

ARTIKEL

Early warning evaluation en de invoeringstoets

Vroege signalering over de koersvastheid van beleid als bijdrage aan de nieuwe bestuurscultuur

Arno F.A. Korsten*

1 Inleiding

Over de uitvoering van overheidsbeleid is in deze kabinetsperiode veel te doen, maar er lijkt een tekort aan goede verbetervoorstellen. Gepoogd wordt in die lacune te voorzien.

Optimalisering van de uitvoering van beleid is aan de orde bij het streven naar een *nieuwe, meer humane bestuurscultuur*. De discussie in kabinetskringen, parlement, samenleving en persmedia over de rampzalige uitvoering van regelingen op het vlak van aardbevingschade is zo heftig dat verantwoordelijk staatssecretaris Vijlbrief zelfs enkele dagen per week in Groningen verblijft om daar zijn werk te doen. Daarnaast zijn er andere affaires geweest en nog lopend, zoals de toeslagenaffaire op het vlak van kinderopvang. De bestuurscultuur moet verschuiven naar meer uitvoeringsgerichtheid, meer en eerder oog hebben voor de mensen die geraakt worden en meer aandacht voor uitvoerders die het beleid kunnen zien derailen. Wat heeft het beleidsonderzoek hier te bieden?

Ik kom met het voorstel van een *early warning evaluation*, waar ik zelf ervaring mee heb opgedaan. Een *early warning evaluation* (EWA) brengt korte tijd na de start van de inwerkingtreding van overheidsbeleid in kaart wat maatregelen doen bij uitvoerders en doelgroepen en hoe het beleid werkt of naar verwachting gaat werken. EWA brengt eerste ervaringen in beeld opdat zo nodig correctie kan plaatsvinden.

Een EWA is een *waarschuwend evaluatie*, een tweede kans voor besluitvormers om 'beleid te corrigeren als het nog niet te laat is'. Deze evaluatie is vooral gericht op het krijgen van *vroege signalen* over de vraag of een beleidsprogramma in de invoering nog steeds *op koers is of blijft*. EWA is derhalve primair een verkenning van de koersbestendigheid. Het is een remedie voor een kabinet om in kaart te brengen of een beleid in de uitwerking wegglijdt van de geplande hoofdkoers of dat dreigt te

* Arno F.A. Korsten is emeritus hoogleraar Bestuurskunde aan de Open Universiteit en Maastricht University. Hij is voormalig lid van de Raad voor het Openbaar Bestuur, was lid van de wetsevaluatiecommissie-Oosting (Wabm) en was betrokken bij de evaluatie van de Wet ketenaansprakelijkheid. Daarnaast was hij (co-)voorzitter van visitatiecommissies ter beoordeling van Vlaamse steden. Hij was ook toezichthouder, informateur en bemiddelaar. Korsten publiceert onder meer over beleidsevaluatie, wetsevaluatie en horizonwetgeving.

Arno F.A. Korsten

doen en tijdig te corrigeren. Tot de producten van een EWA behoort een overzicht van geconstateerde en verwachte knelpunten, afwijkingen en verbeteruggesties.

EWA heeft wat weg van de invoeringstoets, die door het kabinet-Rutte IV is ingevoerd. Maar een invoeringstoets is bescheidener en minder degelijk dan EWA. EWA verschilt door de beoogde toepassing van enkele methodieken. Welke? Een EWA-evaluator moet naast (1) een knelpunteninventarisatie (2) de proces- en/of stappenlogica doordenken die in het gehele beleid, meestal wet- en regelgeving, opgesloten zit. Daarna volgen: (3) de consequentieanalyse en (4) de consistentieanalyse. Een centrale vraag is: als de eerste stappen in een beleidstraject gezet zijn en dit traject wordt verder afgelopen, waar komen de actoren die bij de beleidsuitvoering betrokken zijn dan naar verwachting verderop in de tijd uit en waar is inconsistentie zichtbaar met die lijn?

Dit artikel geeft een overzicht van enkele aspecten van deze nieuwe, nog zelden toegepaste vorm van beleidsevaluatie. Het sluit af met een korte vergelijking met enkele ervaringen uit de toepassing van de invoeringstoets. Daarmee vormt dit artikel een bijdrage aan de discussie over de nieuwe bestuurscultuur in de rijksdienst.

2 Vroege waarschuwingen

De opdracht tot een *early warning evaluation* (EWA) zal doorgaans binnen een jaar na de introductie van wet- en regelgeving bestuurlijk verstrekt zijn. Dat is een gunstig moment om te kijken of *het kompas* van de rijksoverheid de goede richting wijst en naar verwachting *blijft* wijzen. Want daar gaat het om bij EWA: beleid in de goede richting laten werken en tijdig bijstellen als dat nodig is. EWA is gericht op 'leren'. Allerlei vormen van leren zijn uiteraard mogelijk.

Veel overheidsbeleid impliceert de totstandkoming van prestaties of resultaten die moeten uitmonden in gewenste effecten. In werkelijkheid treden in de in- en uitvoering van beleid vaak zodanige belemmeringen en andere problemen op dat beoogde doelstellingen niet bereikt worden. In het rapport *Ongekend onrecht* liet de parlementaire commissie-Van Dam zien dat de uitvoering van de kinderopvang door middel van toeslagen jarenlang niet liep zoals voorzien en gewenst was. Te veel burgers werden daar de dupe van. Was er niet gewaarschuwd? Een topjuriste van de Belastingdienst – Sandra Palmen – schreef enige tijd na invoering van het toeslagenbeleid een memo (bekend geworden als het memo-Palmen) waarin stond dat de dienst laakbaar had gehandeld bij de stopzetting van de toeslag voor burgers bij kinderopvang en dat ouders gecompenseerd moesten worden voor het handelen van de dienst. Het memo werd wel in een beraad van managers van de Belastingdienst behandeld, maar bleef vervolgens ergens binnen de Belastingdienst en het departement van Sociale Zaken hangen en raakte uit het zicht. Het memo spoorde niet met het heersende paradigma achter het beleid. De voorgestelde koerscorrectie werd genegeerd. Memoschrijfster Sandra Palmen heeft nog even overwogen om

klokkenluider te worden, maar heeft dat uiteindelijk nagelaten. Ze kwam er voor zichzelf achter dat klokkenluiders er doorgaans ‘niet gelukkiger op worden’.

Het kabinet trok uit het rapport-Van Dam c.s. de les dat het netwerk van betrokken organisaties, waaronder meerdere diensten en functionarissen van departementen alsook bewindslieden, veel eerder open had moeten staan voor wat er beter kan. Fouten moeten tijdig worden gesignaleerd én opgepakt. Signalen en attenderingen moeten niet in de lijn van een organisatie blijven hangen. Problemen van regelgeving en uitvoering moeten worden opgelost. Hoe daartoe te komen? Het is niet genoeg als verantwoordelijke bewindslieden zich excuseren en/of aftreden. Het volstaat niet als memo’s daadwerkelijk doordringen tot de ambtelijke top. De rijksoverheid dient systematisch en gericht te onderzoeken hoe nieuwe of gewijzigde maatregelen, meestal regelingen, uitpakken voor doelgroepen (dus mensen en organisaties) en ervan leren door behoud van het goede en correctie waar nodig. Het adagium is: ga tijdig op zoek naar wat deugt en niet deugt of suboptimaal is, en kijk naar verbetermogelijkheden, zodat de gewenste gevolgen later alsnog bereikt worden. Pas het beleid als het nodig is inhoudelijk aan, verstrek meer of andere middelen voor de uitvoering, faseer zo nodig en/of verbeter de in- en uitvoering op andere wijze. Een kabinet moet niet voetstoots aannemen dat alles vanzelf wel goed komt.

Een vroege *early warning evaluation* (EWA) kan daaraan bijdragen. Met EWA is jaren terug al enige ervaring opgedaan. Daarom volgt hier een uiteenzetting van wat EWA is en wat de relevante onderdelen zijn. Ook wordt aangegeven in hoeverre EWA verschilt van de *invoeringstoets*, waarmee in de rijkdienst recent eerste ervaringen zijn opgedaan.

De opdracht tot een EWA komt eigenlijk te vroeg als het oog primair gericht zou zijn op het in beeld brengen van prestaties en gewenste en ongewenste beleids-effecten en de bijdrage van beleid daaraan. Beleidseffecten zijn doorgaans dan bij lange na nog niet zichtbaar. Dat houden juristen en rechtssociologen ons voor. Wat dan te doen? Een evaluatiepoging als opdrachtresearch weigeren terwijl beleidsvoerders toch een beeld willen krijgen ‘of het de goede kant op blijft gaan’? De bestuurskundigen Wil Gooren, Arno Korsten en Ignace Snellen werkten samen in een aantal researchprojecten op het terrein van technologisch innovatiebeleid van de rijksoverheid en zochten een oplossing voor dit probleem (zie literatuurlijst). Het ging om de volgende projecten: de evaluatie van de Stichting Technische Wetenschappen (STW); de evaluatie van een onderzoeksprogramma op het vlak van biotechnologie; en de evaluatie van transferpunten op het vlak van technologische innovatie aan universiteiten en technische hogescholen. De vraag bij enkele van die evaluaties was: komt er na verloop van tijd daadwerkelijk technologische innovatie, al of niet via kennistransfer, van de grond en lukt het om daar vanaf het begin van de inwerkingtreding van beleid beleidsmatig naartoe te blijven werken?

EWA staat niet gelijk aan alle evaluatieopdrachten waartoe aan de vroege kant wordt besloten, want dan zou hier geen afzonderlijke aandacht nodig zijn voor dit type evaluatie. EWA is meer. Welk beeld rijst dan op over EWA bij overheidsbeleid?

Arno F.A. Korsten

Ik ga hierop in, ook omdat deze aanpak in gangbare publicaties over evalueren ontbreekt en toepassing urgent is gezien de vele beleidsaffaires (Bekker, 2020; Van der Knaap et al., 2020).

3 Wat Early Warning Evaluation is

Omschrijving

Early warning evaluation (EWA) is een *gebruiksgerichte* analyse van de vraag of opeenvolgende en/of parallelle beleidsgerichte acties na de inwerkingtreding van beleid de *goede richting* hebben en houden en *nog in lijn zijn* en *blijven* met beleidsdoel-einden en beleidsmatige oogmerken en verwachtingen. Blijken moet of de beleidsuitvoering dreigt *af te drijven* van het beleid waardoor de kans groot is dat van het gewenste beleid niets of niet voldoende terechtkomt. EWA stelt het beleid in eerste instantie niet ter discussie.

Als een EWA betrekking heeft op een wet, loopt de vroege evaluatie vooruit op een echte, omvattende *wetsevaluatie*, die meestal enkele jaren later start. Denk maar aan de eerste evaluaties van de milieuwet (Wabm), de wetgeving over openbaarheid (Wob), de Wet universitaire bestuurshervorming (Wub) en de Wet ketenaansprakelijkheid (Wka).

Een wetsevaluatie (ex post) is gericht op een wetstechnische analyse, op de analyse van jurisprudentie, op de uitvoering en de gewenste en ongewenste effecten van beleid en eventueel op de (on)gewenste neveneffecten (Korsten, 1981; Van der Knaap et al., 2020). Een wetsevaluatie kan ook doelstellingsvrij geschieden en de nut en noodzaak van een wet belichten, zoveel jaar na inwerkingtreding.

Er is altijd weer de noodzaak tot specificatie van de probleemstelling van een onderzoek en van een werkwijze, ook bij een EWA. Wat neemt een evaluator mee en wat niet? Elke evaluatie kent een ontwerpdimensie.

Stappenanalyse en logica

EWA is gericht op de analyse van de opeenvolgende stappen van 'beleid tot effect'. Het is een soort '*ketenanalyse met een beoordeling op gevolgen van maatregelen en de consistentie (of logica) van vervolgacties*' waarbij nagegaan wordt of de verdeling van taken, bevoegdheden en middelen gestalte heeft gekregen (organisatiekundige kant) en nog steeds in lijn is met de beleidsinhoud en de bedoelingen achter het beleid; met de voorlichting over de beleidsinhoud; en gedaan is wat beleidsinhoudelijk gezien nodig is.

Consequentieanalyse en consistentietoets

Het gaat er vooral om, om te zien of de uitwerking van doelen, middelenkeuze en andere elementen *consequent* en *consistent* is tot opzichte van de gedachte route van 'beleid tot effect', zodat daarmee *een voorspellend antwoord* gegeven kan worden op de vraag of de doelstellingen überhaupt later bereikt zullen *kunnen* worden. Sporen de al getroffen maatregelen met de veronderstellingen achter het beleid of niet? Of gaan de maatregelen al afwijken of niet? Wat is daarvan dan de reden? Zijn

de bewindspersoon en het parlement daarvan in het kader van de feedback en verantwoording op de hoogte? Hoe zit het met de legitimiteit van het beleid?

Behoeftte aan terugkoppeling

Is er een grote behoefte aan feedback op beleid waartoe besloten is? Die behoefte is volgens de commissie-Bosman (2021) en anderen die zich bezighielden met beleidsuitvoering enorm.

Zijn er voorbeelden van rijksbeleid waarbij overduidelijk een EWA ontbroken heeft of niet goed van de grond is gekomen? Het antwoord valt eenvoudig te vinden in de publicaties van de Algemene Rekenkamer, de Nationale Ombudsman en de Onderzoeksraad voor Veiligheid (OvV). Een voorbeeld is de aanpak van de schades als gevolg van de aardbevingssschade door gaswinning in Groningen. De aanpak was volgens deze drie organisaties lange tijd volkomen inadequaaf. De waarschuwingssignalen kwamen veel te laat (OvV, 2015). Zou een EWA zijn toegepast inzake de afhandeling van aardbevingssschade in Groningen, dan was parlementair onderzoek vermoedelijk niet nodig geweest. Dan had er jaren eerder een beleidscorrectie plaatsgevonden, onder meer in de vorm van herorganisatie van delen van de uitvoering en het nakomen van afspraken. Zo is de positie van de NAM bij de schadeafhandeling in Groningen jarenlang veel te groot geweest.

Een ander voorbeeld zijn de opeenvolgende uitkoopregelingen, gericht op beëindiging van boerenbedrijven. Deze uitkoopregelingen hebben de afgelopen 25 jaren niet gebracht wat gehoopt en verwacht werd. Dat is voorwaar geen snelle evaluatie, maar de casus is inzichtelijk. Een studie van het PBL maakt duidelijk wat, waarom en hoe. De studie laat zien dat de feedback vaak jaren te laat komt.

Wat is dan het voordeel van een snelle evaluatie als voorzien bij EWA? Als een half jaar na inwerkingtreding van beleid een EWA plaatsvindt, zijn de belangen nog niet zodanig gestold dat de ontvankelijkheid voor 'beleidsgericht leren' verdwenen is.

Bij modificerende wetgeving

Een EWA is bijna altijd meer zinvol in geval van modificerende wetgeving en maatregelen (term van de jurist Jimmy Polak) dan van codificerende wetgeving. In dat geval gaat het om het aanbrengen van gedragsbevestigingen of -veranderingen in delen van de samenleving. Mestwetgeving gericht op vermindering van CO₂-uitstoot is een goed voorbeeld van modificerende wetgeving omdat boeren en andere agrariërs hun gedrag moeten afstemmen op de wet.

4 Een referentiemodel van beleidsinvoering

Wat zou een goed vertrekpunt zijn voor de *early warning evaluator*?

Een evaluator moet ergens beginnen en dat is niet direct bij de consequentie-analyse.

Arno F.A. Korsten

Het beste is om een eerste verkenning te plegen van het beleid zelf: het waarom ervan, de inhoud, de belangrijkste actoren en processen, de invoeringsstrategie, waarom het te doen is. Deze actie is niet alleen een kennismaking, maar kan ook nuttig zijn om een EWA *tailormade* te maken.

Om een EWA te kunnen verrichten is een referentie aan wat doorgaans tot de in- en uitvoering van beleid gerekend wordt, handig. Kortheidshalve worden hier op basis van ervaringen met invoering van beleid enkele voorwaarden voor succesvolle beleidsuitvoering geformuleerd. Ik noem dit *het referentiemodel*.

Beleid kent ambitie en richting. Beleidsvoerders willen iets bereiken: wat, waarom, hoeveel, van welke kwaliteit, voor wie, wanneer, hoe? Beleid is gericht op het nastreven en realiseren van geëxpliciteerde doelstellingen met behulp van bepaalde uitvoerders van beleid, bepaalde prestaties (en andere niet), inrichtingsprocessen, budgetten en andere middelen. Voor de realisatie van beleid, in de vorm van prestaties en beoogde effecten, zijn voldoende middelen en tijd gereserveerd. Te ambitieuze doelen kunnen het beleid vertragen of doelbereiking onmogelijk maken. Een memorie van toelichting op een wet kan hierover al veel vaststellen. Een uitvoeringstoets *ex ante* vult dit aan.

De evaluator leert zo mogelijk om de bril met de goede glazen op te zetten.

Beleid is geen blinde sprong in het diepe. Achter een beleidsprogramma – dus een concreet beleid – gaat een geheel van verwachtingen en veronderstellingen schuil over het bereiken van doelstellingen met bepaalde middelen. Die veronderstellingen moeten bij voorkeur juist zijn. Dit heet in het evaluatiejargon *de beleidstheorie*. Opdracht: breng die in kaart.

Invoering en taaktoedeling. Crisisbeleid blijft nog op de plank liggen voor als de pandemie of een andere crisis dreigt. Het meeste beleid wordt na besluitvorming, eventuele pilots en een beslissing over de inwerkingtreding evenwel ingevoerd.

Beleidsinvoering is te omschrijven als het gereed maken van de uitvoerende organisatie voor de toepassing van middelen van een beleid. Dat gebeurt door het vertalen van beleid in taken die zijn gericht op beleidsprestaties, het opsplitsen in deeltaken en het toedelen van taken aan uitvoerders.

Taken gaan gepaard met arbeidsprocessen, inzet van technologie en andere processen. Zo int de Belastingdienst inkomstenbelasting op basis van aangiften, en verstrekt het CBR rijvaardigheidsbewijzen op basis van uitgevoerde autorijexamens.

De uitvoeringsorganisatie. De uitvoering ligt idealiter in de hand van een overheid, één uitvoeringsorganisatie en één toezichthouder en is niet afhankelijk van andere organisaties. Dan is de doorzettingsmacht in één hand en raken taken, bevoegdheden en verantwoordelijkheden niet verdeeld. Het aantal schakels is dan beperkt. Er kan nog overzicht ontstaan.

In werkelijkheid zullen vaak meerdere organisaties actief zijn die afhankelijk van elkaar zijn. Nieuw beleid of een beleidswijziging daalt immers neer in een bepaald stelsel van binnenlands bestuur, waarin al verwant beleid wordt uitgevoerd. Denk maar aan de asielketen met organisaties als het COA, de IND en de gemeenten.

In een beleidsnetwerk is samenwerking nodig die doorgaans moet uitmonden in doelconvergentie, middelenbundeling en mechanismen voor conflictregulering, bemiddeling, verantwoording en toezicht. Complexe brandende beleidskwesties of controversen (*'wicked problems'*) als klimaatpolitiek, energietransitie en asielbeleid worden vrijwel altijd aangepakt in stelsels van meerdere organisaties in een meerlagig bestuur (Korsten, 2019). Dan luistert de afstemming al nauwer. De vraag dient zich aan wie als systeemverantwoordelijke actor optreedt, wie overzicht heeft en de sterkten en zwakten van het stelsel overziet, en zo nodig regie voert door zwakheden in een stelsel te corrigeren.

Procesdoordenking. De verbanden tussen doelstellingen achter beleid en taken voor uitvoerende organisaties zijn bij voorkeur direct en er zijn weinig tussenschakels. Zijn er veel schakels en/of zijn ze complex, dan vergt dat het voorkomen van bottlenecks en knelpunten, wellicht door ketenmanagement.

De doelstellingen, taken, afstemmingen, verbanden, schakels en prestaties betreffen de *operationalisatie* van beleid. Tijdens de invoering van beleid komt naar voren 'hoe beleid gaat draaien', hoe de logica van opeenvolgende stappen en handelingsstromen (bijvoorbeeld typen vergunningen, typen rijbewijzen) is. Daartoe verschijnen er handboeken, draaiboeken, circulaire, handleidingen, instructievideo's, pilots, trainingen, protocollen.

Bij conditionele beleidsprogramma's, bijvoorbeeld op het vlak van burgerzaken – denk aan inschrijving in het bevolkingsregister – is de procesdoordenking overzichtelijker dan bij doelprogramma's.

Bij ongetemde 'brandende kwesties', zoals energietransitie, is de procesdoordenking nog complexer omdat het hier meestal gaat om meerdere beleidsmaatregelen.

Overzicht. Een EWA-evaluator moet op zoek naar de hoofdstructuur van het beleid en moet overzicht krijgen. Veldkennis en gesprekken met uitvoerders en burgers ('storytelling') helpen hierbij. Een invoeringsplan voor een beleid is hiervoor ook behulpzaam. Dat plan bevat veel informatie over beleidsuitvoering en daarmee over doelstellingen, posities, taken, middelen, doelgroepen en sturings-, afstemmings- en controleprocessen. Een invoeringsplan kan wat formeel zijn.

Uitvoeringsproces. De beleidsinhoud en het invoeringsplan geven samen informatie over de gewenste volgorde van stappen bij de invoering van het beleid. De volgorde is expliciet waardoor iets van 'raderwerking' kan ontstaan. Is het beleid niet doelvolgend, maar eerder doelzoekend dan zal er sprake zijn van zekere beleidsvrijheid aan de kant van uitvoerders tot het maken van keuzen en normering door middel van informele regels en toepassing daarvan. Studies naar '*street-level bureaucracy*' illustreren dit.

Acceptatie en steun. Er is volgens het referentiemodel volledig begrip van en overeenstemming over beleid (doelstellingen, middelen, aanpak) onder uitvoerders en doelgroepen en dat blijft zo gedurende het uitvoeringsproces. Is dat de werkelijkheid?

Wie als signalerende evaluator aan het werk gaat, kan aan de genoemde referentiepunten, kortweg het ‘referentiemodel’, al een reeks vragen ontleen om de waardevolle evaluatie te richten.

5 De methodiek van EWA: onderdelen op een rij

Het genoemde referentiemodel biedt bij een EWA een eerste houvast om het beleid waarvan de invoering op gang is gekomen, in kaart te brengen. Omdat het beleid echter nog amper ingevoerd is, dwingt het vroege moment van evalueren tot een bepaalde evaluatiemethodiek. Men kan niet volstaan met een analyse van knelpunten in de uitvoering van beleid, ook al is dat wel praktisch.

Om een EWA uit te voeren dient zich een aantal onderdelen aan:

- a de reconstructie van de bedoelingen en veronderstellingen achter het beleid, dus de beleidstheorie;
- b de operationalisatie van het beleid in invoeringshandelingen;
- c de consequentie-analyse;
- d de consistentietoets;
- e de sturing op verbetering.

De werkzaamheden vanuit het referentiemodel en vanuit a. en b. zijn voorwerk om begrip te krijgen van het beleid en een eerste zicht op wat er eventueel ‘mis’ is. De kern van EWA wordt gevormd door de acties die onder c., d. en e. worden genoemd.

Actie a. De reconstructie van bedoelingen en veronderstellingen. De EWA-evaluator formuleert de doelstellingen en verwachtingen achter beleid en benoemt de cruciale veronderstellingen over het beleid. Dat levert in essentie de beleidstheorie achter het beleid op. Deze informatie is een voorwaarde om überhaupt een EWA te kunnen doen. Er is een vertrekpunt nodig en dat ligt in de beleidsinhoud en de beleidsgeschiedenis. De historie van de discussie in samenleving en parlement geeft zicht op mogelijke prioriteiten en accenten in beleid en de uitvoering en kritische aandachtspunten (‘waar moet ik als evaluator op letten?’).

Actie b. De operationalisatie. De invoering van beleid maakt zichtbaar welke stappen moeten worden gezet in het uitvoeringsproces.

De *stappenlogica* komt in zicht door – zoals bestuurskundigen stellen – een procesmodel van de beleidsuitvoering te maken.

Een *invoeringsplan* kan hier behulpzaam zijn, want zo wordt duidelijk wie de uitvoerders zijn, wat de hoofdcoers is, welke taakverdeling gekozen is, welke afstemingsmomenten aan de orde zijn, en wat de eventuele knelpunten in de vorm van risico’s, barrières of blokkades zijn. Er ontstaat zo een overzicht, wellicht in de vorm van een *stromenschema*.

Actie c. De consequentieanalyse. Het procesmodel uit de operationalisatie-analyse levert het materiaal om te bezien wat de gevolgen zijn van het beleid als de beleidskoers gestalte krijgt. Treedt afwijking op? Is die betekenisvol? Blijven de beoogde effecten in beeld komen of drijft het beleid weg van de doelstellingen? Dreigt misbruik? Dreigt een fiasco?

De centrale vraag van een EWA is naar verwachting in veel gevallen: ‘Als we zo doorgaan met het in- en uitvoeren van het beleid (operationalisatie), bereiken we dan de gewenste resultaten, doelen en effecten?’ Is er consistentie in de aanpak? Zo nee, waarom niet? En wat doen we daar dan aan, zo mogelijk met enig gevoel voor een *evidence-based* oplossing (dus geen slag in de lucht)? Verdere vragen kunnen zijn: Waar kunnen suggesties gedaan worden? Welke partijen moeten hoe gehoord worden? Kan er een machtsvrije dialoog ontstaan over de in- en uitvoering? Kan de in- en uitvoering aangepast worden en zo ja, wat niet of nauwelijks? Hoe zien we het beleid voor ons qua baten en lasten, qua draagvlak en medewerking in het veld en bij de doelgroep? Wat is de opstelling binnen het beleidsveld en de politieke arena’s? Wie gaat hier de regie nemen?

Actie d. De consistentietoets. Zonder consequentieanalyse geen consistentietoets. De consistentieanalyse moet vaststellen of beleid ergens door inconsistentie uit de rails dreigt te lopen of niet. Zo ja, waar en waarom? Een mogelijkheid is dat bijvoorbeeld kortere beslistermijnen voorzien zijn dan in werkelijkheid optreden.

Actie e. Sturing op verbetering. Achter EWA gaat de opdracht schuil om aan te geven waar bijsturing nodig is en in welke vorm. EWA maakt begrijpelijk waarom het gaat zoals het gaat en of een beleid dreigt weg te glijden van gewenste effecten. EWA heeft tot doel om tijdig te waarschuwen of de koers buiten het zicht raakt. EWA wint aan waarde als deze evaluatie niet alleen terugblijkt maar *prospectief* gericht is. Wat kan of moet beter en hoe dan? EWA past dus in een lerende aanpak.

EWA heeft vooral dan betekenis als er een beraad over gevolgen ontstaat tussen bestuurders, beleidsuitvoerders, doelgroepen, zaakwaarnemers van burgers, en politieke representanten. Dat kan omdat de staat een staat met tal van gesprekstafels is: watertafels, klimaattafels, dialoogtafels.

Wat kunnen consequenties zijn van EWA?

Hier dient zich een breed scala van opties uit de studie van beleidsdynamiek van organisatie van beleid aan. Denk aan enkele mogelijkheden: beleid wordt ongewijzigd voortgezet; er volgen nieuwe instructies; nieuwe middelen worden ingezet of het volume aan middelen wijzigt; of het beleid wordt partieel of integraal beëindigd; of beleidsopvolging treedt op (bijvoorbeeld ‘nieuwe schoenen in plaats van oude’). Tal van andere mogelijkheden zijn hier denkbaar.

De kernwoorden rond EWA zijn dus: beoordelen door tijdig signaleren; de vraag stellen: met beleid op koers blijven of wegdrijven?; de beleidstheorie als vertrekpunt nemen; de operationalisatie-, consequentie- en consistentietoets toepassen; de bezinning op verbetering en leren.

6 Hoe EWA zich verhoudt tot andere terugkoppelingen

EWA is een vroege, prospectief gerichte evaluatie-aanpak gericht op de vaststelling van knelpunten in de uitvoering en de koersvastheid van beleid en beleidsinvoering. Evaluatoren bezien of een bijstelling van beleid eventueel nodig is om beleid weer op de ‘goede’ koers te krijgen. EWA volgt wat gebeurt bij de invoering met *een bepaald doel*: is het beleid op koers en is de waarschijnlijkheid heel groot dat dit zo blijft of is correctie nodig in beleid en/of invoering en welke dan? Behoeft de beleidsinhoud verandering, de kwantitatieve of kwalitatieve organisatiecapaciteit, ja wat eigenlijk? Dat is de kern.

Zijn er alternatieven voor een vroeg waarschuwende evaluatie, dus voor EWA? Hoe verhoudt EWA zich tot andere vormen van terugkoppeling, dat wil zeggen vertoont EWA gelijkens met

- 1 *real-time evaluation*,
- 2 een reflectietoets, of
- 3 een invoeringstoets?

Daarop wordt achtereenvolgens ingegaan.

EWA heeft iets van real-time evaluation (RTE)

Een *real-time evaluation* (RTE) is ‘an approach designed to quickly and systematically conduct an evaluation when time and resources are limited’ (Rogers, 2021). Maar EWA gelijk stellen aan RTE doet EWA schromelijk tekort. EWA is geen reguliere evaluatie waarbij de evaluator het slechts moet doen met een sobere opzet en weinig hulpmiddelen of data. EWA vindt plaats in een bepaald kader en kent een bepaalde methodische gerichtheid in de vorm van een operationaliseringsanalyse en een consequentie- en consistentieanalyse. EWA is gericht op beantwoording van de vraag of navigeren naar de beoogde beleidseffecten nog steeds lukt.

Conclusie: we voeren RTE weer af als alternatief voor EWA. Exit RTE.

De reflectietoets dan?

De Tijdelijke Commissie Uitvoeringsorganisaties, voor kenners de TCU of commissie-Bosman, besteedt in het rapport *Klem tussen balie en beleid* (2021) aandacht aan de in- en uitvoering van beleid. Ze onderzocht de Belastingdienst, het Uitvoeringsinstituut Werknemersverzekeringen (UWV) en het Centraal Bureau Rijvaardigheidsbewijzen (CBR).

De commissie stelt vast dat deze organisaties behoorlijk functioneren, maar dat de menselijke maat nogal eens zoek raakt. De commissie stelt verder vast dat er voorafgaand aan de besluitvorming te weinig aandacht wordt besteed aan de uitvoerbaarheid van het voorgenomen beleid. Ook de Tweede Kamer staat daar niet genoeg bij stil.

Een van de conclusies van Bosman c.s. luidt: 'Het komt regelmatig voor dat problemen in de uitvoering niet tijdig worden gesignaleerd en op wel gesignaleerde problemen onvoldoende wordt geacteerd.' Daarom moet er meer oog zijn voor signalen direct na het operationeel worden van wetten, maatregelen en andere beleidsinstrumenten. De commissie bepleit daarom een *'reflectietoets'*. Is dat misschien een praktische invulling van een EWA?

De commissie-Bosman stelt dat het van groot belang is om tijdig in kaart te brengen of nieuwe wetgeving het bedoelde effect heeft. Deze wetgeving kan immers grote gevolgen hebben voor burgers. Volgens de commissie kan een toetsing plaatsvinden een jaar na inwerkingtreding van de wet en dus eerder dan het wettelijk voorgeschreven evaluatiemoment, bijvoorbeeld vier of vijf jaar na inwerkingtreding. De commissie noemt dat een *'reflectietoets'*. De minister of Tweede Kamer zou hiertoe het initiatief moeten kunnen nemen. Alle relevante partijen – waaronder burgers, cliëntenraden en belangenorganisaties – zouden bij een dergelijke toets betrokken moeten zijn. De TCU beveelt een pilot aan.

EWA is meer dan een reflectietoets. EWA kent namelijk een methodisch grondpatroon en de reflectietoets niet. De volgende vijf onderdelen zijn bij EWA gegeven: de beleidstheorie als vertrekpunt (waarheen?); de operationaliseringstoets (stapenlogica, het processchema?); de consequentieanalyse (blijft het beleidsschip op koers?); de consistentietoets (wat is de lijn?); de verbeteringssuggesties voor sturing.

Conclusie: de reflectietoets zou een uitwerking kunnen zijn van de EWA als de onderdelen van EWA gevolgd worden. Deze reflectietoets is uitgewerkt in de invoeringstoets, het volgende alternatief.

De invoeringstoets als een soort EWA-light?

Er dient zich een derde alternatief aan voor EWA: de invoeringstoets. Het kabinet-Rutte III heeft besloten tot invoering van de invoeringstoets bij wet- en regelgeving, en het kabinet-Rutte IV heeft dat initiatief in 2021 overgenomen. Minister voor Rechtsbescherming Franc Weerwind is verantwoordelijk geworden voor de invoering en inkleuring van de invoeringstoets op rijksniveau, zoals dat is gaan heten. In een Kamerbrief van 21 april 2021 voor de Tweede Kamer heeft de bewindsman geschetst wat daaronder dient te worden verstaan.

Het kabinet gaf in een reactie van januari 2021 op het rapport *Ongekend onrecht*, dat handelde over de toekenning van toeslagen voor kinderopvang, aan dat een overheid open moet staan voor wat er beter kan. Signalen over fouten in het overheidsoptreden en problemen van regelgeving moeten in de praktijk actief worden opgepakt door ernaar op zoek te gaan. Een overheid moet nieuwsgierig zijn naar de daadwerkelijke gevolgen van beleid of beleidswijzigingen voor burgers en andere actoren en er minder op vertrouwen dat verwachte uitkomsten vanzelf komen. Daarom is een *invoeringstoets* nodig die feedback geeft over de eerste inzichten betreffende de werking van nieuwe regelgeving en de gevolgen voor de uitvoering en

Arno F.A. Korsten

doelgroep. Een invoeringstoets wordt uitgevoerd op het vroegst mogelijke moment dat iets over de resultaten (prestaties) van wet- en regelgeving gezegd kan worden.

De invoering van de invoeringstoets is een ‘grote operatie’. Achter de invoeringstoets gaan echter ook bescheiden pretenties schuil. De toets kan immers geen uitspraken doen over het succes van het beleid, want effecten blijken dan nog niet. Minister Weerwind stelt in de Kamerbrief van 23 september 2023 dat bij de invoeringstoets ‘bewust gekozen is om juist niet de uiteindelijke doelmatigheid en doeltreffendheid te onderzoeken en te focussen op knelpunten en verbetermogelijkheden’. De toets is daarom ook geen uitgebreide en diepgaande wetsevaluatie, want dan is de pretentie om wel effecten en neveneffecten in beeld te krijgen.

De invoeringstoets gaat in op de vraag wat moet gebeuren op het moment dat bijstelling nog mogelijk is. Een invoeringstoets kan immers – zoveel mogelijk systematisch en precies – signalen van burgers en andere actoren uit de doelgroep en van uitvoerders en ketenpartners naar boven halen. De positie van uitvoeringsorganisaties en decentrale overheden wordt zo sterker. Zij kunnen knelpunten signaleren. Daardoor worden hopelijk blinde vlekken voorkomen. Het rapport van de Werkgroep versterking functies Tweede Kamer (commissie-Van der Staaij, 2021) stelt in *Meer dan de som der delen* dat de invoeringstoets ook de wetgevende, controlerende en vertegenwoordigende functie van de Tweede Kamer kan versterken. Het is om deze reden dat de Commissie voor de Werkwijze van de Tweede Kamer in de loop van 2022 interesse in ervaringen met de invoeringstoets aan de dag legde.

EWA kan een invulling zijn van de invoeringstoets. Dan wordt het *EWA-light*. Waarom? De invoeringstoets is volgens minister Weerwind (kabinet-Rutte IV) in zijn notitie voor de Tweede Kamer gericht op:

- a *Hoe pakt de regeling uit voor de mensen, organisaties of bedrijven voor wie de regeling is bedoeld?*
 - In hoeverre zijn er signalen dat de regeling niet werkt zoals bedoeld?
 - Ontstaan er knelpunten of onbedoelde effecten?
 - Zijn er mogelijk (indirect) discriminerende effecten?
- b *Hoe pakt de regeling uit voor de uitvoering?*
 - Ontstaan er in de uitvoeringspraktijk knelpunten of onbedoelde effecten?
 - In hoeverre is de kwaliteit van de uitvoering geborgd?
 - In hoeverre komt de praktijk overeen met de aannames uit de invoeringstoets (ex ante)?

Wat zijn de overeenkomsten tussen een EWA en de invoeringstoets?

- 1 Beide benaderingen zijn praktisch gericht en worden gestart rond driekwart jaar tot een jaar na de inwerkingtreding van wet- en regelgeving.
- 2 Beide benaderingen hebben aandacht voor de menselijke factor, dus hoe een regeling uitpakt voor burgers, non-profits en bedrijven.
- 3 In beide benaderingen is er aandacht voor signalen over de uitvoering van beleid. De knelpunten in de in- en uitvoering van beleid komen aan bod (zie de Kamerbrief van minister Weerwind van 23 september 2022).

- 4 Ook de aannames achter beleid en beleidsuitvoering worden niet vergeten in de vraagstelling bij de invoeringstoets.

Wat zijn de verschillen tussen EWA en de invoeringstoets?

- 1 Een EWA is uitgebreider dan de invoeringstoets, onder meer omdat de grondlegger spreekt van aandacht voor de beleidstheorie en de andere methodische elementen van de EWA.
- 2 Een EWA kan mogelijk al iets eerder plaatsvinden dan de invoeringstoets, namelijk na de inwerkingtreding van beleid. Een invoeringstoets wordt pakweg een jaar na invoering van een regeling gestart. Niet uitgesloten is dat een EWA eerder plaatsvindt. Het uiteenrafelen van de operationalisering van wet- en regelgeving in een stappenlogica (of proceslogica) en de consequentie- en consistentieanalyse voorzien in het doordenken van de koersvastheid.
- 3 In de EWA-aanpak is er aandacht voor het voorwerk in de zin van een beleidstheorie en voor de operationalisatie in de vorm van de doordenking van de stappenlogica. Vervolgens komen een consequentie- en consistentietoetsing aan bod. Dat betekent dat EWA een doorkijk geeft op de (verwachte) doeltreffendheid van beleid. In de checklist voor de invoeringstoets worden de doelstellingen en verwachtingen als criterium voor beoordeling *niet* genoemd, en vindt niet vanzelfsprekend een consequentieanalyse plaats. Minister Weerwind is daar duidelijk over in de Kamerbrief. Een invoeringstoets is geen evaluatieonderzoek in de zin van een analyse van de doeltreffendheid of effectiviteit van beleid (Kamerbrief 23 september 2022).
- 4 De EWA is ook prospectief en mag zeker verbeter suggesties of voorstellen ‘ophalen’. De invoeringstoets houdt het bij signalen.

Een invoeringstoets is iets meer dan een EWA ‘pretentieloos praktisch’. EWA onderscheidt zich vooral door de methodische gerichtheid op de operationalisatie van beleid, de analyse van de consequenties en van de consistentie van de stappen in de in- en uitvoering (grijpt het raderwerk werkelijk in elkaar om uit te komen op het doel?) en het meenemen van correctie- of verbeter voorstellen.

Conclusie: de vergelijkingen culminereren steeds in de vaststelling dat bij EWA sprake is van een methodische gerichtheid, meer dan voorzien is bij een RTE, een reflectietoets of invoeringstoets. Daar staat tegenover dat de invoeringstoets makkelijker toestaat om deze invoeringstoets op maat te maken. Immers, wet- en regelgeving verschillen onderling nogal.

7 Reflectie vanuit de invoeringstoets

Wat leert de toepassing van de invoeringstoets voor het uitvoeren van een *early warning evaluation* (EWA)? Hoe zou de aanmelding voor een EWA in de rijksdienst geregeld kunnen worden? Ervaringen met de invoeringstoets, zoals geboekstaafd door de directie Wetgeving en Juridische Zaken van het ministerie van Justitie en Veiligheid, bieden inzicht.

Arno F.A. Korsten

Voltooid toetsen

Met invoeringstoetsen is eerst in 2022 een start gemaakt. Er bestaat een overzicht van alle tot september 2022 uitgevoerde, lopende en geplande invoeringstoetsen. Tot de al voltooide invoeringstoetsen behoren onder meer:

- de regeling betreffende het bioptische telescoopsysteem (ministerie van I&W);
- de toepassing van de subsidieregeling verduurzaming (EZK);
- de regeling opvang ontheemden uit Oekraïne – evaluatie van de gemeentelijke toekenning van leefgeld (J&V en SZW + VNG);
- de evaluatie van de werkwijze en uitvoeringpraktijk van de start-up-regeling (J&V, IND);
- de wijziging van de Telecommunicatiewet, invoeren opt-in systeem voor telemarketing (EZK).

De studies van de opvang ontheemden uit Oekraïne heb ik bekeken. Hier gaat het niet om een korte invoeringstoets of Kamerbrief, maar om een verzamelbrief van de bewindspersoon met een aantal bijlagen in de vorm van onder meer een extern onderzoek naar kosten. Qua aanpak van de invoeringsanalyse is het geen ‘klein bier’, z gezegd.

Suggereren

Tal van voorgenomen invoeringstoetsen van rijksbeleid zijn aangemeld. De keuze voor het doen van een invoeringstoets ligt in deze fase van de toepassing van het instrument ‘invoeringstoets’ bij de betrokken uitvoeringsorganisaties, ministeries en/of decentrale overheden. In de praktijk nemen vooral uitvoeringsorganisaties het initiatief om een onderwerp op te pakken en aan te melden. Zij kennen de maatregelen en beschikken over signalen om een wet of regeling door te lichten. Ook de Vereniging van Nederlandse Gemeenten (VNG) heeft aangegeven nut en noodzaak van invoeringstoetsen te zien, maar gemeentebesturen meldden zelf in 2022 nog nauwelijks voorstellen.

Wat er buiten valt

Opvallend is dat tal van actuele discussies over transitie waar het kabinet mee bezig was, zoals het klimaatakkoord, de energietransitie en de stikstofreductie in relatie tot natuurbescherming, nog niet terug te vinden zijn in afgeronde of voorgenomen invoeringstoetsen. Dat heeft meerdere oorzaken:

- a Beleid zit in het stadium van Kamerbrieven en nota’s en is nog niet uitgewerkt in wet- en regelgeving. Er is nog volop debat over. En nota’s buitelen ook over elkaar. Denk maar aan Kamerbrieven over circulaire landbouw.
- b Tal van transitievraagstukken zijn onderwerp geweest van beraad, zoals bij klimaatbureaus, bij het pensioenakkoord, bij een watertafel, bij een landbouwtafel. Wie stelt dan direct al een *early warning evaluation* of invoeringstoets voor?
- c Er bestaat een neiging bij uitvoeringsorganisaties om omvangrijk beleid (met beleidsbundels) niet voor te stellen om te onderwerpen aan een invoeringstoets, maar eerder overzichtelijke maatregelen.

Wicked problems

Betekent dit dat complexer beleid, uit te voeren in meerlagig overheidsbestuur, tot nu toe nog buiten de invoeringstoets blijft? Wat als er meerdere uitvoeringsorganisaties betrokken zijn of sprake is van een beleidsnetwerk – rond een ‘*wicked problem*’, zeg maar een beleidscontroverse – waarin geen enkele actor alle doorzettingsmacht heeft, maar taken en bevoegdheden verspreid zijn en de wederzijdse afhankelijkheden groot?

Dergelijke beleidsproblemen:

- kennen een verdeling van taken, bevoegdheden en middelen over meerdere beleidsactoren; geen actor heeft alle doorzettingsmacht;
- kennen veel onzekerheid over de aard van het probleem en/of de oplossing;
- zijn technisch, juridisch en/of bestuurlijk complex;
- er zijn meerdere overheden betrokken;
- van een definitieve oplossing is geen sprake;
- de overeenstemming tussen actoren is, als die er al is, broos;
- maatregelen zijn per definitie tijdelijk.

Als de systeemverantwoordelijke overheidsactor dan ontbreekt of een rol als ‘regisseur-met overzicht’ niet neemt, dreigt een invoeringstoets juist op het gebied waar veel mis kan gaan (Korsten, 2019), achterwege te blijven.

Er is nog een specifieke factor. Denk aan de reactie op de hoogwatercrisis die in relatie tot de klimaatproblematiek in 2021 speelde in drie gebieden: Zuid-Limburg, Wallonië (Luik e.o.) en Duitsland (Aken e.o.). Ter preventie zijn er financiële middelen vrijgemaakt binnen de bestaande bestuurlijke taakverdeling. Van een invoeringstoets is in dit type gevallen geen sprake, vooral omdat nieuwe wet- en regelgeving ontbreekt. Als nieuwe wet- en regelgeving afwezig is, komt er ook geen invoeringstoets.

Motieven voor toepassing

Welke motieven komen naar voren in de praktijk van toepassing van de invoeringstoets? Een eerste motief betreft het oefenen met de invoeringstoets zelf. Een belangrijk tweede motief om de toets toe te passen ligt bij het maken van een inschatting over de vraag of de kans groot is dat er knelpunten in de uitvoering ontstaan. Een derde motief is gelegen in het vinden van informatie die te gebruiken is voor herziening van de regeling. Een vierde motief ligt in de wens te onderzoeken of de doelstellingen van de regeling zijn of worden behaald. Het koersvastheidsmotief – hier verbonden met *early warning evaluation* – komt in de praktijk dus voor. Een vijfde motief betreft wetgeving die veel impact heeft op een uitvoeringsorganisatie, toezichthouders en/of zorgaanbieders. Denk aan de Wet toetreding zorgaanbieders. Een zesde motief verwijst naar risico’s van wetgeving voor burgers en hun *doenvermogen*.

Deze motieven kunnen ook gelden voor een EWA.

Arno F.A. Korsten

Op de evaluatieagenda

Bij de rijksoverheid is sprake van een voorgeschreven evaluatieagenda. Doel hiervan is te bevorderen dat inzichten in de doeltreffendheid en doelmatigheid van beleid tijdig beschikbaar komen om gebruikt te kunnen worden en gericht bij te sturen. Deze agenda biedt jaarlijks in de begroting van elk ministerie een overzicht van de belangrijkste beleidsthema's, een korte toelichting op de inzichtbehoefte per thema en de daarbij passende agendering van een evaluatie en overig relevant beleidsonderzoek. Op de strategische evaluatieagenda staan daarmee alle voorgenomen of lopende ex ante, ex durante en ex post beleidsonderzoeken. Daartoe behoren ook de invoeringstoetsen van nieuwe regelgeving die voor elk beleidsthema van belang zijn, ook al zijn invoeringstoetsen niet direct gericht op de doeltreffendheid van beleid maar juist op het vaststellen van knelpunten en het geven van verbeteringsuggesties, zoals minister Weerwind in de Kamerbrief van 23 september 2022 aangeeft. De invoeringstoets ligt in het verlengde van de andere onderzoeken waarvan een aantal wel gericht zijn op het in beeld krijgen van de doeltreffendheid van beleid en doelmatigheid.

Wat buiten de aanmelding valt

Uit de inventarisatie van *voorgenomen* invoeringstoetsen blijkt voorts dat tal van onderwerpen al zijn aangemeld. Tot de selectie behoren voornemens uit het belastingpakket en de stimuleringsregeling aardgasvrije huurwoningen. Sommige zijn toegezegd aan de Tweede Kamer. Invoeringstoetsen worden ook aangekondigd in de memorie van toelichting bij een wetsvoorstel.

Zijn er gebieden van regeringsbeleid die makkelijk binnen of juist buiten de invoeringstoets dreigen te vallen en waarvoor hetzelfde kan gelden in geval van een EWA? Stimuleringsregelingen blijken een gemakkelijke 'prooi' voor het doen van invoeringstoetsen.

Het lijkt wat vroeg om definitieve oordelen te geven over wat *buiten* het bereik van een vroege invoeringstoets valt, maar een werkhypothese is zeker te formuleren. Die houdt in dat '*wicked problems*' (brandende kwesties, beleidscontroversen) niet makkelijk aan een invoeringstoets worden onderwerpen omdat beleid gericht op de aanpak van deze weerbarstige problemen transversaal is en daarmee de departementale indeling in beleidsterreinen doorsnijdt.

Wie voelt zich dan de probleem- en beleidseigenaar die een toets laat verrichten? Beleid gericht op *wicked problems* is doorgaans omgeven met onzekerheid, meeromvattend en complexiteit, de overeenstemming is broos, de uitvoering geschiedt in een context van *multi-level government* en de koers is slechts tijdelijk. Dergelijk beleid is bovendien risicovol als het gedragsverandering van burgers vereist. Het beleid omvat bij *wicked problems* doorgaans tal van maatregelen. Wie gaat bij een beleidsbundel om invoeringstoetsen vragen? Ongetwijfeld zal een van de komende parlementaire enquêtecommissies (zoals over aardbevingsschade) hierop wijzen.

Mootjeswerk

Het ziet er nog niet naar uit dat grote complexe kabinetdossiers voorgelegd gaan worden voor een invoeringstoets of EWA, zoals onder meer de klimaatpolitiek, de energietransitie, de pensioenwetgeving, hoogwaterbestrijding en droogteproblematiek. EWA en invoeringstoetsen zijn tot nu toe niet alleen maatwerk, maar ook mootjeswerk. Invoeringstoetsen hebben slechts betrekking op maatregelen, nog niet op de grote wetten, zoals de Wet stikstofreductie en natuurverbetering. Dat is riskant omdat de Tweede Kamer ertoe neigt om een wet niet meer artikelsgewijs te behandelen (mededeling van Kamerlid Omtzigt, januari 2023), en omdat de aanleiding voor invoeringstoetsen een zwarte bladzijde was in de geschiedenis van de rijksoverheid.

8 Enkele conclusies

Early warning evaluation impliceert vanuit een referentiemodel van ‘goede beleidsinvoering’ niet alleen het ophalen van knelpunten in de beleidsinvoering en uitvoering, maar vooral ook de vroege evaluatieve doordenking van de ingezette beleidskoers op een relevante, consequente en consistente wijze.

Centrale vraag bij EWA is: komt het beleid bij de invoering ‘op de rails’ en blijft dat naar verwachting door de vervolgacties zo, waardoor de gewenste doelstellingen in het vizier blijven? EWA moet – vanuit de leertheorie gezien – voor frequentere *feedback loops* zorgen. EWA biedt beleidsmakers een (tweede) kans op herstel.

EWA vervangt als *ex durante* evaluatie latere *ex post* wetsevaluatie niet, want de ambitie achter wetsevaluatie is groter. EWA is geen wetsevaluatie, maar loopt qua lerende optiek wel vooruit op een wetsevaluatie, die meestal een aantal jaren na uitvoering plaatsvindt.

EWA voorkomt bij voorkeur beleidsfiasco’s. Waarom? EWA heeft een signaalfunctie (haalt onder meer naar voren wat een knelpunt is) en is in beginsel prospectief, want deze evaluatie biedt al aanzetten, tips of handreikingen voor beleidsgericht leren. Het is een oplossing als een overheidsbestuur al vroege na de besluitvorming over beleid wil weten of het de goede kant op aan het navigeren is.

Het paradoxale van een EWA is dat een evaluator eigenlijk een voorspelling doet van hoe iets na het begin in de in- en uitvoering van beleid uitpakt, terwijl de evaluator dat nog niet met zekerheid kan weten vanwege het vroege moment en het nog onzichtbaar zijn van de gewenste beleidseffecten. Dat betekent dat de evaluator zich afvraagt: als de uitvoering blijft zoals die nu loopt, komen we dan wel bij het doel uit of drijven we daarvan weg? Een analyse van de stappenlogica na de inwerkingtreding van beleid en de start van de beleidsinvoering ligt dan voor de hand. Een tweede methodische stap is de consequentieanalyse en een derde de consistentieanalyse.

Arno F.A. Korsten

Overeenkomst

Een EWA en een invoeringstoets komen overeen in het volgende:

- a Elk vindt tamelijk snel na de inwerkingtreding van beleid plaats.
- b Elke aanpak moet qua onderzoeksaanpak op maat gemaakt worden in relatie tot de beleidsinhoud en de kenmerken van de beleidsuitvoering.
- c Elke toets moet knelpunten in de uitvoering naar boven halen.
- d Elke reflectie is gericht op vroege signalering en leren.

De ambities achter EWA zijn wat groter dan die bij een invoeringstoets als gevolg van de methodische onderdelen: reconstructie stappenlogica in beleid; consequentie- en consistentieanalyse; beleidsgericht leren.

Lessen

Wat is groot en klein? De lijn is dat een invoeringstoets de kleinste ex durante bezinning is, EWA is als ex durante evaluatie al groter en methodischer, en een wetsevaluatie vier, vijf jaar na de inwerkingtreding van beleid is nog groter, want omvattender. De invoeringstoets let nauwelijks op gewenste effecten, de EWA wel al (er is al uitzicht op het via beleid uitkomen bij effecten), en een wetsevaluatie zet juist in op de vraag welke effecten van een wet uitgaan.

De toepassing van de invoeringstoets in 2022 biedt lessen voor de toepassing van EWA. Eén les is de tendens om de invoeringstoets slechts los te laten op beperkte, operationele maatregelen en niet op grote vraagstukken die veel beleid omvatten. De vraag is hoe een kabinet de dreiging dat het doelen achter grote transities en omvattend beleid niet haalt, tijdig weet te signaleren. Daar was het immers na het rapport van de commissie-Bosman en de commissie-Van Dam om begonnen. Tijdige EWA kan de kans op beleidsfiasco's en latere parlementaire enquêtes bij grote projecten verminderen.

9 Enkele aandachtspunten voor discussie

Deze inzichten leiden tot een aantal aanvullende, puntig geformuleerde aandachtspunten. Hier komen ook enkele punten aan bod die rijp zijn voor parlementaire discussie met minister Weerwind en anderen.

Op rijksniveau komen invoeringstoetsen voor. Invoeringstoetsen moeten in beginsel plaatsvinden bij alle wet- en regelgeving. In de praktijk zal er echter discussie ontstaan over de mogelijkheden en prioriteiten. Er zal worden gezocht naar ontspanningsroutes. Voor kortstondig beleid of voor noodmaatregelen is geen invoeringstoets mogelijk en ook niet nuttig. Dat leert het coronabeleid met onder meer de tijdelijke lockdowns.

Agendering. De bedoeling is dat alle invoeringstoetsen opgenomen worden bij een relevant beleidsthema in de rijksbreed voorgeschreven *strategische evaluatieagenda* die jaarlijks in de begroting van elk ministerie is opgenomen. Die agenda bevat een overzicht van alle beleidsthema's, de inzichtbehoefte en de toelichting daarop en

de agendering van het type beleidsonderzoek, waaronder evaluatieonderzoek en invoeringstoetsen. Op deze wijze staat alle *ex ante* onderzoek, *ex durante* onderzoek (zoals EWA en invoeringstoetsen) en *ex post* beleidsonderzoek open voor debat met het parlement. Kamerleden hebben hiermee een kapstok voor discussie. Ze kunnen voorstellen om onderzoek toe te voegen, en eventueel te bundelen.

Een invoeringstoets vertoont overeenkomst met een *early warning evaluation* (EWA) in zoverre dat beide knelpunten in de in- en uitvoering van beleid naar boven halen en korte tijd na de inwerkingtreding van beleid starten. Achter een EWA gaat de pretentie schuil zicht te krijgen op knelpunten van beleid, de haalbaarheid van effectief beleid en verbetermogelijkheden; invoeringstoetsen zijn slechts gericht op knelpunten, maar zetten zeker in op verbeteringen.

Uit meldingen en voltooide toetsen blijkt dat *de grote dossiers* nog buiten de beoordeling blijven. De neiging om toetsen vooralsnog te beperken tot 'begrensde' maatregelen spoort nog niet helemaal met de door het kabinet-Rutte III en IV uitgesproken enorme ambitie om afscheid te nemen van beleidsfiasco's zoals de toeslagenaffaire.

Een EWA is kompasgericht. Wat zegt het kompas een half jaar tot een jaar na invoering: dat het beleid op koers blijft en de beoogde effecten naar verwachting bereikt zullen worden? Ontbreken de weeffouten die kunnen bewerkstelligen dat van beleid weinig of niets terecht komt? Beoordeling van de STAP-regeling voor levenslang leren leert eind 2022 dat tijdige feedback over wat schuurt, nuttig is. Het gevaar dreigt dat de bureaucratie na onderzoek toeneemt door meer eisen te stellen aan onderwijsgevend en cursussen. Dat dreigt bij alle *ex durante* analyses.

Waar EWA het makkelijkst is. Een evaluatie van conditionele beleidsprogramma's behoort tot de makkelijkste evaluaties vanwege het hoge 'als... dan'-gehalte. Dergelijke programma's zijn evenwel schaars. Doelprogramma's zijn moeilijker te beoordelen omdat gaande de uitvoering doelzoekend gedrag vorm moet krijgen.

Waar EWA vooral nodig is. Waar *early warning evaluation* het meest nodig is, dat zijn de modifierende beleidsprogramma's. De reden is dat hierin de overheid overgaat tot gedragsbeïnvloeding van burgers, bedrijven of andere organisaties. Hier ligt op de loer dat de gedragsveronderstellingen achter beleid niet juist blijken en het zicht op gewenste effecten uitblijft. Welk gebied zich voor EWA leent? Denk aan de komende besluiten over stikstofbeleid, de provinciale plannen voor het landelijk gebied, onderdelen uit het klimaatbeleid.

Verhouding tot ex ante evaluatie. Waarbij zou een EWA mogelijk zijn? Bijvoorbeeld bij een opkoopregeling van agrarische ondernemingen die als piekbelasters de bescherming van de natuur belemmeren. Of bij het doordenken van de invoering van prijscompensatie of een gasprijssplafond. Bemoelijkend is overigens in de laatste twee gevallen dat de werking van deze maatregelen kortstondig is en tussentijdse wijzigingen ongewenst zijn uit een oogpunt van rechtszekerheid. Een aan de besluitvorming voorafgaande (*ex ante*) *uitvoeringstoets* zal daarom misschien nuttiger zijn.

Arno F.A. Korsten

EWA heeft een signaalwerking. EWA waarschuwt of de invoering van beleid op koers is en blijft. De koersvastheid (kompas) staat centraal. Een EWA is toekomstgericht en prospectief en reikt daarmee verder dan beschrijving en verklaring. Deze EWA wijst op knelpunten, hint op continuïteit of verbetering in de koers en reikt aanzetten tot verbetering aan. EWA is hulpmiddel in het (vroege) beleidsonderhoud.

De parlementaire betekenis. Vanuit de Tweede Kamer is aangedrongen op de noodzaak om beleid direct na de inwerkingtreding te beoordelen. Zowel een invoeringstoets als een EWA is nuttig voor het uitoefenen van controle en het tijdig corrigeren van fouten mits de beoordelingen niet slechts betrekking hebben op minuscule regelingen, en grote beleidsprojecten, zoals een nieuwe pensioenwet, niet buiten de toetsing gehouden worden.

Praktische inslag. Een EWA is praktisch in zoverre de belangrijkste knelpunten in de invoering van beleid benoemd kunnen worden. EWA is gericht op hoofdzaken; een echte wetsevaluatie komt later, wellicht vier jaar later. Een knelpuntenlijst kan na een jaar vermoedelijk wel een reguliere vangst zijn, maar mocht een EWA twee maanden na de inwerkingtreding van een maatregel al plaatsvinden, dan is zoiets eerder slechts bijvangst.

Het moment. Een EWA vindt zo spoedig mogelijk plaats als er iets verstandigs te zeggen is over de werking en (consequentiële) richting van een maatregel. EWA start al vroeg, als de bestuurders nog niet zichtbaar in de gaten hebben dat ze mogelijk niet bij het beleidsdoel uitkomen. EWA is te verrichten als het wellicht nog ‘mistig’ is.

Confectie of maatwerk. EWA is het gemakkelijkst uit te voeren in een context van één overheid met doorzettingsmacht en één uitvoeringsorganisatie, maar heeft in veel gevallen feitelijk van doen met een beleidsnetwerk respectievelijk een keten van actoren en handelingen. Daardoor alleen al is een EWA *tailor made*. Bepaalde beleidsprogramma’s kennen overigens een beleidsmix. Dat vraagt om bezinning op de vraag waarop een EWA zich precies richt.

Wie uitvoert. Een EWA kan intern of extern uitgevoerd worden. In alle gevallen is het handig als in een team expertise aanwezig is over zowel de beleidsinhoud, in- en uitvoering van beleid en als evalueren.

Wat het niet is. Monitoring levert cijfers over de beleidsuitvoering, maar zelden biedt het een beoordelingskader en formuleren de onderzoekers conclusies en aanbevelingen. Monitoring is nuttig, maar geen evaluatiemethodiek. EWA graaft dieper dan bij monitoring het geval is, in zoverre dat er sprake is van beoordeling van beleid in relatie tot doelstellingen en veronderstellingen achter beleid. Blijft beleid op koers als de invoering en uitvoering zo blijven gaan als ze gaan? Daarop krijgt men met monitoring geen antwoord.

Geen evaluatie van voorlichting. EWA moet meer zijn dan een vroege evaluatie van een voorlichtingsprogramma rond beleid. Te vroeg heeft geen zin en *alleen* gerichtheid op voorlichting is eveneens ongewenst.

Aard van de oordelen. EWA kan resulteren in ‘ja, mits’-oordelen met een perverse doorwerking. De praktijk laat zien dat een discussie tussen een verantwoordelijke bestuurder van een coalitie en een vertegenwoordigend en controlerend orgaan, zoals het parlement, bij coalitiefracties en daarmee een meerderheid een ‘ja, mits’ kan uitmonden in een ‘ja’. De commissie-Bosman (2021: 148) waarschuwt nadrukkelijk hiervoor omdat het op deze manier negeren van signalen en vermijden van ‘leren’ later ten koste zal gaan van het bereiken van gewenste beleidseffecten. De commissie geeft daarom ter overweging om ook de mogelijkheid van een ‘nee, tenzij’ in het oog te houden.

Heeft EWA beperkingen of begrenzingen? Duidelijk is dat EWA niet een heel uitgebreide evaluatie moet worden die meer tijd in beslag neemt dan de hele beleidsinvoering al geduurd heeft.

Een tweede beperking houdt in dat beleid wel goed ingevoerd kan worden, maar dat een discussiepunt kan blijven of het beleid voldoet aan alle relevante criteria, zoals rechtvaardigheid. Bij een EWA bestaat het gevaar dat ‘dwaas beleid’ buiten schot blijft omdat het wel op koers ligt.

Ten derde besteedt EWA weinig aandacht aan de werkelijke effecten van beleid want die zijn waarschijnlijk nog buiten beeld. EWA is vanwege hetzelfde argument ook niet gericht op neveneffecten van beleid. Een wet aan een EWA onderwerpen is geen (uitgebreide) wetsevaluatie.

Een vierde beperking is dat EWA weinig informatie naar boven kan halen over de vraag of het beleid doelmatig is. Immers, de maatregel is amper ‘in bedrijf’.

Ten slotte, wat gebeurt er met het doordenken van Europese maatregelen die landelijk geïmplementeerd worden?

Ondanks deze beperkingen ziet de auteur van dit artikel echter zeker de meerwaarde van vroege signalementen. Daardoor krijgen regering en parlement een tweede kans.

Literatuur

- Bekker, R. (2020). *Dat had niet zo gemoeten!* Den Haag: Boom bestuurskunde.
- Bekkers, V.J.J.M. (2012). *Beleid in beweging*. Den Haag: Boom bestuurskunde.
- Bressers, J., & Hoogerwerf, A. (red.). (1995). *Beleidsvaluatie*. Alphen aan den Rijn: Samsom.
- Commissie-Van Dam. (2021). *Ongekend onrecht*. Den Haag: Tweede Kamer.
- De Vries, B., Korsten, A.F.A., & Gooren, W. (1983). *Evaluatie van de totstandkoming van het Innovatieprogramma biotechnologie*. Den Haag: Ministerie van EZ/OCW.
- Derksen, W., & Korsten, A.F.A. (red.). (1988). *Uitvoering van overheidsbeleid*. Leiden: Stenfert Kroese.

Arno FA. Korsten

- Gooren, W.A.J., m.m.v. Snellen, I.Th.M., Korsten, A.F.A., & Verstegen, R. (1984). *Onderzoek zaaien, innovatie oogsten – Eindverslag van een evaluatie-onderzoek van de Stichting Technische Wetenschappen*. Tilburg: IVA/KUN.
- Gooren, W.A.J., & Korsten, A.F.A. (1983). *Kennisoverdracht aan kleine en middelgrote ondernemingen via transferpunten in Nederland*. Den Haag: Staatsuitgeverij.
- Gooren, W.A.J., Korsten, A.F.A., & Prakke, F. (1983). *Evaluatie van innovatiebeleid – Een overzicht over uitvoering en wenselijkheid van evaluatie van beleidsvoornemens uit de Innovatienota*. Den Haag: Ministerie van O&W/Economische Zaken.
- Hoogerwerf, A., & Herweijer, M. (red.). (2014). *Overheidsbeleid*. Deventer: Kluwer.
- Hupe, P.L. (2007). *Overheidsbeleid als politiek*. Assen: Van Gorcum.
- Korsten, A.F.A. (1981). Evalueren van wetten. *Bestuurswetenschappen*, 73, 462-484.
- Korsten, A.F.A. (1985). Uitvoeringsgericht ontwerpen van overheidsbeleid. *Bestuur*, (8), 12-20.
- Korsten, A. (1997). Effectrapportage. *Beleidsanalyse*, 26(1/2), 4-7.
- Korsten, A.F.A. (2010). *Deugdelijk bestuur*. Heerlen: Open Universiteit.
- Korsten, A.F.A. (2013). Wat evaluatoren van kritiek kunnen leren. *Beleidsonderzoek Online*. december. doi:10.5553/Beleidsonderzoek.000030
- Korsten, A.F.A. (2019). Omgaan met 'wicked problems'. *Beleidsonderzoek online*, maart. doi:10.5553/BO/221335502019000002001
- Korsten, A.F.A., De Jong, P., & Breed, C.J.M. (2010). *Regeren met programma's*. Den Haag: Boom Lemma.
- Korsten, A., & Leers, G. (2005). *Inspirerend leiderschap in de risicomaatschappij*. Utrecht: Lemma.
- Korsten, A., & Leeuw, F. (1997). Handleidingen in het openbaar bestuur. *Bestuurskunde*, 6(8), 332-350.
- Korsten, A.F.A., & Pröpper, I.M.A.M. (1993). Beleidsevaluatie. In A. Bekke et al. (red.), *Management in overheidsorganisaties* (B1160, pp. 1-14). Alphen aan den Rijn: Samsom.
- Korsten, A.F.A., & Snellen, I.Th.M. (1988). Openbare aanbesteding en uitbesteding: evaluatie van de wet ketenaansprakelijkheid. In W. Derksen, A.F.A. Korsten, & A.F.M. Bertrand (red.), *De praktijk van onderzoek* (pp. 473-491). Groningen: Wolters-Noordhoff.
- Leeuw, F.L. (1989). Beleidstheorieën: veronderstellingen achter beleid. In A. Hoogerwerf (red.), *Overheidsbeleid* (pp. 91-108). Alphen aan den Rijn: Samsom.
- Leeuw, F.L., & Van de Graaf, H. (1988). Onderzoek naar beleidstheorieën: uitgangspunten, methodische aspecten en praktische relevantie. *Beleid en Maatschappij*, 15(1), 1-7.
- Mazmanian, D., & Sabatier, P. (1981). *Effective policy implementation*. Lexington: Lexington Books.
- Ministerie van Justitie en Veiligheid – Directie Wetgeving en Juridische zaken. (2022). *Inkleuring en stand van zaken invoeringstoets*. Den Haag, 21 april.
- Ministerie van Justitie en Veiligheid – Directie Wetgeving en Juridische zaken. (2022). *Reactie op aanvullende verzoeken inzake de invoeringstoets*. Den Haag, 23 september.
- Morgan, G. (1998). *Images of organization*. London: Sage.
- Onderzoeksraad voor Veiligheid (OvV). (2015). *Aardbevingsrisico's in Groningen*. Den Haag.
- Parlementaire Onderzoekscommissie Uitvoeringsorganisaties. (2021). *Klem tussen balie en beleid*. 25 februari (commissie-Bosman).
- Rogers, P. (2021). Why do we need more real-time evaluation. 15 april (open source, internet).
- Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach*. Newbury Park: Sage.
- Snellen, I.Th.M. (Ed.). (1985). *Limits of government: Dutch experiences*. Amsterdam: Kobra.

- Van den Berg, C., & Korsten, A.F.A. (2020). *Meer greep op natuurbeleid: Een handreiking voor de Tweede Kamer*. Den Haag: Vaste Commissie voor LNV.
- Van der Knaap, P., Pattyn, V., & Hanemaayer, D. (2020). *Beleidsevaluatie in theorie en praktijk*. Den Haag: Boom bestuurskunde.
- Van Dijk, J.W.A., & Hegeman, J.H. (1983). Invoeren van beleid. *Beleid en Maatschappij*, 9(9), 239-246.
- Van Hoesel, P., Mevissen, J., & Dekker, B. (red.). (2015). *Kennis voor beleid: Beleidsonderzoek in Nederland*. Assen: Van Gorcum.
- Weiss, C.H. (1972). *Evaluation research*. Englewood Cliffs: Prentice Hall.
- Werkgroep versterking functies Tweede Kamer. (2021). *Meer dan de som der delen*. Kamerstukken II 2021/22, 35992, nr. 1 (commissie-Van der Staaij).
- Winter, H.B., & Klein Haarhuis, C.M. (2007). Wetsoverstijgende evaluaties. *Tijdschrift voor Gezondheidsrecht*, 31(8), 620-626.