

# Systematisch leren van evalueren

## Waarden, effectiviteit, onafhankelijkheid en kwaliteit als pijlers voor de brug tussen wetenschap en politiek\*

*André Knottnerus, Peter de Goede & Peter van der Knaap*

*Beleidsvaluatie heeft twee functies: verantwoording afleggen en leren. Deze bijdrage gaat in op de vraag hoe beleidsvaluatie beter benut kan worden, waarbij het leren centraal staat. Daarbij nemen wij als vertrekpunt het spanningsveld tussen wetenschappelijke en politieke rationaliteit en daarmee samenhangende barrières voor de doorwerking van kennis.*

*Binnen genoemd spanningsveld moet beleidsvaluatie stevig verankerd zijn om haar onpartijdige rol te kunnen vervullen. Dat is van belang voor de validiteit en toegevoegde waarde van evaluaties, maar ook voor het publiek vertrouwen in het bestuur en het gevoerde beleid. Dit artikel schetst het belang van waarden, van inzicht in de effectiviteit van interventies en van de onafhankelijkheid van evaluatie-onderzoekers. Ook een relevante vraagstelling, hoogwaardige kwaliteit, responsiviteit, een goede timing en een heldere rapportage zijn essentieel om systematisch te kunnen leren van beleidsvaluaties. Dit zijn geen abstracte maar concreet vorm te geven waarden.*

*Een op waarden en effectiviteit gestoelde agenda voor beleidsvaluatie kan een brug slaan tussen het wetenschappelijke en het politieke domein. Daarbij moet gestreefd worden naar een evenwichtig evaluatieportfolio, waarin individuele evaluaties worden gezien als deel van een geheel. Dat biedt perspectief voor een overheid die niet alleen casus-specifiek maar vooral ook systematisch wil leren.*

\* Wij danken dr. Petra Jonkers, senior wetenschappelijk medewerker van de WRR, in het bijzonder voor haar inbreng met betrekking tot een systematische kennisopbouw en een evenwichtige evaluatieportfolio. Dit artikel berust ten dele op teksten die gepubliceerd zijn in een artikel in *Beleidsonderzoek online* (Van casus-specifieke beleidsvaluatie naar systematische opbouw van kennis en ervaring. *Beleidsonderzoek online*, 2016, DOI: 10.5553/BO/221335502016000005001. Prof. dr. J.A. Knottnerus is voorzitter van de Wetenschappelijke Raad voor het Regeringsbeleid en hoogleraar huisartsgeneeskunde aan de Universiteit Maastricht. Van 2001 t/m 2010 was hij voorzitter van de Gezondheidsraad, het wetenschappelijk adviesorgaan van de regering op het gebied van de volksgezondheid, knottnerus@wrr.nl. Dr. P.J.M. de Goede is Senior wetenschappelijk medewerker bij de Wetenschappelijke Raad voor het Regeringsbeleid. Eerder was hij hoofdredacteur en redactiesecretaris van 'Bestuurswetenschappen' (1999-2009) en werkte hij als senior adviseur bij de Raad voor het openbaar bestuur en de Raad voor de financiële verhoudingen, goede@wrr.nl. Dr. P. van der Knaap is directeur-bestuurder van de Stichting Wetenschappelijk Onderzoek Verkeersveiligheid en voorzitter van de beroepsvereniging van toezichthouders, inspecteurs, handhavers en evaluatoren, VIDE, Peter.van.der.knaap@swov.nl.

## Introductie: evalueren is nodig maar leren lastig

Evaluatie van beleid kost energie, tijd en geld, maar is onmisbaar voor democratische controle en verantwoording en om te kunnen leren van meer en minder geslaagd beleid (WRR, 2006; Bovens en Schillemans, 2009). Beleidsevaluatie is in essentie het waarderen van de effecten en/of uitvoering van beleid en richt zich vaak op doeltreffendheid (effectiviteit) en doelmatigheid (kosteneffectiviteit) van maatregelen. Toetsing van effectiviteit en doelmatigheid, bijvoorbeeld van interventies in de gezondheidszorg en van het energiebeleid, kunnen immers belangrijke input leveren voor nieuw of bij te stellen beleid.

Aan beide functies van evalueren – verantwoorden en leren – wordt in Nederland gehecht:

- de Regeling periodiek evaluatieonderzoek (RPE) van de minister van Financiën stelt eisen ten aanzien van de kwaliteit en frequentie van onderzoek naar doeltreffendheid en doelmatigheid, beleidsdoorlichtingen, en subsidie-evaluaties;
- in veel wettelijke regelingen is bepaald dat beleidsevaluaties moeten worden uitgevoerd;
- wetenschappelijke instituten, de Algemene Rekenkamer, de Nationale Ombudsman en andere evaluatoren kijken ‘van buiten’ naar het handelen van departementen en hun uitvoeringsdiensten;
- daarnaast vinden evaluaties plaats op initiatief van de ministeries of op verzoek van de Tweede Kamer.

Al met al gaat het om honderden evaluaties per jaar (vergelijk Algemene Rekenkamer, 2012). Uitgangspunt bij al deze evaluatie-activiteiten is dat de overheid dient te streven naar zo effectief mogelijk beleid en daarover ook verantwoording aflegt. Is het beleid succesvol in het bereiken wat voor ogen stond? Werken interventies conform oogmerk en zonder te hoge kosten en negatieve effecten?

Om zulke vragen te kunnen beantwoorden is goed evaluatieonderzoek nodig. Daarbij gaat het ook om vertrouwen en geloofwaardigheid: wie laat zien lessen te trekken uit fouten of teleurstellende resultaten, bouwt aan vertrouwen (Jonkers, 2013). De overheid is, in de woorden van Stern, ‘accountable for learning’: verantwoording afleggen gaat niet alleen over resultaten, maar ook over hoe tussentijdse lessen zijn gezocht en geleerd (Stern, 2011). Verantwoording afleggen heeft niet zozeer ‘afrekenen’, maar vooral verbetering als hoofddoel (Van der Knaap, 2006a).

Leren in een politiek-bestuurlijke omgeving blijkt echter vaak lastig te zijn. Parlementaire enquêtes en onderzoeken in Nederland gaan dan ook vaak over de vraag waarom onvoldoende geleerd is van eerdere waarschuwingen en signalen. Dit ‘hardleerse’ karakter van de overheid kent verschillende oorzaken. Gebrek aan tijd en middelen om de grote hoeveelheid informatie te verwerken is er een van: kennis uit evaluaties is slechts één van de vele informatiebronnen om beleid te ontwikkelen of te herzien. Ook kan het zijn dat een evaluatie geen antwoord geeft op

belangrijke vragen, te abstracte aanbevelingen bevat of eenvoudigweg te laat beschikbaar komt voor beleidsmakers of politici. Druk vanuit belangengroepen, snelle oordelen van of in de media en politieke pressie kunnen daarbij een rol spelen. Deze ontwikkelingen kunnen leiden tot een fenomeen dat ook wel de 'kennisparadox' wordt genoemd: er is kostbare kennis beschikbaar, maar deze wordt niet of maar in beperkte mate benut (vergelijk Rouw, 2011).

Daar ligt de uitdaging waarop het vervolg van dit artikel zich toespitst: het te midden van barrières voor leren zoeken naar mogelijkheden om dat leren te bevorderen. Daarbij is gebruik gemaakt van de wetenschappelijke en vakliteratuur en van bevindingen in het kader van diverse WRR-studies. Wij staan achtereenvolgens stil bij:

- de paradoxale spanning tussen politisering enerzijds en de steun voor met 'evidence' onderbouwd beleid anderzijds;
- de kans om naast specifieke lessen voor beleidsprogramma's ook algemenere inzichten over effectief beleid op te doen en de rol die breder geldende waarden daarbij kunnen vervullen;
- de spanning tussen onafhankelijkheid en betrokkenheid van onderzoekers, en;
- het belang van hoge evaluatiekwaliteit, van vraagstelling tot en met rapportage.

### **Politieke en wetenschappelijke rationaliteit**

Het streven naar een compacte overheid heeft de kennisfunctie op ministeries onder druk gezet. Bezuiniging op kennisdirecties en de roulatie van (top)ambtenaren verminderen zowel de doorlopende dossierkennis bij individuele functionarissen als het 'collectieve geheugen' binnen departementen. Er lijkt ook sprake te zijn van een dalende trend wat betreft het belang dat gehecht wordt aan inhoudelijke zaakkennis bij ambtenaren (vergelijk 't Hart, 2014; Bekker, 2014). Ook in de competenties zoals geformuleerd door de Algemene Bestuursdienst scoren politieke sensitiviteit en strategische vaardigheden hoog.<sup>1</sup>

Velen ervaren een politisering van de ambtelijke cultuur, gekenmerkt door een gerichtheid op de korte termijn, het uit de wind houden van de minister en het liever vooruit dan achteruit kijken (Nieuwenkamp, 2013). Deze politisering kan de ruimte voor evaluatoren beperken en de aandacht voor hun bevindingen verminderen.

Tegenover deze ontwikkeling staat een brede roep om 'evidence-informed' of zelfs 'evidence-based' beleid, die een belangrijke impuls kreeg met het parlementair onderzoek onderwijsvernieuwingen in 2008. Bij goed bestuur hoort immers inzicht in de effectiviteit van beleidsmaatregelen. Dat geldt zowel bij het inzetten

1 Zie bijvoorbeeld het competentieprofiel van de ABD-manager: [www.algemenebestuursdienst.nl](http://www.algemenebestuursdienst.nl). (geraadpleegd op 10 mei 2016).

van nieuw beleid als bij het verbeteren van bestaand beleid. De zowel nationaal<sup>2</sup> als internationaal<sup>3</sup> toenemende aandacht van overheden voor ‘verwetenschappelijking van beleid’, waarbij het belang van ‘evidence’ voor de effectiviteit van interventies benadrukt wordt, is daar onlosmakelijk mee verbonden.

Men kan ook hier dus spreken van een paradox: enerzijds zijn er tekenen van politisering van de ambtelijke cultuur en anderzijds is er aandacht voor verwetenschappelijking van de beleidsontwikkeling. Deze spanning tussen politieke en wetenschappelijke rationaliteit is inherent aan overheidsbeleid (Snellen, 2007). In 1979 stelde Wildavsky al dat de belangrijkste functie van evaluatoren bestaat uit het *Speaking truth to power* (Wildavsky, 1979). Anno 2016 kan men dit breder trekken: het gaat om het aanreiken van goed onderbouwde, onafhankelijke inzichten aan de samenleving en het politiek bestuur. De voor het beleid relevante vraag is dan onder welke voorwaarden een ‘evidence-informed’ benadering in het geschetste spanningsveld houvast kan bieden voor beleidsgericht leren en te maken keuzes.

### **Overbruggen: waarden en breed inzicht in effectieve interventies**

Politiek gaat over hoe burgers in een bepaald verband het beste kunnen samenleven. Daarbij staan waarden centraal: de idealen die door de samenleving en regerende partijen of coalities als wenselijk en dus als nastrevenswaardig worden beschouwd. Voorbeelden zijn vrijheid, gelijkheid, veiligheid, rechtvaardigheid en welzijn.

Deze publieke waarden bieden ook relevante uitgangspunten voor beleidsevaluatie: zij gelden als de uiteindelijke maatstaven voor legitiem beleid en bestuur en specifieke doelen van overheidsbeleid. Beleidsdoelen staan immers zelden op zichzelf, maar worden doorgaans gesteld om een ‘algemenere waarde’ te realiseren. Anders gesteld: effecten ontlenen hun relevantie aan hun betekenis voor maatschappelijke waarden, zoals gelijke kansen of duurzaamheid. Vaak zijn deze waarden in het doel van een wet of beleidsinterventie verankerd. Zo moeten evaluaties van de Algemene Wet Gelijke Behandeling inzicht verschaffen in de mate waarin naleving en handhaving van de wet bijdragen aan gelijke behandeling en het bestrijden van discriminatie. Zulke waarden kunnen soms ook uit doelen van wetgeving en beleid worden afgeleid als ze daarin niet expliciet zijn vermeld.

Aangezien beleidsevaluatie in essentie het waarderen van beleid is, kunnen (1) het vooraf in beeld brengen van belangrijke waarden en (2) het bepalen van de behartiging of realisatie daarvan dus sturing geven aan evaluatieonderzoek (Leh-

2 Zie bijvoorbeeld de Kabinetsreactie op de adviesrapporten van Rli, RMO en WRR over de benutting van gedragswetenschappelijke kennis in beleid: Den Haag: Ministerie van Economische Zaken, 4 december 2014.

3 Wij verwijzen bijvoorbeeld naar de ‘Blueprints for informed policy decisions: A review of laws and policies requiring routine evaluation’, een rapport van het Noorse ‘Knowledge Centre for the Health Services’. No 16–2012 Review. Oslo, 2012.

ning, 1991; Fischer, 1995). Hierbij kunnen verschillende perspectieven van gebruikers en doelgroepen op onderliggende waarden aan het licht komen. Vooral bij complexe dossiers waarbij sprake is van veel onzekerheden of risico's is het zinvol om expliciet te onderzoeken welke waarden gediend worden met het beleid en of zich al dan niet ongewenste neveneffecten voordoen. Een voorbeeld is de evaluatie van de impact van grote interventies als stelselherzieningen en decentralisatieoperaties: in welke mate komt rechtsgelijkheid in het geding bij het streven naar maatwerk en doelmatigheid, en in hoeverre is dat acceptabel?

Naast achterliggende waarden bieden ook algemene mechanismen en wetmatigheden kansen om bruggen te slaan tussen politiek en wetenschap. Veel evaluaties leveren niet alleen onderwerp-specifieke kennis op over specifieke beleidstheorieën, zij leiden ook tot meer generieke kennis over de effectiviteit van (typen) interventies en werkwijzen via het identificeren van algemene mechanismen. Zulke mechanismen zijn vaak breder toepasbaar dan alleen op het specifiek bestudeerde beleid (Klein Haarhuis en Niemeijer, 2008; Veerman, Mulder en Meijsing, 2013).

Het is in dit verband van belang meer aandacht te besteden aan het vooraf verkrijgen van inzicht in de potentiële werkzaamheid van beleidsinterventies (Hane-maayer, 2013). Daarbij is het vooraf stellen van de vraag waarom iets zou kunnen werken, en voor wie en onder welke omstandigheden, cruciaal (Leeuw en Mevis-sen, 2015; Leeuw, 2008; Nelen, 2008; Nelen et al., 2012; vergelijk Pawson en Tiley, 1997). Ex ante evaluatie biedt de gelegenheid om systematisch na te gaan in hoeverre verwacht mag worden dat essentiële publieke belangen en waarden daadwerkelijk gediend zullen zijn met voorgenomen interventies. Dit is ook de eerste kernvraag die de overheid zich steeds moet stellen bij de voorbereiding en evaluatie van het risico- en veiligheidsbeleid (WRR, 2014).

Omdat bij beleidsevaluatie dus gestuurd kan worden op bijdragen aan meer gene-rieke kennisontwikkeling, dient steeds ook de vraag te worden gesteld hoe – naast de kennis die nodig is voor het specifieke beleidsdossier – een evaluatie kan bij-dragen aan een bredere systematische opbouw van kennis over interventiemecha-nismen. Het kan daarbij helpen evaluatierapporten te voorzien van een specifiek en een breder geldend algemeen deel.

In het rapport *Met kennis van gedrag beleid maken* pleitte de WRR voor de opbouw van capaciteit voor het opdoen, vasthouden en toepassen van (gedrags)weten-schappelijke kennis in beleid (WRR, 2014). Systematische reviews van eerder uit-gevoerd beleidsonderzoek kunnen daarbij van grote waarde zijn (Pettigrew en Robert, 2006). Internationaal is de Campbell Collaboration Library of Systematic Reviews daarvoor een belangrijke bron.<sup>4</sup> Deze dekt overigens nog maar een beperkt deel van de diverse beleidsvelden, en bijdragen aan (continue) aanvulling hiervan vormt belangrijke uitdaging voor evaluatie-onderzoekers. Ook in Neder-land wordt op dit gebied veel werk verzet. Zo bouwt Movisie een databank 'Effec-

4 Zie voor een overzicht: <http://www.campbellcollaboration.org>.

tieve sociale interventies' op met informatie over de onderbouwing en effectiviteit van interventies.

## Het belang van onafhankelijkheid en onpartijdigheid

Onafhankelijkheid en onpartijdigheid zijn essentieel voor zowel de validiteit en toegevoegde waarde van evaluatie, als voor het publiek en politiek-bestuurlijk vertrouwen in de resultaten ervan (WRR, 2013). Dat wil echter niet zeggen dat de evaluator altijd daadwerkelijk onafhankelijk is, dan wel in de gelegenheid is als zodanig te opereren. Zo concludeerde de KNAW in haar rapport *Wetenschap op bestelling* dat 'de meest concrete voorbeelden van ongewenste druk op de onderzoeker worden gemeld bij opdrachten door de overheid' (KNAW, 2005).

Scherp blijven op onafhankelijkheid van evaluaties is dus geen overbodige luxe. Daarbij moeten we ons realiseren dat onafhankelijkheid geen abstract gegeven is: zij krijgt concreet betekenis door het vermijden van de specifieke afhankelijkheden die het adequaat en onpartijdig onderzoeken en beoordelen van een beleids-terrein en het overheidsoptreden kunnen belemmeren. Dit kan geborgd worden door een passende positionering van de evaluator. Maar onafhankelijkheid schuilt ook in de onpartijdige houding van personen die de evaluatie uitvoeren en de wijze waarop wordt omgegaan met hun belangen en relaties met beleidsactoren.

Naast de onafhankelijkheid van de evaluator zelf zijn van belang:

- de evaluatieagenda: wie bepaalt wanneer wat moet worden geëvalueerd;
- de bepaling van de probleemstelling en onderzoeksvragen alsmede de methoden en technieken van de evaluatie;
- de samenstelling en opdracht van begeleidingscommissies;
- de inrichting van het onderzoeksproces (wordt bijvoorbeeld, ook door begeleidingscommissies, de onafhankelijkheid van de onderzoekers te allen tijde gerespecteerd?);
- de procedures inzake belangenverstrengeling;
- de toegang tot alle relevante informatie;
- de mogelijkheid om onderzoeksresultaten vrijelijk en volledig te publiceren, en
- de publieke transparantie van al deze aspecten.

Onafhankelijkheid moet echter niet verworden tot een toestand van 'splendid isolation': een te grote afstand tussen evaluatieonderzoekers en beleidsmakers kan namelijk juist weer een barrière vormen voor de bruikbaarheid en acceptatie van resultaten van evaluatiestudies. Van belang is ook dat niet altijd een externe evaluator hoeft te worden ingeschakeld: wij zijn, in lijn met Korsten (2013), voorstander van een getrapte benadering. Net als een slager als eerste moet willen weten of zijn vlees van goede kwaliteit is, ligt de primaire evaluatieverantwoordelijkheid voor beleidsmaatregelen bij het politiek bestuur. Dat dient te zorgen voor een adequate beleidsevaluatie, die uitgevoerd of aangestuurd wordt door een daartoe goed geëquipeerde onafhankelijke eenheid. Als de resultaten of de onpar-

tijdigheid daarvan twijfels oproepen, of als sprake is van bijzondere risico's of controverses, moeten externe onafhankelijke evaluatoren worden ingeschakeld.

Daarnaast kan via periodieke algemene reviews met behulp van onafhankelijke evaluatoren extern getoetst worden of de overheid zijn evaluatieprocessen op orde heeft. Ook kan – conform de *Regeling periodiek evaluatieonderzoek* – op deze wijze op het niveau van belangrijke beleidsdoelen van tijd tot tijd de balans worden opgemaakt.

### **Evaluatiekwaliteit: van vraag tot rapportage**

De kwaliteit van beleidsevaluatie wordt door veel elementen bepaald. In het kader van deze bijdrage richten wij ons op vraagstelling, responsiviteit, methodologie, timing en benodigde tijd, draagvlak en rapportage.

#### *Vraagstelling*

Om met voldoende zekerheid iets te kunnen zeggen over doelbereiking en over werkzame mechanismen, is het wenselijk dat evaluaties een zo hoog mogelijke kwaliteit hebben. Goede kwaliteit begint met een adequate en duidelijke vraagstelling: wat willen we waarom weten? Zoals gezegd staat de doeltreffendheid (effectiviteit) en doelmatigheid bij veel evaluaties voorop. Het gaat dan om de vraag in welke mate en tegen welke kosten het beleid ertoe leidt dat de beoogde doelstellingen worden gehaald. Bij de beoordeling van deze aspecten is het aantrekkelijk de aandacht vooral te richten op reeds beschikbare of relatief gemakkelijk meetbare doelen en daaraan gekoppelde indicatoren: het leidt tot focus in de onderzoeksaanpak en tot feedback-informatie over de mate van succes.

Tegelijkertijd verklaart de associatie van beleidsevaluaties met een focus op wat gemakkelijk meetbaar is ('meten is weten') waarschijnlijk ook dat in ambtelijke kring soms een zekere scepsis bestaat. Er wordt ook wel gesproken van 'management by measurement' (Noordegraaf en De Wit, 2012; Noordegraaf en Abma, 2003). Beperkt of niet-meetbare waarden, die voor beleidsmakers en doelgroepen van belang zijn, kunnen daarmee namelijk onderbelicht blijven (zoals subjectief welzijn of rechtvaardigheid). Ook bestaat het risico dat het gedrag van degenen die geëvalueerd worden zich in de loop der tijd vooral op bekende, goed meetbare indicatoren gaat richten. Men duidt dit ook wel aan als 'pervers effect' van een dergelijke indicatorbenadering (vergelijk Gezondheidsraad 2006, 2013).

#### *Responsiviteit*

Daarom is het van belang dat evaluaties niet alleen een focus op kwantificeerbare indicatoren hebben, maar doelen ook relateren aan achterliggende waarden. Dit vergt bovendien dat evaluatoren responsief zijn ten opzichte van de beleids- en uitvoeringspraktijk die zij beoordelen (vergelijk Van der Knaap, 2006b). Evaluaties die vooral het top-down perspectief belichten (heeft men gedaan en bereikt wat in het beleid is afgesproken en heeft dat niet te veel gekost?) en onvoldoende aandacht besteden aan ervaringen, zienswijzen en kennis uit het werkveld dragen

het risico in zich dat onwenselijke neveneffecten onderbelicht blijven en dat men aan onhaalbaar beleid blijft vasthouden (Guba en Lincoln, 1989).

Ook om zowel bruikbare als kansrijke aanbevelingen te kunnen doen is gedegen kennis nodig van het praktijkveld, de gewoonten van de doelgroep en de invloed van de omgeving (Korsten en Mevissen, 2015). Zo moet men bij het ontwerpen van beleid met betrekking tot uitkeringsfraude inzicht hebben in de motieven van betrokkenen, en is preventiebeleid gericht op duurzame gedragsverandering alleen kansrijk als men rekening houdt met omgevingsfactoren (Brummelkamp, Kerckaert en Engelen, 2014; Van den Berg, Post, Hamberg-van Reenen, 2014).

### *Methodologie*

Er is geen principiële tegenstelling tussen het via methodologisch optimale evaluatiedesigns verkrijgen van kennis en inzichten enerzijds en het alledaagse praktijkleren van professionals anderzijds. Integendeel, deze vormen van leren vullen elkaar aan. Dat past ook bij de *mixed-methods* benadering waarin diverse onderzoeksdesigns, zowel kwantitatief als kwalitatief van aard, complementair worden ingezet. Naast vergelijkende studies, *surveys*, analyses van grote databestanden, documentanalyses en casestudies, kunnen interviews, focusgroepen, en praktijkobservaties een belangrijke bijdrage leveren. Het op onafhankelijke en systematische wijze 'oogsten' van praktijkervaring speelt daarbij dus een essentiële rol.

Bij al deze methoden geldt dat er systematisch gebruik moet worden gemaakt van de mogelijkheden die de veelheid aan informatiebronnen anno nu biedt. De informatie die beleidsdirecties en communicatie-afdelingen van ministeries, toezichthouders en planbureaus continu verzamelen over de wijze waarop beleid en regelgeving ontvangen worden is van onschatbare waarde voor de evaluatieonderzoeker. Ook de betekenis van de 'digital society' en de daaraan gerelateerde mogelijkheden en vereisten op het gebied van beschikbaarheid en gebruik van gegevens zal naar verwachting alleen nog maar toenemen (Algemene Rekenkamer, 2014; Janssen, 2015; WRR, 2016).

### *Timing*

Voor goede doorwerking van evaluaties in de beleidspraktijk is adequate *timing* van groot belang. Een evaluatie kan bijvoorbeeld te vroeg komen, met name wanneer kort na de invoering van een maatregel redelijkerwijs nog geen goede uitvoering of effecten mogelijk, laat staan meetbaar, zijn. Zo'n te vroege evaluatie levert niet alleen een onvolledige evaluatie op, het kan ook schadelijk zijn voor het beleid in kwestie. Berichten over tegenvallende resultaten ('nog niemand voert de wet uit' of 'er is nog geen verbetering waar te nemen') kunnen de motivatie om beleid of regelgeving uit te voeren ondermijnen. Ryan sprak treffend over 'death by evaluation' (2003).

Evaluaties kunnen ook te laat verschijnen om nog van invloed te kunnen zijn. Een geëvalueerde wet of regeling kan bijvoorbeeld inmiddels zijn opgegaan in een omvangrijk wijzigingsprogramma: de resultaten van een evaluatie van de 'oude



wet' worden dan al gauw niet meer relevant geacht. Daarmee kunnen echter generieke leerpunten die voor de toekomst wel degelijk van belang zijn verloren gaan.

#### *Benodigde tijd*

Gedegen evaluatie, gericht op wetenschappelijk betrouwbare en valide conclusies, vraagt ook de nodige tijd. Temeer omdat ook gekeken moet worden naar de geschiktheid van bestaande dataverzamelings- en monitoringsystemen: zijn deze bij de tijd en voldoende dekkend en doelmatig? Ook de vraag of opvolging is gegeven aan eerdere aanbevelingen en in hoeverre dat tot verbetering heeft geleid kost tijd, maar is onmisbaar voor een werkelijk lerende overheid. Dit alles betekent ook dat onredelijke tijdsdruk van de kant van opdrachtgevers, resulterend in voor aanvaardbare kwaliteit te korte doorlooptijden, niet geaccepteerd moet worden.

#### *Betrokkenheid en draagvlak*

Het is aan te bevelen beleidsmedewerkers in een vroeg stadium bij evaluaties te betrekken. Dat bevordert de erkenning van het belang van evaluatie, komt het draagvlak ervoor ten goede, en vergroot de kans dat de bevindingen in de beleidsvorming worden meegenomen. Wanneer medewerkers tijdens een door hen in gang gezette evaluatie vertrekken, kan er een probleem ontstaan wanneer de uitvoering van evaluaties binnen een departement afhangt van de inzet van slechts enkele toegewijde personen. Het veiligstellen van continuïteit en tijdige overdracht van evaluatiedossiers aan opvolgers, bij voorkeur gekoppeld aan persoonlijke kennismaking met de evaluatoren, kijkt in zulke gevallen extra nauw.

#### *Rapportage*

Soms lijkt het of er meer geschreven dan gelezen kan worden. In zo'n wereld is behoefte aan heldere en beknopte rapportage. Gegeven aanbevelingen voorzien daar lang niet altijd in. Zo bevatten veel evaluaties nogal algemene aanbevelingen als het aanpakken van 'coördinatieproblemen' of de noodzaak om 'eerder op te schalen'. Tot de hinderpalen voor bruikbaarheid behoren ook formuleringen die gevoelige snaren mijden. Het herschrijven van stukken totdat ze aanvaardbaar zijn voor alle betrokkenen, kan leiden tot zodanige algemeenheden dat het onduidelijk is of aangedragen oplossingen wel passen bij de onderliggende problemen. In dit verband moet gewaarschuwd worden voor 'writing that is written by many hands' (Drenth von Februar, 2006; vergelijk Vining, 1986).

### **Conclusie en discussie: naar een systematische, op waarden en effectiviteit georiënteerde evaluatieagenda**

Het belang van evaluatieonderzoek wordt breed onderkend, maar er is een paradoxaal spanningsveld tussen toenemende aandacht voor verwetenschappelijk van beleid enerzijds, en de groeiende vraag naar politieke sensitiviteit en strategische vaardigheden van ambtenaren anderzijds.

Binnen deze context moet beleidsevaluatie stevig verankerd zijn om haar onafhankelijke en onpartijdige rol te kunnen vervullen. Dat is van belang voor de validiteit en toegevoegde waarde van evaluaties, maar ook voor het vertrouwen in het openbaar bestuur en het gevoerde beleid. Ook publieke transparantie over het oogmerk, de onafhankelijkheid en kwaliteit van evaluatieonderzoek is daarvoor essentieel. In het geschetste spanningsveld biedt hoge evaluatiekwaliteit houvast, inclusief goede bruikbaarheid van aanbevelingen. Die bruikbaarheid dient te gelden voor zowel het politiek-bestuurlijke als professionele praktijkveld.

Beleidsvaluatie kan niet los worden gezien van bredere publieke doelen en maatschappelijke waarden als vrijheid, gelijkheid, veiligheid, rechtvaardigheid en welzijn, waartoe veel beleid en wetgeving immers te herleiden is. Dat inzicht biedt houvast als het gaat om de vraag wat bij beleidsvaluaties de hoofd- en bijzaken zijn. Niet alleen vormt dit een nuttig tegenwicht tegen mogelijke vertekening ten nadele van minder gemakkelijk meetbare maar essentiële waarden en uitkomsten, het kan ook bijdragen aan het verbinden van politieke en wetenschappelijke rationaliteit.

Binnen een op systematische kennisontwikkeling gerichte, evenwichtig evaluatieportfolio is het mogelijk breder geldende – eventueel vanuit het parlement aangevraagd – waarden extra aandacht te geven, ook gespreid over meerdere evaluaties. Zo'n bredere benadering houdt bovendien in dat beleidsvaluaties niet alleen van betekenis zijn voor het specifieke onderwerp waarop zij primair gericht zijn, maar kunnen worden geplaatst in een bredere context van systematische en evenwichtige kennisopbouw ten behoeve van voortgaande beleidsverbetering. Dat helpt ook om onnodig en overmatig evalueren in vergelijkbare dossiers tegen te gaan.

Dit geldt ook voor ex ante beleidsevaluatie: kennis over werkingsmechanismen van beleidsinterventies is vaak van toepassing op meer terreinen en draagt daarmee bij aan een doelmatige inzet van nieuw beleid. Zulke kennis maakt het ook beter mogelijk de evaluatie-energie te concentreren op de essentiële kwesties waarover vooraf discussie of onzekerheid bestaat. Het is immers onmogelijk en vaak ook niet nodig om alle elementen van een beleidsinterventie wetenschappelijk te evalueren.

Een goed gekozen evaluatieportfolio kan voorts bijdragen aan een kennisklimaat dat via een 'pars pro toto effect', zowel intra- als interdepartementaal, kwaliteitsbevorderend werkt richting minder belichte terreinen. Ook dat draagt bij aan een doelmatige benutting van beschikbare kennis en expertise.

Systematische opbouw van kennis en ervaring dient uiteraard ook kennis en leerpunten over minder geslaagd beleid te omvatten, want juist ook daarvan is veel te leren. Dat geldt te meer naarmate de redenen waarom iets niet werkte kan worden verbonden met actuele beleidsuitdagingen. De waarde die alom gehecht wordt aan *best practices*, hoe belangrijk en terecht ook, mag daarom niet leiden tot

het wegzijden van ineffectieve regelgeving en beleidspraktijken en de factoren die daaraan hebben bijgedragen.

Dat evaluaties niet op zichzelf moeten (blijven) staan betekent niet alleen dat zij moeten worden gerelateerd aan evaluatie-uitkomsten in andere beleidsdomeinen maar ook aan de internationale stand van kennis. Dat kan gepaard gaan met vergelijkende analyse gericht op overeenkomsten en verschillen met Nederland. Juist daar waar beleidsmakers in de 'heat of the moment' niet gauw naar de laatste wetenschappelijke gegevens zullen grijpen, kan zo'n informatievoorziening belangrijke uit de literatuur gedestilleerde inzichten en conclusies aanreiken. Welke wetenschappelijke benaderingen zijn er? Welk type kennis leveren ze op? Wat is daarvan bruikbaar in de nationale beleidscontext?

Een evenwichtige evaluatieportfolio, waarin individuele evaluaties worden gezien als deel van een geheel – vergroot de kans dat nieuwe evaluaties input kunnen leveren voor het geheel aan kennis (*body of knowledge*) op verschillende niveaus van beleid: feitelijke details (bijvoorbeeld ten behoeve van de verbetering van specifieke interventies), grote lijnen (zoals bijvoorbeeld wetgeving of beleidsprogramma's), en de samenhang tussen grote lijnen – de bredere beleidsstrategie. Die ontwikkeling kan worden bevorderd door aan elke specifieke evaluatie ook een standaarddeel te koppelen dat gericht is op een meer generieke kennisopbouw. Het is dan zaak er bij de methodologische inrichting en de rapportage van evaluaties ook rekening mee te houden dat ze valide en reproduceerbare input kunnen leveren voor systematic reviews.

Al met al zien wij een belangrijke agenda voor een overheid die niet alleen casus-specifiek maar vooral ook systematisch wil leren. De overheid moet blijven investeren in gedegen beleidsevaluatie, een zowel qua expertise als positionering kwetsbaar kernelement van de beleidscyclus. Om haar voldoende aantrekkelijk te houden voor talentvolle onderzoekers en beleidsmedewerkers is voldoende aandacht, prioriteit en waardering voor beleidsevaluatie cruciaal, evenals daadwerkelijke benutting van haar resultaten.

Een en ander vraagt om goede verankering en positionering binnen de kennisinfrastructuur van de overheid en effectieve samenwerking met ten aanzien van beleidsevaluatie actieve kennisinstellingen. Ook is van groot belang dat de politiek-bestuurlijke leiding er geen twijfel over laat bestaan dat een goed evaluatievermogen een essentiële pijler is van professioneel openbaar bestuur en een adequaat bemenste organisatie met continuïteit vereist (Klein Haarhuis en Parapuf, 2016). Het interdepartementaal verbinden van evaluatie-expertise helpt daarbij.

## Literatuur

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

Algemene Rekenkamer (2014). *Tendrapport open data*. Den Haag: Sdu Uitgevers.

- Algemene Bestuursdienst (2016). *Competentieprofiel ABD-manager*, [www.algemeenbestuursdienst.nl](http://www.algemeenbestuursdienst.nl). (geraadpleegd op 10 mei 2016).
- Bekker, E.R. (2014). Afscheidscollege 'Een toekomst zonder ministeries'. Albeda Leerstoel Universiteit Leiden/CAOP: <http://www.deleerstoelen.nl/fileadmin/caop/data/Actueel/2014/Verslag-afscheidscollege-Roel-Bekker-en-Roel-Nieuwenkamp.pdf> (geraadpleegd op 10 september 2014).
- Berg, M. van den, Post, N.A.M., Hamberg-van Reenen, H.H, Baan, C.A. & Schoemaker, C.G. (red) (2014) *Preventie in de zorg, Themarapport Volksgezondheid Toekomst Verkenning 2014*. Bilthoven: RIVM.
- Bovens, M, Schillemans, T. (2009). *Handboek publieke verantwoording*. Den Haag: Lemma.
- Brummelkamp, G.W., Kerckhaert, A.C. & Engelen, W.M. (2013). *Het verhaal achter notoire uitkeringsfraude*. Zoetermeer: Panteia.
- Drenth von Februar, M. (2006). De rapporterende burger. Over incidentalisme versus cerebralisering. In: P.L. Meurs, E.K. Schrijvers, G.H. de Vries - *Leren van de praktijk. Gebruik van lokale kennis en ervaring voor beleid*. Wetenschappelijke Raad voor het Regeringsbeleid. Amsterdam: Amsterdam University Press.
- Fischer, F. (1995) *Evaluating Public Policy*. Chicago: Nelson-Hall.
- Gezondheidsraad (2006). Vertrouwen in verantwoorde zorg? Effecten van en morele vragen bij het gebruik van prestatie-indicatoren. In: *Signalering ethiek en gezondheid, jaargang 2006, nr. 1*. Den Haag: Centrum voor ethiek en gezondheid.
- Gezondheidsraad (2013). *Publieke indicatoren voor kwaliteit van curatieve zorg. De stand van de discussie*. Den Haag: Gezondheidsraad.
- Guba, E.G. & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. London: Sage.
- Hanemaayer, D. (2013). *Ex ante evaluatie in Nederland*. Den Haag: Boom Uitgevers.
- 't Hart, P. (2014). *Ambtelijk Vakmanschap 3.0: Zoektocht naar het handwerk van de overheidsmanager*. Utrecht: Universiteit Utrecht & Nederlandse School voor Openbaar Bestuur.
- Janssen, R. (2015). *The art of audit*. Amsterdam: Amsterdam University Press.
- Jonkers, R. (2013). Zet transparantie liever in voor bekritiseerbaarheid dan voor vertrouwen. In: Broeders D, Prins JEJ, Griffioen H, Jonkers P, Bokhorst M, Sax M (red.). *Speelruimte voor transparantere rechtspraak*. Wetenschappelijke Raad voor het Regeringsbeleid. Amsterdam University Press, 2013: pp 449-471.
- Klein Haarhuis, C.M. & Niemeijer, E. (2008). *Wet en werkelijkheid. Bevindingen uit evaluaties van wetten*. Den Haag: Boom Juridische uitgevers.
- Klein Haarhuis, C.M. & Parapuf, A. (2016). *Evaluatievermogen bij beleidsdepartementen. Praktijken rond uitvoering en gebruik van ex post beleids- en wetsevaluaties. Concept Eindrapportage*. Den Haag: WODC.
- Knaap, Peter van der (2006a). Theoriegerichte beleidsevaluatie. In: *Bestuurskunde*, jaargang 2006, nr. 4, pp. 78-87.
- Knaap, P. van der (2006b). Responsive evaluation and performance management - overcoming the downsides of policy objectives and performance Indicators. In: *Evaluation*, Vol. 12, No. 3, pp. 278 - 293.
- KNAW (2005). *Wetenschap op bestelling. Over de omgang tussen wetenschappelijk onderzoekers. en hun opdrachtgevers*. Amsterdam: KNAW.
- Korsten, A.F.A. & Mevissen, J. (2015). Risico's en remedies. Kritiek op beleidsonderzoek en wat ervan te leren, in: P.H.M. van Hoesel, J. Mevissen en L.H. Dekker (red.). - *Kennis voor beleid - Beleidsonderzoek in Nederland*. Assen: Van Gorcum.
- Korsten, A.F.A. (2013) *Wat ervan terecht komt. Zicht op beleidsevaluatie*. <http://www.arnokorsten.nl/PDF/Onderzoek/ABC%20van%20beleidsevaluatie.pdf> (geraadpleegd op 8 april 2014).

- Leeuw, F.L. (2008). *gedragsmechanismen achter overheidsinterventies en rechtsregels*. Universiteit Maastricht.
- Leeuw, F. & Mevissen, J. (2015). Beleidstheorieën, gedragsmechanismen en evaluatieonderzoek. In: Van Hoesel, P.H.M., J. Mevissen en L.H. Dekker - *Kennis voor beleid. Beleidsonderzoek in Nederland*. Assen: Van Gorcum.
- Lehning, P.B. (1991), *Beleid op niveau. Over de architectuur van overheidsinterventie*. Meppel/Amsterdam: Boom.
- Nelen, H. (2008) *Evidence maze; het doolhof van het evaluatieonderzoek*. Universiteit Maastricht.
- Nelen H, Leeuw, H.B.M., Bakker, F. & Herrenberg, T. (2012). *In de war en uit de bocht; ex-ante evaluatie van de aanpak van solistische dreigers binnen de pilot Dreigingsmanagement*. Den Haag: Boom/Lemma.
- Nieuwenkamp, R. (2013). *Schaduwpolitici, bontkragen en blokkendozen. Een onderzoek naar de politiek-ambtelijke betrekkingen in Nederland*. Amsterdam: Universiteit van Amsterdam.
- Noordergraaf, M. & Wit, B. de (2012). *Van maakbaar naar betekenisvol bestuur. Een achtergrondstudie naar (keten)governance en (nieuw) publiek management en de gevolg voor toezicht en evaluatie*. WRR-webpublicatie nr. 63. Den Haag: WRR.
- Noordegraaf, M. & Abma, T.A. (2003). Management by measurement? Public management practices amidst ambiguity. In: *Public Administration*, Vol. 2003/81, Nr. 4, pp. 853-871.
- Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.
- Petticrew, M. & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. Malden: John Wiley & Sons.
- Rouw, R. (2011). *Gevoel voor bewijs. Naar vloeiende verbindingen tussen kennis en beleid*. Den Haag: NSOB.
- Ryan, B. (2003). 'Death by evaluation'?: Reflections on monitoring and evaluation. In: *Australia and New Zealand. Evaluation Journal of Australasia*, 2003/ Vol. 3.; nr. 1, pp. 6-16.
- Snellen, I.Th.M. (2007). *Grondslagen van de bestuurskunde. Een essay over haar paradigma's*. Den Haag: Lemma.
- Stern, E. (2011). Editorial. In: *Evaluation, jaargang 2011, Vol. 17, nr. 3, pp. 213-15*.
- Veerman, G.J., m.m.v. R.J. Mulder & E.S.M. Meijsing (2013). *Een empathische wetgever. Meta-evaluatie van empirisch onderzoek naar de werking van wetten*. Den Haag: Sdu Uitgevers.
- Vining, J. (1986). *The Authoritative and the Authoritarian*. Chicago: University of Chicago Press.
- Wildavsky, A. (1979). *Speaking truth to power. The art and craft of policy analysis*. New Brunswick: Transaction Publishers.
- WRR (Wetenschappelijke Raad voor het Regeringsbeleid) (2006). *Lerende overheid: een pleidooi voor probleemgerichte politiek*. Amsterdam: Amsterdam University Press.
- WRR (2013). *Toezien op publieke belangen. Naar een verruimd perspectief op rijkstoezicht*. Amsterdam: Amsterdam University Press.
- WRR (2014a). *Consistent maatwerk – handreikingen voor dossieroverstijgend risico- en veiligheidsbeleid*. Amsterdam: Amsterdam University Press.
- WRR (2014b). *Met kennis van gedrag beleid maken*. Amsterdam: Amsterdam University Press.
- WRR (2016). *Big Data in een vrije en veilige samenleving*. Amsterdam: Amsterdam University Press.