

Evaluëren = Leren

Deskundigen aan het woord



Evaluëren = Leren

Deskundigen aan het woord



Verslag van drie gesprekken die de Zuidelijke Rekenkamer voerde met externe deskundigen en marge van haar vervolgonderzoeken naar beleidsevaluatie in de provincie Limburg en Noord-Brabant

Eindhoven, juni 2012

Inleiding

En marge van haar vervolgonderzoeken naar de stand van zaken van het gebruik van het instrument beleidsevaluatie door de provincie Limburg en Noord-Brabant heeft de Zuidelijke Rekenkamer in het voorjaar van 2012 gesprekken gevoerd met drie externe deskundigen. Doel van deze gesprekken was om van hen te vernemen wat de stand van het denken over beleidsevaluatie is en voor welke uitdagingen provincies staan die de effectiviteit en efficiëntie van hun functioneren en presteren willen onderzoeken.

De neerslag van deze gesprekken bieden wij hierbij als een aparte publicatie aan onder de titel 'Evalueren = Leren'. Wij vertrouwen erop dat de vele suggesties, kanttekeningen en commentaren van de geïnterviewden betrokken zullen worden bij de verdere vormgeving van beleidsevaluatie door de beide provincies.

De gesprekken vonden plaats rondom de volgende vragen:

- Wat zijn recente ontwikkelingen en trends op het terrein van beleidsevaluatie?
- Hoe veranker je (de aandacht voor) beleidsevaluatie in je organisatie?
- Hoe zorg je ervoor dat de evaluaties kwalitatief aan de maat zijn, hoe geef je een meerjarig onderzoeksprogramma vorm, hoe ga je om met vraagstukken van uitbesteding versus zelf uitvoeren?
- Wat moet je doen om de uitkomsten van evaluaties echt te benutten?
- Wat is de rol voor de volksvertegenwoordiging op het terrein van beleidsevaluatie, hoe organiseer je de aandacht van de volksvertegenwoordiging/ Provinciale Staten voor het onderwerp, hoe kunnen zij sturing geven, hoe zie je het terug in de formele documenten van planning & control?
- Welke vorm van beleidsevaluatie past bij een beleid dat voortdurend in beweging is en waar naast de provincie vele andere actoren in de ontwikkeling en uitvoering van gezamenlijke programma's en projecten actief zijn (Europa, rijksoverheid, gemeenten, bedrijfsleven, maatschappelijke organisaties)?

De drie deskundigen hebben een langdurige expertise op het terrein van beleidsevaluatie. Hun deskundigheid bestrijkt de wereld van onderzoek, wetenschap en publieke sector.

Het zijn Peter van der Knaap, Froukje Boonstra en Frans Leeuw.

Peter van der Knaap: directeur doelmatigheidsonderzoek bij de Algemene Rekenkamer. In een eerdere functie was hij onder meer hoofd Beleidsevaluatie en Instrumentatie van het ministerie van Financiën en verantwoordelijk voor de rijksbrede kaderstelling en uitvoering op het vlak van evaluatieonderzoek en resultaatgericht begroten (VBTB). In 1997 promoveerde hij aan de Erasmus Universiteit Rotterdam op het proefschrift 'Lerende overheid, intelligent beleid', een onderzoek naar de bijdragen van beleidsonderzoek aan het leervermogen van de overheid.

Froukje Boonstra: gepromoveerd op een onderzoek naar besluitvorming in gebiedsgericht beleid aan de toenmalige Katholieke Universiteit Nijmegen. Was coördinator beleidsondersteunend onderzoek van het toenmalige ministerie van Landbouw, Natuur en Voedselkwaliteit voor het thema ILG (Investeringsbudget Landelijk Gebied) en gebiedsprocessen. Momenteel senior bestuurskundige bij Alterra, kennisinstituut voor de groene leefomgeving van Wageningen University & Research centre, waar zij zich richt op het ondersteunen van bestuurlijke vernieuwing in het landelijk gebied, via onderzoek, advies en procesfacilitatie.

Frans Leeuw: directeur van het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het ministerie van Veiligheid en Justitie. Daarnaast is hij hoogleraar 'Recht, Openbaar Bestuur en Sociaal-wetenschappelijk Onderzoek' aan de Maastricht University. In eerdere functies was hij onder meer hoogleraar aan de Universiteit Utrecht met als leeropdracht 'Onderzoek van maatschappelijke effecten van overheidsbeleid', directeur doelmatigheidsonderzoek bij de Algemene Rekenkamer en Hoofdinspecteur, Inspectie van het Onderwijs.

Peter van der Knaap

10 mei 2012

We leven in een netwerksamenleving waarin beleidsinterventies van allerlei partijen over elkaar heen buitelen. Dat heeft consequenties voor onderzoek naar maatschappelijke effecten van beleid en vraagt om een 'slimme' evaluatiebenadering. Een benadering waarbij je steeds opnieuw voor de vraag komt te staan wat voor soort evaluatieonderzoek, in welk stadium, gegeven de specifieke omstandigheden, het meest geëigend is.

Als je voor je beleid afhankelijk bent van andere actoren, dan kun je in ieder geval beginnen met de vraag hoe de processen zijn ingericht, hoe de overdracht van bijvoorbeeld dossiers tussen de betrokken instanties is geregeld. Dat soort onderzoek moet eigenlijk vast onderdeel van het beleidsproces zijn. Juist bij veel schakels in de keten moet je alert zijn op de verhouding tussen de schakels onderling. De zwakste schakel bepaalt immers de kracht van de ketting. Dat betekent dat je aan de voorkant goed nadenkt hoe je het beleid inricht en er op bedacht bent welke flessenhalzen je kunt tegenkomen. En als je het beleid vervolgens aan het uitrollen bent, moet je in de gaten houden of, en waar de knelpunten zich in de praktijk voordoen. Dat soort inzichten heb je nodig om te kunnen bijsturen. Heb je dat niet, dan vraag je om problemen.

Een aantal jaren geleden zijn we bij de Algemene Rekenkamer begonnen met een ander soort onderzoek waarmee we achteraf heel gericht kijken naar verklaringen voor behaalde resultaten. Daaruit hebben we geleerd dat dit soort informatieproblemen in netwerken en samenwerkingsverbanden een belangrijke verklaring vormen voor achteraf tegenvallende resultaten. Dat onderstreept dus het belang van een *ongoing evaluation* of evaluatie *ex durante*. Je hebt dat nodig om goed te kunnen sturen. Begrijp me goed, ik vind niet dat het in de plaats mag komen van effectevaluaties. Het gaat er alleen aan vooraf, als een noodzakelijke voorwaarde voor goede sturing.

Heilige graal

Er is natuurlijk veel op te merken over voor- en nadelen van verschillende vormen van evaluatie. De heilige graal, dat zijn *randomized controlled trials*. Je hebt een voormeting en een nameting en je hebt een experimentele groep en een controlegroep. In tegenstelling tot wat veel mensen denken kun je de beleidspraktijk vaak wel zo inrichten, het ministerie van Onderwijs, Cultuur en Wetenschap is wat dat betreft goed bezig. Maar dat kan niet altijd en het kan ook kostbaar zijn. Je moet in veel gevallen wachten voordat je weet wat eruit komt. En je kunt erover discussiëren of het altijd maatschappelijk aanvaardbaar is. Er zijn tal van argumenten om het niet te doen, maar toch kom je het soms tegen en dat getuigt van moed en visie. Als Algemene Rekenkamer doen wij het niet, omdat wij geen beleid maken.

Wat wij wèl doen, is achteraf kijken, bijvoorbeeld naar de verschillen in kosten-effectiviteit van beleidsinstrumenten. Zo hebben we laatst onderzoek gedaan naar de kosteneffectiviteit van verschillende maatregelen op het gebied van energiebesparing.

Geïnformeerd

Het is verstandig om kleinschalig te beginnen. De commissie Dijsselbloem, die het parlementair onderzoek deed naar de onderwijsvernieuwingen, kwam ook met dat advies. Maar al te vaak wordt nieuw beleid heel breed uitgerold, zonder dat ook maar iets bekend is over de vraag of het werkt of niet. De conclusie die je daaraan moet verbinden, is: ga in ieder geval goed om je heen kijken wat er over een onderwerp al bekend is. Dat heet met een sjiek woord *evidence informed* beleid. Zodat wat je doet minimaal is gestoeld op aannames die geworteld zijn in kennis over wat wel en niet werkt. En richt het dan als het even kan zo in dat je zinvolle opmerkingen kunt maken over de effectiviteit.

Op dat punt kan er veel meer dan vaak wordt gedacht. Een voorbeeld: je hebt een regeling die van toepassing is op mensen van 55+ en je weet na verloop van tijd hoeveel mensen daar gebruik van hebben gemaakt. Dan kun je vervolgens, tenminste als je een grote populatie hebt, met wat correcties nagaan hoe het zit met 53-jarigen en 54-jarigen. En kun je uitspraken doen over de kans dat ook zij van die regeling gebruik zouden willen maken. Het is een vorm van *impact evaluation*. Hessel Oosterbeek van de Universiteit van Amsterdam is iemand die daarin is gespecialiseerd. Eigenlijk zijn het natuurlijke experimenten, weliswaar niet als zodanig opgezet, maar met een onderzoeksbil kun je er wel zo naar kijken. Het is ook een vorm van evaluatie die veel meer toepasbaar is dan de klassieke vooropgezette controlegroep of experimentele groep; minder wetenschappelijk gevalideerd, maar je kunt er een heel eind mee komen en ook nog eens tegen minder kosten. Hetzelfde geldt voor *syntheseonderzoek*, dat verschillende (deel)onderzoeken combineert. Dan kijk je wat er elders al gedaan is, wat de ervaringen zijn met programma's die wel of niet tot succes hebben geleid.

Misschien kun je zeggen dat dat voor provincies allemaal wat hoog gegrepen is. Maar vergis je niet, op het gebied van ruimtelijke ordening en openbaar vervoer bijvoorbeeld, is er op provinciaal niveau zeer veel informatie beschikbaar. De inzichten die je daar uit kunt halen, zou je al dan niet kunnen bundelen binnen het Interprovinciaal Overleg of in andere samenwerkingsverbanden. Vandaag de dag hoeft je ook niet meer de fysieke bibliotheek in, je hoeft het niet allemaal zelf te onderzoeken of in een experimentele vorm te gieten. Het is van het grootste belang om vooral in de voorkant van het beleid te investeren. Dat is de beste manier

om enige mate van zekerheid te krijgen dat de interventie die je wilt plegen in ieder geval aangrijpt op de belangrijkste factoren.

Verkeersveiligheid is in dat verband een aardig voorbeeld. We weten uit onderzoek dat decentrale wegbeheerders niet altijd de meest moderne inzichten toepassen bij de inrichting van rotondes. Terwijl die informatie er wel is. Wat mij betreft is er geen excuus om die beschikbare kennis niet toe te passen. Je hoeft echt geen tonnen uit te geven om aan dat inzicht te komen, dat is gewoon beschikbaar.

Zinvolle evaluaties beginnen met serieuze aandacht voor wat je wilt bereiken. *Factfull politics*, geïnformeerde besluitvorming en het zodanig vormgeven en uitwerken van je beleid dat je je niet achteraf vertwijfeld hoeft af te vragen wat nou eigenlijk ook alweer het doel van het beleid was en hoe je dat kunt operationaliseren. Dat moet je voor een belangrijk gedeelte al uit je bestaande stuurinformatie kunnen halen. Dat betekent dat als het over verkeersveiligheid gaat, dat je vanaf het begin ongevallen bijhoudt en dat je op basis daarvan kunt zeggen of een bepaald beleid heeft geleid tot een significante daling van ongelukken.

Een andere belangrijke vraag is hoe je beter kunt nagaan wat het specifieke effect is van de inzet van jouw middelen als provincie, temidden van middelen die door andere partijen zijn ingebracht, zoals Europa, rijksoverheid, gemeenten, bedrijfsleven en maatschappelijke organisaties. Succes heeft vele vaders. Als het fout gaat, ligt dat anders. Tegelijkertijd, als je pas achteraf ontdekt dat andere ketenpartners niet hebben geleverd en dat je daarom bepaalde doelen niet hebt gehaald, dan heb je zelf je werk al die tijd niet goed gedaan. Dan is het alsof je aan de lopende band staat en zegt: ik breng de spatborden aan en of de auto aan het eind van de productielijn ook echt kan rijden, daar ga ik niet over. Modern bestuur kan zo niet werken. Ook al ben je niet de regisseur van de keten - wat de provincie overigens vaak wel is - dan nog heb je een verantwoordelijkheid om in de gaten te houden hoe het met het geheel zit. Op het moment dat de andere ketenpartners verzaken of het beoogde effect niet wordt bereikt, dan zul je toch echt je hand moeten opsteken. En al helemaal als je de regie wèl hebt; dan kun je niet achteraf zeggen: we constateren nu vier of vijf jaar later dat het niet is gelukt, maar by the way ... het ligt niet aan ons maar aan 'hullie'. Dat kan niet. Je zult zeker moeten stellen dat de keten in zijn geheel goed werkt, anders ben je in je reguliere sturing al niet goed bezig.

Focus op beleidsresultaat

Langzamerhand zien we veranderingen in het soort onderzoek dat we doen. Dat heeft te maken met veranderingen in de beleidsomgeving. Dat er één partij is die

op basis van zo goed mogelijke informatie kan besluiten wat er moet gebeuren, dat vervolgens in zijn eentje kan uitvoeren en achteraf ook nog zelf vaststelt of het heeft gewerkt, dat kom je in de beleidspraktijk bijna niet meer tegen.

Maar deze ontwikkeling lijkt ook wel een 'excustruus' te worden voor overheden en bestuurders waar ze zich achter kunnen verschuilen. Hoe meer partijen een gedeelde verantwoordelijkheid hebben, hoe minder vaak er eentje is die zegt: ik ga er over. Op allerlei terreinen waar sprake is van decentralisatie, lijkt de rijksoverheid zich terug te trekken op een soort systeemverantwoordelijkheid: wij zijn er voor de voorlichting, voor het overmaken van het geld en misschien nog voor het maken van afspraken. Maar wat er uiteindelijk bereikt wordt, dat is de verantwoordelijkheid van andere overheden, van Brussel of van uitvoerende instanties in het veld.

Dat is natuurlijk een vreemde figuur. In de Tweede Kamer gaat het niet over een convenant dat je hebt afgesloten, maar over het beleidsresultaat zelf, over zwerfafval of over voedselveiligheid. Dan kun je nog zo hard roepen dat zwerfafval iets is voor de gemeenten, maar als je er geld aan uitgeeft en je hebt ambities en je kiest voor dat systeem, dan mag je ook verwachten dat je daar verantwoording over aflegt. Vaak hoor je dan: als we de samenwerking evalueren of we evalueren de mate waarin partners tevreden zijn, dan is dat ook voldoende. Nee dus, dat is nooit voldoende. Ewald Engelen, de financieel geograaf, zegt in dat verband: hoe ingewikkelder het systeem, hoe meer 'beleidslulkoek', waarbij het alleen nog maar gaat om het afsluiten van convenanten en het beleggen van meetings, terwijl het echte onderwerp, zoals bijvoorbeeld het revitaliseren van de lokale economie, naar de achtergrond verdwijnt. En als je dan ook nog eens de procesevaluatie zo inricht dat het op de achtergrond blijft, dan is iedereen heel tevreden, maar je weet één ding zeker: de burger is al lang afgehaakt. En je weet ook niet, en dat is misschien nog wel kwalijker, of wat je doet enig effect heeft. Je zult dus altijd de resultaten moeten kunnen beoordelen.

Dat zien wij als een ontwikkeling voor ons als Algemene Rekenkamer. Dat wij blijven hameren op effecten en daar onderzoek naar doen. En voor onszelf, als instelling die evaluatieonderzoek doet, is de belangrijkste ontwikkeling dat we steeds meer de ontvanger van beleid centraal stellen. Of dat nou bedrijven zijn of burgers, of de natuur, Waarbij de centrale vraag is: wat is het probleem en sluit het beleid aan op dat probleem?

Een voorbeeld: vorig jaar hebben we onderzocht of maatregelen, die waren bedoeld om de veiligheid van winkeliers in het midden- en kleinbedrijf te verbeteren,

of die aansluiten bij wat die winkeliers zelf belangrijk vinden. Dat is een nieuw type onderzoek. We hebben niet zozeer gekeken naar de formele doelstellingen van het Rijk of van veiligheidspartners. Nee, we hebben gekeken hoe het zit met de ontvanger. Kan hij ermee uit de voeten? En vervolgens: wat kunnen we zien wat er in de praktijk gebeurt? Daalt het aantal overvallen? We kijken dus niet alleen of het aansluit, maar ook naar de feiten die daarbij horen. We pellen niet meer de beleidstoren van boven naar beneden af, zoals we altijd hebben gedaan, maar we draaien het om. We proberen naar het beleid te kijken door de ogen van de mensen waar het beleid zich op richt. Dat is een vorm van *responsieve beleidsevaluatie*. De doelgroep van beleid heeft natuurlijk maling aan convenanten en kwaliteitsafspraken. Hen gaat het om het effect. Het is een minder formalistische evaluatie, waarbij je ook open moet staan voor dingen die misschien niet zo bedoeld zijn maar in de praktijk wel goed blijken te werken. Voor rekenkamers en traditionele evaluatoren blijkt dat overigens een lastig waar te maken ambitie.

Organiseren van aandacht

Hoe organiseer je de aandacht voor evaluatie in de bredere zin? Ik denk dat het helemaal onderaan begint bij de kwaliteit van beleidsmedewerkers en een morele verantwoordelijkheid van bestuurders. Soms gaat het daar al mis. Dan zie je bestuurders die vooral geïnteresseerd zijn om de wereld te laten zien waartoe ze zo al in staat zijn en die geen behoefte hebben om na te gaan of datgene wat ze doen ook echt iets uithaalt.

Wat betreft de beleidsmedewerkers kom je bij de vraag of je op de wereld bent om problemen te helpen beheersen of op te lossen, of dat je er bent om je bestuurder te dienen. Idealiter ligt dat in het verlengde van elkaar. Dan is er ook een goede voedingsbodem om je aan de voorkant al bezig te houden met een inhoudelijke vraag als: wat is er eigenlijk bekend – als dit het probleem is dat we willen beheersen of kleiner willen maken – over instrumenten die succesvol zouden kunnen zijn? Om dat te kunnen doen heb je voldoende capaciteit op niveau nodig. Daar maak ik me wel eens zorgen over, zeker bij decentralisaties. Op het terrein van natuurbescherming constateerden we bijvoorbeeld een aantal jaren geleden dat provincies en gemeenten heel weinig mensen hadden met voldoende verstand van natuurbescherming, terwijl zij wel projectaanvragen moeten beoordelen. Dan staat de kwaliteit van de beoordeling onder druk en loopt de natuurbescherming gevaar. Je *standard operating procedures* moeten natuurlijk wel op orde zijn.

Durven leren

Wat ontzettend helpt, is het gebruik van nieuwe media en het organiseren van transparantie. Zodat iedereen met een paar simpele muisklikken kan opvragen wat

er over een onderwerp al bekend is. Het begint bij *evidence informed* besluitvorming. Als je dat niet op orde hebt, dan kun je organiseren wat je wil, je kunt zaken verplicht stellen, maar dan blijf je altijd aan de achterkant bezig. Dat is misschien wel dankbaar werk voor een rekenkamer, maar het is natuurlijk niet de goede manier om duurzaam aandacht voor evaluatie te krijgen.

Het organiseren van interne tegenspraak is in dat verband ook erg belangrijk. Dat gaat ons over het algemeen niet gemakkelijk af. Er is heel lang geoefend met goed verantwoording afleggen. We zijn inmiddels weg uit de fictie dat beleid altijd succesvol zal zijn. Maar naast verantwoording afleggen is het ook van belang dat overheden leren. Ik denk dat de volgende stap is dat je verantwoording zo interpreteert dat verantwoording afgelegd moet worden over het leerproces. Dat betekent ook dat op het moment dat je niet weet of iets werkt of je weet het misschien wel maar in de praktijk blijkt het toch anders te gaan, dat dat dan niet per definitie als een brevet van onvermogen wordt uitgelegd. Bestuurders moeten zich ook verantwoorden over hoe ze ervoor zorgen dat ze dingen die mis gaan tijdig in beeld krijgen. Dat ze niet na vijf jaar volledig verrast worden door een onverwachte boodschap, als ze een evaluatierapport van een extern bureau op hun bureau krijgen.

We zien dat organisaties de neiging hebben om weg te blijven uit een scherpe verantwoording. Dat leidt tot koppensnellen en daar houden we niet van. Een kritische vraag aan een collega: 'wat een mooi project, maar helpt het nou ook?' wordt meestal ook niet gesteld. Terwijl je je daarvoor juist open zou moeten stellen vanuit een grondhouding: 'we weten het niet, maar we proberen het want het is belangrijk dat we er wat aan doen. En als blijkt dat wat we doen niet werkt, dan zijn we daar open in en gaan we het anders proberen. Maar wees er zeker van, beste inwoners van onze provincie: wat we hebben gedaan was wel geïnformeerd, het was niet zomaar iets. Het was *state of the art*.'

Dat is een manier om uit die kramp te komen. Dat vergt ook veel van toezichthouders, rekenkamers en evaluatoren. Die beginnen altijd met iets dat tegenvalt en merken daar zelden bij op dat het ook erg complex is. En dan constateren ze vervolgens dat de informatievoorziening niet op orde was. Als de informatievoorziening wel op orde is en je kunt laten zien: 'dit is wat we wisten en toen hebben we dat gedaan om te zorgen dat er dit en dit gebeurde' dan sta je ook veel sterker als provinciaal bestuur. Beleidsevaluatie kan dan helpen om heel gericht verantwoording af te leggen, omdat je kunt laten zien wat je hebt gedaan om ervoor te zorgen dat je zo goed mogelijk bezig bent met het belang van de burger.

Programmeren

Overheden worstelen met het formuleren van langjarige onderzoeksprogramma's. De Amerikanen spreken in dat verband vaak over *monitoring and evaluation*. Verantwoord bestuur begint bij beleidsprogramma's die hun informatiehuishouding op orde hebben: wat is het probleem? hoe groot is het probleem? wie zijn er door geraakt? wat zijn de oorzaken? wat kan ik doen om te zorgen dat ik op die oorzaken ingrijp? wat moet dat dan zijn ten opzichte van de probleemgrootte? hoe ontwikkelt zich dat van kwartaal tot kwartaal en van jaar tot jaar? Op zo'n moment vormt een evaluatie een logisch slotstuk. Met natuurlijk als moeilijkste punt: is de verandering exclusief aan de interventie toe te schrijven?

In de praktijk gaat het meestal anders. Dan blijkt er een evaluatieverplichting te zijn, soms pas op het allerlaatste moment ontdekt. Dan is het helemaal erg als ook nog eens de doelstelling moet worden gereconstrueerd en men tot de ontdekking komt dat allerlei informatie niet voor handen is waar men al lang en breed over had moeten beschikken als ze er gewoon normaal van dag tot dag mee bezig waren geweest. Er worden dan in allerlei interviews met *stakeholders* gehouden, vaak over proceszaken, en dat is het dan.

Een verantwoord bestuurder moet zich niet alleen afvragen: 'doe ik het zo goed mogelijk?', maar ook: 'doe ik het goede?' Wat ik doe, haalt dat wat uit? Je kunt natuurlijk heel goed een subsidieregeling uitvoeren, keurig netjes de beschikkingen op tijd rond krijgen, met een mooie klachtenprocedure erom heen, waardoor iedereen op dat punt tevreden is. Maar als het weinig of niets uithaalt en misschien zelfs contraproductief werkt, dan kun je wel zeggen dat je het goed doet, maar je doet ondertussen wel het verkeerde. En als je dan drie maanden voor het aflopen van de periode nog eens ontdekt, o ja we moeten ook nog evalueren op het nut, dan ben je geen knip voor je neus waard.

Als je wilt programmeren, dan moet je aansluiten bij je bestuurlijke prioriteiten en van daaruit vaststellen wat je in welke volgorde gaat evalueren. Het helpt wel als je een checklist hebt met regelingen die aan de beurt zijn, maar de intrinsieke interesse moet zijn: wat vind ik als provinciebestuur het belangrijkste om te realiseren? Op die terreinen wil ik ook graag laten zien dat ik de middelen zo goed mogelijk inzet, dat ik dat monitor en dat ik daar de effecten van evalueer. Daarmee komt het uit de verplichtingsfeer en past het veel meer bij het verhaal dat je richting burger wilt vertellen.

Uitbesteden

Ik ben een groot voorstander van het in eigen beheer uitvoeren van evaluaties. Ik

heb dat op het ministerie van Financiën zelf gezien. Het interne leereffect is veel groter dan wanneer je ze uitbesteedt. Het is een stimulans voor mensen om dingen zelf te doen. Het geeft een mooie kans voor benutting van de resultaten zowel tijdens als na het onderzoek. Ook de kosten kunnen veel lager zijn. Natuurlijk hangt dat wel af van de manier waarop je het organiseert. Het standaard in huis hebben van allerlei expertise die je eens in de zoveel jaar gebruikt, dat is natuurlijk onzin. Voor bepaalde methodologische aspecten is het verstandig om zaken in te huren of gebruik te maken van kennis en expertise die al beschikbaar is.

Te vaak wordt bij uitbesteding uitgegaan van de fictie dat het onafhankelijk is. Dat klinkt natuurlijk aantrekkelijk. Maar ik heb ook vaak evaluaties uitbesteed en dan krijg je soms het voorstel om maar vast op te schrijven wat er uit moet komen. Hoezo onafhankelijkheid? Wie betaalt, bepaalt dan in hoge mate. Goede onderzoekskantoren zullen dat overigens nooit doen of in ieder geval nooit toegeven. Maar zo groot is Nederland niet en wat dat betreft hebben wij ook wel ervaren dat het lastig is om onderzoeksbureaus te vinden die niet al eerder betrokken zijn geweest bij het beleid dat je wilt evalueren.

Je kunt externe onafhankelijkheid ook organiseren met een begeleidingscommissie of een *second opinion*. Je kunt je prima voorstellen dat je een beleidsdirectie of een organisatie zelf laat evalueren, maar dat je daar een begeleidingscommissie boven op zet, bestaande uit mensen van onbesproken gedrag, kennis, deskundigheid en afstand. Die kijken als eerste of de goede vragen gesteld worden. Als je alleen wilt weten of partners een bepaald beleid waarderen, dan is dat heel wat anders dan wanneer je naar effecten vraagt. Een begeleidingscommissie kan naar de opzet van de evaluatie en de vraagstelling kijken, kan de methode beoordelen, maar kan ook wegen of wat er uit komt als valide beschouwd kan worden en zich uitspreken over evenwichtigheid in de rapportage.

Benutting

Zelf evalueren biedt een goede kans op benutting van de resultaten tijdens en na het onderzoek. Daarbij moet je wel in de gaten houden of het belangrijk is dat de uitkomsten via de volksvertegenwoordiging moeten worden geleid. De directe procesbenutting heb je bij een extern bureau veel minder. Daar wordt vaak afgewacht tot het rapport is verschenen, vervolgens komt het op de agenda en wordt er nog weer later eens een keer naar gekeken. Dan is het niet van jezelf – *not invented here* – en ben je minder geneigd ermee aan de slag te gaan.

Veel reacties op evaluaties, zeker als het kritische evaluaties zijn, zijn afhoudend. Het is niet goed onderzocht, te snel, te laat, te duur, enzovoort. Terwijl als je dan

later naar die organisatie kijkt, dan zie je dat ze sommige lessen wel degelijk hebben geleerd. Dat ze dingen zijn gaan doen die ze in reactie op de evaluatie eerst met de grootste mogelijke stelligheid hebben verworpen. Op het moment dat ze even in de spotlights stonden werd vanuit een verdedigingsstrategie gezegd dat de aanbeveling niet deugde. In het openbaar zeggen dat je niets hebt aan de aanbevelingen, dat past in het patroon dat je wilt beschermen in het verantwoordingsdebat. Terwijl als de druk er af is, dan kun je wel degelijk je voordeel doen met een evaluatie.

Wat wij als Algemene Rekenkamer steeds meer doen, is niet aan de achterkant een rapport afscheiden en dan daar alle benutting van laten afhangen. We gaan gaandeweg het proces ook tussentijds uitkomsten presenteren en bediscussiëren. Op sommige punten testen we zelfs aanbevelingen. Waarmee je dus degene die geëvalueerd wordt of van wie zijn programma wordt geëvalueerd al in het onderzoeksproces meeneemt. En hem daarmee 'verleidt' om, al was het maar voor een deel, probleemeigenaar te worden.

De Ecologische Hoofdstructuur is daar een voorbeeld van. We waren met dat onderzoek begonnen, maar moesten het halverwege stopzetten omdat we een tussentijds verzoek van de Tweede Kamer kregen. Toen is het ministerie met die tussenstand aan de gang gegaan voor het verder verbeteren van het beleid. Als onderzoeker of rekenkamer denk je dan misschien wel even: 'een beetje jammer, dan kunnen we straks minder in ons rapport kwijt'. Maar als je er vanuit het publieke belang naar kijkt, en dat moet je natuurlijk doen, dan is dat een hartstikke goede zaak.

Wat ook werkt, is het omdraaien van je veronderstelling. Vaak is de gedachte: beleid werkt, tenzij het tegendeel wordt bewezen. Draai het nu eens om: het werkt niet, tenzij bewezen. Dat levert interessante nieuwe perspectieven op. We hebben het meegemaakt bij ons onderzoek naar subsidies. Dan wordt de stelling: als je na vijf jaar subsidies te hebben verstrekt nog steeds niet hebt onderzocht of het werkt, dan moet je er mee stoppen. Dat is misschien een wat minder sympathieke vorm om te zorgen voor benutting, maar het werkt wel. En we zijn er best trots op dat het Kabinet onlangs zo'n horizonbepaling heeft ingesteld voor alle subsidie-regelingen. Dus vertaald naar het provinciaal niveau: als je bij Provinciale Staten autorisatie vraagt voor het uitgeven van publiek geld, dan geldt voortaan de conditie: prima, maar als je na vijf jaar nog niet hebt onderzocht of het ook iets uithaalt, dan stoppen we er mee. Dat zorgt er natuurlijk ook voor dat je er belang bij hebt om de effecten te laten zien en dat je stuurt op benutting.

Controle

Over de rol van de volksvertegenwoordiging. Als de Tweede Kamer zelf een onderzoek wil instellen of dat gaat doen, dan is dat meestal toch een reactie op het feit dat de minister dat niet of niet goed genoeg heeft gedaan. Eigenlijk zou het anders moeten gaan. Het zou natuurlijk zo moeten zijn dat Provinciale Staten, gemeenteraad of Tweede Kamer tegen de uitvoerende macht zeggen: prima om daar geld aan uit te geven, maar leg ons eens uit waarom dat zo'n goed idee is en wat de onderbouwing is voor de aannahme dat het gaat werken. Of, als het een nieuw en ingewikkeld probleem is: laten we maar eens proberen of het kan werken, maar we willen wel geïnformeerd worden of we waar voor ons geld krijgen. Het begint dus bij het autoriseren van het uitgeven van geld, bij de start van nieuw beleid, bij het lastig vallen van burgers en bedrijven. Daar past natuurlijk bij dat je zegt: 'we verwachten wel dat je duidelijk maakt wat het oplevert'.

Dat moet je ook terugzien in de planning en controldocumenten. Ik weet wel dat onderzoek vaak is 'los gezongen' van de formele planning en controlcyclus. Van het meeste beleid kun je geen effecten op korte termijn verwachten. De cyclus kan niet elk jaar gesloten worden. Sowieso zit er altijd een uitvoeringsjaar tussen, daar heb je ook last van. Dus moet je meer in programma's en meerjarige verantwoordingstermijnen denken als je het hebt over resultaten en effecten.

Dat vraagt om geduld en dat maakt het lastig. We willen heel snel resultaten (van experimenten) zien, zeker tegen de achtergrond van de bestuurlijke cycli van verkiezingen en gekozen worden. Maar als je reëel bent, ontcom je er niet aan om zaken in meerjarig perspectief te plaatsen. Dat kan gewoon niet anders. Ik denk alleen wel dat je dan in de tussentijd kunt monitoren en de vinger aan de pols houden. De basisregel blijft: bij het autoriseren van beleidsuitgaven hoort altijd interesse voor de achterkant.

Tot slot

Afsluitend, is er ook een vraag die we onszelf als evaluatoren mogen stellen: wat is onze toegevoegde waarde? Niet alleen in termen van wat we adviseren of wat de uitkomsten zijn van onze onderzoeken en wat er mee gebeurt, maar heel direct ook over ons eigen presteren: wat kunnen we zeggen over de kosten-baten-verhouding van onze inspanningen? Dat is een hele interessante vraag, waar we nog te weinig aan toe komen. Een vraag die wel belangrijk is, al was het maar om onszelf scherp te houden.

Mijn beeld is dat evaluaties te vaak nog een ondergeschoven kindje zijn, waar niet iedereen het nut van inziet. Volgens mij komt dat doordat de insteek van de meeste evaluaties vooral gericht is op verantwoording. Een evaluatie wordt te veel gezien als verplicht nummer en te weinig als een kans om te leren van de beleidsuitvoering. Dan is het natuurlijk ook niet zo verwonderlijk dat mensen niet staan te springen om een evaluatie prominent op de agenda te krijgen. Het heeft ook te maken met timing. Vaak wordt er alleen pas achteraf iets over het beleid gezegd. Tegen de tijd dat dat gebeurt - omdat er een evaluatierapport op tafel ligt en men daar iets van moet vinden - is het beleid zelf al afgerond en spelen er weer allerlei andere onderwerpen. En voor een deel ligt het ook aan de kwaliteit van evaluaties. Die bovendien meestal ook nog eens alleen ingaan op één specifiek beleidsprogramma, waardoor er nauwelijks lessen uit getrokken worden voor andere onderdelen van een organisatie. Met dat soort evaluaties slaag je er niet in om de materie te 'overstijgen' en op een hoger niveau conclusies te trekken.

Vaak zie je dat de aandacht voor een beleidsevaluatie is neergelegd bij een aparte eenheid in de organisatie. Dat draagt een afbreukrisico in zich. Als het onderwerp meer geïntegreerd wordt in de beleidscyclus en programmamanagers er zelf verantwoordelijk voor zijn en op aangesproken kunnen worden, dan ontstaat een inhoudelijk 'eigenaarschap'. Dan leer je ook dat als je een evaluatie op een bepaalde manier organiseert, het onderwerp meer gaat leven en je echt iets aan de uitkomsten kunt hebben.

Dat pleit er volgens mij voor om een onderzoekseenheid binnen een ministerie, provincie of grote gemeente vooral een faciliterende rol te geven. Een eenheid die anderen in de organisatie helpt om beleidsevaluaties op te zetten of uit te voeren. Zoals ik het zie, moet de uiteindelijke verantwoordelijkheid in de lijn worden neergelegd. Bij de medewerkers die aan het eind van de rit ook met de uitkomsten aan de slag moeten.

Verplichting

De evaluatie van de Reconstructiewet (gericht op herinrichting van het landelijk gebied) is op dat punt een interessant voorbeeld. De aanleiding was heel traditioneel. Er stond nu eenmaal een bepaling in de wet dat er geëvalueerd moest worden. En zo werd het ook gevoeld door alle partijen. Dat gaf reacties in de trant van: 'Wat moeten we daar nou mee?' en: 'De inkt van de reconstructieplannen is nauwelijks droog en nu moeten we alweer kijken hoe het er voor staat.' Dat zorgde voor een negatieve, bijna vijandige sfeer. Dat was één van de redenen dat men op zoek ging naar een andere manier van evalueren, waarbij gebruikers en betrokken partijen overtuigd konden worden van de meerwaarde van de evaluatie.

Het ministerie heeft er vervolgens voor gekozen om de evaluatie niet zelf uit te voeren, maar uit te zetten bij Alterra, onderdeel van Wageningen University & Research centre. Het opdrachtgeverschap werd neergelegd bij de lijnmanagers die de reconstructie vormgaven en niet bij de directie van het ministerie die zich met kennis en evaluatie bezighoudt. Dat was een hele bewuste keuze en vormde uiteindelijk een belangrijke succesfactor.

Het had vooral met de invulling van de opdrachtgeversrol te maken. De medewerker in kwestie wist van de hoed en de rand wat betreft de reconstructie, maar had tegelijkertijd feeling met de onderzoekskant en een lerende houding. Hij was zich er terdege van bewust dat als het ministerie, provincies en andere stakeholders iets aan de evaluatie wilde hebben, dat hij dan voldoende afstand moest houden tot de onderzoekers om hen in intellectuele vrijheid hun werk te kunnen laten doen. Juist daardoor is het goed gegaan.

Een verplichting om te evalueren is geen garantie voor succes. Je kunt het onderwerp er wel mee op de agenda krijgen en ervoor zorgen dat mensen er aandacht aan besteden. Maar of dat dan ook op een goede manier gebeurt, zodat je er echt iets aan hebt, dat is nog maar de vraag.

Moet echt alle beleid geëvalueerd worden? En op welk niveau dan en hoe ver ga je daarin? Dat je prestaties en voortgang meet is voor alle beleid een goede zaak. Wat mij betreft is daar weinig discussie over nodig. Maar het optuigen en uitvoeren van doe-evaluaties voor alle typen beleid, dat is een wel erg zwaar en kostbaar middel. Ik betwijfel sterk of dat altijd nut heeft. Ik kan me wel voorstellen dat je centrale beleidsprogramma's consequent met beleidsevaluaties belegt. Waarbij je natuurlijk ook weer een discussie kunt krijgen wat centrale beleidsprogramma's zijn en wat niet. Ik denk dat de volksvertegenwoordiging daar een belangrijke rol in kan en moet spelen.

Veel onderzoek wordt trouwens als beleidsevaluatie aangemerkt terwijl het dat, als je er goed naar kijkt, niet is. Waar het in veel gevallen op vastloopt, is dat er in het begin geen duidelijke beleidsdoelen zijn geformuleerd. Dan is het sowieso al heel lastig om daar achteraf criteria voor te ontwikkelen. Ook zie je nogal eens dat evaluatoren niet duidelijk maken aan de hand van welke criteria zij evalueren. Dan is er geen duidelijke norm waartegen de beleidsuitvoering wordt afgezet. Ook vind ik dat veel beleidsevaluaties te kort schieten op het punt van verklaringen: je krijgt dan geen zicht op waarom beleid wel of niet heeft gewerkt, hoe het complexe spel van beleidsuitvoering in elkaar zit. Wat vervolgens leidt tot simplistische conclusies als 'er moet meer geld bij' of 'de regels moeten strenger'.

Als lid van de Provinciale Raad voor het Omgevingsbeleid in Gelderland kreeg ik laatst uitkomsten van een uitgebreide monitoring van het streekplan voorgeschoteld. Dat bevat een hele hoop indicatoren die lang geleden zijn ontwikkeld. Maar veel van de doelen die in het beleidsplan staan, zijn inmiddels niet meer relevant door alle beleidswijzigingen die er hebben plaatsgevonden. Zijn dit dan nog de zaken die we willen weten? En bovendien, misschien nog wel belangrijker: Wat betekenen al die gegevens? Wordt de ontwikkeling die uit de monitoring naar voren komt inderdaad veroorzaakt door het provinciaal beleid, of is er sprake van een autonome trend?

Als evaluator kun je proberen om daar zelf een antwoord op te formuleren, maar ik denk dat je beter de dialoog daarover kunt organiseren met mensen uit het veld. Zij hebben verstand van het onderwerp. Dan gaat het over vragen als: Wat betekenen deze trends die we zien? Is dit een wenselijke ontwikkeling of niet? Wat is het effect van provinciaal beleid op deze ontwikkeling? De antwoorden hierop zijn niet altijd eenduidig en daarom is het belangrijker dat je daar een gesprek met elkaar over hebt want dat bevordert het 'beleidsleren'.

Programmeren

Het is ingewikkeld om in lange termijnperspectief een evaluatieprogramma te ontwikkelen. Bij ministeries zie je dat alle onderzoek meerjarig wordt geprogrammeerd, niet alleen evaluatieonderzoek. Op het oog zit zo'n systeem heel goed in elkaar. Het begint met een agenderingsfase, daarna wordt het verfijnd en uitgezet en vervolgens brengen onderzoeksbureaus offertes uit voor de uitvoering. Dat is een jaarlijks terugkerend proces. En dan heb je nog een orgaan dat in de gaten houdt of zo'n kennisprogramma een beetje op koers ligt of dat er bijstellingen nodig zijn. Waarbij als het goed is ook nog geprobeerd wordt om in te spelen op zaken die op wat langere termijn spelen.

Maar de keerzijde van een dergelijke, op papier, fraaie aanpak is dat je ongemerkt steeds verder van de beleidspraktijk komt te staan. Dat andere directies denken: 'Er is een aparte club die het organiseert, dus wij kunnen ons met de lopende zaken bezighouden'. Dat is een spanning waar je in evaluatieonderzoek ook mee te maken hebt. Beleidsambtenaren, de goede niet te na gesproken, houden zich over het algemeen vooral bezig met de vragen van vandaag en niet met die van morgen. Ze hebben wel vragen, ook evaluatieve vragen, maar die bestrijken zelden een langere periode. Terwijl je met je onderzoek en je evaluatie juist iets wilt neerleggen waar het beleid op langere termijn ook wat aan heeft.

Uitbesteden

Het grote risico van alleen uitbesteden is dat het onderwerp niet binnen de organisatie beklijft. Daar kun je als extern onderzoeker overigens wel het nodige aan doen. Ik denk juist dat het voordelen kan hebben om een evaluatie door een externe te laten doen. Je hebt dan in ieder geval een ander perspectief van waaruit je naar het onderwerp kijkt. Belangrijk is wel dat het onderzoek zelf goed ingebed is in de organisatie. Dat betekent dat de externe onderzoeker samen met de opdrachtgever op zijn minst heel scherp kijkt naar de probleemstelling en vraagstelling van het onderzoek. Ook tussentijdse resultaten moeten met mensen in de organisatie worden doorgesproken. Als dat niet goed georganiseerd is, dan dreigt het een ver-van-mijn-bed-show te worden.

Zeker in de beginfase moeten opdrachtgevers dus flink investeren in de relatie met de onderzoekers. Zij moeten hen meenemen in het onderwerp, de context schetsen en zo helder mogelijk aangeven wat de oorsprong van de vraag is. Vaak gaat het daar fout. Dan heeft de organisatie daar te weinig tijd voor ingeruimd. Het onderzoek is immers uitbesteed, zodat de lijnorganisatie het 'eigenlijke' werk kan blijven doen.

Een andere, eveneens foute gedachte is: de externe onderzoeker moet zijn werk onafhankelijk kunnen doen, dus dan moet ik hem ook maar niet al te veel vertellen over de dingen die ons bezighouden, want dat zou hem beïnvloeden. Terwijl de onderzoeker die inside-informatie juist nodig heeft om scherp te krijgen waar de vraag van de opdrachtgever zit. Dat betekent dat de opdrachtgever ook bereid moet zijn om de onderzoeker een kijkje in de keuken te gunnen.

Het begint natuurlijk al met de keuze van een goed onderzoeksbureau en een goede onderzoeker. Maar ook dat je waardering kunt opbrengen voor een kritisch geluid. Als je te krampachtig bent als opdrachtgever en de vraagstelling zo beperkt houdt dat je alleen maar iets op de vierkante centimeter kunt doen, dan gaat het natuurlijk nooit goed. Dan heb je het eigenlijk over het verschil tussen de vraag: 'Doen we de dingen goed?' en de vraag: 'Doen we de goede dingen?'. Vaak zie je dat men zich beperkt tot het eerste, maar dat men de onderliggende vraag, op het niveau van het beoogd maatschappelijk effect, uit de weg gaat. Aan die diepere vraag komt men dan niet toe. En daar kunnen juist externen of meer wetenschappelijk georiënteerde instituten mee aan de slag. Puur het technische meten van prestaties en die langs een maatstaf leggen, dat kan een ministerie of provincie heel goed zelf. Maar als het meer de strategische kant opgaat en er ook een lange termijnperspectief gevraagd wordt, dan is het zinnig om de buitenwereld daarbij te betrekken. Goed opdrachtgeverschap houdt dus investeren in tijd

in, maar ook erkennen dat er mensen zijn die vaker met dat bijltje hebben gehakt.

Wat ik nogal eens zie, is dat men vooral op basis van de vraagstelling handelt. Die wordt dan van alle kanten bekeken en helemaal doorgeakkerd, terwijl men geen goed beeld heeft wat er uit een onderzoek zou kunnen komen. Dat kan tot verrassende discussies aan het eind van een onderzoek leiden. Dan heeft de opdrachtgever ineens allerlei extra vragen die hij onvoldoende in het onderzoek beantwoord ziet. Het blijkt voor opdrachtgevers lastig zich puur op basis van de vraagstelling een beeld te vormen wat er uit kan komen. Het is dus zaak de opdrachtgever gedurende het onderzoek goed te blijven informeren over keuzes en tussenresultaten.

Breder verband

Op het punt van de benutting van onderzoek is nog veel winst te boeken. Vaak wordt alleen gekeken naar de één-op-één relatie met de opdrachtgever. En dat is dan meestal ook nog maar eens één persoon binnen een bepaalde directie. En dan moet je maar hopen dat die ene ambtenaar binnen de kolom waarin hij werkzaam is zijn relaties op orde heeft en verbindingen weet te leggen.

Ik vind dat je daar als onderzoeker en onderzoeksbureau ook een verantwoordelijkheid in hebt om dat te organiseren. Of, als dat een brug te ver is en als te bemoeizuchtig overkomt, om het in ieder geval bespreekbaar te maken. Mijn ervaring als onderzoeker is dat wanneer jij het niet doet, het meestal helemaal niet gebeurt. Daarom moet je eigenlijk in je onderzoeksplan al aandacht vragen voor het grotere geheel. En dus bewust breder gaan dan die ene afdeling waar je het onderzoek voor doet. Dan moet je er natuurlijk wel in slagen om een goede aansluiting te vinden bij datgene waar die andere eenheden mee bezig zijn.

Bij het onderzoek naar de reconstructie in Noord-Brabant was de vraag vanuit Provinciale Staten (PS) ook duidelijk: Wat kunnen we vanuit de reconstructie leren voor andere grote transitieprocessen? Dat ging dus verder dan alleen kijken naar het onderwerp. Het onderzoek is later om andere redenen helemaal de mist ingegaan, maar dat laat onverlet dat de vraagstelling op zich goed was. Een brede invalshoek dus, die het voor de hele organisatie en de politiek relevanter en interessanter maakt.

Overigens lukt het ons als onderzoekers niet altijd om die brede scope erin te krijgen. Om zaken praktisch toepasbaar te houden, moet je het vaak juist klein houden. De truc is dan om te proberen om over het onderwerp bijvoorbeeld een gesprek met de gedeputeerde te hebben. Die heeft veel meer overzicht. Op die

manier kun je soms toch verbreden en onderwerpen naar een hoger niveau tillen. Als er een onderzoekseenheid is en deze voldoende statuur heeft, zou zij die rol ook uitstekend op zich kunnen nemen. Maar je zult volgens mij altijd ook een bestuurlijke hoeder moeten hebben.

Leren

Waar wij steeds meer op uitkomen, is het combineren van verantwoorden en leren. Dat gaat dan onder de noemer van lerende evaluaties. Daar zijn natuurlijk allerlei gradaties in. Essentieel is dat je gebruikers mee laat denken, sowieso over je onderzoeksopzet, maar ook dat je regelmatig terugkoppelt met tussenresultaten en dat je probeert de organisatie te betrekken bij het geven van betekenis aan de resultaten.

Het nadeel van klassieke evaluaties is: je voert je onderzoek uit, presenteert je resultaten en dan hoop je maar dat het werking gaat hebben. Met lerende evaluaties probeer je het zo in te richten dat die werking al veel eerder begint. En overigens kunnen die tussentijdse momenten ook als effect hebben dat het onderzoek wordt bijgesteld. Mensen gaan dingen eerder in de praktijk toepassen, waardoor ook jouw object van onderzoek gedurende de rit verandert. Iemand die het denken daarover in het bijzonder heeft gevoed is Jurian Edelenbos, hoogleraar Bestuurskunde aan de Erasmus Universiteit. In zijn visie zijn gebruikers geen mensen die pas aan het eind van de rit in beeld komen. Je praat dan over echt *ex durante* onderzoek.

Wat er ook erg in zit, is het benadrukken van het doorlopen van leercycli: het gezamenlijk betekenis geven aan bevindingen zodat mensen daarvan kunnen leren en ze kunnen toepassen in hun praktijk. Idealiter maak je in een lerende evaluatie verschillende leercycli door. Bij *ex post* onderzoek is het eigenlijk eenmalig. Als de focus op leren is gericht, ben je langere tijd betrokken, zodat je zo'n leercyclus na verloop van tijd ook kunt herhalen. Het is niet een scherp omschreven methodiek, maar meer een inspiratiebron. En afhankelijk van de situatie probeer je daar elementen van toe te passen.

Als je ervan uitgaat dat de werkelijkheid complex is en er verschillende perspectieven zijn, geef dat dan ook een plek in je evaluatie. Probeer dat niet plat te slaan in één kader. Geef betrokkenen een stem in het normenkader. Dat is altijd een heel cruciaal moment. Waar ga je op evalueren? Soms kom je tegen dat de beleidsdoelen niet helder zijn of dat er nog geen overeenstemming is over de doelen. Ook kun je te maken krijgen met meerdere perspectieven of met doelen die in de loop van de tijd zijn veranderd. Of de provincie heeft geen zeggenschap over de

doelen, iets wat je bij gedecentraliseerd beleid nogal eens tegenkomt. Probeer daar dan op zijn minst verschillende partijen over te horen. Dan krijgt de evaluatie bijna een soort beleidsontwikkende rol. Want dan ga je je nog eens een keer heroriënteren op de vraag: Waarom deden we dit eigenlijk ook alweer? Het kan invloed hebben op de beleidstheorie achter het beleid. Is die nog wel zo passend? Dan ga je schuiven. Ik vind dat een interessante ontwikkeling.

Controle

Over de rol van de volksvertegenwoordiging. Ik denk dat evaluaties zowel voor Gedeputeerde Staten (GS) als uitvoerende partij, als voor de controlerende rol van PS belangrijk zijn. Overigens kunnen PS in dat verband hele andere vragen stellen dan GS. Juist voor PS zijn vragen op meta-niveau belangrijker. Dus niet of het allemaal volgens de planning verloopt, maar meer: Wat zijn de maatschappelijke effecten en wat betekent dat nu allemaal voor de wat langere termijn, voor de komende beleidsontwikkeling?

Het onderzoek in Brabant naar de reconstructie was daar een goed voorbeeld van. Daar stelde de commissie Beleidsevaluatie de goede vragen. Nogmaals, dan heb ik het niet over wat er daarna allemaal fout ging. Zelf deed ik een andere ervaring op in Overijssel, eveneens rond de reconstructie. Daar vraag ik me in alle oprechtheid af of men iets aan het onderzoek heeft gehad. PS hadden een tijdelijke constructie voor dat onderzoek opgetuigd en vervolgens hebben ze lang onderhandeld over de precieze vraagstelling. Het onderwerp was van meet af aan heel politiek. Toen wij als onderzoekers vervolgens aan de slag gingen, was in ieder geval één ding helder: wat er ook gebeurde, aan de vraagstelling kon niet worden getornd. Die was immers de uitkomst van een langdurig en uiterst delicaat onderhandelingsproces. Maar als onderzoeker vond ik die vraagstelling behoorlijk beperkt. De lange termijn vragen zaten daar onvoldoende in. Het onderzoek had meer op kunnen leveren.

Beleidsevaluatie is ook echt een vak, daar hoeft je niet omheen te draaien. Dat doe je er niet zo maar even bij, daar moet je je echt in verdiepen wil je er uit halen wat er in zit. Als PS zou je een soort 'kop' bovenop de door GS uit te voeren evaluaties kunnen zetten met specifieke aandacht voor vragen die samenhangen met de kaderstellende en controlerende rol.

Een voorbeeld is Gelderland waar men nu bezig is met een nieuwe structuurvisie. Daar wordt een heel interactief proces rondom heen ingericht, waar de politiek overigens op onderdelen moeite mee heeft omdat zij zich soms buitenspel gezet voelt. Maar als je het vanuit een andere invalshoek bekijkt, is het heel logisch om te zeggen: 'We weten eigenlijk nog helemaal niet wat het effect van het vorige

streekplan is geweest. We kunnen nu natuurlijk wel een nieuw proces in gang zetten, maar het is toch niet verkeerd om eerst eens stil te staan bij wat het eerdere proces heeft opgeleverd en wat we daaruit kunnen leren. Waar blijft de beleidsevaluatie?' Dat zou ik typisch een vraag vanuit PS vinden om op een soort meta-niveau de vinger aan de pols te houden of dat soort dingen gebeuren. Als je zoals PS voor een lange reeks van jaren grote sommen geld moet autoriseren voor een beleid, dan moet je toch willen weten, om een goede afweging te kunnen maken, wat er van het vorige beleid op hoofdlijnen terecht is gekomen. En daarmee GS dus ook controleren op wat zij gerealiseerd hebben. In die zin is er voor PS een taak.

Partijen

Beleid komt meestal tot stand in een samenspel van meerdere partijen. Dan is het moeilijk om de relatie tussen de interventie en het effect te leggen. Dat levert methodische problemen op. Het probleem doet zich met name voor bij de verantwoording, want daar is het juist die ene actor die zich moet verantwoorden tegenover de volksvertegenwoordiging.

Voor het leren is het minder een probleem, dan kun je aangeven dat je een gezamenlijke verantwoordelijkheid hebt. Natuurlijk, ieder vanuit een eigen rol en verantwoordelijkheid, maar uiteindelijk is het toch een samenspel waarbij het er om gaat hoe je je processen zo goed mogelijk kunt inrichten. En daar kun je over leren, hoe je dat beter kunt doen. Zeker als je van meet af aan de partijen die ermee te maken hebben, bij een dergelijk evaluatieonderzoek betreft. Dan kun je dat ook onderwerp laten zijn van die evaluatie. En biedt de evaluatie als je die goed aanpakt een uitstekende kans aan partijen om even uit de arena te treden waarin zij elkaar steeds tegenkomen en op een andere manier met elkaar te praten over waar ze mee bezig zijn. Als onderzoeker geeft dat ook veel voldoening als je dat gesprek kunt organiseren en partijen niet in hun 'normale' vechthouding zitten.

Er zijn helaas geen gemakkelijke antwoorden. Maar dat wil niet zeggen dat je helemaal niks kunt zeggen. Je kunt verschillende vormen van onderzoek inzetten om meer zicht te krijgen op wat bepaalde interventies teweeg brengen. Alleen kun je het niet altijd rechtstreeks en heel eenduidig 'terugredeneren' naar één actor en één interventie. Je moet af van het sturingsmodel dat uitgaat van één centrale actor. Dat betekent ook dat de evaluatie dat niet moet suggereren. Dan kun je overigens nog steeds uitspraken doen over wat meer of minder effectief is. De gesprekken die je daarvoor moet hebben, kun je voeden met goed onderzoek. Waarbij je laat zien hoe dingen werken, welke patronen en mechanismen er zijn. Dat soort onderzoek is ook geschikt om de heersende beleidstheorie nog eens

tegen het licht te houden. Op zo'n manier kan ook het beleid weer een stukje verder gebracht worden.

Ontwikkeling

Een belangrijke trend in evaluatieonderzoek, zoals hiervoor al opgemerkt, is het steeds meer combineren van evalueren en leren. De meest uitgesproken vorm daarvan is de *responsieve evaluatie*, waarin nauwelijks verantwoordingselementen zitten en die vooral dient als ondersteuning van het leerproces.

Een andere stroming gaat meer uit van de *kritische theorie*. Daarbij wordt beleid niet zozeer langs de meetlat gelegd van de eigen beleidsdoelen, maar wordt vooral gezocht naar het antwoord op de vraag wat de functie van het onderzochte beleid is. Lost het de maatschappelijke problemen op? Welke machtsrelaties spelen een rol? Wie wint, wie verliest?

Een andere trend heeft te maken met het monitoren van allerlei ontwikkelingen. Er wordt heel veel gemeten, terwijl het evaluatieve aspect vaak achterwege blijft. Dat is het zoeken naar de verklaring achter al die gegevens. Ligt het aan het beleid? Ligt het aan andere dingen? Als je niet weet waardoor de dingen komen, dan kun je het wel van allerlei kanten bekijken, maar dan weet je nog steeds niet wat je moet doen. Dat zie ik ook wel steeds meer ontstaan. Dus een beweging weg van het pure monitoren, meer in de richting van het zoeken naar verklaringen.

Een continue zoektocht naar inzicht in hoe beleid werkt. En daarmee draag je misschien ook een beetje bij aan een meer bescheiden rol van de overheid. Die overheid is niet almachtig.

Op die manier kan er van onderzoek een opvoedende rol richting overheid uitgaan. 'Trek niet een te grote broek aan, het zit allemaal heel ingewikkeld in elkaar. Je bent maar één van de radertjes in het grotere geheel.'

Frans Leeuw

31 mei 2012

Als ik gevraagd wordt om trends op het gebied van evaluatiestudies aan te geven, wil ik allereerst een zeer recente ervaring delen. Gisteren woonde ik de viering bij van het 35-jarig bestaan van de IOB, de Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie. Ik ben voorzitter van het panel dat de kwaliteit, bruikbaarheid en gebruik van de evaluaties van de IOB bekijkt. Kees Vendrik van de Algemene Rekenkamer en Marjolein van Asselt van de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) spraken daar ook. De WRR heeft een groot project 'Lessen van evaluatie'. Eén van de kernpunten uit hun bijdragen was: vroeger keek men bij evaluaties *ex post* naar beleid en dan werd er over resultaten gerapporteerd. Die resultaten werden vervolgens vrij snel in verband gebracht met het beleid en als er geen resultaten waren dan lag de verklaring voor de hand: dat kwam door het slechte beleid. Methodologisch allemaal niet erg geavanceerd. De kwaliteit van effectonderzoeken is de laatste tien jaar ten gevolge van de discussie over het belang om robuuste effectstudies te doen, waarbij al snel gedacht wordt aan *experimentele* en *quasi experimentele* onderzoeken, steeds meer op de agenda aan het komen. Het is interessant dat ook de IOB op dat gebied vooruitgang boekt.

Serieus

Dat thema was dan ook één van de rode lijnen in dit jubileumcongres van de IOB. Met name bij ontwikkelingssamenwerking speelt al een aantal jaren de vraag hoe op robuuste wijze tot uitspraken te komen over de effecten van ontwikkelingsinterventies, programma's en beleid. De discussie daarover is vooral aangejaagd door de publicatie van het rapport *When will we ever learn?*¹ uit 2006. Daarin wordt de balans opgemaakt van het vele evaluatiewerk en komt men tot de vernietigende conclusie: 'We weten het nog steeds niet'. Dat heeft een serieuze boost gegeven om, net zoals de medische wetenschap dat doet, effectstudies veel serieuzer op te pakken. Het heeft trouwens ook geleid tot uiterst felle discussies tussen methodologen. Met aan de ene kant de 'randomistas' die *randomized trials* bepleiten (de 'experimentele effectevaluaties') en aan de andere kant de meer rekkelijken, soms wel 'relativistas' genoemd. Die discussie is inmiddels enigszins beslecht. Ik heb daar een bescheiden bijdrage aan geleverd met een handboek dat ik samen met een Vlaamse collega voor de Wereldbank heb geschreven en waarin de voor- en nadelen van verschillende methoden van onderzoek op een rij worden gezet en wordt aangegeven wanneer welk evaluatiedesign het beste wel en niet kan worden ingezet.²

¹ Center for Global Development: *When will we ever learn? Improving lives through impact evaluation*. Rapport van de Evaluation GAP Working Group, mei 2006.

² Frans Leeuw en Jos Vaessen, *Impact Evaluations and Development - NONIE Guidance on Impact Evaluation*, Network of Networks on Impact Evaluation, 2009.

Ik beschouw dat als de *eerste trend*: dat we in een aantal maatschappelijke sectoren afstappen van de slordige, door iedereen op een achternamiddag te verrichten ex post studietjes die resultaten te snel toe wijzen aan beleid. Dat toewijzen kan alleen als dat methodologisch verankerd en zorgvuldig gebeurt. Dat betekent dus ook dat niet iedereen zomaar in het wilde weg kan roepen dat hij bezig is met effectonderzoek.

Steeds meer dringt het besef door dat als we willen weten of instrumenten werken, we er dan niet mee weg komen door wat tevredenheidsmetinkjes te doen en aan mensen in een twintig-seconden-survey op internet te vragen: 'Vindt u ook niet dat enzovoort.....?'.

Ex ante

Trend twee, het WODC, het Wetenschappelijk Onderzoek- en Documentatiecentrum van het ministerie van Veiligheid en Justitie, brengt daar over enige tijd een rapport over uit: de sterke groei van *ex ante* evaluaties. Het zijn soms een soort van uitvoeringstoetsen gericht op de vraag: 'Zal dit gaan werken?'. Die trend is in de eerste plaats een gevolg van het feit dat ex post evaluaties, of ze nu slecht of goed zijn, doorgaans veel tijd kosten. Een bestuurder kan daar niet op wachten, die moet morgen of overmorgen al wat weten. Dat geeft een druk richting ex ante evaluatie. En op dat vlak begint er ook steeds meer serieuze methodologie te komen.

Kennisschatkamers

Een derde trend is het werken met zogenaamde *repositories*. Dat zijn kennisschatkamers, portals waar systematische onderzoeksbevindingen naar de werking van interventies maar ook van andere zaken te vinden zijn. De organisaties die deze 'systematic reviews' en syntheses maken doen vaak niet zelf evaluaties, maar beoordelen de kwaliteit van de aangetroffen evaluaties en andere studies, trekken daar conclusies uit en – als het kan – rekenen opnieuw de 'effect maten' uit door gebruik te maken van de – vaak – tientallen onderzoeken die ze samenvatten. Die samenvattingen worden per onderwerp bij elkaar gebracht. Wetenschappers fileren honderden evaluaties, scheiden het kaf van het koren en maken vervolgens deze 'systematic reviews' van enkele tientallen pagina's. En daarbij brengen ze, als het kan en wetenschappelijk verantwoord is, ook de effecten van verschillende interventies in beeld. Je ziet steeds meer van die repositories komen, bijna altijd zijn ze openbaar en gratis toegankelijk. Deze ontwikkeling komt veertig jaar nadat de medische wetenschap daarmee is begonnen. De eerste repository was die van Archie Cochrane. In zijn 'bibliotheek' zitten resultaten naar honderdduizenden effectstudies van therapieën, pillen, organisatieveranderingen in medische 'edifices' enzovoort. Als je bijvoorbeeld wilt weten wat de beste manier is om nierstenen te

verwijderen, dan zie je in tientallen studies samengevat dat als je het op de ene manier doet je dit effect hebt en doe je het op een andere manier dan zie je die en die effecten optreden. Deze manier van werken slaat nu sinds zo'n tien/vijftien jaar over naar de wereld van overheidsinterventies en – ook – interventies door de private sector.

Ritualisering

De vierde trend zou ik willen omschrijven als: 'Vrienden, waar zijn wij mee bezig?'. Het is een trend van ritualisering. Of, om het anders te zeggen: geef iemand een hamer en alles moet gespijkerd worden. Er is een ongelooflijk grote hoeveelheid organisaties die allemaal bezig zijn met M&E, *monitoring and evaluation*. Het houdt niet op. Ik zeg wel eens tegen mijn studenten of in een spreekbeurt: 'Wanneer is bij u het lokale café toe aan een sous-chef Monitoring and Evaluation?'. Let wel, het idee van M&E is natuurlijk helemaal niet slecht. Maar als de hele wereld bezig is met het opzetten van allerlei monitoringsystemen, dan moet je natuurlijk niet raar staan te kijken als het op een gegeven moment niet meer serieus wordt genomen. Hadden we vroeger de situatie dat er niets was – Hoogerwerf vroeg daar aandacht voor in zijn Twentse oratie in 1977, ik heb het zelf in 1993 in mijn Utrechtse oratie laten zien – nu zijn we een paar decennia verder en is het een *mer à boire* geworden met over het algemeen een beperkte kwaliteit. Peter Dahler-Larsen, oud-voorzitter van de European Evaluation Society, sprak in dat verband over *evaluation machines*.

Verankering

Wat kun je als ministerie of provincie doen om niet in die valkuil van ritualisering te trappen? Met andere woorden: hoe veranker je op een goede manier de aandacht voor evaluatie? Ik zie een aantal antwoorden op die vraag. Allereerst moet je een klein clubje hebben dat in de gaten houdt of, met name bij de grote programma-uitgaven, de mensen die daarvoor verantwoordelijk zijn zelf bijhouden wat de doelen zijn en informatie over die doelen vergaren. Zo'n club moet je niet te groot maken, het moet geen 'directie' of zo worden. Wat mij betreft moet je het ook zoveel mogelijk integreren met *financial audit*, ook al is dit echt ander werk.

Als je zo'n club hebt, dan moet je nog steeds oppassen voor ritualisering. Want voor je het weet, worden ze 'ingepakt' door beleidsdirecties of hoe die mogen heten. Officieel zeggen die hun medewerking toe, maar als de deur dicht is, haalt men de schouders op en gaat men vaak over tot de orde van de dag. Je zult je dus heel slim moeten organiseren.

Echter, vrienden, robuuste effectstudies voor de grote programma's, dat is te hoog

gegrepen voor zo'n klein clubje. Die kunnen een tevredenheidsonderzoek doen of een quick scan. Serieuze effectstudies moet je niet zelf doen – denk aan de eerste trend – die moet je uitbesteden en dan niet zomaar aan de eerste de beste, maar aan absolute specialisten. Dat moet je alleen incidenteel doen, je hebt niet elke dag een groot beleidsprogramma dat op die manier tegen het licht gehouden moet worden.

Het eigen 'clubje' kan wel de selectie van het onderzoeksbureau voor zijn rekening nemen en de aanbesteding organiseren. Mits er natuurlijk één of enkele mensen in zitten die een stevige verankering in dat werk hadden en nog hebben. Anders wordt het weer amateurisme. Door zo te werken ben je ook van het probleem af dat je geen impact hebt, want in deze situatie heb je op jaarbasis misschien vijf of zes grote rapporten en die zullen zeker niet ongemerkt in de la verdwijnen. De organisatie zal zich er wel degelijk om bekreunen. Het kost veel geld, maar je hebt wel uitstekend onderzoek met uiterst relevante resultaten in handen.

Daar staat tegenover dat er mensen zijn die zeggen dat je het onderzoek juist wèl in huis moet uitvoeren, omdat je er alleen dan van verzekerd bent dat het onderwerp beklijft. Omdat je de ambtelijke organisatie vanaf het begin meeneemt in wat er daarna moet gebeuren. Ik ben het daar niet mee eens. In mijn visie zorgt de kleine onderzoeksclub voor de agendering en voor tijdigheid van rapportages. Bovendien beschikt het over enige expertise om bij de aanbesteding de goede keuzes te kunnen maken. Maar het onderzoek zelf doen ze niet. Daarvoor hebben ze ook geen materiedeskundigen of methodologen in huis. Zij zijn alleen de 'lopers' die het organiseren.

Grote organisaties beschikken wèl over eigen onderzoekscapaciteit. Zoals het SCP (het Sociaal en Cultureel Planbureau), het WODC bij het ministerie van Veiligheid en Justitie en het KIM, het Kennisinstituut voor Mobiliteitsbeleid van het ministerie van Infrastructuur en Milieu. Bij het WODC hebben we 100 fte en een budget van € 17 miljoen. Daar werken hoogleraren en methodologen. Dat is ook zo bij de paar andere 'Haagse' onderzoeksinstituten/planbureaus. Maar provincies ... dat is een andere schaal, die kunnen dat niet aan. Als een provincie toch een SCP-achtige constructie wil maken, dan ben je er niet met 3,9 fte en een beetje geld. Dan moet je daar fors in investeren. Ik raad dat af. Kennis kan ook gedeeld worden (via de repositories bijvoorbeeld of in samenwerkingsarrangementen).

Kwaliteit

Een belangrijke vraag is vervolgens hoe je de kwaliteit van evaluaties borgt. Ik vind dat in dat kleine clubje – waar ik dus voorstander van ben – enkele fte's aanwezig moet zijn van mensen die uit de onderzoekswereld komen. Die zijn te vinden.

Zij beschikken over een netwerk en weten de weg als ze op zoek moeten naar kwalitatief hoogstaande onderzoekers en onderzoeksbureaus. Als het iemand is die dat niet in zich heeft en in zijn eentje op de 17e verdieping van het hoofdgebouw zit te 'studeren', dan wordt het niks.

Over de samenstelling van zo'n clubje is nog meer op te merken. Je zult er een paar 'piskijkers' bij moeten hebben. Eigenlijk zoals mijn rol bij de ontwikkelings-samenwerking. De IOB is een intern verzelfstandigde eenheid van het ministerie, met een hoogleraar als directeur die met zijn positie en bevoegdheden buiten het schootsveld van de minister ligt. Maar toch is – op verzoek van toenmalig minister Koenders – een extern adviespanel georganiseerd, bestaande uit drie Nederlanders en een buitenlander, dat vier keer per jaar scherp kijkt naar wat zij aan het doen zijn. Niet als een grote toezichthouder, daarvoor is dat panel te klein, maar als een buitenboordmotor die de organisatie mede op koers houdt en toetert als het onzes inziens de verkeerde kant op gaat.

Als provincie kun je evaluatieonderzoek en zeker dat naar effecten maar beter niet in eigen beheer houden, want dan word je de slager die zijn eigen vlees keurt. Het kost je nauwelijks iets zo'n panel van mee- en piskijkers, want experts op dit gebied vinden het doorgaans spannend om het te doen en met een beetje vacatiegeld ben je rond. Mensen moeten er ook niet al te veel geld voor krijgen want anders worden ze belanghebbend. Er is zoveel evaluatieonderzoek in Nederland en er zijn zoveel uitmuntende onderzoekers die daar al jaren mee bezig zijn, dat je zo een aantal geschikte mensen daarvoor kunt vinden. Wat in ieder geval nooit mag gebeuren, is dat de organisatie een paar mensen, die men elders liever kwijt dan rijk is, in zo'n clubje parachuteert. Overheden hebben nog wel eens de neiging zo hun personele problemen op te lossen. Ook daar moet gekeken worden naar kwaliteit.

En er is nog een belangrijk punt om je kwaliteit te borgen. Elke beslissing rondom aanbesteding komt wat mij betreft in een databestand. Ik geef een voorbeeld hoe dat bij het WODC gaat. Wij werken al 25 jaar met aanbestedingsprocedures en beschikken over een bestand waar 220 instituten in zitten. Dat zijn bureaus die allemaal op enig moment geoffreerd hebben en waarvan er een aantal in de loop der jaren ook producten hebben afgeleverd. In dat bestand staat precies hoe goed ze waren, wat voor narigheid ze bezorgd hebben, of er geld terug is gevorderd en nog veel meer. Dat is de basis van je lerend vermogen. Wat je bij veel provincies en andere overheden ziet, is dat die informatie alleen in de hoofden van sommige medewerkers is opgeslagen. En dus van de ene op de andere dag verloren kan gaan. Dat werkt niet. Je moet dat externaliseren, buiten jezelf plaatsen en transparant maken. Je kunt dat gewoon met open source software organiseren. Je moet

alleen 0,4 fte hebben om het in te kloppen. De eerste twee jaar zullen mensen zeggen: 'Wat hebben we hier in hemelsnaam aan?' Maar dat geluid zal verstommen, als men er achter komt dat je door hierin te investeren je eigen kwaliteit op niveau brengt.

Een klein kwalitatief hoogstaand clubje dus, lerend vermogen en een paar grote 'klappers' van onderzoeken. Onderzoeken die allemaal in de openbaarheid komen. Je kunt evaluatieonderzoek niet achter de gordijnen uitvoeren. Dat wil niet zeggen dat je het aan de betreffende geleerde overlaat om het tijdstip te bepalen waarop het rapport openbaar wordt. Je kunt daar hele goede afspraken over maken. Bijvoorbeeld: het rapport wordt door de externe onderzoeker afgerond, een begeleidingscommissie met in meerderheid onderzoekers/geleerden geeft haar fiat en vervolgens krijgt de provincie drie maanden de tijd om te bepalen wat ze ermee wil gaan doen. En na die periode wordt het openbaar gemaakt.

Kwaliteit heeft ook te maken met je informatiehuishouding, met de manier waarop je met data omgaat. Ik geef een voorbeeld op het terrein van Veiligheid en Justitie. De politie heeft tal van bestanden waar ze mee werkt: bedrijfsvoeringssystemen, bestanden met gegevens over verdachten die worden opgepakt, etc. Het Openbaar Ministerie heeft weer andere bestanden met gegevens over bijvoorbeeld: hoeveel verdachten van de politie bij het OM komen, wat daar mee gebeurt en hoe lang zaken duren. En daarnaast zijn er nog veel meer organisaties met allerhande gegevens: de rechterlijke macht, de Raad voor de Kinderbescherming, Slachtofferhulp Nederland, Bureau HALT. Deze informatie bestaat al jaren en wordt beheerd door de Justitiële Informatiedienst van het ministerie van Veiligheid en Justitie in Almelo.

Aan het begin van deze eeuw kwamen we er achter dat het met moderne technieken als *dataspaces* en relatiemanagers heel goed mogelijk is om data uit die verschillende bestanden met elkaar te verbinden.

Vanuit evaluatieoogpunt kunnen wij nu dus laten zien hoe het zit met doorlooptijden, wachttijden, verdwenen zaken, verdwenen respondenten, doden in politieregisters – er zitten in Nederland 8.000 mensen in politieregisters die al lang dood zijn –, enzovoort. We kunnen zo op gezette tijden de *flow* en *stock* van het hele terrein van Veiligheid en Justitie (in dit geval betreffende criminaliteit en strafrecht) laten zien. Dan denk je misschien: 'Dat zal wel gigantische hoeveelheden geld kosten'. Helemaal niet, die data zijn er al en met een paar fte (en relevante software) kom je een heel eind.

Ergens in het provinciehuis moet een eenheid zijn – vroeger zou je het misschien

registratuur en statistiek hebben genoemd – die zich daarmee bezighoudt. Het gaat om het ontsluiten van akten, prestaties, activiteiten enzovoort in een digitale wereld. Bij provincies kun je dan denken aan gegevens over: subsidies, vergunningen, inspecties, heffingen, onderhoud van wegen, incidenten, meldingen, maar ook uitspraken van bestuurders en toezeggingen die ze hebben gedaan.

Je hoeft dus zelf geen onderzoek te starten om die gegevens boven water te krijgen. Mijn advies voor de verankering van evaluatieonderzoek is dan ook om een programma te ontwikkelen en een plan te maken voor aanwezige registerdata, hoe je die kunt opschonen en veredelen. Je kunt dan ook *early warnings* gaan formuleren. Niemand kan dan meer aankomen met het argument dat je opmerkingen gebaseerd zijn op een veel te kleine steekproef. Nee, ze zijn gebaseerd op duizenden data over een reeks van jaren: een *compelling force of evidence*.

Een laatste manier om de kwaliteit te verankeren is het openbaar maken van de veronderstellingen die aan het beleid ten grondslag liggen. Er is vandaag de dag heel veel informatie over beleidstheorieën – zeker een paar honderd gezaghebbende studies inmiddels – en als onderzoekseenheid binnen een provincie zul je de beleidsmensen moeten 'verleiden', misschien zelfs dwingen om zich daartoe te verhouden en de assumpties die aan het beleid ten grondslag liggen openbaar te maken.

In dat verband herinner ik mij mijn eerste contact ooit met een lid van de Tweede Kamer, dat was Gerd Leers, eind jaren 80/begin jaren 90. Naar aanleiding van het centraal thema onderzoek 1991 van de Algemene Rekenkamer naar de stand van de beleidsevaluatie wilde hij een motie opstellen en indienen waarin, analoog aan de Memorie van Toelichting bij wetgeving, er ook een Memorie van Toelichting zou moeten komen bij elk nieuw beleid. Alle nieuw beleid van enige omvang zou in zijn gedachte de beleidstheorie, waar de voorstellen op gestoeld waren, moeten articuleren. De motie heeft het op een paar stemmen na niet gehaald maar is volgens mij nog steeds heel relevant. Je zou bij provinciaal beleid dus een oplegnotitie over de beleidstheorie kunnen afdwingen voordat de politieke besluitvorming erover plaatsvindt. Ik ben ervan overtuigd dat dat het debat aanscherpt en de kwaliteit van het beleid ten goede komt.

Programmering

Het formuleren van meerjarige onderzoeksprogramma's is heel ingewikkeld. We zijn als Panel van Advies voor de IOB daar recent mee bezig gegaan. Waar wij tegen aanliepen is dat er wel een heel keurig interactief proces voor ontwikkeld is en dat er aan het eind van zo'n rit een fraai programma uitrolt, maar een echt

intelligente onderbouwing zit daar nog niet in. We wilden weten hoe we dat op een betere manier kunnen doen en hoe dat in andere landen gaat. Overigens geldt deze bevinding zeker niet alleen voor deze inspectie, maar breder.

Eén van de gedachten is dat je, los van het programma dat je *bottom up* in co-creatie met andere partijen maakt, als onderzoeksinstituut moet zeggen: wij willen in ieder geval dat deze twee of drie projecten hoe dan ook altijd plaatsvinden. Die zijn van zo'n vitaal belang dat ze buiten de programmering staan en altijd doorgaan. Voor het WODC is dat bijvoorbeeld de recidivemonitor. Wij weten, vanaf begin jaren 90, van alle Nederlanders hoe vaak ze met politie en justitie in aanraking zijn gekomen. We volgen hen door de jaren heen. Als iemand zou zeggen: we moeten maar ophouden met die monitor dan is dat hetzelfde als wanneer je tegen het KNMI zou zeggen: de komende vijf jaar verzamelen we geen gegevens meer over het weer. Een ander onderzoek dat we altijd blijven doen is onderzoek naar de manier waarop Nederlanders hun geschillen oplossen. Dat is maatschappelijk zo belangrijk dat het moet worden uitgevoerd. Essentiële onderwerpen moet je dus altijd buiten de programmering houden. Doe je dat niet, dan kunnen ze onderdeel worden van allerlei nut- en noodzaakdiscussies, met alle risico's van dien. Je moet wel heel zorgvuldig kiezen voor welke onderwerpen je dat wilt doen. Als je 25 onderwerpen zo wilt behandelen, dan kan dat niet.

Wat we ook wel hebben geprobeerd is om de programmering in één plaatje samen te brengen, gerangschikt naar de kerndoelen en subdoelen van het beleid. Zodat je kunt zien of je spreiding van onderzoeksonderwerpen enigszins op orde is. Dat geeft enig houvast, maar grosso modo blijft een meerjarige onderzoeks-programmering best ingewikkeld.

En dan moet je daarbij ook nog eens bedenken dat Nederland een geschiedenis heeft in programmering. In de jaren 70 is het geprobeerd met behulp van de zogenaamde nationale programma's. Die waren bepalend voor de onderzoekswereld. Vervolgens is het co-creatie geworden, waar je partijen met elkaar het gesprek liet voeren over wat er moest gebeuren en een gezamenlijke verantwoordelijkheid gaf voor het eindresultaat. Als je me vraagt: Is er ook nog een derde weg? Ik zie hem niet.

Burgers

Wat de rol van de volksvertegenwoordiging is bij evaluaties? Allereerst wil ik daarbij opmerken dat een flink deel van de onderzoeken die wij bij het WODC uitvoeren, rechtstreekse verzoeken zijn van de Tweede Kamer. Dus de link met de volksvertegenwoordiging is duidelijk.

Maar misschien moet je ook nog wel een andere, meer principiële vraag stellen: Is de volksvertegenwoordiging het volk? In dat verband is het goed om stil te staan bij iets dat *civic evaluation* wordt genoemd. Daarbij gaat de burger zelf als een evaluator te werk: hij bestudeert de doelen van het beleid, verzamelt en analyseert gegevens die hij ophaalt uit bronnen van allerlei instituten en confronteert daarna de overheid met zijn bevindingen. Via het open databeleid dat in Europa is afgesproken komen steeds meer evaluaties van onderzoeksinstituten ter beschikking. Dat geldt ook voor statistische gegevens van de nationale bureaus. Eén van de effecten daarvan is dat burgers, al dan niet georganiseerd, elkaar steeds vaker vinden rond thema's die voor hen belangrijk zijn en in de meest uiteenlopende netwerkconstructies voor hen relevante onderwerpen bij de kop pakken en uitspitten. Het is het begin van een nieuwe traditie aan het worden in Europa.

Ik wil in dat verband ook wijzen op het essay van Marc Chavannes in de publicatie *'Turven, tellen en toetsen'*, het jubileumboek uit 2007 van Vide, de beroepsvereniging voor toezichthouders, inspecteurs, handhavers en evaluatoren. Chavannes laat daarin met bijzonder fraaie voorbeelden zien hoe het establishment in de Verenigde Staten niet meer om de burger heen kan. Het zijn juist de burgers die zich roeren en in toenemende mate onrechtmatigheden aan het licht brengen.

In de App-wereld zie je de ene na de andere toepassing op de markt komen waarmee gegevens op alle mogelijke terreinen verzameld kunnen worden. Ik ken schitterende studies bijvoorbeeld waarin via burgers het beleid op het gebied van geluidsoverlast wordt getoetst. Met hun smartphones verrichten ze geluidmetingen die *realtime* worden doorgestuurd en opgeslagen in een centraal gegevensbestand, van waaruit weer analyses gemaakt kunnen worden. Dat is een volstrekt nieuwe benadering. Nu is het meestal zo dat een gemeente na klachten van burgers een buurtonderzoek naar geluidsbeleving houdt ('Hebt u veel of weinig last van de disco?') en als dat niet voldoende gegevens oplevert misschien één of twee keer met een geluidswagen uitrukt om de decibellen te meten. In zo'n nieuwe setting brengen honderden burgers geheel vrijwillig duizenden metingen bij elkaar en zorgen daarmee voor een gigantische hoeveelheid gegevens waar de gemeente niet omheen kan.

Niemand weet op dit moment nog precies hoe je dit soort acties van de burger moet incorporeren in de formele evaluatieprocessen, maar dat het een ontwikkeling is waar enorme mogelijkheden liggen, is voor mij duidelijk. Wat ook duidelijk is, is dat je op twee manieren met zo'n ontwikkeling kunt omgaan: of je laat de burger links liggen of betreft hem erbij. Als je echt wilt innoveren, vind ik dat je in deze laatste richting verder moet zoeken. Ik vind dat een fascinerend iets. Ook de

volksvertegenwoordiging zal moeten aantonen dat zij echt de burger representeert. Want slaagt ze daar niet in, dan zal de burger steeds meer zijn eigen weg gaan en met nieuwe media het bestuur gaan bestoken.

Actoren

Provincies voeren hun beleid uit in samenspel met andere overheden, bedrijfsleven en maatschappelijke organisaties. Bij evaluaties loop je dan tegen het probleem aan hoe je kunt laten zien wat de interventie van die ene actor te midden van al die andere actoren heeft opgeleverd. Met andere woorden: Hoe doeltreffend en doelmatig was de euro die de provincie heeft ingezet, naast de euro's van die andere partijen?

Daar is eigenlijk maar één oplossing voor. Je moet daarvoor de cruciale actoren redelijk vroeg in het traject bevragen: 'Wat is uw verwachting van dit programma? Waarom doet u mee? Wat is uw bijdrage in termen van middelen en resultaten?'. Je brengt dus in kaart hoe de actoren die erover gaan het gewicht van hun eigen bijdrage wegen. Dan krijg je bijvoorbeeld van de zeven belangrijkste actoren een soort overzichtsplaatje. Je kunt door middel van *infographics* laten zien hoe partijen zich tot elkaar verhouden. En wanneer je in een *netwerkevaluatie* naar de resultaten van een programma hebt gekeken, dan ga je de partijen confronteren met wat zij in het begin hierover hebben gezegd. Dan wordt hun bijdrage aan het geheel ook in het goede perspectief geplaatst. En dan kan de muis misschien wel roepen dat hij zo lekker samen met de olifant heeft gestampt, maar de andere deelnemers zullen hem dan toch tot de orde roepen en zijn verhaal in perspectief plaatsen.

Wanneer het alleen gaat om uitvoering, dan is het wel mogelijk om ieders bijdrage scherp in beeld te krijgen. Er is namelijk altijd iemand verantwoordelijk voor (een deel van) de uitvoering. En als het daar mis is gegaan, dan kun je in ieder geval uitspraken doen over de aansturing.

De enige andere manier om de mate van doeltreffendheid en doelmatigheid in kaart te brengen, is om elke afzonderlijke bijdrage te kwantificeren in termen van zijn betekenis. Ik denk niet dat dat haalbaar is. Met een project van € 600 miljoen kun je het misschien proberen, maar als je niet aan die schaal komt – en voor de provincie is dat doorgaans het geval is – dan is het een heel inefficiënte manier van onderzoek doen. Meer aandacht voor de *cost and benefit* van evaluaties is in dat verband ook een belangrijke zaak.

Colofon

Interviews: Zuidelijke Rekenkamer

Landschapsfotografie: Henk van den Berg, Oss

Vormgeving: Henk van den Berg Creatieve Communicatie, Oss

Drukwerk: Drukkerij Wihabo, Geffen

Artikelen of delen daarvan mogen alleen uit deze publicatie worden overgenomen met bronvermelding en goedkeuring van de Zuidelijke Rekenkamer.

ISBN 978-90-8768-035-0



9 789087 680350



Postbus 2000
5600 CA Eindhoven
Keizersgracht 5
5611 GB Eindhoven
tel. 040 - 232 93 38
www.zuidelijkerekenkamer.nl

