

## **Evaluerend leren: vrijheid, maar geen vrijbrief**

Appreciatie stuurgroep ex durante evaluatie pilot Lerend evalueren

14 april 2022

## **Inleiding**

VWS is in 2018 gestart met de pilot Lerend evalueren. Een belangrijk doel van de pilot is om werkendeweg het inzicht in de doeltreffendheid en doelmatigheid van het beleid te vergroten en om op basis daarvan het beleid te verbeteren.

In 2019 is er een ex ante evaluatie uitgevoerd om de situatie voor de start van de pilot in beeld te brengen. Deze startevaluatie<sup>1</sup> is in april 2019 aan de Kamer aangeboden. Hierbij is gekeken naar de factoren en omstandigheden die invloed hebben op de kwaliteit van de evaluatie en het leren omtrent en verbeteren van beleid.

Dit onderzoek betreft de ex durante evaluatie van de pilot Lerend evalueren. In het onderzoek zijn twee lijnen onderscheiden, namelijk (1) de kwaliteit van de individuele evaluaties en (2) de veranderkundige aspecten van VWS als lerende organisatie. De resultaten uit beide lijnen zijn belangrijke informatie om de kwaliteit en het gebruik van lopende en toekomstige lerende evaluaties te verbeteren. Het onderzoek is uitgevoerd door het team van externe en interne onderzoekers. Het externe onderzoeksbureau Regioplan is daarbij verantwoordelijk voor het onderdeel kwaliteit van de individuele evaluaties. Voor het onderdeel lerende organisatie zijn VWS-onderzoekers verantwoordelijk. De beide rapporten zijn toegevoegd.

Het onderzoek is begeleid door een stuurgroep met onafhankelijke deskundigen en een onafhankelijke voorzitter (zie bijlage). Op basis van beide onderzoeken komt de stuurgroep tot de volgende appreciatie en aanbevelingen voor VWS.

## **APPRECIATIE**

### **Algemeen: een wisselend beeld**

Algemeen is er sprake van een wisselend beeld. Zo heeft de pilot ervoor gezorgd dat VWS het beleid vaker is gaan evalueren. Tegelijkertijd zijn de eerste elf evaluaties die zijn uitgevoerd in de eerste jaren van de pilot - en die onderdeel zijn van deze ex durante evaluatie - wel zeer divers. De reikwijdte is groot: de evaluaties beslaan het brede terrein van VWS. Het doel om evaluaties niet alleen te beperken tot de middelen die op de begroting staan, maar ook breder te kijken naar de middelen die vallen onder het uitgaven plafond zorg (UPZ), heeft hieraan bijgedragen.

De aandacht voor evalueren is met de pilot eerder in het beleidsproces komen te liggen. Dat is positief: er worden meer ex ante en ex durante onderzoeken uitgevoerd. In de uitvoering van de evaluaties zijn lerende factoren meegenomen om de bruikbaarheid van de evaluaties te vergroten, zoals (1) het betrekken van verschillende perspectieven, (2) het combineren van kwalitatief en kwantitatief onderzoek en (3) het aanstellen van een begeleidingscommissie inclusief een onafhankelijke voorzitter.

Niet positief is dat de meest evaluaties geen onderzoek naar de impact van beleid betroffen (doeltreffendheid en doelmatigheid). Soms was er zelfs geen sprake van beleidsevaluerend onderzoek als zodanig: de pilot is ook aangegrepen om verkenningen uit te voeren (preventie jeugd, langer thuis ouderen) en een analyse van een uitvoeringsproces (subsidies). Bij Maatschappelijke

---

<sup>1</sup> TK, 2018-2019 31865 nr. 150

diensttijd (MDT) is een ex durante evaluatie gebruikt om een beleidstheorie te ontwikkelen en het programma vorm te geven conform de wensen van het kabinet.

De vragen die in de evaluaties binnen de pilot zijn gesteld, worden over het algemeen goed beantwoord en geven inzicht in bijvoorbeeld hoe verschillende betrokkenen tegen beleid en praktijk aankijken. Deze inzichten vergroten de handelingsopties, zoals te zien in het bijsturen van beleid door VWS op basis van (tussentijdse) uitkomsten van de onderzoeken. In een aantal evaluaties hebben ook stakeholders hun handelen aangepast op basis van de resultaten van de evaluatie.

De pilot heeft - zoals blijkt uit het actieonderzoek - een positieve sfeer gecreëerd om binnen VWS meer reflectief te kijken naar beleid.

#### **Uitvoering onderzoeken: adequaat**

De stuurgroep is van oordeel dat de beide onderzoeken correct zijn uitgevoerd. Regioplan heeft ondanks de zeer diverse onderzoeken vijf algemene aanknopingspunten voor verbetering opgesteld. De interne VWS onderzoekers geven op basis van het actieonderzoek veel concrete aanbevelingen om lerend evalueren meer te borgen binnen VWS.

Beide onderzoeken komen met aanbevelingen om (1) het inzicht in doelmatigheid te verhogen, (2) het gebruik van de beleidstheorie te bevorderen, (3) vooraf de leerbehoeften of leervragen te onderbouwen en (4) te komen tot een evaluatieprogrammering. Het interne onderzoek geeft ook adviezen voor de ondersteuning van beleidsmedewerkers. Hieronder geeft de stuurgroep op deze aanbevelingen een appreciatie.

De stuurgroep adviseert VWS om de aanbevelingen uit te werken en ze SMART te formuleren om zodoende bij de ex post evaluatie van de pilot Lerend evalueren de voortgang te kunnen toetsen.

#### **Inzicht in doeltreffendheid en doelmatigheid: teleurstellend**

Door het stellen van uiteenlopende onderzoeksvragen en bij elkaar brengen van verschillende bronnen (kwalitatief en kwantitatief) worden uiteenlopende uitspraken gedaan over de effecten van het beleid. Daarnaast worden in de studies soms wel maar vaker geen concrete verbeteringsuggesties/aanbevelingen gedaan om de doeltreffendheid van het beleid en de uitvoering te verbeteren.

Het is opvallend dat het inzicht in de doelmatigheid van beleid slechts bij één evaluatie (Transitie Autoriteit Jeugd) is onderzocht. Bij tien evaluaties is de vraag naar doelmatigheid niet opgenomen in vraagstelling. Bij verkenningen en onderzoek naar processen (n=3) is de doelmatigheidsvraag nog niet aan de orde. Bij evaluaties van decentraal beleid (n=2) is dit in sommige gevallen elders belegd.

Deze bevinding stelt teleur. Hoewel het vanuit de ex durante insteek deels verklaarbaar is (niet alle effecten zijn tijdens de uitvoering van beleid al onderzoekbaar), is inzicht in doelmatigheid van beleid wel essentieel: publieke middelen zijn immers schaars en moeten daarom doelmatig worden besteed. Beleidsevaluatie is een onmisbaar instrument om dit te realiseren. Meer vrijheden bij evalueren mogen niet afleiden van (gebrekig) inzicht in doeltreffendheid en doelmatigheid of zelfs het voorkomen van het stellen van relevante vragen hierover. Het lijkt erop dat 'leren' en 'doelmatigheid evalueren' als uitsluitende categorieën worden gezien. Dat is jammer en een misverstand: inzicht in doelmatigheid is niet alleen nodig om te verantwoorden dat publieke middelen goed worden besteed, maar juist ook om te kunnen leren.

Het uiteindelijke doel van beleid is maatschappelijk toegevoegde waarde voor de burger. Door de burger vooraf bij beleidsontwikkeling te betrekken bestaan er ook achteraf meer mogelijkheden om dit 'burgerperspectief' ook te betrekken bij de evaluatie. Dit is tot nu toe te weinig gedaan. Bij nieuw beleid is het goed om – voor zover mogelijk en zinvol - vooraf vast te leggen wat relevante indicatoren zijn voor de beoogde effecten van het beleid. Dit helpt om de maatschappelijke impact te onderzoeken. Het is ook zinvol om de doelen van beleid te koppelen aan monitoring en aan tussendoelen, zodat ex durante de voortgang kan worden getoetst. Ex durante evaluaties kunnen daarbij ingezet worden om bestaande doelen en indicatoren te verbeteren of aan te vullen.

#### *Aanbeveling*

*Maak een strakke richtlijn met minimumeisen waaraan een evaluatie altijd moet voldoen om inzicht in – verklaringen voor mee- of tegenvallende – doelmatigheid en doeltreffendheid op te leveren. Zorg dat de beleidsdirecties dit gaan toepassen en laat de directie FEZ daarbij ondersteunen.*

#### **Een adequate beleidstheorie: conditio sine qua non**

Uit de onderzoeken blijkt dat, in tegenstelling tot wat in de ex ante evaluatie van de pilot wordt aanbevolen, een beleidstheorie veelal ontbreekt in het onderzochte beleid. Een beleidstheorie is, ook bij onvolkomen kennis, de rode draad voor beleidsontwikkeling en beleidsgericht leren en daarmee onmisbaar. Achteraf reconstrueren van een beleidstheorie is mogelijk, maar het zou veel beter zijn om een beleidstheorie van meet af aan te ontwikkelen en laten door ontwikkelen zodat deze kan functioneren als rationele toetssteen voor het beleidsproces annex beleidsgerichte leerproces. Zonder (te werken aan) een gedeelde beleidstheorie is het niet goed mogelijk om te toetsen of de veronderstelde mechanismen tot effecten leiden en dus ook om aanknopingspunten voor het verbeteren van de doeltreffendheid te vinden. Theorieën over succesvolle interventies 'stuwen' als het ware het leren: aannames die onjuist of onvolledig waren kunnen worden bijgesteld, aannames die blijken te kloppen worden uiteraard gehandhaafd en waarschijnlijk komen er nieuwe bevindingen bij die de beleidstheorie verrijken.

#### *Aanbeveling*

*Werk voor ieder beleid een beleidstheorie uit en leg deze vast, ook als deze nog niet helemaal 'uitontwikkeld' is. Dit is de rode draad voor beleidsontwikkeling en beleidsgericht leren en maakt op leren gericht evalueren aanzienlijk beter uitvoerbaar en relevanter. Als er nog geen (vastgestelde) beleidstheorie aanwezig is, wordt de theorie gereconstrueerd om zodoende een ijkpunt te hebben voor de evaluatie. Op basis van de evaluatie wordt de beleidstheorie verder verrijkt.*

#### **Leren van evalueren: ongerichtheid moet plaatsmaken voor een bewuste leerstrategie**

In alle evaluaties is in de ogen van de betrokkenen, in enige mate en op enige wijze geleerd, vooral door VWS. Dit is positief: voorafgaand aan de pilot werd vooral door externe bureaus geleerd. Tevens worden bij de opzet stakeholders betrokken: door de opzet gezamenlijk te bespreken is er ook sprake van reflectie bij anderen over de maatschappelijke impact van beleid en hoe die valt te vergroten.

De keerzijde is dat er ongericht wordt geleerd en daarmee kansen worden gemist. Beleidsevaluaties en -onderzoek staan in de onderzochte casus op relatief grote afstand van het beleidsproces. Dit heeft te maken met een afstandelijke uitvoering van de onderzoeken en dat bemoeilijkt het leerproces. Door evaluaties en onderzoek wel onderdeel te maken van het beleidsproces, kan er veel meer van worden geleerd: op veel meer momenten dan de rapportage alleen en ook door veel meer actoren. Zie bijvoorbeeld ook hoe R&D-processen in het bedrijfsleven verlopen.

Rapportages van beleidsevaluaties zijn voor de politiek nogal eens confronterend wanneer er geen sprake is van 'continu werken aan verbeteringen': het beeld is er dan een van voortdurende

tegenvallers. Dit geldt helemaal wanneer het beleid van meet af aan niet goed in elkaar is gezet of onderschat is hoe complex en taai een op te lossen probleem in werkelijkheid is. De politiek zou met een werkelijk op leren gericht continu proces van verbeteringen enorm geholpen zijn: bij het afleggen van verantwoording kan dan immers gewezen worden op geleerde lessen, daarmee ingezette positieve lijnen en bij gevolg niet steeds weer moet worden gereageerd op vervelende incidenten zonder dat duidelijk is waarom die incidenten zich voordoen. In een lerende evaluatie komen op die manier leren en verantwoorden samen. Door de werking en voortgang van (sturing op) beleid in beeld te brengen wordt verantwoording afgelegd ('of het beleid werkt') en kan worden geleerd ('van wat niet of juist heel goed werkt').

Beleidsmedewerkers en andere betrokkenen bij de evaluaties leggen de nadruk eenzijdig op leren (in plaats van verantwoorden). Zij beschrijven de leeropbrengsten in termen van legitimiteit, het verkrijgen van een breder perspectief op een vraagstuk, het versterken van het netwerk en de samenwerking tussen stakeholders, het vergroten van onderling vertrouwen, inzicht in werkzame elementen, agendering van (nieuwe) thema's en het vergroten van het gevoel van eigenaarschap. De stuurgroep vermoedt dat het 'minder bedreigende karakter' van lerende evaluaties hierbij een grote rol speelt. De keerzijde daarvan is dat er ongericht en zelfs vrijblijvend wordt geleerd. Een bewuste leerstrategie kan dat helpen verbeteren.

Uit de ex durante evaluatie blijkt dat de leerbehoeften voorafgaand aan een evaluatie niet heel specifiek zijn geïnventariseerd: wat moet er eigenlijk door wie geleerd worden? Ook was de opzet van het onderzoek niet altijd gericht op leren.

Het leren in het evaluatieproces kan breder worden verankerd door perspectieven van meerdere stakeholders in te brengen. Hiermee kan worden voorkomen dat beleidsactoren vervallen in 'groupthink' en worden gezorgd dat zij een beter zicht krijgen op de werkelijke implicaties van hun beleid. Door vooraf de leerbehoeften van de betrokkenen te benoemen, worden de onderzoeksvragen meer gericht op leren en het handelingsperspectief van de betrokkenen. De leervraag kan verschillen voor de verschillende betrokkenen: waar voor sommigen de inhoudelijke waarde voorop staat, gaat het bij anderen om het onderling vertrouwen of het legitimeren van bepaalde keuzes.

#### *Aanbeveling*

*Formuleer een bewuste leerstrategie en leg voorafgaand aan de evaluatie de leerbehoeften van de verschillende betrokkenen vast, dit bevordert de benutting van de resultaten van het onderzoek.*

#### **Evaluatieprogrammering: essentieel hulpmiddel om leren 'in de tijd' te organiseren**

Bij leren kan onderscheid worden gemaakt tussen instrumenteel, reflectief of systeemleren. Eenmalige lerende evaluaties van beleidsinstrumenten of (kortlopende) beleidsprogramma's kunnen bijdragen aan instrumenteel leren (kennis die bijdraagt aan beter uitvoeren bestaand beleid). Om met lerend evalueren reflectief (kennis die bijdraagt aan de onderbouwing van beleid, en eventueel bijsturen) en systeemleren (kennis die inzicht geeft in wat nodig is om de context van de het systeem te veranderen, bijvoorbeeld aanpassing in regelgeving) te bewerkstelligen, is het zinvol om lerend evalueren te koppelen aan een meerjarig beleids(vormings)traject. Meerdere evaluatiemomenten kunnen dan worden 'ingebouwd' en als prikkel dienen om te reflecteren op de doeltreffendheid van het beleid op de lange termijn.

Voor het leerproces in de loop van de tijd is het programmeren van onderzoek een essentieel hulpmiddel. Onderzoekprogrammering legt de verbinding met de bestaande kennis, laat kennislacunes zien, agendeert nieuw onderzoek, zorgt voor adequate onderzoekdesigns, zorgt voor

begeleiding, zorgt voor benutting, zorgt dat de nieuwe kennis wordt opgeslagen in de kennishuishouding en zorgt voor metastudies.

Het ideaal is een onderzoekprogrammering waarin lerende evaluaties en andere (meer reguliere) evaluaties een plek hebben, zodat verschillende typen evaluaties elkaar kunnen verrijken. Maak helder voor welke typen vragen een lerende evaluatie geschikt is. De strategische evaluatieagenda is hier zeer geschikt voor.

#### *Aanbeveling*

*Investeer in een goed onderbouwde Strategische Evaluatie Agenda (SEA) en organiseer daarmee een onderzoekprogrammering waarmee 'beleidsgericht leren en verantwoorden in de tijd' wordt georganiseerd.*

Door gestructureerd per beleidsthema de inzichtbehoefte vast te stellen (welke vragen moeten worden beantwoord om het beleid zinvol te evalueren?) kan goed worden onderbouwd welke eventuele plaats daar voor lerende evaluaties is weggelegd, naast andere vormen van ex-ante, ex-durante en ex-post onderzoek. De SEA kan goed worden gecombineerd met de (strategische) kennisagenda van VWS.

#### **Ondersteuning vanuit de pilot: voorziet in duidelijke behoefte, maar betrokkenheid van 'de departementale top' is eveneens geboden**

Vanuit de pilot is ondersteuning geboden aan de beleidsmedewerkers die een evaluatie uitvoeren of begeleiden. Het gaat bijvoorbeeld om een draaiboek, leerateliers, hulp bij aanbesteden onderzoek en een toolbox. Er is daarnaast een Review Board waarin zowel vooraf (opstellen onderzoeksvragen, nadenken over organisatiestructuur voor het onderzoek) als tijdens het onderzoek kritisch wordt meegedacht.

De ondersteuning en leerateliers worden positief gewaardeerd door de beleidsmedewerkers. Zij ervaren ook belemmeringen bij het evalueren: het ontbreekt aan tijd en capaciteit om een evaluatie uit te voeren, de uitvoering is complex (veld betrekken maakt evalueren ook complexer), langdurig proces, onvoldoende expertise van evalueren en de Review Board werd niet altijd als meedenkend ervaren eerder als het afleggen van verantwoording.

Lerend evalueren is bekend bij de trekker van evaluaties binnen de pilot maar er is nog geen 'spin off' richting de rest van de organisatie. De cultuur bij VWS is nog niet echt gericht op leren. Een beleidsmedewerker wordt zelden beloond op lerend evalueren.

#### *Aanbeveling*

*Ga door met lerend evalueren en trek voldoende middelen voor uit de ondersteuning, begeleiding en benutting ervan.*

De betrokkenheid van (top)management is belangrijk om lerend evalueren bredere bekendheid te geven in de organisatie. Door te investeren in het borgen en monitoren van de uitkomsten van evaluaties krijgt lerend evalueren meer bekendheid.

## Stuurgroep ex durante evaluatie pilot Lerend evalueren

Peter van der Knaap (voorzitter)	Ministerie Buitenlandse Zaken
Peter van Hoesel	Erasmus Universiteit
Eva Kunseler	Planbureau voor de leefomgeving
Thomas Niaounakis	Ministerie Financiën
Hans Leenders	Ministerie VWS
Margriet Germing (secretariaat)	Ministerie VWS