.¹¹

Uitdagingen voor evaluatie: attributie, contributie en nieuwe betekenis voor klassieke termen

Bestuurders en beleidsmakers worden steeds vaker geconfronteerd met complexe systemen. Tegelijkertijd worden de formele beleidstheorie en aannames van beleid nog vaak in de taal en het conceptueel raamwerk van ‘ingewikkelde systemen’ uitgedrukt. Dat brengt een aantal uitdagingen met zich mee.

In een als *complex* te typeren systeem is het lastig om *attributie* aan te tonen; het is lastig om te zien of dynamiek in het systeem een gevolg is van interventies – of dat het gaat om beweging die er anders ook in meer of mindere mate was geweest – en ‘van wiens acties’ de effecten een gevolg zijn. In complexe systemen reageren elementen in het systeem op elkaar en ontstaan als gevolg van die interacties nieuwe verhoudingen en relaties. Effecten zijn het gevolg van wisselwerking en daarmee per definitie moeilijk te herleiden tot één enkelvoudig beleid van één actor.

In de meest eenvoudige vorm is reflexiviteit te zien als een gevolg van handelen; de interventie van één actor resulteert in leereffecten, mogelijk met als gevolg dat anderen ook leren en zich anders gaan gedragen. Je zou dan kunnen zeggen dat de verandering in het systeem het gevolg is van de interventie die het leervermogen heeft aangesproken. Als een ruimhartige subsidieregeling massaal leidt tot de aanschaf van zonnepanelen, omdat partijen door hebben dat die regeling voor hen heel gunstig uitpakt, dan is dat in zekere zin ‘bedoelde reflexiviteit’ in het systeem.

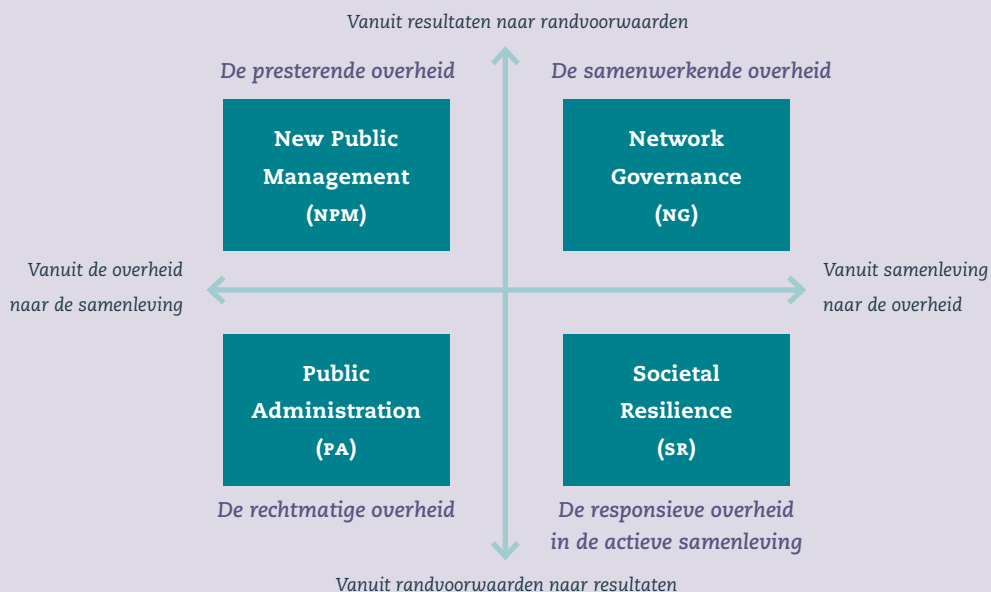
Maar vaak is reflexiviteit minder eenduidig, minder voorspelbaar en ook minder helder te herleiden tot één interventie. Partijen reageren op elkaar en op de dynamiek in het systeem. Ze passen hun handelen op elkaar aan en zo ontstaan nieuwe patronen in het systeem. Welke kant ze daarbij op bewegen valt nog te bezien. Beleid werkt in die veranderde omgeving onverminderd door, maar het is steeds lastiger om beweging geloofwaardig te relateren aan interventies vanuit het beleid. Zo is *attributie* temidden van complexiteit moeilijk vast te stellen, omdat onduidelijk is welk deel van het effect precies het gevolg van het beleid en welk deel vanuit het systeem zelf komt, en/of er anders ook ‘gewoon’ geweest was. *Attributie* is in complexe systemen niet afwezig, maar krijgt in complexe systeem-dynamiek andere betekenis. Dat probleem is veel minder aan de orde ten aanzien van *contributie*, maar dat wordt vanuit de traditionele aanpak van beleidsevaluatie gezien als een ‘zwakkere’ en laagwaardigere norm; pas als

attributie niet aantoonbaar is, vervallen we naar een lagere standaard, namelijk die van contributie.

3.2 Netwerkende strategieën voor het realiseren van maatschappelijke meerwaarde

Vier perspectieven voor overheidssturing

Beleidsdoelen kunnen op verschillende manieren gerealiseerd worden. In de literatuur over sturing kunnen we onderscheid maken in sturing van en door de overheid zelf en sturing waarin een grotere rol voor maatschappelijke partijen ligt. Daarnaast is er onderscheid mogelijk tussen sturing die zich primair richt op het borgen van condities en randvoorwaarden en sturing waarin de realisatie van concrete prestaties centraal staat. Als we deze beide elementen tegenover elkaar plaatsen ontstaat een model met vier kwadranten waarmee perspectieven op overheidssturing te ordenen zijn (zie [figuur 2](#)). We zullen elk van de vier perspectieven in dit kwadrantenmodel kort bespreken.¹²



Figuur 2 Vier perspectieven op sturing, die ieder andere waarden centraal stellen (Van der Steen et al., 2015).

‘Public administration’: de rechtmatige overheid (linksonder)

Vanuit het perspectief van een *rechtmatige overheid*¹³ is verankering in het democratisch systeem en het recht de basis voor alles wat overheidsorganisaties doen.¹⁴ Rechtmatigheid, rechtsgelijkheid en rechtsstatelijkheid vormen de basis van bestuur.¹⁵ Het perspectief van de rechtmatige overheid stelt die legitimiteit en rechtmatigheid centraal. De verhouding met de samenleving is gericht op het borgen van rechten en plichten, die als maatschappelijke ordening fungeren.

‘New public management’: de presterende overheid (linksboven)

In het perspectief van de *presterende overheid* gaat overheidssturing primair om het op een effectieve en efficiënte manier bereiken van de vooraf afgesproken resultaten. De overheid is zelf verantwoordelijk voor het doelmatig realiseren van afgesproken prestaties. Dit perspectief komt voort uit het denken vanuit New Public Management (NPM), een stroming die sinds de jaren '90 in publieke organisaties is geïntroduceerd.¹⁶ Vanuit dit perspectief bezien ontleent de overheid zijn legitimiteit aan het vermogen om prestaties te leveren: niet alleen zorgvuldig en juist (rechtmatig), maar ook efficiënt en effectief (doelmatig).

‘Network governance’: de samenwerkende overheid (rechtsboven)

Bij de *netwerkende overheid* staat het principe centraal dat doelen vaak worden bereikt in samenwerking met anderen.¹⁷ Om de eigen doelen te realiseren, moet de organisatie de weg naar buiten kunnen bewandelen. Ze gaat op zoek naar partners die kunnen bijdragen aan die doelen, omdat ze dezelfde, gelijksoortige of deels overlappende belangen en doelen hebben of omdat doelen niet door de overheid alleen gerealiseerd kunnen worden. Het bij elkaar brengen van die uiteenlopende belangen vormt de kern van het perspectief van de netwerkende overheid. Dat gaat om het formeren van coalities van partijen, die vanuit gedeelde belangen de benodigde middelen inbrengen om een probleem op te lossen.

‘Societal resilience’: de participerende overheid (rechtsonder)

In het perspectief van de *participerende overheid* staat maatschappelijke energie centraal.¹⁸ Publieke waarde wordt niet alleen door de overheid ‘gemaakt’, maar komt evengoed van anderen: burgers, bedrijven en het maatschappelijk middenveld. Op de manier die zij willen, vanuit de doelen die zij zichzelf stellen.¹⁹ De overheid gaat niet naar buiten om daar bondgenoten te werven voor de eigen uitvoeringsagenda, maar komt in aanraking met maatschappelijke dynamiek die er al is. Soms is dat een initiatief

dat aansluit bij de eigen agenda en dat past binnen kaders die de overheid hanteert. Vaak zijn het initiatieven die gedeeltelijk passen binnen de doelen en kaders van de overheid. De vraag is dan hoe om te gaan met praktijken die niet helemaal passen binnen de overheidsdoelen, maar wel bijdragen aan maatschappelijke belangen. Dat vereist politieke afweging: soms betekent het 'gewoon niets doen en laten bestaan', soms is de uitkomst dat het initiatief moet stoppen, omdat het andere belangen schaadt.

Publieke waarde realiseren in passende combinaties van perspectieven

Elk van de perspectieven legt de nadruk op specifieke waarden van goed bestuur. Het model is nadrukkelijk niet bedoeld als keuzepallet voor één van de perspectieven, maar expliciteert wel de accenten die bij overheidssturing worden gelegd. Denk bijvoorbeeld aan het *Klimaatakkoord* en aan de regionale agenda's die nu worden afgesproken in het kader van het *Interbestuurlijk Programma*. Dat zijn voorbeelden van sturing die primair is ingericht volgens de principes van de netwerkende overheid, met daarin tevens een poging om maatschappelijke veerkracht en zelforganisatie (societal resilience) zoveel mogelijk te mobiliseren. Tegelijkertijd gaat het ook om majeure inzet van overheidsmiddelen, zowel in financiële zin in subsidies en incentives, als in regulatieve zin, in de vorm van wetgeving en normstelling of juist de uitzondering daarop. De nadruk in de sturing ligt dus op het perspectief van de *netwerkende overheid*, maar daarbij hoort ook het op adequate wijze organiseren van monitoring, verantwoording en het laten zien van de doelmatigheid en effectiviteit van het beleid (passend bij het denken vanuit New Public Management). En hierbij hoort ook het organiseren van deugdelijke politieke processen, borging van rechtstatelijke waarden en de borging van zwakke belangen, zoals het de rechtmatige overheid betaamt. Zo is het steeds de kunst om enerzijds de nadruk op één van de perspectieven op een goede manier invulling te geven en anderzijds ook de waarden van de andere kwadranten goed in te vullen. Daarbij is de keuze voor de aanpak van de sturing afhankelijk van de opgave, in het hedendaagse discours aangeduid als 'opgavegericht werken'.²⁰

De strategie voor doelrealisatie is afhankelijk van de aard van het vraagstuk en de constellatie van partijen waarin het vraagstuk en de te realiseren waarde(n) spelen. Dat kan betekenen dat bijvoorbeeld de warmtetransitie vorm krijgt met veel meer en meer directe sturing vanuit de overheid, of in een meer beperkt netwerk van direct betrokkenen; het gaat hier om een transitie waarin infrastructuur een centrale rol speelt en op een schaal die moeilijk in kleinschalige bottom-up initiatieven of lokale coöperaties te

realiseren is. Voor opwekking van duurzame energie geldt dat veel minder. Daar kan decentrale opwekking, in lokale samenwerkingsverbanden juist een belangrijk deel van de oplossing vormen en is sturing mogelijk meer gericht op het aanjagen en ondersteunen van bottom-up initiatieven. De idee achter deze manier van kijken naar sturing is dat opgaven steeds meer en meer bewust worden aangepakt volgens de sturingsstrategie die daar het best bij past, *waarna* het zoeken is naar de verbinding met de andere perspectieven die evenzeer van belang zijn. Daarbij geldt voor de linker-kwadranten dat de bestaande interpretaties, mechanismen en werkwijzen vaak minder goed van toepassing zijn op de werkpraktijk die ‘rechts’ in het schema ontstaat.

Observaties ten aanzien van sturing: werken volgens ‘rechts’ vereist herinterpretatie van ‘links’

In de departementale beleidspraktijk wordt al steeds vaker gewerkt vanuit het principe dat sturing ‘opgavegericht’ is en dat per opgave gezocht wordt naar de best passende aanpak. In veel opgaven wordt gekozen voor meer netwerkende strategieën, met een grote en eigenstandige rol voor maatschappelijke partijen. Voorbeelden hiervan zijn de aanpak van de energietransitie, maatschappelijke verbinding van het natuurbeleid, en de vormgeving van regionale energiestrategieën. Hierbij wordt expliciet gewerkt vanuit strategieën die de nadruk sterk leggen op combinaties van ‘rechtsboven’ en deels ook ‘rechtsonder’. Dat vereist vervolgens wel dat de waarden van linksonder en linksboven passende invulling krijgen; rechtmatigheid, rechtsgelijkheid en rechtszekerheid, verantwoording, effectiviteit en efficiency zijn met deze aanpak niet weg, maar moeten een invulling krijgen die past bij de aard van de opgave. Die uitdaging wordt in het beleid dagelijks gevoeld, in die zin dat er weinig concreet instrumentarium voorhanden is om die verbinding goed te maken.²¹

Uitdagingen voor beleidsevaluatie: evalueren op een manier die past bij een netwerkende aanpak

Voor steeds meer vraagstukken werken overheden vanuit een perspectief dat meer ‘rechts’ in het schema ligt. Ze maken verbinding met andere partijen in een netwerkaanpak, waarbij door samenwerking beleid tot stand en/of tot uitvoering komt. Dat maakt additionaliteit op zich vast te stellen, hoewel het lastig blijft om tot attributie te komen. Als politie, om, gemeentelijke diensten, de FIOD, sociale wijkteams, jeugdleiders, sportverenigingen, groepen ouders en ‘peers’ van jongeren samen een aanpak ontwikkelen om de betrokkenheid van jongeren bij hennepcultuur en handel te bestrijden,

hoe is dan vast te stellen wie welk deel van het effect heeft veroorzaakt? Het is misschien wel mogelijk om meerwaarde vast te stellen, maar die waarde is lastig toe te bedelen aan individuele actoren of aan precies afgebakende interventies. Het beleid is als een cluster van handelingen van een groep actoren te evalueren, maar dieper dan dat lukt niet.

Ingewikkelder wordt het naarmate strategieën meer richting ‘rechtsonder’ (aansluiten bij maatschappelijke veerkracht) gaan en het beleid inhoudelijk ook meer indirect wordt; denk aan het scheppen van voorwaarden, het bieden van een platform, of het ondersteunen van verbindingen die in het netwerk ontstaan. Hier is additionaliteit lastig vast te stellen, omdat het steeds de vraag is of de beweging er niet ook ‘gewoon’ zonder verbindende, ondersteunende of ‘platformende’ handelingen van de overheid zou zijn geweest. Creëert indirecte sturing de beweging, of is het meer een bevestiging van wat er al was? En wat zijn manieren om dat vast te stellen? Hetzelfde geldt voor attributie; niet alleen is de meerwaarde van het scheppen van de gelegenheid moeilijk vast te stellen, maar de vraag is ook wie precies de gelegenheid maakt? Is dat een overheid die een netwerk bij elkaar brengt door bijeenkomsten te organiseren, of zijn het de aanwezigen die met elkaar aan de slag willen en daarbij nut en noodzaak van de aanwezigheid van de overheid inzien? Hoe is wat er gebeurt tot het handelen van individuele actoren te herleiden?

Het meest principieel is misschien nog wel de vraag wat en van wie het beleid in deze aanpak is en wat evaluatie in deze context betekent. Immers, de essentie van de netwerkende aanpak is dat beleid met anderen tot stand komt en dat het eigenaarschap van de sturing in zekere mate gedeeld zijn. Dat is min of meer expliciet de bedoeling van het beleid. Soms is de overheid daarin de organiserende en verbindende partij, maar netwerken kunnen ook van onderop ontstaan, waarin de overheid zich er op enig moment als extra of ontbrekende schakel bijvoegt. Beleid is dan geen ‘therapie’ die een overheid in het systeem injecteert, maar veeleer de uitkomst van een beweging waarin de overheid samen met niet-statelijke actoren participeert.²² Wie is dan de ‘evaluator’ en in welke relatie staat die dan tot de anderen?

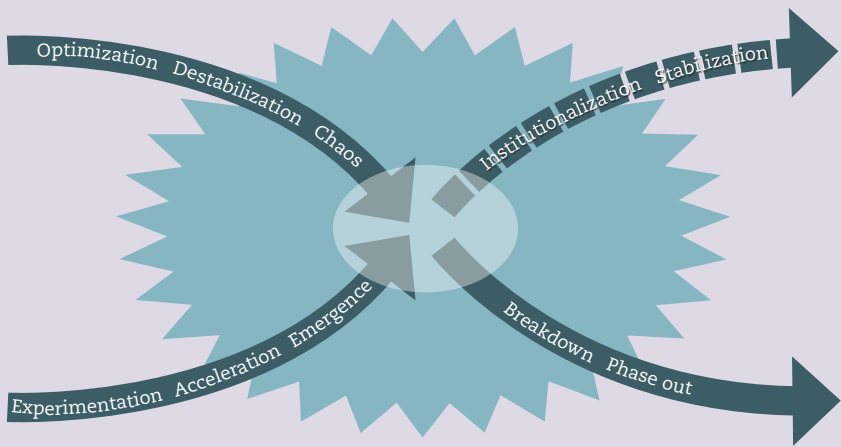
Zo liggen er specifieke uitdagingen voor beleidsevaluatie in een netwerkcontext, waarbij partijen in allianties samenwerken om gedeelde doelen te realiseren en opgaven op een passende manier op te lossen. Die uitdagingen gaan over de aard van evaluatie (wat wordt er geëvalueerd?), over de

rolverdeling en verantwoordelijkheid, en over het gebruik van de evaluatie als middel tot verantwoording of als middel om over beleidstrajecten te leren en deze waar nodig bij te stellen.

3.3 Beleidsrealisatie in transitiedomeinen: sturing voor het onbekende

Voor veel beleidskwesties geldt dat het beleid niet alleen gaat om intensivering of in ieder geval continuering van het bestaande pad, maar om een transitie; om een ‘verandering van regime’, zoals het in de transitieliteratuur wordt genoemd. Het idee van transitiebeleid is dat een min of meer uitgekristalliseerd systeem uit balans raakt en nieuwe configuraties van partijen ontstaan.

Transitiebeleid bestaat altijd uit twee gelijktijdige ontwikkelingen²³: enerzijds het kanaliseren van de dynamiek in het nieuwe pad, in de beoogde richting en met de bedoelde snelheid, anderzijds het in beweging brengen van het bestaande pad, wat ook het ‘afbreken’ van het bestaande regime kan betekenen.²⁴ **Figuur 3** brengt die dynamiek in beeld.



Figuur 3 De dynamiek van transities (Loorbach et al., 2017).²⁵

Transitieopgaven zijn in deze zin nog ‘gewoon’ doelgericht: er zijn transitiedoelen, vaak zelfs uitgedrukt in heel concrete percentages of ‘telbare’ eenheden (reductie van uitstoot, penetratie van andere modaliteiten, ‘van het gas af’), en meestal ook concrete jaartallen waarin dat bereikt moet zijn (‘2050’). Vaak worden daarbij *transitiepaden* vastgesteld, die tussentijdse mijlpalen en te behalen waarden markeren. In het geval van het Deltaprogramma zijn dergelijke mijlpalen en tussentijdse waarden zelfs verbonden aan ‘automatische’ beleidsaanpassingen die in werking treden als waarden niet gehaald worden. Tot zover lijkt er ten aanzien van beleidsevaluatie weinig aan de hand te zijn: via doelstellingen en waarden is het mogelijk om te doorzien of het beleid op koers ligt en de transitie zich voltrekt op de vooraf bedachte manier en met de daarvoor aangegeven snelheid.

Tegelijkertijd beperkt de dynamiek in transities zich niet tot de uitgezette paden. De doelen en paden zijn in transitiedomeinen vooral bedoeld om in beweging te komen en de richting aan te geven. Neem het voorbeeld ‘van het gas af’. Dat is een ingrijpende transitie waarvan het einddoel benoemd is, maar tegelijkertijd geldt dat het doel net zozeer een middel is, gericht op een veel breder benoemde energietransitie. Door het doel uit te spreken ontstaat beweging in het systeem en komt een nieuw regime in gang. Dankzij die beweging komen interventies, instrumenten en partijen in ontwikkeling die er mogelijk voor gaan zorgen dat het ook daadwerkelijk gebeurt. Hoe dat er uit ziet, hoe snel het gaat, wie het doet, wat er gebeurt, et cetera is inherent onzeker. Het gaat er daarbij niet zozeer om dat de doelen of het pad precies gevolgd wordt, maar dat de beweging daadwerkelijk in gang is gezet, zodat doel en pad zich in de praktijk gaan vormen. Een energiesysteem zonder gas is daarmee niet alleen precies dat, maar vooral ook een indicator voor een veel omvattender transitie naar verduurzaming van dat energiesysteem.

Beleidsrealisatie in transitiedomeinen heeft daarmee een wat paradoxale opbouw. Sturing beoogt radicale verandering, maar de terminologie verhoudt de intensiteit van die verandering ook enigszins. Het spreken van ‘nieuwe regimes’ klinkt gecontroleerd en beheersbaar. In werkelijkheid zijn de doelen en de trajecten ernaartoe vaak vaag en onbestemd. Vaak weten we wel wat we niet willen, maar is nog niet erg duidelijk wat we wel willen. Daarom zijn veel transities negatief geformuleerd als ‘van het concrete één af, naar het nog onbekende andere’. Zo is beleid voor een transitie ook beleid voor onontgonnen terrein. Het is beleid bedoeld voor een nog onbekende constellatie, met doelen, middelen en ambities die

tamelijk vanzelfsprekend veranderlijk zijn. Dat maakt transitiebeleid ‘doelgericht open’, zonder vooraf helderheid over wat er op welke manier – waarbij ook de constellatie van het relevante netwerk voor de transitie vooraf lang niet duidelijk is.

In de beleidspraktijk betekent dit dat beleid wel geformuleerd kan worden als transitie of systeemverandering, maar in de praktijk van onzekerheden en complexiteit is de uitvoering doorgaans een zoekproces, vol van navigatie en bijsturing. Sturing is bedoeld om in beweging te zetten, en het bestaande regime te ontregelen, zodat er nog onbekende andere zaken voor in de plaats komen. Waar beleid en sturing dan weer verder op in kunnen spelen. Zo krijgt sturing een sterk *ex durante* en incrementeel karakter. De paradox is dat idealen van transitie niet anders kunnen worden uitgevoerd dan als een stapsgewijs proces van adaptief bestuur. Een belangrijke valkuil is ook dat een dergelijke incrementele aanpak leidt tot continue optimalisatie, zonder dat de ambitieuze doelstellingen gehaald kunnen worden.²⁶ Tegelijk is er geen alternatief: een al te technocratische benadering – ‘social engineering’ – leidt al snel tot blauwdrukdenken, waarbij complexe vraagstukken worden verward met ingewikkelde vraagstukken (in de taxonomie van Snowden²⁷, zoals uitgelegd in 3.1).²⁸

Het beeld van sturen en continu bijsturen is één van de grote uitdagingen voor beleidsevaluatie. Bij transities wordt beleid gemaakt voor een onzekere en niet goed voorspelbare toekomst. Dat beleid is opgavegericht en doordacht, maar ook inherent open en onvoorspelbaar, zowel in de opbrengsten als in de dynamiek die beleid in gang zet. Dat maakt dat het moeilijk is om attributie vast te stellen, hoewel contributie al beter in te passen is in deze benadering. Is transitiedynamiek het gevolg van concreet aanwijsbare interventies, of was het toch wel gebeurd? Welk deel van de beweging is herleidbaar tot welk beleid? En hoe gaat men in een beleidsevaluatie om met de bewegende doelen, die ook bijgesteld kunnen worden onderweg? Dat is vanuit beleidsoptiek logisch, maar het is lastig in te passen in een traditionele methodiek van beleidsevaluatie. Als sturing bedoeld incrementeel en *ex durante* vorm krijgt (op een manier die niet perse het vooraf ‘vastgestelde’ transitiepad volgt) dan vereist dat evaluatievormen die met die aangroeiende dynamiek mee kunnen bewegen.

3.4 Opgavegericht evalueren: evaluatie die past bij de opgave

We hebben gezien dat beleid en sturing onder bepaalde omstandigheden specifieke gevolgen hebben voor beleidsevaluatie. Beleidsrealisatie in complexe systemen, in netwerkkarrangementen, en/of in transitiedomeinen biedt specifieke uitdagingen voor beleidsevaluatie (zie tabel 2). Als beleid op deze manier vorm krijgt, dan zijn er specifieke dilemma's voor beleidsevaluatie die zich voordoen, naast de hiervoor al benoemde inherente dilemma's van beleidsevaluatie. Voor veel beleid zal gelden dat het met tenminste één en waarschijnlijk twee of drie van de specifieke omstandigheden en bijpassende dilemma's te maken heeft. Het EZK-programma *Samen in Beleid* en de nadruk op de opgave, complexiteit en transitiedoelstellingen in de uitlatingen van topambtenaren en politiek bestuurders van het departement laten zien dat dat de komende periode eerder zal toenemen dan afnemen. De hier benoemde omstandigheden die specifieke dilemma's veroorzaken zullen vaker de 'normaaltoestand' worden in de aanpak van opgaven.

Het ligt daarom voor de hand om voor beleidsevaluatie hetzelfde pad te volgen als nu met beleid en sturing al is ingezet: door minder te denken in termen van beleid of sturing als generieke term, maar vooral *opgavegericht*, waarbij de invulling van sturing en beleid de aard en context van de opgave volgt. Wij noemen dat hier *opgavegericht evalueren*: beleidsevaluatie die past bij de aard en de context van de opgave waarop sturing is gericht. Evaluatie gaat dan primair over het realiseren van de best mogelijk oplossing voor een opgave, met rekenschap dat een overheid probeert in democratische processen bekrachtigde doelen te realiseren, via beleid, via arrangementenbouw, via sturing, en door uitvoering. Het democratisch gestelde *doel* staat centraal en dat doel komt voort uit een maatschappelijke opgave. Vanuit de contouren van het doel en de maatschappelijke opgave ontwerpen betrokkenen een bepaalde manier van sturing, die de kans op de aanpak van een opgave zo groot mogelijk maakt. Vervolgens *matchen* ze daar een manier van evalueren bij die past bij deze manier van opgavegericht werken: *opgavegericht evalueren*.

	Observaties t.a.v. sturing	Uitdagingen voor evaluatie
<i>Beleidsrealisatie in complexe systemen</i>	Meervoudige wegen naar vaak kwalitatief gestelde maatschappelijke doelen. Context van veranderende relaties, emergentie en reflexiviteit. Voorspelbaarheid en kenbaarheid van het systeem is beperkt, dus beleidsrealisatie is onzeker: 'getting to maybe'. Controle is beperkt en sturing is een navigatieproces van voortdurende bijsturing.	Attributie is lastig vast te stellen, door gebrek aan stabiele causaliteit en door meervoudigheid van mogelijke oplossingen. Contributie gaat over (relatieve) beïnvloeding van stromen, processen en patronen. ²⁹
<i>Netwerkende strategieën voor beleidsrealisatie</i>	De overheid staat naast andere partijen en komt met die anderen tot doelen en interventies. Verdeling van rollen en van verantwoordelijkheden (deels) afhankelijk van beleidsarrangement.	Beleidsinterventie als gedeelde verantwoordelijkheid van het arrangement; vaak meer gericht op proces dan op specifiek doel. Doelstellingen en kaders kunnen gedurende proces worden bijgesteld door betrokken partijen
<i>Beleidsrealisatie in transitiedomeinen</i>	Transitie gaat over disruptieve processen, waarbij een bestaand regime wordt vervangen door een, nog onbekend, ander regime. Contouren, partijen, technieken en doelen zijn a priori nog onbekend. 'Transitiedoelen' of 'transitiepaden' zijn doorgaans ambities en inzet op inherent onzekere processen.	Ankerpunten voor evaluatie staan niet vast, maar zijn net zo goed onderwerp van transitie. Evaluatie is niet (alleen) gericht op het vaststellen van doelbereik, maar (tevens) op het vaststellen van de richting en snelheid van de bedoelde beweging en de kracht of dichtheid van de stroom.

Tabel 2 Specifieke dilemma's voor beleidsevaluatie.

4 Omgaan met dilemma's: ontwikkelingen in de evaluatiepraktijk

In dit hoofdstuk belichten we drie voorbeelden van opgavegericht evalueren, om zo meer zicht te krijgen op de praktische dilemma's van het evalueren in verweven dynamiek. Aan de hand van deze voorbeelden illustreren we vijf ontwikkelingen in het denken over beleidsevaluatie, binnen een context van complexiteit. De vraag die daarbij centraal staat is: *hoe wordt beleids-evaluatie in een context van complexiteit eigenlijk gedaan en welke lessen kunnen we daaruit meenemen voor de ontwikkeling van een nieuwe standaard?* Dat is een combinatie van herkenning van dilemma's, maar ook een weergave van manieren om er mee om te gaan in de praktijk, aan de hand van voorbeelden die ieder op hun eigen manier evalueren in verweven dynamiek in beeld brengen. Hiermee halen we geen compleet of finaal beeld op van de praktijk, maar krijgen wel gevoel voor de kwesties waar evaluatoren in hun praktijk mee kampen.

4.1 Meervoudig evalueren in de praktijk

We werken hier drie voorbeelden uit van beleidsevaluatie die rekening houden met complexe beleidscontexten. Deze voorbeelden zijn: de evaluatie van de Topsectorenaanpak (door Dialogic), de Lerende Evaluatie van het Natuurpact (uitgevoerd door het Planbureau voor de Leefomgeving in samenwerking met Wageningen University en Research) en de evaluatiepraktijk van de Directie Internationaal Onderzoek Buitenlandbeleid (IOB, ministerie van BZ). De selectie van deze praktijken vond pragmatisch plaats: in een verkennende gespreksronde werden wij door diverse betrokkenen gewezen op deze praktijken als interessante vindplaatsen van de praktijkdilemmas waar we naar op zoek waren. We introduceren de voorbeelden, alvorens we ingaan op ontwikkelingen binnen het denken over evalueren te midden van complexe beleidscontexten.

Evaluatie Topsectorenaanpak

In 2011 heeft het Ministerie van Economische Zaken de zgn. Topsectorenaanpak gelanceerd. De Topsectorenaanpak beoogt de versterking van het innovatie- en verdienvermogen van sterke sectoren in Nederland, middels het stimuleren van samenwerking tussen verschillende partijen. De aanpak bouwt op de ‘gouden driehoek’ bestaande uit bedrijven, kennisinstellingen en overheden, die gezamenlijk visies afstemmen om de behoefte van het bedrijfsleven en de expertise van de kennisinstellingen beter af te stemmen en om knelpunten weg te nemen die in de weg staan voor het verwezenlijken van collectieve plannen voor innovatie. Doel is om het ondernemingsklimaat van bedrijven te stimuleren en te helpen bij het duurzaam en innovatief ondernemen. De Topsectorenaanpak is gericht op negen topsectoren: Tuinbouw & Uitgangsmaterialen, Agri & Food, Energie, Chemie, Logistiek, Water, High Tech Systemen & Materialen, Life Sciences & Health en Creatieve Industrie. In al deze sectoren heeft Nederland wereldwijd een vooraanstaande positie.

Dialogic heeft het Topsectorenbeleid geëvalueerd.³⁰ Specifiek richt de evaluatie zich op de vraag of de aanpak effect sorteert: helpt de aanpak om op basis van collectiviteit systeemveranderingen in gang te zetten? Deze hoofdvraag is uitgewerkt in vier subvragen die betrekking hebben op de beleidsimpact, de doelmatigheid, het beleidsontwerp (herleidbaarheid van uitkomsten tot de manier waarop de aanpak is vormgegeven) en beleidsaanbevelingen. De evaluatiemethode omvat zowel een raamwerk voor beoordeling van impact als een ontwerpraamwerk. Met name dit tweede onderdeel is vernieuwend, omdat het niet alleen gericht is op uitkomsten, maar ook op het beleidsproces. De evaluatiemethodiek is door Dialogic nieuw ontwikkeld op grond van academische literatuur en in samenwerking met de Harvard Kennedy School for Public Policy, met als uitgangspunt om beperkingen van reguliere evaluatiemogelijkheden voor het evalueren van beleid te verkleinen.³¹ Omdat het Topsectorenbeleid gekenmerkt kan worden als ‘transformatief beleid’, dat als uitgangspunt heeft dat overheden ondersteuning bieden bij het ontwikkelen van breed gedragen innovatietrajecten, wordt er beleid gevoerd dat *selectief* (gefocust op kansen en uitdagingen), *adaptief* (aanpassingsgericht) en *multi-instrumenteel* (op basis van een beleidsmix) is. Inherent daaraan is dat uitkomsten vooraf niet goed voor te stellen zijn; het gaat immers om innovaties. De gebruikte evaluatiemethodiek is op deze kenmerken en omstandigheden afgestemd, en richt zich met name op de vraag in hoeverre het beleid passend is om het vermogen van actoren om te innoveren te optimaliseren.

Lerende Evaluatie van het Natuurpact

In 2011 is de verantwoordelijkheid voor het Nederlandse natuurbeleid vanuit het Rijk overgedragen naar provincies. In 2013 kreeg dit uitwerking in het Natuurpact, een bestuurlijk akkoord over de ambities en financiering van het Nederlandse natuurbeleid tot aan 2027, waarin het Rijk en de provincies afspraken hebben gemaakt over ambities, financiering en de verdeling van middelen. Deze afspraken zijn formeel vastgelegd in de wet Natuurbescherming. Rijk en provincies hebben in het Natuurpact afspraken gemaakt over de ‘ambities’ in het natuurbeleid. Er zijn 3 hoofdambities: (1) versterken biodiversiteit; (2) versterken maatschappelijke betrokkenheid bij natuur; en (3) versterken van de relatie tussen natuur en economie.

De complexiteit van deze opgave zien we terug in de abstracte lange-termijndoelstellingen, de bestuurlijke gelaagdheid, de verschillende betrokken actoren en het gegeven dat natuurbeleid is belegd in verschillende wettelijke regelingen, beleidsregels en akkoorden. Dat maakt het Natuurpact lastig op traditionele wijze te evalueren. De doelstellingen in het Natuurpact zijn bovendien breed geformuleerd, met gaandeweg ruimte voor nadere uitwerking. Er zijn wel kaders: de provincies werken het beleid op twaalf verschillende manieren uit, binnen wettelijke regelingen en akkoorden van de Rijkskaders. Daarmee is het niet alleen relevant om te weten *hoe* je dit het beste evalueert, maar vooral ook wat een nuttige opbrengst is van de evaluatie voor de verdere ontwikkeling van het beleid.

Het Planbureau voor de Leefomgeving (PBL) is gevraagd om het Natuurpact te evalueren en daarbij ook handelingsperspectieven voor een effectiever en efficiënter natuurbeleid voor te stellen. De hiervoor ontwikkelde ‘lerende evaluatie’ vindt driejaarlijks plaats; de eerste evaluatie is gepubliceerd in 2017. Dit onderzoek is uitgevoerd in samenwerking met Wageningen University en Research (WUR)³², volgens een methodiek ontwikkeld door PBL, WUR en de Vrije Universiteit Amsterdam (VU).³³ Lerend evalueren betekent dat de evaluatie is uitgevoerd samen met de diverse groepen betrokkenen en gericht is op het formuleren van lessen om beleid tijdig bij te sturen, indien voortschrijdend inzicht dit noodzakelijk maakt. Dit is een voortdurend (*ex durante*) proces, waarin de evaluatie kennis levert die door beleidsmakers kan worden gebruikt voor verdere beleidsontwikkeling. Zo wordt gaandeweg lerend vermogen opgebouwd in het beleidsveld.

De Directie Internationaal Onderzoek en Beleidsevaluatie (IOB)

De Directie Internationaal Onderzoek en Beleidsevaluatie (IOB) is een onafhankelijke evaluatiedienst van het Ministerie van Buitenlandse Zaken dat rapporteert over resultaten van het Nederlandse buitenlandbeleid op de langere termijn. Alle beleidsartikelen op de begroting van het ministerie worden periodiek en systematisch geëvalueerd, zoals de Regeling Periodiek Evaluatieonderzoek (RPE) voorschrijft. De directie doet verschillende typen evaluatieonderzoek, waaronder beleidsanalyses, syntheses studies en periodiek en doorlopend evaluatieonderzoek. Voorbeelden van onderzoeken die de directie uitvoert zijn: een review van het Nederlandse voedselzekerheidsbeleid in ontwikkelingspartnerlanden (o.a. in hoeverre is bijgedragen aan betere toegang tot gezonde voeding en hogere landbouwproductie); een beleidsdoorlichting naar hoe effectief en efficiënt steun van Nederlandse maatschappelijke organisaties is aan lokale *counterparts* in het buitenland; en een meervoudig landenonderzoek naar de effecten van bezuinigingen op ontwikkelingssamenwerking (in Bolivia, Burkina Faso, Guatemala, Tanzania en Zambia).

Het IOB hanteert expliciet de dubbeldoelstelling ‘verantwoorden en leren’ als uitgangspunt voor de evaluaties. De beleidsdoorlichtingen die de Directie IOB uitvoert zijn gericht op het in kaart brengen van de doeltreffendheid, doelmatigheid en relevantie van het beleid; impactstudies zijn specifiek gericht op effecten van beleid op de lange termijn en in syntheses studies worden overige vragen verkend. De IOB hanteert de volgende evaluatiecriteria, die grotendeels overeenkomen met de criteria van het internationaal gebruikte model ontwikkeld door de Development Assistance Committee van de OESO³⁴:

1. Doelmatigheid (mate waarin resultaten opwegen ten de kosten van gekozen middelen);
2. Doeltreffendheid (mate waarin de directe resultaten van activiteiten bijdragen aan realisatie van geformuleerde beleidsdoelstellingen);
3. Relevantie (mate waarin de effecten van uitgevoerde activiteiten bijdragen aan de realisatie van het uiteindelijke doel);
4. Duurzaamheid (in welke mate het effect van de activiteit blijvend is);
5. Coherentie (mate waarin doelen en resultaten van het beleid niet worden aangetast door andere vormen van beleid, waaronder consistentie van beleidsmaatregelen die tegenstrijdigheid vermijden tussen verschillende beleidsterreinen, alsook beleidssamenhang).



Met name het criterium van duurzaamheid verbreedt de horizon van het evaluatieonderzoek; er is daarmee niet alleen aandacht voor direct meetbare effecten – *output* – maar ook voor indirecte gevolgen en bredere uitkomsten – *outcomes*. Het perspectief van de IOB is steeds meer verschoven van evaluator van projecten, tot evaluator van grotere programma's en brede beleidsopgaven. Daarmee komen ze op een terrein dat minder goed traditioneel te evalueren valt.

4.2 Ontwikkelijnen in de evaluatiepraktijk

We zien grofweg vijf ontwikkelijnen in de praktijkvoorbeelden die pogen meer recht te doen aan complexiteit. Deze hebben achtereenvolgens betrekking op de *functie* van het evaluatieonderzoek, de *timing*, de *focus* die centraal staat, de *rol* van de evaluator en de *motivatie* achter de evaluatie.

- **Verantwoorden, maar ook leren en bijsturen**

“In het evaluatiedenken krijgt complexiteit de laatste jaren voet aan de grond; daarbij wordt leren centraler gesteld.”³⁵

Functie

Het doel waarmee evaluatieonderzoek wordt uitgevoerd is onderhevig aan verandering en daarmee verbreedt de functie van beleidsevaluatie. Naast de traditionele functie van verantwoording is er steeds meer aandacht gekomen voor leren en bijsturen als direct doel. Het gaat daarbij niet om een drastische verandering – verantwoording afleggen blijft nog steeds de meest dominante functie van evaluatie – maar om een verschuiving; het accent ligt niet meer uitsluitend op verantwoorden, maar leren en bijsturen zijn er als expliciete doelen bijgekomen. Dit betekent bijvoorbeeld dat uitkomsten van evaluatieonderzoek niet alleen resultaten zijn, maar dat er bijvoorbeeld ook handelingsperspectieven ontwikkeld worden die door beleidsmakers meegenomen kunnen worden voor de verdere ontwikkeling van beleid. Beleidsevaluatie wordt hierin erkend als een logisch moment van aandacht voor heroverweging en eventuele bijstelling van beleid.

Dit zien we terug in de Lerende Evaluatie van het Natuurpact, waarin leren uitdrukkelijk tot een van de hoofddoelen is gemaakt. Onderzoekers en beleidsbetrokkenen ontwikkelen een gezamenlijk leerproces waarin nagedacht wordt over nieuwe handelingsperspectieven en beleidsopties. Naast ‘standaardelementen’ van een evaluatie, zoals efficiëntie, effectiviteit en

transparantie, is er ook aandacht voor responsiviteit en leervermogen en besteedt de lerende evaluatie veel aandacht aan wat in de praktijk werkt en wat niet. Ook is er veel oog voor interactie met stakeholders; input van betrokkenen wordt uitvoerig uitgewisseld en zoveel mogelijk meegenomen in de lessen. Dit krijgt bijvoorbeeld vorm in leer- of werksessies waarbij telkens personen aanwezig zijn uit alle betrokken ‘niveaus’, bijvoorbeeld vanuit het Rijk, de provincies en maatschappelijke organisaties. Het is daarmee niet alleen een evaluatie, maar ook een beleidsontwikkelingsmiddel. Ook het IOB voert evaluaties uit met een specifiek opgenomen leerfunctie, bijvoorbeeld onderzoek gericht op het verkrijgen van inzicht in effecten van interventies in het kader van diplomatie.

- **Ex post, maar ook ex ante en ex durante**
“Ik zou graag zien dat we niet achteraf evalueren omdat het moet, maar op momenten dat het ertoe doet en je er ook echt wat mee kan.”³⁶

Timing

In het verlengde hiervan ligt een tweede ontwikkeling, namelijk een verschuiving in de *timing* van evaluatieonderzoek. Hier liggen verschillende vragen aan ten grondslag, bijvoorbeeld: wanneer kunnen we beleid het beste evalueren? Wat is een passend moment om de ‘thermometer’ erin te steken? Deze vragen hangen nauw samen met de functie of het doel van evaluatieonderzoek. In het klassieke perspectief, waarin verantwoordelijk nog het voornaamste doel was van evaluatie, ligt het voor de hand om na afloop – *ex post* – te evalueren wat beleidsopbrengsten zijn. Wanneer het expliciete doel is om ervan te leren, of zelfs om het beleid bij te sturen, ligt het veel meer voor de hand om dit al vanaf het begin, *vooraf* (*ex ante*), of *gedurende* (*ex durante*), de beleidsperiode te doen. Na afloop van een beleidscyclus is er immers geen beleid meer om bij te sturen. Tegelijk verschuift in zekere zin ook de vormgeving van beleid; in de context van complexiteit is beleid immers ‘nooit af’. Na een beleidsinterventie volgt altijd een volgende stap, die middels een gestructureerd leerproces kan voortbouwen op de lessen van het voorgaande. Deze verschuiving in timing heeft gevolgen voor verschillende elementen van evaluatieonderzoek, waaronder de focus en de evaluatieprogrammering.

Eindopbrengsten én tussenopbrengsten

Evaluatieonderzoek dat rekening pogt te houden met complexe beleidscontexten is veel meer gericht op *tussenopbrengsten* dan op *eindopbrengsten*. Dit is een groot verschil met klassiek evaluatieonderzoek, dat typisch

ex post plaatsvindt en daarom per definitie gericht is op eindresultaten. In de Lerende Evaluatie van het Natuurpact zien we juist een focus op *tussenopbrengsten*. Het doel van de evaluatie is specifiek om lessen te trekken uit het huidige beleid, zodat er gaandeweg bijgestuurd kan worden in de aanpak. Om die reden is er in de lerende evaluatie gekozen voor een speciaal ontwikkelde methode waarin leren het zwaartepunt is.

Transitiebeleid

Dit sluit aan op *transitiebeleid*, dat doorgaans wel (tussen)doelen kent, maar deze zijn bij aanvang vaak abstract en kwalitatief geformuleerd, als stip op de horizon voor een maatschappelijk proces, gericht op de lange termijn. Dit type doelstellingen is vaak moeilijk meetbaar en laat zich dan ook lastig vangen in klassieke beleidsevaluaties. Binnen de context van complexiteit zijn ze juist het focuspunt. Het concretiseren van doelen gebeurt gaandeweg. Het is veelal een uitkomst van processen van onderhandeling, strijd en soms zelfs conflicten, waarna samen optrekken en uitproberen uitkomsten biedt. Sturing van complexe systemen volgt zelden een vooropgesteld plan, maar is juist de uitkomst van adaptatie en bijsturing, op grond van continue waarneming en monitoring. Beleid in de context van complexiteit doorloopt als het ware telkens een cyclus van twee verschillende fasen: in de eerste fase wordt het systeem of worden huidige structuren ter discussie gesteld, waarna er in een tweede fase nieuwe structuren ontstaan en er allerlei nieuwe regels worden geformuleerd. De eerste fase is daarom niet geschikt om te evalueren; het is dan nog te vroeg om te zeggen of doelen zijn bereikt of niet. Doelen en middelen lopen in complexe contexten bovendien regelmatig in elkaar over. Of het doel verschilt per jaar omdat er geen overeenstemming over is. Soms zit achter een doel ook een ander doel, een zogenaamd *proxy-doel*, waarbij een aansprekend doel wordt geformuleerd, teneinde iets anders te bereiken. Een voorbeeld daarvan is het bewerkstelligen van zalm in de Rijn. Dat is niet het ultieme doel, maar het is wel een indicator voor een ecologisch schonere rivier, te bereiken door een beleidsmix van het tegengaan van vervuiling en ecologisch beheer. Bij dergelijke complexe contexten past bij uitstek een *ex durante* evaluatie-aanpak.

Er gaat in complexiteitserkennende evaluaties ook meer aandacht uit naar het startpunt van beleid. Bij de Directie IOB hecht men bijvoorbeeld veel waarde aan dit *ex ante* perspectief; de directie zou graag meer aandacht willen besteden aan het vooraf goed beschrijven wat er aan de hand is in complexe beleidscontexten. Dit stelt in staat om meer vanuit het vraag-

stuk te denken, in plaats van vanuit de ‘toolbox’ van standaardevaluatiemethoden die voorhanden is. Vanuit die focus kom je vermoedelijk op andere evaluatiestrategieën uit, die beter rekening houden met situationele en contextuele factoren van beleid. Dit vraagt om sensitiviteit voor de empirische situatie die voorhanden is. Een dergelijke focus stelt je daardoor in staat flexibeler om te gaan met complexe beleidsevaluaties.

Evaluatieprogrammering

Evaluatieprogrammeringen die pogen recht te doen aan complexe beleidscontexten zijn idealiter gericht op verbeterprocessen die *ongoing* zijn. Een dergelijke aanpak zien we terug bij de Directie 106 die de eigen beleidsterreinen periodiek en doorlopend evalueert, aan de hand van een meerjarige evaluatieprogrammering, waarin de meeste beleidsevaluaties een vast onderdeel vormen. Hierdoor worden evaluaties van beleidsprogramma’s automatisch breder in opzet en minder projectmatig. De evaluatie van het Natuurpact is in haar opzet ook periodiek en langetermijngericht: het beleid zal driejaarlijks worden geëvalueerd tot aan 2027.

- **Attributie, maar ook contributie**

“Een standaardopbrengst van evaluaties met betrekking tot de attributievraag is dat de conclusie altijd luidt: het was niet mogelijk geweest dit ook met minder middelen te doen. Evaluatie is zo eigenlijk een rechtvaardiging van het beleid in plaats van om ervan te leren.”³⁷

Focus

In een beleidsevaluatie onderzoekt de evaluator of de empirische praktijk zich heeft ontwikkeld volgens de verwachtingen die vooraf aan het beleid gesteld waren: zijn de verwachte effecten opgetreden en waren deze er zonder het beleid niet geweest? Daarin is de *attributievraag* van belang: zijn gevonden effecten een direct gevolg van een aanwijsbare concrete interventie, of zijn effecten mogelijk het gevolg van andere factoren? Deze ontwikkellijn manifesteert zich op het niveau van deze *vraag*. We zien een verschuiving in evaluatieonderzoek optreden van een eenzijdige focus op *attributie* naar een breder perspectief, waarin ook aandacht is ontstaan voor *contributie*: heeft het geheel van beleidsactiviteiten bijgedragen aan de richting waarin het veld zich ontwikkelt? Het principe van contributie legt minder nadruk op geïsoleerde specifieke actoren of maatregelen, die een eveneens te isoleren effect tot gevolg hebben gehad. Als een interventie bijdraagt aan de *richting* is dat vanuit dit perspectief een relevant resultaat.

Naarmate beleidsvraagstukken complexer worden, doelen en middelen, oorzaken en gevolgen door elkaar gaan lopen en de inzet van verschillende actoren al dat niet betrokken bij beleid effect kunnen sorteren, omdat zij allemaal (bedoeld of onbedoeld, gericht of ongericht) invloed hebben op het vraagstuk, is het niet altijd voorspelbaar wat er precies gaat gebeuren. Als gevolg daarvan zijn resultaten niet zo gemakkelijk toe te schrijven aan preciezerbare actoren. De beleidswereld rondom het vraagstuk is überhaupt niet zo makkelijk in kaart te brengen. Deze ontwikkeling vraagt daarom om een ander blikveld, en in het verlengde daarvan ook om andere meetmethodes die in aanvulling op de standaardmethodes kunnen worden gebruikt, zodat andere opbrengsten – ‘bijvangsten’ in het attributieperspectief – ook gemeten worden. Dit omdat de onverwachte of onbedoelde effecten en bijvangsten soms interessantere uitkomsten zijn dan de geplande of verwachte uitkomsten.

Blikveld

In complexe systemen is het lastig om attributie aan te tonen, omdat complexe systemen veel meer gekenmerkt worden als het summum van verschillende componenten en hun onderlinge interacties. In complexe systemen reageren elementen in het systeem op elkaar en ontstaan als gevolg van die interacties nieuwe verhoudingen en relaties. Effecten zijn het gevolg van wisselwerking en daarmee per definitie moeilijk te herleiden tot een enkelvoudig beleid van één actor. Het is in deze reflexiviteit daarom lastig om te zien of dynamieken in een systeem een gevolg zijn van specifieke interventies, of dat de in gang gezette beweging ook zonder de maatregel tot stand was gekomen; ‘opbrengsten’ of ‘resultaten’ zijn lastig toe te schrijven aan specifieke actoren.

Beleidsmakers weten vaak niet of beleid effectief is, of de bedoelde toegevoegde waarde er is en of die waarde te herleiden is naar de gepleegde inspanningen. Er wordt in evaluatieonderzoek daarom gekeken naar bredere effecten die bereikt zijn en daarom worden er ook methoden gebruikt om bredere resultaten in kaart te brengen die zich maatschappijbreed manifesteren. In klassiek evaluatieonderzoek zouden dit eerder ‘bijvangsten’ zijn, terwijl ze in de nieuwe stijl veel meer de kern van het blikveld vormen. Dit past ook bij de trend dat andere partijen dan de overheid, waaronder ook bedrijven, maatschappelijke organisaties en burgers, in toenemende mate betrokken zijn bij het realiseren van maatschappelijke meerwaarde. Het past niet om dan alleen naar de overheidsinzet te kijken; toch is dat wel de kern van klassiek evaluatieonderzoek dat vertrekt

vanuit het startpunt van de attributievraag: in hoeverre zijn resultaten toe te schrijven aan het beleid – de inzet van overheidsinstanties gefinancierd uit publieke middelen.

Dit werpt wel de vraag op van wie het beleid eigenlijk is en wat evaluatie eigenlijk betekent; de essentie van netwerkende manieren van werken, typisch voor complexe beleidsvraagstukken, is immers dat beleid in samenspel met andere actoren tot stand komt en dat het eigenaarschap van sturing dus gedeeld is. De verschuiving van *attributie* naar *contributie* is een manier om hiermee om te gaan en de ontwikkeling komt deels ook voort uit een wens naar onzekerheidsreductie. De onzekerheid rond transitievraagstukken, die vaak gericht zijn op de lange termijn, moet gaandeweg nog uitkristalliseren in concrete(re) doelen en middelen. Dit brengt een informatietekort met zich mee, terwijl informatie wel nodig is voor de bepaling van effecten van het beleid.

Een ruimer blikveld heeft ook consequenties voor de inzet van middelen voor het evalueren van beleid. De vraag is of de overheid verantwoordelijk is voor het evalueren van de inzet van anderen. Vertrekkend vanuit de attributievraag is het antwoord eenduidig – nee, dat is ze niet – maar vanuit een breder perspectief over contributie ligt het wel weer voor de hand om ook de inzet van anderen (bedrijven, burgers) te evalueren, zodat zinvolle lessen getrokken kunnen worden op basis van de kennis en expertise ook van anderen. We zien de blikverruiming ook terug in de criteria van evaluatie, die bovendien over een (veel) langer tijdspad beoordeeld moeten worden. Zo heeft de Directie IOB naast doelmatigheid, doeltreffendheid, relevantie en coherentie, ook duurzaamheid – de mate waarin een effect van een activiteit blijvend is – specifiek als evaluatiecriterium opgenomen. Dit illustreert dat er niet alleen aandacht is voor de directe effecten van beleid, maar ook in hoeverre er daadwerkelijk impact gerealiseerd wordt in een breder perspectief. Dit zien we ook terug bij het Topsectorenbeleid; in de evaluatie is rekening gehouden met een aantal elementen waardoor dit beleid gekenmerkt wordt; de evaluatie onderkent dat er sprake is van *transformatief* (innovatie maakt onderdeel uit van de beoogde strategie), *adaptief* (aanpassingsgericht) en *multi-instrumenteel* (op basis van een beleidsmix) beleid.

Aanvullende meetmethodes

Dit heeft vanzelfsprekend ook betekenis voor de gebruikte meetmethodes. In lijn met de ontwikkelingen van *public administration* en *new public management* naar *network governance* en *societal resilience*, waarbij de *compliance audits*

(zijn de procedures juist gevolgd?) en *performance audits* (zijn de doelen bereikt?) plaats hebben gemaakt voor *public value audits* (wordt er publieke meerwaarde gecreëerd?). Hierbij passen alternatieve meetmethodes die focussen op andere zaken dan droge uitkomsten van beleid. Het gaat veel meer om *outcomes* dan om *output*.

Zo is de evaluatiepraktijk van de IOB de laatste jaren verschoven van doorlichter van projecten naar evaluator van grotere programma's en brede beleidsopgaven, vaak ook met lange termijn doelstellingen. Dit heeft de evaluatiepraktijk complexer gemaakt, waarvoor traditionele evaluatiemethoden minder geschikt zijn. Hoewel de attributievraag – die de directie specificeert als: kunnen de waargenomen effecten toegeschreven worden aan het beleid? – voor de IOB centraal staat in al het evaluatieonderzoek en zelfs als standaard uitgangspunt wordt gehanteerd, suggereren de gehanteerde methoden toch ook dat er in toenemende mate interesse is ontstaan voor alternatieve of overige opbrengsten, die wat minder goed meetbaar zijn met traditionele methoden (zoals een kosten-batenanalyse of kosten-effectenanalyse). Te denken valt aan procesevaluaties (met een focus op 'pathways', bij gebrek aan causale relaties) of impactstudies en tevredenheidsonderzoeken. De resultaten die hiermee gemeten worden brengen niet noodzakelijkerwijs of uitsluitend in beeld wat opbrengsten zijn die toegeschreven kunnen worden aan overheidsinzet; ze veranderen daarmee de attributievraag. Dit sluit aan bij een ander principe dat deel uitmaakt van denken vanuit complexiteit: het feit dat je geen effecten meet, betekent niet dat ze er niet zijn. Dit geeft ook een relativering van de zekerheden die een evaluatie zou kunnen bieden; evaluaties van complexe vraagstukken draaien veel meer om de vraag wat je naar alle waarschijnlijkheid kunt zeggen, dan om wat je met zekerheid kunt vaststellen.

Ook wordt er meer gebruik gemaakt van *mixed methods*: zowel kwantitatieve als kwalitatieve methoden vormen deel van evaluaties van complexe beleidsvraagstukken. In de Lerende Evaluatie van het Natuurpact zijn naast kwantitatieve gegevens ook resultaten uit casestudy's en interviews (kwalitatieve meetmethodes) gebruikt. Het evaluatieonderzoek was in hoge mate gericht op interacties tussen verschillende actoren; het evaluatiekader dat gebruikt werd is bijvoorbeeld al gezamenlijk ontwikkeld middels interviews en werksessies. Doelen van de evaluatie, de afbakening en onderzoeksvragen zijn zo bijvoorbeeld samen met stakeholders bepaald. Het verzamelen van gegevens gebeurde eveneens in gezamenlijkheid, namelijk op basis van groepsgesprekken, casestudy's, bilaterale consultaties

en interviews. Daarnaast werd er op basis van overkoepelende werksessies en groepsreviews betekenis gegeven aan onderzoeksbevindingen, met veel aandacht voor praktijkervaringen van direct betrokken en in onderlinge interactie. Ten slotte werden resultaten van de evaluatie in leersessies gepresenteerd.

Voor de Topsectorenaanpak geldt dat er een evaluatiekader is gebruikt dat gericht is op beleidsnoodzaak en beleidsimpact. Omdat beleidsdoelen heel abstract waren en dus lastig om te evalueren, is er meer gekeken naar het ontwerp van beleid en of de samenwerking goed was opgezet. Vervolgens is er gereflecteerd op de vraag welke afspraken er zijn gemaakt en of deze worden nageleefd. Ook is er gekeken naar belemmeringen van systeemveranderingen, waaronder structurele veranderingen zoals economische activiteiten die actoren ontplooiën. Hiervoor is de vraag beantwoord wie er voorafgaand aan het beleid met elkaar samenwerkte en wie niet. Gebruikte methoden om data te verzamelen zijn ook gemengd: zowel deskresearch als ook interviews met betrokken stakeholders hebben plaatsgevonden.

Toch is nooit hard te maken dat deze resultaten (uitsluitend) zijn toe te schrijven aan de aanpak. Het blijft lastig om abstracte brede opgaves echt goed meetbaar te maken; hoe meet je zaken als maatschappelijke ontwikkeling, brede welvaart of democratische betrokkenheid? Wat zijn hiervoor zinvolle indicatoren (direct of indirect) en wat zijn daar dan handreikingen voor? Tegelijkertijd moet er gewaakt worden voor complexiteit als dekmantel, om verantwoording uit de weg te gaan. Als over doelen en middelen valt te twisten, kan er namelijk onduidelijkheid ontstaan. Is een doelstelling een norm die behaald moet worden, of slechts een streefdoel? Gaat het om een katalyserende doelstelling en om richting, of zijn er wel degelijk iets concretere doelstellingen in de opgavegerichte benadering te onderscheiden?

- **Een onafhankelijk onderzoeker op afstand, maar ook betrokken**
“De vraag of je afstand moet hebben tussen beleidsvoorbereiding en evaluatie komt neer op de vraag: wil je effect hebben of niet?”³⁸

Rol van de onderzoeker

Ontwikkelingen in de evaluatiepraktijk hebben invloed op de rol en positie van evaluatoren – de onderzoekers die het evaluatieonderzoek uitvoeren. Bovenstaande quote drukt dat uit: of hij nu gelijk heeft of niet, de geïnterviewde drukt uit dat hij nabijheid zoekt bij het proces van beleidsvorming en hij ziet dat als een verandering met ‘vroeger’. Blijkbaar is er een ervaren

dilemma in de mate waarin het evaluatieonderzoek afstand moet houden van de beleidspraktijk, of daar juist dicht bij moet opereren.

We zien een voorzichtige verschuiving in de houding en insteek van onderzoekers. Klassiek evaluatieonderzoek heeft een sterk wetenschappelijk ingebedde insteek, waarin waarden als objectiviteit, onafhankelijkheid en distantie tot het studieobject hoog in het vaandel staan. De evaluatie gaat 'over' het beleid, maar staat verder 'buiten' het beleidsproces. Het ontbreken van betrokkenheid bij het gehele proces heeft volgens een gesprekspartner echter als keerzijde dat evaluatoren op één meetmoment langskomen om de thermometer in een beleidsproces te steken en te meten hoe het er op dat moment voor staat, terwijl complexe beleidsvraagstukken dynamische beleidsprocessen kennen, die constant in ontwikkeling zijn.³⁹ Een enkel meetmoment zegt niet zoveel voor beleid dat constant wordt bijgesteld, en kan ook niet veel meerwaarde opleveren in termen van leereffecten of het bijstellen van beleid, zeker niet als dit meetmoment pas laat in of na afloop van de beleidscyclus plaatsvindt. Überhaupt is het een vraag of er in het geval van complexiteit daadwerkelijk sprake is van een beleidscyclus; in veel gevallen is eerder sprake van een beleidscontinuüm. Dit pleit voor evaluatieonderzoek met *ongoing* en actieve betrokkenheid van evaluatoren, maar dit vraagt tegelijkertijd om een scherpe blik op het risico dat de evaluator zijn onafhankelijke positie in de loop van het proces verliest. De vraag om meer betrokkenheid vertaalt zich zowel op het niveau van de individuele onderzoeker (wat voor type onderzoeker voert het evaluatieonderzoek uit?) alsook op de evaluatiegemeenschap (wie zijn erbij betrokken?), omdat betrokkenheid in een complexiteitsgerichte evaluatiestijl vraagt om interactie en samenwerking met anderen. Althans, dat is de spanning die gesprekspartners benoemen: zij ervaren een noodzaak tot meer nabijheid en betrokkenheid bij het bredere beleidsproces, omdat ze het gevoel hebben dat dat hun evaluatie verbetert. Aan de andere kant zien ze ook dat juist afstand tot het vaak politieke en grillige beleidsproces een belangrijke waarde is voor hun evaluatie. Sommigen maken dat expliciet, door aan te geven dat ze zoeken naar nabijheid met behoud van distantie. Het is interessant om nader te verkennen hoe dat er uit ziet.

Type onderzoeker

Het type onderzoeker dat evaluaties van complex beleid uitvoert is bij een lerende evaluatie idealiter nauw betrokken bij het object van analyse, zodat de complexe context optimaal kan worden meegenomen in het onderzoek. Goed op de hoogte zijn is van belang om complexe contexten

te doorgronden en vraagt om een bepaalde mate van situationele kennis en begrip. Hierdoor heeft de zogenaamde *boundary researcher* zijn intrede gedaan in het evaluatieonderzoek: de onderzoeker die zelf wat dichterbij de beleidsmaker zit en als het ware op de grens tussen wetenschap en beleidspraktijk opereert.

Onderzoek heeft aangetoond dat het gros aan evaluaties in de Nederlandse beleidspraktijk tot nu toe wordt uitbesteed aan onderzoeksbureaus, wetenschappelijke instituten en universiteiten.⁴⁰ Dat past goed bij de gedachte dat evaluatieonderzoekers ver van de beleidspraktijk af zouden moeten zitten. Het IOB is in feite ook een onafhankelijke evaluatiedienst van het ministerie van Buitenlandse Zaken. Daarnaast zijn zowel het Natuurpact als het Topsectorenbeleid geëvalueerd door externe onderzoekers, namelijk respectievelijk het Planbureau voor de Leefomgeving (PBL) in samenwerking met Wageningen University en Research (WUR) en Dialogic. Toch is er ook een ontwikkeling gaande waarin de onderzoeker idealiter juist wat dichterbij de beleidspraktijk opereert, bijvoorbeeld doordat er meer betrokkenheid gevraagd wordt bij het onderwerp van studie. Dit vraagt niet alleen om een andere rol van de onderzoeker – namelijk meer betrokken – maar ook om een ander type onderzoeker: een die goed op de hoogte is van wat er zich in het beleidsproces allemaal voordoet.

De onafhankelijkheid van de onderzoeker moet hierdoor wel op een andere manier worden ingericht en ook is het zoeken naar de optimale rol van de evaluator: dat is idealiter iemand die op zoek is naar verschillende bronnen van kennis, zowel in de wetenschap als in de buitenwereld, maar die ook een bescheiden houding heeft ten aanzien van de kennis en informatie waarover hij beschikt; niet iemand die komt vertellen 'hoe het zit', maar iemand die zich bescheidener opstelt en tegelijkertijd ook fundamenteel onderzoeker is, en dus niet op de stoel van de beleidsmaker probeert te zitten. Hier komen uiteraard ook dilemma's om de hoek kijken. Rollen van belanghebbenden en evaluatoren kunnen door elkaar lopen: wie legt verantwoording af aan wie? Dit vraagt om duidelijke en heldere grenzen tussen belanghebbenden en evaluatoren; resultaten moeten niet afhankelijk zijn van de mening van belanghebbenden, maar wel echt objectief zijn. Vraag is ook of resultaten van een evaluatieonderzoek nog zo gemakkelijk geaccepteerd worden, zowel door de politiek alsook door alle andere stakeholders als er onvoldoende distantie is tussen onderzoeker en onderzoeksobject om een bepaalde mate van onafhankelijkheid en objectiviteit te

waarborgen. Het verwijt ‘de slager keurt zijn eigen vlees’ is gauw gemaakt als er een te nauwe betrokkenheid is van de evaluator. Daarom blijft onafhankelijkheid onverminderd van belang, maar wel in combinatie met een betrokken houding.

De evaluatiegemeenschap

Een ander centraal kenmerk in complexe beleidsaanpakken is dat de governance rondom een vraagstuk tot stand komt doordat relevante partijen gezamenlijk tot een gedeelde beweging komen, die tot publieke waarde moet leiden. Typisch daarbij is een mix van interventies, die bij verschillende partijen belegd zijn. De overheid gaat daarbij op zoek naar partners van buiten om mee samen te werken en actoren werken daarbij steeds vaker in netwerkende strategieën en arrangementen.

Er ontplooiën zich momenteel allerlei initiatieven rondom evaluatieonderzoek waarbij verschillende betrokken actoren bijeen worden gebracht, waaronder zowel onderzoekers uit de wetenschappelijke wereld alsook beleidsmedewerkers van departementen om in zogenaamde ‘communities of practice’ gezamenlijk ervaringen met beleidsevaluatie uit te wisselen. Het Planbureau voor de Leefomgeving (PBL) experimenteert momenteel met dergelijke structuren, waarin bevindingen ook cross-sectoraal worden uitgewisseld; ervaringen met evaluatieonderzoek in zeven verschillende sectoren worden in communities of practice met elkaar gedeeld. De Directie IOB werkt op zijn beurt met zogenaamde referentiegroepen om evaluaties voor te bereiden. Dit zijn groepen waarin zowel beleidsmakers als wetenschappers plaatsnemen en afstemmen over de zogenaamde ‘points of reference’ van het voorgenomen evaluatieonderzoek.

Omdat beleidsevaluatie wordt afgestemd op de specifieke omstandigheden en context, zodat er een optimale aansluiting is, zien we ook steeds bredere evaluatiegemeenschappen ontstaan, bestaande uit een veelheid van actoren. In aanloop naar het Natuurpact zijn op verschillende plekken in Nederland bijvoorbeeld lokale en/of regionale samenwerkingsverbanden ontstaan waarin brede coalities van burgers, ondernemers en maatschappelijke organisaties zich verenigen en gezamenlijk inzetten om resultaten te behalen. Om resultaten van een dergelijke netwerkaanpak goed in kaart te brengen, is het zaak om aansluiting te vinden bij dergelijke groepen, en visies van alle betrokken stakeholders mee te nemen in het evaluatieonderzoek. Dit om voldoende zicht te hebben op een diverse set aan op-

brenghen, gerealiseerd door een veelheid aan betrokkenen. De betrokkenheid van de evaluator bij dergelijke initiatieven resulteert zo automatisch in een evaluatiegemeenschap. In de Lerende Evaluatie van het Natuurpact was dan ook veel oog voor interactie met stakeholders, die zich op allerlei ‘niveaus’ manifesteerden: het Rijk, de provincies en maatschappelijke organisaties. Input van betrokkenen werd uitvoerig uitgewisseld en zoveel mogelijk meegenomen in de lessen die als opbrengst van de evaluatie zijn gepresenteerd.⁴¹

Een vergelijkbaar uitgangspunt geldt voor de evaluatie van de Topsectoren-aanpak; dit beleid stimuleert samenwerking tussen bedrijven, publieke onderzoeksinstituten en de overheid. Om die reden wordt er ook gesproken van een Triple Helix samenwerking en wordt er stevig ingezet op samenwerking, het verstevigen van relaties, het wegnemen van barrières tussen verschillende actoren en het creëren van structuren waarin collectief en integraal ontwikkelingspaden worden gecreëerd. Om inzicht te verkrijgen in dergelijke doelstellingen en deze op een zinvolle manier te evalueren, is betrokkenheid bij en verbondenheid met het netwerk van actoren vereist.

- **Evalueren voor politiek en bestuur, maar ook voor de ambtelijke praktijk**

“Het taboe op fouten maken staat in de weg van het lerend vermogen van de overheid. Als je in het bedrijfsleven tien dingen uitprobeert en één werkt, dan ben je een succesvol ondernemer. Bij de overheid geldt het omgekeerde: één fout en je ligt eruit als bewindspersoon.”⁴²

Motivatie

Een andere verschuiving in evaluatieonderzoek komt voort uit een bijgestelde motivatie. Deze vijfde ontwikkelijn komt neer op het ontstaan van een nieuwe motivatie voor evaluatieonderzoek, namelijk het verbeteren van de (eigen) werkpraktijk. Naast de politieke en bestuurlijke logica's van evaluatie, is daarin ook aandacht voor de ambtelijke praktijk.

Politieke logica

De politieke logica komt terug in de politieke motivatie voor evaluatieonderzoek. Politici gebruiken evaluaties om hun controlerende functie uit te voeren en te weten te komen waaraan publieke middelen besteed zijn, of dat zinvol was, wat opbrengsten waren en in hoeverre deze zijn toe te schrijven aan beleid. De politieke logica schuilt ook in het instrumenteel

maken van een evaluatie aan politieke doelen. Zo kan met de opzet van een evaluatie (termijn, indicatoren, afbakening, etc.) gestuurd worden naar een politiek gewenste uitkomst.

Bestuurlijke logica

De bestuurlijke logica kent een instrumentelere en strategischere grondslag, hoewel die in het verlengde ligt van de politieke. Bewindspersonen en bestuurders willen daarbij inzicht in resultaten die bereikt zijn, zodat zij enerzijds op de hoogte zijn van wat er gebeurt, maar anderzijds ook weten welke bijdrage zij hebben geleverd gedurende hun ambtstermijn. De achtergrond van de evaluatie hangt hier vooral samen met een vorm van 'good governance': een bestuurlijke wens om het beleid zo effectief en doelmatig mogelijk in te richten.

Ambtelijke logica

Het leidend motief van evaluatieonderzoek in een complexe beleidscontext lijkt daarnaast ook een veel praktischer aspect in zich te hebben: de vraag wat nuttige uitkomsten zijn van evaluatieonderzoek opdat de ambtelijke organisatie ermee aan het werk kan om beleidsprocessen bij te sturen. De bedoeling van de evaluatie is er veel meer in gelegen dat er iets van opgestoken wordt. Het is in die zin minder van belang of er positieve of negatieve resultaten bereikt zijn; het is al waardevol als er in elk geval iets van geleerd wordt, zodat aanpassingen gemaakt kunnen worden gedurende de beleidscyclus, zodat er in ieder geval niet onnodig lang wordt doorgegaan met aanpakken die geen vruchten afwerpen. Ook is de termijn waarin resultaten geboekt worden minder van belang. Als er bijgedragen wordt aan langetermijndoelen is dat een prima resultaat, ook als dat resultaat pas na verloop van lange tijd geogst kan worden. Deze ambtelijke logica is fundamenteel anders dan de politieke, waarin de nadruk ligt op resultaten binnen een electorale termijn. Er is daarom ook een groot verschil met klassiek evaluatieonderzoek waarin deze politieke en bestuurlijke logica's dominant zijn. Daarin is het juist van belang om zo snel mogelijk resultaten te behalen, zowel om politici tevreden te stellen als om resultaten toe te kunnen schrijven aan bewindspersonen, die dit als een belangrijk middel zien om hun prestaties aan te tonen.

Een voorbeeld waarin we deze praktisch ingestoken ambtelijke benadering terugzien is de evaluatie van het Topsectorenbeleid. Het evaluatiekader bestaat uit een *impactraamwerk* en een *ontwerpraamwerk*, die beide tot doel hebben om een oordeel te vellen over de beleidsimpact en het beleidsont-

werp en tegelijkertijd hands-on aanbevelingen te doen die het beleid in de toekomst moeten versterken. Opvallend daarbij is dat de aanpak niet zo zeer resultaatgericht is, als wel gericht op impact. Ook wordt er heel specifiek aandacht besteed aan beleidsontwerp; dat past binnen een ambtelijke logica. De vraag die centraal staat is niet zo zeer in hoeverre bepaalde doelstellingen daadwerkelijk gerealiseerd zijn, maar eerder in hoeverre de aanpak als zodanig geschikt is als instrument om deze doelstellingen te bereiken. Concreet richt de evaluatie zich op de vraag in hoeverre het Topsectorenbeleid partijen bijeen kan brengen om draagvlak te creëren voor het collectief uitwerken van een innovatiekoers. Het raamwerk richt zich daarom op verschillende elementen: beleidsorganisatie (waaronder sturingsvermogen, accountability, transparantie en leiderschapscompetenties) en beleidsoriëntatie (waaronder grensverleggende vernieuwing, openheid, het creëren van spilovers en inputs in plaats van subsidies). Eén van de uitkomsten van de evaluatie is bovendien dat er aanbevelingen worden gedaan voor verbeteringen die de beleidsaanpak kunnen versterken, ondanks dat dit geen expliciete doelstellingen waren van het beleid.

5 Inbedding: opgavegericht evalueren vereist bewust organiseren

Weten is nog geen doen: waarom anders evalueren moeilijk gaat

De analyse van de beschikbare methoden en technieken voor complexiteitsgerichte evaluatie en het beeld van praktijkervaringen laat zien dat de worsteling met op een andere manier evalueren slechts ten dele een inhoudelijke kwestie is. Het is niet zo dat er geen enkele methode voor evaluatie beschikbaar is, of dat er geen enkele ervaring mee is opgedaan. Integendeel, er wordt al mee gewerkt. De kennis van het probleem is aanwezig en dat geldt ook voor mogelijke oplossingsrichtingen. Dat roept de vraag op waarom de methoden slechts beperkt gebruikt worden en geïnterviewden aangeven dat het toch lastig blijkt om dergelijke methoden te benutten. Het *zeggen dat het anders zou moeten* betekent nog niet dat het ook daadwerkelijk anders gaat. Of, met een knipoog naar het WRR-rapport (2017) over eigen kracht van burgers: *weten is nog geen doen*. Zelfs de praktijkmensen die aangeven op een andere manier te willen evalueren geven ook aan dat het in de praktijk lastig blijkt. Daarom is het nodig om te onderzoeken of de omgeving en inbedding waarin evalueren plaatsvindt wel passen bij een meer opgavegerichte benadering.

Om deze vraag te beantwoorden is het belangrijk om eerst de lens scherp te stellen van waaruit we kijken naar beleidsontwikkeling en beleidsevaluatie: hoe komt beleid tot stand en welke rol speelt evaluatie daarin? Daarvoor zijn verschillende perspectieven beschikbaar, zoals onder andere uitgewerkt in een klassieke studie naar de besluitvorming rond de Cubaanse raketten-crisis (1962).⁴³ In “*Explaining the Cuban Missile Crisis*” zien we dat je vanuit verschillende perspectieven naar een besluitvormingsproces in een crisis kunt kijken en dat vanuit die verschillende perspectieven heel andere zaken opvallen. Drie perspectieven staan centraal: (a) een rationeel perspectief, (b) een politiek perspectief en (c) een institutioneel perspectief.

Het rationele perspectief veronderstelt dat beleid en besluitvorming het product zijn van het verzamelen, analyseren en wegen van informatie. Het besluitvormingsproces volgt een cognitief-rationele logica, waarin feiten objectief gewogen worden. De feiten leiden hier de besluitvorming. Als de feiten op tafel liggen dan nemen bestuurders de besluiten die de feiten voorschrijven. Daarin is uiteraard een zekere bandbreedte mogelijk, maar uiteindelijk gaat het om objectieve afwegingen die een cognitieve rationaliteit van besluitvorming kennen. Goed beleid is dan evidence-based, of in ieder geval evidence-driven. Bij het inrichten van beleidsprocessen gaat het erom te zorgen voor een voldoende goed gevulde informatiebasis en het vermogen om die informatie te vertalen in een analyse van het probleem en concrete oplossingen.

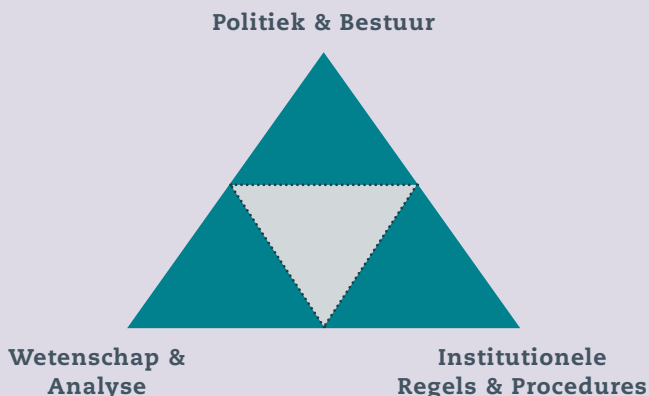
In het *politieke perspectief* op besluitvorming is de dynamiek anders. Hier is besluitvorming eerst en vooral een gevolg van politieke afwegingen en van politieke dynamiek. Dat gaat om macht en belangen die tegenover elkaar komen te staan en om het vermogen om krachtige coalities van macht en belangen te sluiten. Een besluit volgt als er voldoende politieke macht voor is gemobiliseerd en als het lukt om weerstand van andere belanghebbenden te minimaliseren. Hierbij wordt besluitvorming vaak voorgesteld als een 'barrière-model', waarin om tot besluiten te komen steeds een aantal barrières overwonnen moet worden, zoals in een hordenloop. Als er voldoende politieke steun voor een besluit is, kan het besluit de volgende barrière over en door naar de volgende fase: eerst agendering, dan een besluit over de analyse van het probleem, dan steun voor een bepaalde oplossing, dan een besluit, gevolgd door steun voor een implementatiestrategie en tenslotte voldoende steun voor een evaluatie en het gebruik daarvan. In elk van die stappen moet een beleidsmaker steun verwerven, waarbij inhoud en feiten een rol spelen, maar instrumenten in het proces van politieke draagvlakontwikkeling – of bij het uitschakelen van politieke tegenstand – ook. Feiten doen er nog steeds toe, maar als één van de grondstoffen voor het politieke proces, naast andere grondstoffen.

In het *institutionele perspectief* zijn regels, protocollen, procedures en 'rules of the game' bepalend voor wat er in besluitvorming gebeurt. Er is kennis en er is politiek, maar die krijgen vorm binnen de kaders en context die institutionele patronen daarvoor bieden. Dit is fraai verwoord in de ondertitel van het boek *Rediscovering institutions: how decisions happen*.⁴⁴

Ze beschrijven hoe besluiten letterlijk ‘gebeuren’, omdat er in institutionele patronen, regels en kaders momenten zijn ingebouwd waarop een besluit ‘moet’ en ‘hoort’. Dit perspectief zet voor het begrijpen van besluitvorming en beleidsvorming de schijnwerper op de institutionele hardware van de organisatie die bepaalt – letterlijk – hoe de hazen binnen een besluitvormingsproces lopen. Zo hebben veel organisaties regels over het soort informatie en pre-advies dat zij moeten inwinnen alvorens een besluit genomen ‘mag’ worden; die procedures lijken niet meer te zijn dan context waarbinnen ‘alles kan gebeuren’, maar in de praktijk zijn ze heel sturend voor hoe processen verlopen. Het protocol schrijft voor welke analyse op welk moment plaatsvindt, en wat daarvan de rol en positie is.

Beleidsmakers en ook bestuurders die besluiten nemen denken weliswaar dat zij controle hebben over die procedures – dat ze er boven staan – maar zijn in de praktijk volgend aan het institutionele patroon. Ze doen waar het institutionele patroon naartoe leidt en blijven daardoor ook per definitie binnen het bestaande: het via institutionele regels, kaders en protocollen uitgezette pad. Binnen dergelijke processen vragen betrokkenen zich bij grote en kleine situaties steeds af⁴⁵: “*wat doet iemand zoals ik in een situatie zoals deze?*”. Daarvoor gebruiken ze institutionele kaders, patronen, protocollen, regels en gewoonten als aanwijzingen. Het protocol schrijft voor dat een evaluatie de ‘Theeuwes-norm’ moet halen. ‘Dus’ kiezen actoren in situaties consequent voor instrumenten waarvan zij denken – op basis van wat regels, kaders en patronen in hun omgeving voorschrijven – dat deze passen bij de Theeuwes-norm. Zo reproduceert zich een bepaalde vorm van handelen, die zichzelf bovendien versterkt: door bij herhaling volgens de regel te handelen verdiept de idee dat de regel van kracht is.

Als we de drie perspectieven bij elkaar nemen dan ontstaan drie logica’s voor het begrijpen van processen van besluitvorming en beleidsvorming (zie [figuur 4](#)). Deze logica’s kunnen we vervolgens gebruiken voor het beantwoorden van de vraag hoe de praktijk van evaluatie meer rekening kan houden met complexiteit.



Figuur 4 Drie logica's van besluitvorming en beleidsvorming.

In de verschillende interviews die we hebben gevoerd viel op dat veel betrokkenen aangeven dat methoden *inhoudelijk* wel voorhanden zijn: het probleem is niet dat de methoden niet bestaan, of dat er geen deugdelijke alternatieven voor de meer traditionele manier van evalueren bestaan. Daar komt bij dat evaluatoren zich bewust zijn van complexiteit en in hun rapportages goed kunnen wijzen op de grenzen aan hun onderzoek en de keuzes die ze onderweg maken inzichtelijk kunnen maken.

Het probleem is ook niet direct politiek. Dat wordt weliswaar in veel interviews gezegd, “de politiek wil hier nog niet aan”, maar het is de vraag hoe we die opmerking moeten duiden. Immers, in een puur politiek model zal het de besluitvormer niet uitmaken hoe de kennis tot stand komt, als deze maar inzetbaar is om coalities te smeden en barrières te bedwingen. De kwaliteit van de kennis maakt vanuit een puur politiek perspectief op besluitvorming niet uit. De enige begrenzing is hier een politiek-ethische begrenzing, bijvoorbeeld of men bereid is te liegen of gebruik te maken van ‘fake news’ of ‘alternative facts’. Maar dat is niet de status van de alternatieve methoden voor evaluatie. Die vallen heel goed binnen de ethische bandbreedte van informatie die in een politiek proces als ‘fair game’ geldt. Hoe PBL het natuurbeleid evalueert zou bij puur politieke besluitvorming niet uit moeten maken: zolang er maar een rapport of onderzoek is waarmee partijen in de politieke strijd aan de slag kunnen. Evaluatie is hier niet meer dan een instrument – als dat instrument er is, in welke vorm dan ook – dan kunnen de politieke strijders er mee aan de slag. Dat geldt

voor het politieke domein van bestuurders en volksvertegenwoordigers, maar ook voor het domein van bureaupolitiek en de strijd om verdeling van schaarse middelen en relevantie binnen de organisatie, of tussen organisaties (bijvoorbeeld tussen directies of ministeries onderling).

Veel logischer is het om⁴⁶ dergelijke opmerkingen van geïnterviewden te plaatsen in het perspectief van institutionele regels en kaders. Hun observatie dat bestuurders om een bepaalde vorm van evaluatie vragen klopt waarschijnlijk, maar om een andere reden: in de institutionele patronen in de organisatie waarin beleidsvorming plaatsvindt wordt een bepaald soort evaluatie als 'passend' gezien en een afwijking van die standaard als 'zwak' en 'moeilijk uit te leggen'. Daarom is de vanuit dit perspectief logische vervolgstap dat actoren vragen om de methoden waarvan ze vooraf 'weten' dat die herkend worden als krachtig en valide. Daarom zoeken ze geen afwijkende vorm van evaluatie, volgens een 'andere', 'alternatieve' methode, maar zoeken ze juist naar de 'gewone methode', volgens de geëigende vorm, met het gebruikelijke en gekende soort analyse en product. De bestuurders en beleidsmakers in kwestie handelen niet vanuit een rationele of politieke preferentie, maar vanuit een institutionele logica: ze doen datgene waarvan ze denken dat het binnen de situatie waarin ze menen te verkeren past.

Diezelfde logica zien we terug in veel interviews: zo heeft het CPB bijvoorbeeld een krachtige positie in het institutionele patroon waarin beleid tot stand komt. Veel beleid moet op enig moment door het CPB 'bekrachtigd' worden en dat maakt dat bij het opstellen van beleid die normen ook gevolgd worden. Een beleidsevaluatie 'telt mee' als aan bepaalde voorwaarden voldaan wordt. Die voorwaarden zijn institutionele regels, die weliswaar veranderbaar zijn maar die tegelijkertijd wel dwingend voelen voor de betrokkenen.

Institutionele kaders en patronen kunnen soms 'hard' zijn, in die zin dat ze in dwingende organisatieprotocollen of procedures zijn vastgelegd, bijvoorbeeld in organisatieregels die ambtenaren in hun beleidsvoorbereiding moeten volgen. Institutionele kaders kunnen ook voortkomen uit informele regels, gewoonten, gebruiken, of in de organisatie gedeelde beelden over hoe het hoort en wat er zou moeten gebeuren. Van belang is hierbij ook de professionele achtergrond en het krachtig en dwingend institutioneel kader dat daaruit voortvloeit.⁴⁷ Als ambtenaren zich als econoom aangesproken voelen, zullen ze zich 'economisch gedragen': ze pakken economische

theorieën en modellen op, gebruiken economische begrippen en halen economen aan als gezag en autoriteit; ze hechten veel waarde aan het CPB en zien weinig waarde in methoden en theorieën uit andere wetenschapsgebieden. Andersom gebeurt overigens hetzelfde – in omgevingen waar economie veel minder als primaire discipline wordt gezien is het juist heel ‘ongepast’ om op economische argumentatie terug te grijpen. Dat kan en mag uiteraard wel, maar het is een afwijking van de institutionele norm en wordt door anderen ook als zodanig herkend.

Als we vanuit deze overwegingen kijken naar de interviews die we hebben gevoerd en de ervaringen met ‘andere vormen van evalueren’ analyseren, dan gaat het vooral om frictie in de ‘rechtsonderkant’ van de driehoek in [figuur 4](#) (de hoek van de institutionele regels en procedures). Alternatieve methoden sluiten onvoldoende aan bij de institutionele kaders en patronen die er zijn en worden daarom per definitie als afwijkend en ‘zwak’ gezien. Er is een vanzelfsprekende vorm van evalueren, die voorbij reflectie en bewuste afweging is verankerd in institutionele patronen, waarover letterlijk niet meer gepraat hoeft te worden en die ‘gewoon’ is. Een andere vorm van evalueren kan alleen systematisch toegepast worden als er in de institutionele patronen en kaders die *normaliserend werken* ruimte voor ontstaat. Bijvoorbeeld als er een ‘Planbureau voor Maatschappelijke Meerwaarde’ is, of een ‘Commissie Evalueren in Verweven Dynamiek’ (die een standaard voor alternatieve vormen van evalueren aanreikt), of een ‘standaardmethode’ voor de evaluatie van transitieprocessen. Niet als uitzondering voor ‘afwijkende dossiers’, die zo bijzonder en anders zijn dat ze een geheel andere en daarmee ook onzekere benadering vereisen, maar als gebruikelijke, geëigende en ‘gewone’ manier van werken om een specifiek en herkenbaar type dossier goed aan te pakken. Daarmee ontstaat *institutionele verevening van het speelveld*: verschillende aanpakken van evalueren staan normatief naast elkaar, in plaats van boven of onder elkaar op een ladder van adequaatheid. Daarom is het ook bijzonder interessant om te zien hoe het Planbureau voor de Leefomgeving, maar ook uitvoeringsorganisaties als rvo, zich verder ontwikkelen. Dit zijn organisaties die een sterke institutionele positie innemen én zich nadrukkelijker manifesteren als dragers van een breed perspectief op opgavegericht werken en daarbij ook zoeken naar andere manier van evalueren (en verantwoorden). Ook programma’s als ‘realistisch en bruikbaar evalueren’ en discussies binnen de domeinen van de Algemene Rekenkamer en interne auditors, zoals de Eenheid Audit en Advies van de Provincie Zuid-Holland⁴⁸ zijn voorbeelden. Dergelijke organisaties kunnen het institutionele gewicht beïnvloeden dat

de bedding voor het dagelijkse handelen van praktijkmensen vormt; van daaruit beantwoorden zij de vraag: “wat doet iemand zoals ik in een situatie zoals deze?”

Het voorgaande heeft een paradoxaal gevolg voor de ontwikkeling (of bredere bekendmaking) van andere manieren van evalueren. Het is belangrijk om goede evaluatievormen te ontwikkelen; het is belangrijk dat ze er zijn en dat mensen ervan af weten. Tegelijkertijd is dat niet de oplossing voor het probleem: weten is nog geen doen. Bepalend voor het handelen is net zozeer de institutionele bedding waarin dagelijkse keuzes over wat passend is plaatsvinden. Om de nog te ontwikkelen of al beschikbare methoden in gebruik te nemen is het daarom eerder nodig om in het institutionele repertoire veranderingen aan te brengen, die meer *vanzelfsprekend* leiden naar *opgavegerichte manieren* van evalueren. Daarvoor is een proces van **normalisering** nodig, meer dan een ontwikkelproces van nieuwe methoden van evalueren. Nieuwe methoden zijn echter niet *vanzelfsprekend* voorhanden. Het is dan ook de uitdaging om de evaluatietechniek te vinden die past bij de complexe maatschappelijke realiteit en daarbij horende beleidspraktijk van nu.

6 Conclusie: opgavegericht evalueren vereist normalisering

We begonnen dit essay met een aantal beleidsonderwerpen op de agenda van het Ministerie van EZK. Hoe evalueer je beleid voor de energietransitie, elektrisch rijden of circulaire economie? Onderwerpen waar veel publiek geld in omgaat, de overheidsinzet groot is en de maatschappelijke gevolgen vaak majeur zijn. Het belang van een adequate beleidsevaluatie staat daarom buiten kijf.

Tegelijkertijd is die vraag moeilijk te beantwoorden. De genoemde dossiers dragen kenmerken die beleidsevaluatie lastig maken. We hebben laten zien dat die ingewikkeldheid het gevolg is van twee soorten kwesties. Deels door inherente dilemma's die bij de ontwikkelde 'gouden standaard' voor beleidsevaluatie horen en die altijd wel in enige mate aan de orde zijn – maar waarvoor ook vaak best praktische oplossingen gevonden kunnen worden. Deels zijn er ook meer specifieke dilemma's, die voortkomen uit bepaalde omstandigheden die horen bij een opgavegerichte aanpak. Deze specifieke dilemma's doen zich krachtig voor in veel onderwerpen die hoog op de beleidsagenda staan, maar ze zijn lastig op te lossen vanuit de bestaande standaard voor 'goed evalueren'.

Eén van de problemen bij het omgaan met die dilemma's is het frame waarin het denken over evalueren terecht is gekomen: namelijk dat van een gouden standaard die de norm is, waar evaluatoren in reactie op concrete omstandigheden als het ware stappen omhoog en omlaag kunnen doen op een trap met treden van criteria naar steeds betere evaluaties, richting de gouden standaard. Dit 'trap op-trap af'-beeld van evalueren heeft in zich dat het aanpassen van de evaluatie op concrete omstandigheden zou leiden tot minder goede evaluaties; wel evalueren, maar niet zo goed, op een lagere trede van kwaliteit. Deze metafoor doet echter geen recht aan het evalueren van beleid voor inherent complexe vraagstukken. In dit essay hebben we gewezen op het belang van een andere generatieve metafoor voor het denken over evalueren: niet de trap op of af, maar ver-

schillende trappen die naast elkaar bestaan en ‘trap-over’ stappen die neerkomen op van de ene naar de andere trap bewegen, naar gelang het specifieke vraagstuk en het specifiek doel daar om vragen.

Uiteindelijk is beleidsevaluatie een toets op ideeën vanuit de werkelijkheid, weten wat werkt en wat niet werkt. Met rekenschap van complexiteit gaat evaluatie niet meer (alleen) over toetsing en *accountability* ten aanzien van normbereik, maar veel meer over de organisatie van sociale vernieuwing richting visies en publieke waarden als doelen.

Het beeld dat daarbij past is *opgavegericht evalueren* ten aanzien van beleidsprogramma’s die zijn gericht op maatschappelijke opgaven en grootschalige transitie. De uitkomsten zijn niet altijd gewis, maar er zijn wel ambities ten aanzien van de voorliggende opgave. Beleid krijgt dan vorm als *opgavegerichte aanpak*; dat kan via een beleidsprogramma, maar ook door allianties met partners aan te gaan, of zelfs door zelf bewust af te zien van een beleidsinterventie. Daartoe past een breed evaluatieperspectief, niet één gouden standaard, maar *meerdere standaarden*, afgestemd op verschillende opgavegerichte benaderingen. Passend evalueren komt neer op een manier die aansluit bij opgavegericht werken en die op de eigen manier goed is. Niet één trap op of af, maar meerdere trappen naast elkaar, met elk een eigen aanpak die aansluit bij de aard van het vraagstuk en de beleidspraktijk.

Tegelijkertijd hebben we in de praktijkgesprekken gezien dat het *weten* dat het zo werkt niet betekent dat het ook gebeurt in de toepassing. Dat vereist meer dan methoden, namelijk *normalisering*. De ‘gouden standaard’, en de onderliggende waarden en assumpties, zitten diep verankerd in gedeelde manieren van denken, doen en voelen, in institutionele regels (formeel en informeel), in taal en routines. Zolang die onderliggende patronen en systemen niet mee veranderen met de zich ontwikkelende manieren van opgavegericht werken, blijft het beeld van goed of minder goed evalueren van kracht. Dit essay opent de gerichte zoektocht naar *opgavegerichte evaluatie* en de arrangementen van *normalisering* die dat vereist.

Eindnoten

- ¹ Algemene Rekenkamer (2016). *Inzicht in publiek geld. Uitnodiging tot bezinning op de publieke verantwoording*. Den Haag: Algemene Rekenkamer. Zie ook het 15^e rapport van de Studiegroep Begrotingsruimte (2016). Het Regeerakkoord kondigt hierom een 'Operatie Inzicht in Kwaliteit' aan.
- ² Ministerie van Economische Zaken (2017). *Sturen in een verweven dynamiek*. Den Haag: Ministerie EZ.
- ³ Expertgroep Effectmeting (2012). *Durf te meten. Eindrapport expertgroep effectmeting*. Den Haag: SEO (publicatie 2012-90).
- ⁴ Hoogerwerf, A. & M. Herweijer (2008), *Overheidsbeleid*, Alphen aan de Rijn: Wolters Kluwer Nederland.
- ⁵ o.a. Ten Heuvelhof, E.F. (2016), *Strategisch gedrag in netwerken*, Den Haag: Boom Bestuurskunde.
- ⁶ De Bruijn, H. (2007), *Een gemakkelijke waarheid: waarom we niet leren van onderzoekscommissies*, Den Haag: NSOB.
- ⁷ Snowden, D.J., & Boone, M.E., (2007). 'A leader's framework for decision making', *Harvard Business Review* (november 2007), <https://hbr.org/2007/11/a-leaders-framework-for-decision-making>.
- ⁸ Snowden, D.J. & M.E. Boone (2007), *A leader's framework for decision making*. Cambridge: Harvard Business Review.
- ⁹ M. Mitchell (2009), *Complexity, a guided tour*, Oxford, Oxford University Press.
- ¹⁰ Dit hangt samen met de oorsprong van het kennisveld: chaostheorie is de oude benaming voor (wiskundige) kennis van non-lineaire dynamische systemen (vaak geïllustreerd door zgn. fractals). Er is ook wel overlap tussen chaotische en complexe systemen, bijvoorbeeld t.a.v. de gevoeligheid van een systeem voor de beginsituatie.
- ¹¹ WRR (2011), *Evenwichtskunst, over de verdeling van verantwoordelijkheid voor fysieke veiligheid*, Den Haag, Wetenschappelijke Raad voor het Regeringsbeleid.
- ¹² Dit model vormde een basisbestanddeel in het programma Samen in Beleid (SIB) en is ook veelvuldig gebruikt in gesprekken en sessies met de Bestuursraad en verschillende directies. We bespreken het model daarom relatief kort. Een uitgebreide bespreking van het model is te vinden in Van der Steen et al (2014) en Van der Steen et al (2018).
- ¹³ Weber, M. (1947), *The theory of social and economic organization*, New York: Oxford University Press en Wilson, J.Q. (1989), *Bureaucracy: what government agencies do and why they do it*, New York: Basic Books.

- ¹⁴ Frissen, P.H.A. (2010), *De staat van verschil: een kritiek van de gelijkheid*, Amsterdam: Van Genneep.
- ¹⁵ Bovens, M., 't Hart, P. & van Twist, M. (2012), *Openbaar bestuur: beleid, organisatie en politiek*, Deventer: Kluwer.
- ¹⁶ Pollit, C. & Bouckaert, G. (2011), *Public management reform: a comparative analysis: new public management, governance and the neo-weberian state*, Oxford: Oxford University Press.
- ¹⁷ Klijn, E.H. & Koppejan, J.F.M. (2012), Governance network theory: past, present and future. In: *Policy & Politics*, 40(4), 187-206.
- ¹⁸ Hajer, M. (2011), *De energieke Samenleving: op zoek naar een sturingsfilosofie voor een schone economie*, Den Haag: PBL; WRR (2012), *Vertrouwen in burgers*, Amsterdam: Amsterdam University Press; Van der Steen, M., Hajer, M., Scherpenisse, J., Van Gerven, O.J. & Kruitwagen, S. (2014), *Leren door doen: overheidsparticipatie in een energieke samenleving*, Den Haag: NSOB.
- ¹⁹ ROB (2012), *Loslaten in vertrouwen: naar een nieuwe verhouding tussen overheid, markt én samenleving*, Den Haag: Raad voor het openbaar bestuur; Van der Lans & De Boer (2011), *Burgerkracht: de toekomst van het sociaal werk in Nederland*, Den Haag: RMO.
- ²⁰ o.a. De Baas, J.H. (2018), *Voorbij de eeuw van bureaucratie*, Den Haag: Boom Bestuurskunde.
- ²¹ o.a. Scherpenisse, J. & M. Van der Steen (2018), *Gepast geregeld: tijdig financieren in netwerken*, Den Haag: NSOB; De Baas, J.H. (2018), *Voorbij de eeuw van bureaucratie*, Den Haag: Boom Bestuurskunde; Den Ouden, W. & J.E. Van den Brink (2016), *De subsidie nieuwe stijl: publiek geld verplicht?*, *Nederlands Juristenblad*, 39(91), 2889-2897.
- ²² Driessen, P., C. Dieperink, F. van Laerhoven, H.A.C. Runhaar en W.J.V. Vermeulen (2012), 'Towards a conceptual framework for the study of shifts in modes of environmental governance -experiences from The Netherlands', *Environmental Policy and Governance* 22: 143-160.
- ²³ Volgens: Loorbach, D. (2014), *To Transition! Governance Panarchy in the new Transfor-mation*, drift. Rotterdam: Erasmus University Rotterdam; en Loorbach, D., N. Frantzeskaki & L.R. Huffenreuter (2015), *Transition management: taking stock from governance experimentation*, *Journal of Corporate Citizenship*, 58: 48-66.
- ²⁴ Zie ook: Loorbach, D., N. Frantzeskaki & F. Avelino (2017), 'Sustainability transitions research: transforming science and practice for societal change', *Annual Review of Environment and Resources* 42: 599-626.
- ²⁵ Loorbach, D., N. Frantzeskaki & F. Avelino (2017), *Sustainability transitions research: transforming science and practice for societal change*, *Annual Review of Environment and Resources*, 42: 599-626.

- ²⁶ Loorbach, D.A. (2018), Adaptief sturen in transitie, In: P. Mettau & Z. Hulsenboom (Red.), *Adaptief bestuur: essays over adaptiviteit en openbaar bestuur*, Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, pp. 63-72.
- ²⁷ Snowden, D.J. & M.E. Boone (2007), *A leader's framework for decision making*. Cambridge: Harvard Business Review.
- ²⁸ Faber, A. & A. Idenburg (2017), Sturen op transitie: van utopie naar stapsgewijze voortgang, in: *Sturen in een verweven dynamiek*, Den Haag: Ministerie van Economische Zaken, pp. 24-33.
- ²⁹ Deze relativisering zien we overigens ook in 'reguliere' evaluaties, vanuit het besef dat de invloed van een beleidsinstrument op een maatschappelijk doel a) indirect of getrap is, en b) onderdeel is van een veel breder palet aan beïnvloedende factoren. Zo kunnen we van innovatiesubsidies zeggen dat ze uiteindelijk beogen bij te dragen aan maatschappelijke spillover van kennis, maar dit verloopt via R&D, innovatie en kennisuitwisseling. Het instrument beoogt dus bij te dragen aan iets dat meer is dan in een evaluatie naar voren zal komen, want dan gaat het bijv. over de extra R&D-investeringen door bedrijven. Dus: de invloed van het instrument op een stroom, die uiteindelijk (gebaseerd op literatuur) van invloed zou zijn op het doel.
- ³⁰ Dialogic (2017). *Evaluatie topsectorenaanpak – deel I: hoofdrapport*, Dialogic, Utrecht.
- ³¹ Zie ook: M. Janssen (2018). 'What bangs for your buck? Assessing the design and impact of Dutch transformative policy', *Technological Forecasting & Social Change* (in press).
- ³² PBL (2017), *Lerende evaluatie van het Natuurpact. Naar nieuwe verbindingen tussen natuur, beleid en samenleving*, Den Haag: PBL.
- ³³ Verwoerd, L., R. de Wildt-Liesveld & B.J. Regeer (2017). *The value of reflexive evaluation. A review of the Natuurpact evaluation (2014-2017)*. Amsterdam: Athena Institute, VU University Amsterdam.
- ³⁴ Zie: OECD (1991). DAC Principles for Evaluation of Development Assistance. Paris: OECD, via <http://www.oecd.org/dac/evaluation/50584880.pdf> en OECD (2002). Glossary of Key Terms in Evaluation and Results Based Management, via <http://www.oecd.org/dac/evaluation/2754804.pdf>.
- ³⁵ Citaat van een deelnemer van de focusgroeptijdenkomst.
- ³⁶ Citaat van een deelnemer van de focusgroeptijdenkomst.
- ³⁷ Citaat van een deelnemer van de focusgroeptijdenkomst.
- ³⁸ Citaat van een deelnemer van de focusgroeptijdenkomst.

- ³⁹ Patton, M.Q.M. (2011), *Developmental evaluation, applying complexity concepts to enhance innovation and use*, Guilford Press. Zie ook Westley et al. (2006), *Getting to maybe. How the world is changed*, Vintage Canada.
- ⁴⁰ Klein Haarhuis, C.M. en A. Parapuf (2016), *Evaluatievermogen bij beleidsdepartementen*, Den Haag: wodc, Ministerie van Veiligheid en Justitie.
- ⁴¹ Folkert, R., L. Verwoerd & F. Verwest (2018), *Lerend evalueren: navigeren tussen verantwoord en leren*, Beleidsonderzoek Online.
- ⁴² Citaat van een deelnemer van de focusgroepbijeenkomst.
- ⁴³ Allison, G.T. (1971), *Essence of decision: explaining the Cuban missile crisis*, Boston: Little, Brown and Company.
- ⁴⁴ March, G.J. en J.P. Olsen (1989), *Rediscovering institutions*, New York: The Free Press.
- ⁴⁵ March, G.J. en J.P. Olsen (1989), *Rediscovering institutions*, New York: The Free Press en March, G.J. en J.P. Olsen (1995), *Democratic Governance*, New York: The Free Press.
- ⁴⁶ Zoals Allison (1971) en ook March & Olsen (1989; 1995) doen.
- ⁴⁷ Noordegraaf, M. en B. de Wit (2012), *Van maakbaar naar betekenisvol bestuur: een achtergrondstudie naar (keten)governance en (nieuw) publiek management en de gevolgen voor toezicht en evaluatie*, Den Haag: WRR (nr. 63).
- ⁴⁸ <https://www.zuid-holland.nl/overons/missie-visie/artikel/Netwerkvitaliteitsmeter>.



Over de NSOB

Nederlandse School voor Openbaar Bestuur

De Nederlandse School voor Openbaar Bestuur (NSOB) verzorgt sinds 1989 hoogwaardig postacademische opleidingen. In dat jaar besloten de Universiteit Leiden en de Erasmus Universiteit Rotterdam tot de oprichting van de Nederlandse School voor Openbaar Bestuur. Sinds juli 1995 participeren behalve de Universiteit Leiden en de Erasmus Universiteit Rotterdam ook de Universiteit Utrecht, de Universiteit van Amsterdam, de Technische Universiteit Delft, de Universiteit van Tilburg en de Vrije Universiteit Amsterdam in de NSOB. Vanaf haar aanvang streeft de NSOB ernaar om met geavanceerde en uitdagende opleidingen voor het topsegment van het management in openbaar bestuur en publieke sector bij te dragen aan een hoogwaardig openbaar bestuur.

De opleidingen van de NSOB onderscheiden zich door geavanceerde didactische concepten, een excellent docentenkorps van top-wetenschappers en vooraanstaande professionals uit de praktijk en intensieve en kleinschalige onderwijsvormen. De opleidingen bieden een mix van cognitieve verdieping, theoretische en professionele reflectie, oefening in professionele en persoonlijke competenties, toepassing van kennis en inzichten in complexe advies- en onderzoeksopdrachten. De opleidingen zoeken de grenzen van weten en kunnen op en dagen de deelnemers uit hetzelfde te doen.

De NSOB is sinds 2006 niet alleen een hoogwaardig opleidingsinstituut voor de publieke sector, maar tevens een *denktank*. In die functie wil de NSOB bijdragen aan kennisontwikkeling voor en over openbaar bestuur en publieke sector. Het gaat daarbij om strategische vragen over beleidsinhoud en sturingsrelaties, over verschuivende verhoudingen tussen private, publieke en politieke domeinen en over de vormgeving en instrumentering van veranderingen in die domeinen.

De NSOB werkt aan vragen die worden aangereikt door opdrachtgevers uit openbaar bestuur en publieke sector, maar ook aan vragen die voortkomen uit autonome wetenschappelijke en professionele reflectie. De NSOB biedt ruimte en inspiratie voor gasten uit de praktijk en de wetenschap, tijdens en na hun loopbaan. De NSOB organiseert publieke debatten en verzorgt wetenschappelijke en professionele publicaties.





Nederlandse School voor Openbaar Bestuur

Lange Voorhout 17

2514 EB Den Haag

(070) 302 49 10

www.nsob.nl

info@nsob.nl

