

*In deze rubriek gaat het om verhalen over en/of reflecties op de ervaring met methodische, praktische, politieke, morele en wie weet esthetische problemen waar onderzoekers in de praktijk mee te maken krijgen. Redacteur voor deze rubriek zijn Griet Verschelden: griet.verschelden@ugent.be en Anneloes van Staa: a.van.staa@hro.nl.*

## Lerend evalueren; een expeditieverslag

*Jurian Edelenbos en Arwin van Buuren*

### Inleiding

In 2003 hebben wij met enkele andere onderzoekers het Stimuleringsprogramma Burger en Milieubeleid van het Nederlandse Ministerie van Volkshuisvesting Ruimtelijke Ordening en Milieu (VROM) geëvalueerd met behulp van een lerende evaluatieaanpak (zie Edelenbos e.a., 2003; Edelenbos & Van Buuren, 2005). Dit programma moest het milieubeleid meer ‘van, voor en door de burger’ maken. We hebben dit onderzoek ‘anders dan anders’ aangepakt, met name om een grotere doorwerking van de studie te bewerkstelligen. Daarom hebben we gaandeweg een ‘lerende evaluatie’ vormgegeven, die was gericht op daadwerkelijke gedragsverandering en leermomenten voor het betrokken programmteam. Tegelijk was de evaluatie klassiek, in de zin dat de politiek-bestuurlijke top en de volksvertegenwoordiging moesten kunnen beoordelen of het programma moest worden voortgezet. In deze bijdrage gaan we vooral in op het proces van de lerende evaluatie en wat de ervaringen zijn die we hebben opgedaan bij de uitvoering van dit type van evalueren.

Het Stimuleringsprogramma Burger en Milieubeleid was een direct uitvloeisel van een motie die door de Tweede Kamer was aangenomen bij de behandeling van de begroting van milieu voor 2002. Daarin werd gesteld dat het milieubeleid meer ‘van, voor en door de burger’ zou moeten worden. Burgers zouden een meer coproducerende rol moeten krijgen bij de totstandkoming van het Nederlandse milieubeleid (Tweede Kamer, vergaderjaar 2001-2002, 28000 XI, nr. 21).

Het Stimuleringsprogramma kende verschillende projecten die door ons moesten worden geëvalueerd. Die projecten bevonden zich op de terreinen milieu, transitie en duurzame ontwikkeling. Zo waren er projecten als afvalbeleid, biodiversiteit in de provincie Zuid-Holland en omgang met veiligheidsrisico's. De evaluatie moest echter ook projectoverstijgend zijn; zij moest bijdragen aan het tijdig kunnen bijsturen van het totale programma. Het Stimuleringsprogramma werd getrokken door een programmteam. Dit team was ons aanspreekpunt voor het uitvoeren van de lerende evaluatie en het terugkoppelen van (tussen)resultaten.

## Evaluatie van het Stimuleringsprogramma

### *Oprichting aan evaluatieteam*

Er werd reeds in een vroeg stadium nagedacht over de vormgeving van het evaluatieonderzoek. Het programmateam wilde een evaluatie ‘tijdens de rit’. Daarmee wilde het bereiken dat de evaluatie ook kon dienen om tijdens de begrotingsbehandeling voor 2004 een discussie over de middelen voor het programma te ondersteunen. Het programmateam wilde met name een evaluatie die hen in staat zou stellen om tijdens de uitvoering van het programma bij te sturen.

De centrale evaluatievraag was in feite heel klassiek: ‘In hoeverre is het Stimuleringsprogramma Burger en Milieubeleid erin geslaagd om milieubeleid meer van, voor en door de burger te laten worden en welke adviezen kunnen worden gegeven om het Stimuleringsprogramma ten opzichte van deze hoofddoelstelling (nog) beter te doen functioneren?’

Het onderscheidende karakter van het onderzoek zat niet zozeer in de te beantwoorden evaluatievragen, maar veel meer in de wijze waarop dit onderzoek werd uitgevoerd: sterk gericht op het tijdens de rit laten doorklinken van aanbevelingen, waarbij het evaluatieteam ook lette op de wijze waarop het programmateam omging met deze geluiden (hun lerend vermogen).

### *Methoden van onderzoek*

In het evaluatieonderzoek werden verschillende onderzoeksmethoden gebruikt:

- twee telefonische enquêtes onder medewerkers van het directoraat-generaal Milieu van het Ministerie van VROM om de bekendheid van het Stimuleringsprogramma te testen en de burgeroriëntatie te achterhalen;
- een schriftelijk enquête onder gesubsidieerde projecten van maatschappelijke organisaties om de opzet, uitvoering en geboekte resultaten te achterhalen;
- een analyse van beleidsdocumenten, memo’s, projectvoorstellen, enzovoort;
- interviews met projectleiders en leden van het programmateam;
- observaties van een scala aan programmagerelateerde bijeenkomsten;
- zes leersessies, voornamelijk met het programmateam.

### *Evaluatieproces*

De evaluatie betrof met name de startfase van het Stimuleringsprogramma: januari tot en met mei 2003, en het halfjaar daaraan voorafgaand (augustus-december 2002) toen het programma ambtelijk in de steigers werd gezet. Het definitieve evaluatierapport werd uiteindelijk opgeleverd in oktober 2003.

#### **1. Eerste ronde (januari-februari 2003): verkenning**

Een evaluatie ‘tijdens de rit’ met expliciete aandacht voor tussentijdse terugkoppeling sprak de opdrachtgever erg aan. Dit was een vruchtbare basis voor het evaluatieteam om met elkaar na te denken over de vormgeving van dit evaluatietraject. De evaluatoren hebben veel tijd

gestoken in het expliciteren van de logica en ambities van het Stimuleringsprogramma. Deze verkenning en voorbereiding was nodig om met de opdrachtgever mee te kunnen denken in de verdere ontwikkeling van het Stimuleringsprogramma.

## **2. Tweede ronde (maart-juni 2003): toewerken naar een aanpassing van het programma**

Na de verkennende fase startte het evaluatieteam zijn evaluatiewerkzaamheden. Een eerste ronde interviews onder de projectleiders van de verschillende in gang gezette burger- en milieuprojecten en de uitkomsten van de eerste 'leersessies' met het programmateam leidde al snel tot de conclusie dat het team weliswaar veel initiatieven had ontwikkeld, maar dat het effect daarvan nauwelijks zichtbaar was. In het kort kwam de boodschap erop neer dat het programma te gefragmenteerd was en de projecten te weinig aansloten bij de agenda van de burger. Op deze wijze zou het Stimuleringsprogramma zijn algemene doelstelling om milieubeleid meer van, voor en door de burger te maken, niet halen.

Het programmateam heeft op deze interventie van het evaluatieteam gereageerd met een notitie waarin het pleitte voor een aantal aanpassingen. Daarbij werd met name ingegaan op een sterkere oriëntatie van de agenda op de burger (vraaggericht beleid).

## **3. Derde ronde (juli-september 2003): ver(ant)woording onderzoeksresultaten**

In deze derde ronde is voor het eerst een concepteindrapportage met de leden van het programmateam besproken. Deze rapportage is, vergeleken met het intensieve contact met de opdrachtgever in de eerste twee rondes, in relatieve distantie geschreven. Het conceptrapport veroorzaakte veel commotie bij de leden van het programmateam. Er kwam veel kritiek op de wijze waarop een en ander was verwoord. De evaluatoren hadden steeds meegedacht met het programmateam, maar stelden zich nu wat afstandelijker op en kwamen 'opeens' met confronterende analyses en conclusies.

De twistpunten waren niet zozeer inhoudelijk, maar vooral semantisch van aard. De wijze van verwoorden werd vaak kort door de bocht en weinig subtiel gevonden. Er is toen besloten om het rapport meer in samenspraak met de programmateamleden te schrijven. Het evaluatieteam heeft gekozen voor deze strategie van gezamenlijke verwoording, omdat het dit nodig achtte voor een goede doorwerking van de onderzoeksresultaten. Het evaluatierapport is uiteindelijk zonder veel ophef in het management van het ministerie en in de Tweede Kamer behandeld. Besloten werd om het Stimuleringsprogramma voort te zetten.

## **Evaluatie van de 'lerende evaluatie'**

De startcondities voor het uitvoeren van een lerende evaluatie waren gunstig. De opdrachtgever accepteerde de open insteek en het explorerende karakter van de lerende evaluatie. Met name het idee dat de evaluatie gericht was op vervlechting met het programma en de projecten die daaronder vielen, sprak de opdrachtgever aan. Dit uitgangspunt gaf het programmateam de mogelijkheid om het programma bijtijds bij te sturen.

### **Co-evolutie**

Centraal element van de lerende evaluatie is het *leervermogen* van actoren die een centrale rol spelen bij de uitvoering van (beleids)programma's. Het gaat in het geval van het geëvalueerde programma om de *wil* en de *mogelijkheid* om de verkregen inzichten te vertalen in bijsturing van het programma en de projecten daarbinnen. Lerende evaluaties gaan minder uit van afrekenen (een doelstelling is niet gehaald), maar benadrukken vooral leren (hoe kan een doelstelling wel worden gehaald?). Een leerevaluatie is daarmee sterk aan verandering onderhevig. Indien een leerpunt door het programmteam werd opgepakt, veranderden hiermee tegelijkertijd weer de focus en locus van de evaluatie. Onze boodschap dat de projecten meer vanuit de beleving van de burgers moesten worden opgezet en uitgevoerd, werd opgepakt. Daarna gingen we kijken hoe dat precies in praktijk werd gebracht.

Lerende evaluaties zijn sterk co-evolutionair: de evaluatie verandert mee met de ontwikkelingen in het te bestuderen object. Leeronderzoek staat bloot aan grote dynamiek, het moet de ontwikkelingen in de beleidspraktijk actief volgen en de waarnemingen en conclusies zijn beperkt houdbaar.

### **Strategie van interactie**

Bij aanvang van de lerende evaluatie was het de bedoeling om frequent terug te koppelen naar het programmteam om het de gelegenheid te geven om het Stimuleringsprogramma bij te kunnen sturen. Het programmteam gaf aan dat de terugkoppeling van de onderzoeksresultaten frequenter had gekund. Het programmteam had graag meer tussentijdse resultaten gezien om 'de vinger aan de pols' van de projecten te kunnen houden.

In de verkennende ronde van het evaluatietraject hebben we veel tijd geïnvesteerd in het inzichtelijk krijgen van de dominante beelden van de leden van het programmteam over het Stimuleringsprogramma en hun bedoelingen met dit programma. Er was toen nog geen gelegenheid voor terugkoppeling van onderzoeksbevindingen. Daar hadden we meer tijd voor kunnen en moeten inruimen. Toen wij uiteindelijk in de derde ronde van het evaluatietraject intervenueerden, kwam dat tamelijk onverwacht. Ook was de evaluatietijd (vijf maanden!) te kort om ook de bijgestuurde acties van het programmteam intensief te kunnen monitoren en hierover terug te kunnen rapporteren.

### **Evalueren in het hart en niet aan de zijlijn van het beleidsproces**

Evalueren is een activiteit die niet los kan worden gezien van de praktijk zelf waarin ervaringen cumuleren. In het proces zijn reflectie en actie in de context van een praktijkprobleem voortdurend onderling op elkaar betrokken. Evalueren is niet zozeer het systematisch langs een van tevoren uitgezet traject een gewenst doel bereiken, als wel het vanuit een open grondhouding, vertrekkend vanuit globale onderzoeksdoelen, op een intelligente en creatieve wijze omgaan met de onzekerheid en onvoorspelbaarheid in een voortdurend veranderende werkelijkheid (betreffende de bestaande en beoogde situatie). We opereerden als *reflective practitioners* (Schön, 1983); we poogden een grote mate van situationele responsiviteit (*situated responsiveness*, Alkin, 1990:74) ten toon te spreiden. Evalueren zagen we niet zozeer als een gestandaardiseerde ontleding van het probleem in fasen en delen, maar meer

als een voortdurende iteratie, het lerend zoeken, het inpassen van creatieve vondsten in het grotere geheel.

Tijdens de evaluatie vervulden we diverse rollen. Soms waren we beoordelaars (daar vroeg de opdrachtgever vaak expliciet om) door ons positief en soms negatief uit te spreken over het verloop van het programma en de afzonderlijke projecten. Dan traden we weer op als expert op het vlak van burgerbetrokkenheid door hen te wijzen op de mogelijkheden om burgers te betrekken bij de vorming van milieubeleid. Dan weer namen we de rol aan van coach/facilitator door leden van het programmteam actief te begeleiden in hun gedeelde zoektocht naar ideevorming over het programma.

Eigenlijk was er sprake van een dubbel traject: een leertraject gericht op interactie met de centrale doelgroep en een evaluatietraject waarin de evaluatoren zich meer op afstand opstelden en de tijd namen waarnemingen te verzamelen. De beide sporen zijn tijdens het evaluatieproces lichtelijk van elkaar vervreemd geraakt. De waarneming in het tweede traject hadden we meer moeten inbrengen in het eerste traject van de evaluatie. Het was moeilijk om de balans te vinden tussen een afstandelijke houding (in onze rol van beoordelaars) en een betrokken houding (in onze rol van coach, facilitator, adviseur). Vooral in het begin namen we een meedenkende rol aan: adviseren, coachen, faciliteren. Maar aan het eind traden we meer op als beoordelaars, toen we onze ervaringen en bevindingen moesten rapporteren. In die fase namen we ook meer afstand van het programmteam.

Het zou goed geweest zijn als we binnen het evaluatieteam een rolverdeling hadden doorgevoerd waarbij enkele leden meer de adviserende en coachende rol hadden gespeeld en de andere leden meer de klassieke evaluerende of beoordelende rol. Dit had ook meer duidelijkheid geschapen naar onze opdrachtgever en had verkeerde verwachtingen en miscommunicatie over het verloop en de uitkomsten van de lerende evaluatie kunnen voorkomen.

### **Conclusie: randvoorwaarden en positionering**

Dit artikel wilde voorzien in een empirische beschrijving van wat wij een lerende evaluatie hebben genoemd. Deze vorm van evalueren wijkt sterk af van de meer klassieke vormen van effectevaluatie, omdat hij oploopt met het te evalueren (beleids- of besluitvormings)proces en gaandeweg de rit terugkoppelt naar de gebruikers van de evaluatie. De lerende evaluatie moet dan ook niet als een sluitstuk van het beleidsproces worden gezien, maar als een steeds terugkerende schakel in het beleidsproces.

We zijn van mening, op basis van verkregen inzichten, dat minimaal de volgende condities aanwezig moeten zijn, wil toepassing van de lerende evaluatie tot de mogelijkheden behoren:

– *Beleidsprogramma in de ontwikkelingsfase*

De lerende evaluatie heeft de grootste kans van slagen als een beleidsprogramma nog in ontwikkeling is en zich nog moet bewijzen, omdat deze dan het meest vatbaar is voor adviezen van buitenaf. De opdrachtgevers/gebruikers van de lerende evaluatie moeten bereidheid tonen om hun beleidsprogramma tijdens de rit naar verkregen inzichten te willen aanpassen.

– *Organisatie van continue feedback*

Bij de lerende evaluatie staat de wisselwerking tussen de evaluatie en het beleidsproces centraal. Deze wisselwerking ontstaat echter niet vanzelf, maar moet georganiseerd worden. Onze eigen leerervaring was dat wij in de uitvoering van de lerende evaluatie eerst te weinig hadden ingezet op die wisselwerking. Op een gegeven moment kwam het programmteam zelf met de wens de wisselwerking te intensiveren, maar tegelijkertijd gaven zij aan dat de verantwoordelijkheid voor de organisatie ervan bij het evaluatieteam lag.

– *Vaardigheden en rollen van de evaluator*

Een lerende evaluatie vergt veel van de evaluator. Deze moet betrokkenheid tonen (bij het beleidsprogramma) om relevante adviezen te kunnen geven, maar moet tevens voldoende distantie behouden om onafhankelijk te blijven. Het is verstandig om binnen het team van evaluatoren tot een rolverdeling te komen: de ene groep richt zich vooral op het geven van adviezen (betrokken rol), de andere groep richt zich meer op het beoordelen van het te evalueren programma (onafhankelijke rol). Het is van belang om de interactie binnen de groep goed te verzorgen en de communicatielijnen kort te houden. De rolverdeling moet ook goed naar de opdrachtgevers worden gecommuniceerd om geen verkeerde verwachtingen te wekken. Aan hen moet duidelijk worden gemaakt wie in welke rol zit, maar ook op welke fase van het evaluatieproces het accent ligt. Onze eigen ervaring is dat wij hier aanvankelijk te onduidelijk over waren naar de opdrachtgevers toe.

Terugkijkend kunnen we concluderen dat de lerende evaluatie een hybride vorm van evaluatiestudie is, die tegelijkertijd valt te benoemen als:

- een rationele evaluatie (Provus, 1971), omdat ze focust op de vraag of (eventueel tijdens de rit gewijzigde) doelstellingen daadwerkelijk zijn gerealiseerd;
- een constructivistische evaluatie (Fischer & Forester, 1993) omdat ze het beleidsproces benadert als het resultaat van een proces van vormgeving en zingeving door actoren met verschillende interpretatiekaders;
- een responsieve evaluatie (Stake, 1983) omdat ze erop gericht is voortdurend informatie met gebruikers van de studie uit te wisselen en daarmee kiest voor een open, adaptieve aanpak van de evaluatiestudie;
- een participatieve evaluatie (House & Howe, 2000) omdat ze ernaar streeft participanten een stem te geven in de vormgeving en uitvoering van de evaluatiestudie om daarmee de bruikbaarheid en de kwaliteit van de studie te vergroten;
- een op doorwerking georiënteerde evaluatie (Patton, 1997) omdat ze focust op het maximaliseren van de bruikbaarheid van de conclusies, adviezen en aanbevelingen voor de beleidspraktijk.

Daarenboven heeft de lerende evaluatie een eigen waarde, omdat ze ook het leervermogen en het adaptatievermogen van de betrokken actoren tracht te waarderen. Daarnaast is zij uniek omdat zij niet voorafgaand of na afloop van een beleidsprogramma wordt uitgevoerd, maar

tijdens de rit. Als de lerende evaluatie erin slaagt de verschillende kenmerken van de diverse evaluatietypen vruchtbaar te combineren, resulteert dat in een spannende, maar tegelijk ook relevante evaluatiebenadering, die in potentie voor zowel de praktijk als de wetenschap meerwaarde heeft.

## Literatuur

- Alkin, M.C. (1990). *Debates on Evaluation*. London: Sage.
- Edelenbos, J., M.W. van Buuren, M. van As, M. Duijn, A. Slob, & W. Vermeulen (2003). *Milieu beleid meer van, voor en door de burger. Resultaten van de lerende evaluatie van het Stimuleringsprogramma Burger en Milieu beleid*. Erasmus Universiteit Rotterdam.
- Edelenbos, J. & A. van Buuren (2005). The Learning Evaluation. A Theoretical and Empirical Exploration. *Evaluation Review*, 29(6), 591-612.
- Fischer, F. & J. Forester (eds.) (1993). *The Argumentative Turn in Policy Analysis and Planning*. Durham/London: Duke University Press.
- Guba, E. & Y. Lincoln (1989). *Fourth Generation Evaluation*. Newbury Park: Sage.
- House, E.R. & K.R. Howe (2000). Deliberative democratic evaluation. *New Directions for Evaluation*, 85, 3-12.
- Patton, M.Q. (1997). *Utilization-focused Evaluation. The next century text*. 3rd edition. Thousand Oaks: Sage Publications.
- Probus, M.N. (1971). *Discrepancy evaluation*. Berkeley, CA: McCutcheon.
- Schön, D.A. (1983). *The reflective practitioner: how professionals think in action*. New York: Basic Books.
- Tweede Kamer, 2001-2002. Motie Feenstra, Klein Molenkamp en Augusteijn-Esser, 28 000 XI: 21.
- Stake, R.E. (1983). Program evaluation, particularly responsive evaluation. In: G.F. Madaus, M. Scriven, & D.L. Stufflebeam (eds.). *Evaluation models*. Boston: Kluwer-Nijhoff, 287-310.