

Evalueren als leerproces

Een nadere kennismaking met de 'lerende evaluatie'

Jurian Edelenbos & Arwin van Buuren

Dr. J. Edelenbos

Erasmus Universiteit Rotterdam

Centrum voor Publiek Management

Postbus 1738, Kamer M8-06

3000 DR Rotterdam

edelenbos@fsw.eur.nl

Drs. M.W. van Buuren

Erasmus Universiteit Rotterdam

Centrum voor Publiek Management

Postbus 1738, Kamer M8-05

3000 DR Rotterdam

vanbuuren@fsw.eur.nl

Referentie:

Edelenbos, J. & M.W. van Buuren (2005) Evaluatie als leerproces. Een nadere kennismaking met de 'lerende evaluatie'. In: *Bestuurskunde*, 14 (6), pp. 2-12.

1. Inleiding

Evaluatiestudies zijn in het openbaar bestuur aan de orde van de dag. Ze zijn er in alle soorten en maten. Recent wordt nagedacht over 'lerende' evaluaties: evaluatiestudies waarbij de nadruk niet zozeer wordt gelegd op het inventariseren van successen en mislukkingen, maar waarbij het vooral gaat om het gezamenlijk door evaluatoren en geëvalueerde te komen tot verbeteringen van het betreffende project. Beoordelen is een element van deze evaluatie, maar daarnaast gaat het er vooral ook om een leerproces te organiseren. Deze evaluatiemethode stelt uiteenlopende eisen aan evaluatoren en genereert veel rolconflicten voor hen. Genoeg reden om bij dit fenomeen stil te staan.

Dit artikel wil in de eerste plaats licht werpen op de plaats van de lerende evaluatie temidden van de stortvloed aan wetenschappelijk onderzoek naar evaluatiestudies. Aan de hand van een analyse van een door onszelf uitgevoerd lerend evaluatieproject komen wij tot een nadere plaatsbepaling van de lerende evaluatie. Onze ervaringen kunnen de theorievorming rond de lerende evaluatie maar ook de praktijk van dergelijke evaluatiestudies verder brengen.

In 2003 hebben wij met enkele anderen het Stimuleringsprogramma Burger en Milieubeleid van het Ministerie van VROM geëvalueerd met een lerende evaluatieaanpak (zie Edelenbos, e.a., 2003; Edelenbos & Van Buuren, 2005). De door ons uitgevoerde evaluatiestudie was een ontdekkingsreis voor onszelf in de wereld van evaluatieonderzoek. We wilden geen klassieke, maar een 'leerevaluatie'. Het moest in ieder geval een evaluatievorm worden dat gericht was op (1) doorwerking van de resultaten uit de evaluatie; (2) het bereiken van gedragsverandering; (3) vergroting van het leervermogen van het team dat het Stimuleringsprogramma trok.

In paragraaf 2 geven we een korte beschrijving van tendensen en ontwikkelingen in de wereld van evalueren. Vervolgens bespreken we onze ervaringen met de vormgeving en de uitvoering van de evaluatie van het Stimuleringsprogramma. In paragraaf 4 evalueren we de evaluatieaanpak die zich gaandeweg heeft ontwikkeld. In paragraaf 5 vatten we de kenmerken van een lerende evaluatie samen en positioneren deze in de bestaande literatuur over evaluatiestudies.

2. Evaluatieonderzoek in vogelvlucht

In de jaren tachtig zien we een omslag in het denken over evaluaties. Deze omslag speelt zich af parallel aan een ontwikkeling in het denken over de overheid en over de wijze waarop beleid tot stand komt en wordt uitgevoerd. De klassieke visie op de overheid, de wijze waarop zij de samenleving stuurt, wordt meer en meer losgelaten. Er komen andere conceptualisering van het beleidsproces in zwang. Daarmee wijzigen ook de perspectieven op beleidsevaluatie.

Een belangrijke ontwikkeling is de democratisering van het evaluatieonderzoek. Daarvan is sprake in tal van verhandelingen over participatieve, constructivistische en responsieve vormen van evaluatie (zie bijvoorbeeld: Scriven, 1974; Stake, 1983; Schwandt, 1984; Guba & Lincoln, 1989; Abma, 1996; House & Howe, 2000). In verschillende gradaties worden evaluatiestudies in samenspraak met betrokkenen uitgevoerd. Een eerste stap is het betrekken van direct betrokkenen bij de uitvoering van het beleidsprogramma (Patton, 1997), zoals de medewerkers, maar ook de directe klanten of afnemers. Daarnaast kan men stakeholders, niet zijnde de klanten, maar personen en organisaties die een belang hebben bij de evaluatie, betrekken (Guba & Lincoln, 1989). Tot slot kan participatie zich uitstrekken tot alle geïnteresseerde burgers (wat vaak wordt genoemd bij zaken als 'community care'). Er wordt dan ook wel gesproken van deliberatieve evaluatie (House & Howe, 2000).

In participatieve evaluatievormen staat de notie dat mensen zelf betekenis geven aan de wereld om zich heen, centraal. In het kennen van de werkelijkheid is intersubjectieve overeenstemming het hoogst haalbare ideaal. Elke representatie van de werkelijkheid is normatief. Noch de beleidsmaker, noch de evaluator heeft de waarheid in pacht. Om een zinvol evaluatietraject uit te voeren is het bereiken van overeenstemming tussen de diverse betrokkenen cruciaal.

Veel traditionele evaluatiestudies hebben een ex post karakter: een programma wordt achteraf beoordeeld op haar uitkomsten. Daarmee kunnen politici en bestuurders worden afgerekend op hun daden. Erkenning van de complexiteit waarin beleidsprocessen zich afspelen doet nadenken over mildere vormen van evaluatie, waarbij verbetering van het functioneren van programma's een rol speelt. Daarbij passen evaluaties met een meer ex durante karakter: een programma wordt tijdens de uitvoering nauwgezet gevolgd en van reflecties voorzien die kunnen leiden tot aanpassingen.

Vierde generatie evaluatieonderzoek

Veel van bovenstaande ontwikkelingen krijgen een plaats in wat Guba & en Lincoln (1989) vierde generatie evaluatieonderzoek hebben genoemd. Deze wijze van evalueren betekent dat de rol van de evaluator anders wordt (Guba & Lincoln, 1989; Abma, 1996; Edelenbos en Van Eeten, 2001). De evaluator is geen controller meer, maar samenspeler. Hij tracht samen met de geëvalueerde te komen tot betekenisvolle inzichten. Daarbij onderwijst hij en tracht hij zijn kennis over te dragen en gezamenlijk tot gedeelde kennis te komen. Tot slot is hij ook een verandermanager. Hij neemt verantwoordelijkheid om het programma te verbeteren. Hij is een sleutelfiguur in het creëren van een 'more sophisticated reality' (Guba & Lincoln, 1989, p.262).

Lijnen voor hedendaagse evaluatiearrangementen

Uit bovenstaande blijkt dat zowel het in interactie met gebruikers en belanghebbenden uitvoeren van het evaluatieonderzoek, als ook het insteken op de praktische toepasbaarheid van de studie, belangrijke insteken zijn voor het hedendaagse evaluatieonderzoek. Dit hangt samen met de democratisering en fragmentatie van beleidsprocessen en de meer dienende rol die wordt weggelegd voor evaluatoren (zie Teisman et.al. 2002). Daarbij is ook ruimte voor methodologische vernieuwing door het toepassen van vormen van actieonderzoek waarbij wetenschappers opereren als verandermanager door het doorlopen van korte cirkels van reflectie, observatie, discussie, theoretiseren, concluderen, plannen, handelen, reflecteren, observeren, enzovoorts (Wadsworth, 2001, p.52).

3. Project Stimuleringsprogramma Burger en Milieubeleid

3.1 Achtergronden van het project

Ontstaan van het Stimuleringsprogramma

De start van het Stimuleringsprogramma Burger en Milieubeleid was het gevolg van de door de Tweede Kamer bij de begroting van Milieu voor 2002 aangenomen motie van de leden Feenstra, Klein Molekamp en Augusteijn-Esser (Tweede Kamer,

vergaderjaar 2001-2002, 28 000 XI, nr. 21) dat het milieubeleid meer ‘van, voor en door de burger’ zou moeten worden. Met het amendement werd beoogd om de burgers veel meer in een coproducerende rol bij de totstandkoming van het milieubeleid te betrekken. Dit besluit sloot aan op een politiek breed geuite wens om het openbaar bestuur dichterbij de burger te brengen.

Doelstellingen

Het Stimuleringsprogramma zou in het eerste jaar het accent leggen op het realiseren van een sterkere oriëntatie in de beleidsvoorbereiding op wat de burger beweegt en belangrijk vindt. Beleid moest beter gaan aansluiten bij de leef- en denkwereld van de burger. Daar waar mogelijk zouden ook ervaringen met burgeroriëntaties in de andere fasen (besluitvorming, uitvoering, handhaving/naleving, monitoring en evaluatie) van het beleidsproces opgedaan worden.

De eerste doelstelling was het versterken van de ‘burgeroriëntatie’ in milieu- en duurzaamheidsbeleid. Dit element was gericht op beleidsmakers. Zij moeten meer rekening leren houden in beleidsvorming met de effecten van beleidsmaatregelen op het dagelijkse leven van mensen.

De tweede doelstelling van het Stimuleringsprogramma was het betrekken van burgers bij de beleidsvorming: de coproductie van beleid. Onder ‘coproductie’ werd in het kader van dit programma een breed scala van vormen van burgerparticipatie in beleidsvoorbereiding en besluitvorming verstaan: van consultatie en advisering tot aan meebeslissen.

3.2 Evaluatie van het Stimuleringsprogramma

Evaluatieopdracht

Er werd reeds in een vroeg stadium nagedacht over de vormgeving van het evaluatieonderzoek. Het programmateam dat het Stimuleringsprogramma uitvoerde wilde een evaluatie ‘tijdens de rit’, zodat zij tijdens het verloop van het programma, projecten kon bijsturen op basis van tussentijdse inzichten. Ook wilde zij dat de evaluatie kon dienen om tijdens de begrotingsbehandeling voor 2004 een stand van zaken op te leveren om de discussie over de middelen voor het programma te kunnen ondersteunen. Het programmateam zag vier doelen voor de evaluatie:

1. het opstellen van een terugrapportage naar Tweede Kamer over de stand van zaken met betrekking tot de uitvoering van de motie – Feenstra c.s.;
2. een inventarisatie van de knelpunten van het Stimuleringsprogramma, gericht op de verbetering van dit programma;
3. een analyse van de potenties van de aanpak van het stimuleringsprogramma;
4. het verkrijgen van inzicht in de wijze waarop burgerbetrokkenheid vorm krijgt (oriëntatie van ambtenaren, participatie van burgers en initiatie van burgers).

De centrale evaluatievraag luidde: “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.”

Mix van methoden

Het eigene karakter van het evaluatieonderzoek zat vooral in de wijze waarop het onderzoek zou worden uitgevoerd. Het programmateam wilde een evaluatie die sterk gericht was op het tijdens de rit laten doorklinken van aanbevelingen, waarbij het evaluatieteam ook zou letten op de wijze waarop het team om zou gaan met deze geluiden in het eventueel bijstellen van het programma. Het evaluatieteam was niet alleen bezig met het verzamelen van gegevens over de wijze waarop het Stimuleringsprogramma en het Programmateam functioneerden, maar was ook nadrukkelijk bezig om het Programmateam verder te helpen. Daarmee kreeg deze evaluatie het karakter van zowel een beoordelende als een adviserende evaluatie.

Evaluatieproces

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

1. Voorbereiding en eerste verkenning (januari – februari 2003)

Van de kant van de opdrachtgever was enthousiasme over de vernieuwende aanpak van het evaluatieonderzoek. Een evaluatie ‘tijdens de rit’ met expliciete aandacht voor

tussentijdse terugkoppeling sprak erg aan. Dit was een vruchtbare basis om met elkaar na te denken over de vormgeving van dit evaluatietraject. De evaluatoren hebben veel tijd gestoken in het duidelijk krijgen van de logica en ambities van het Stimuleringsprogramma. Een aantal leersessies was hiervoor nodig. Toen bleek dat de teamleden volstrekt niet op één lijn zaten. De samenstelling van het programmateam bleek een belangrijke verklarende factor. Een aantal mensen van het Instituut voor Publiek en Politiek (IPP) waren ingehuurd om een bijdrage te leveren aan de personele bezetting van het team. In meerderheid hadden zij verdergaande ambities met betrekking tot burgerparticipatie dan de VROM-medewerkers.

2. Evaluatie, bijstelling, evaluatie (maart – juni 2003)

Min of meer gelijktijdig met de organisatie van deze leersessies, startte het onderzoeksteam haar evaluatiewerkzaamheden. Een eerste ronde interviews onder de projectleiders van de verschillende 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. Er waren tal van kleinere en grotere projecten binnen het departement opgezet (zoals het project zwerfafval, het project externe veiligheid en dergelijke), maar de projecten werden nog erg gedomineerd door de agenda van ambtenaren en professionals; de burger kwam er in die fase bekaaid vanaf. In een leersessie in maart 2003 heeft het evaluatieteam aan de rem getrokken met de boodschap dat het Stimuleringsprogramma haar doelstelling om milieubeleid meer 'van, voor en door de burger' te maken niet zou halen als de ingezette koers zou worden gehandhaafd. In het kort kwam het erop neer dat het programma te gefragmenteerd was en er te weinig vanuit de agenda van de burger projecten werden opgezet en uitgevoerd. Deze boodschap leidde tot een tussentijdse koerswijziging en een intensivering van het Stimuleringsprogramma.

Het programmateam schreef in reactie op deze noodkreet een notitie voor de Directeur-generaal waarin zij pleitte voor een aantal intensiveringen in de geest van het amendement Feenstra c.s. Deze 'intensiveringsnotitie' heeft geresulteerd in het voornemen enkele aanpassingen in het programma door te voeren:

- burgers moesten een belangrijke plaats worden gegeven bij de agendavorming; er moest sterker van 'buiten naar binnen' worden gewerkt. Burgers actief betrekken

bij de agendavorming wekte verwachtingen die ook waargemaakt moesten kunnen worden;

- meer nadruk op burgerinitiatieven: om te realiseren dat meer 'van buiten naar binnen' zou worden gewerkt zouden meer mogelijkheden moeten worden gecreëerd voor burgerinitiatieven;
- geen prioriteit bij meebeslissende rol van burgers.

Na deze schriftelijke koerswijziging werd de evaluatie vervolgd. Nieuwe leersessies werden georganiseerd, de projectleiders werden nogmaals geïnterviewd, voor de tweede maal werd een steekproefsgewijze enquête onder DGM-medewerkers uitgezet en werden projecten geobserveerd.

3. Ver(ant)woording onderzoeksresultaten (juli – september 2003)

Na een intensieve periode van zo'n drie maanden, met veel leersessies, onderzoek en interactie, heeft het evaluatieteam zich teruggetrokken voor het schrijven van een eerste concept van het evaluatierapport. Dit eerste concept werd in juli 2003 in het programmateam besproken.

Het eerste conceptraapport riep veel weerstand op bij de programmateamleden. Hoewel de ingehuurd medewerkers van het IPP positiever waren dan de uit de ambtelijke organisatie afkomstige teamleden, kwam er veel kritiek op de wijze waarop de conclusies waren verwoord.

De twistpunten waren vooral taalkundig 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 onderlinge interactie tussen onderzoekers en programmateamleden betekenis te geven. Evaluatoren leverden conceptversies van hoofdstukken aan en programmateden gaven daar commentaar op. Het evaluatieteam heeft expliciet gekozen voor deze strategie van gezamenlijke verwoording, omdat zij dit nodig achtte voor een goede doorwerking van de onderzoeksresultaten. Het evaluatierapport is uiteindelijk geruisloos de Tweede Kamer gepasseerd. Voortzetting van het Stimuleringsprogramma werd nodig geacht.

4. Evaluatie ‘lerende evaluatie’

In deze paragraaf blikken we terug op het verloop van het evaluatieproces. We passen een ‘zelfevaluatie’ toe op een door onszelf vormgegeven en uitgevoerde evaluatie. Daarnaast hebben we gebruik gemaakt van commentaar van het programmteam dat in een evaluatiegesprek in oktober 2003 met hen naar voren kwam.

Een welwillende opdrachtgever als vertrekpunt

De condities voor het uitvoeren van een lerende evaluatie waren gunstig. Het was mogelijk om een vorm van lerende evaluatie uit te voeren, vanwege de positieve houding van het programmteam (dat als opdrachtgever van de evaluatiestudie optrad). Met name het idee dat de evaluatie gericht was op regelmatige terugkoppeling van informatie uit de evaluatie, zodat bijsturing van de projecten mogelijk werd, sprak de opdrachtgever aan. Een open insteek en een explorerend karakter van de lerende evaluatie accepteerden zij als uitgangspunt.

Lerende evaluatie: steeds veranderende focus en locus

Een belangrijke norm bij lerende evaluatie is het *leervermogen* van actoren die een centrale rol spelen bij de uitvoering van het programma Burger en Milieubeleid: in hoeverre hebben zij de *wil* en de *mogelijkheid* om de verkregen inzichten te vertalen in bijsturing van het programma en de verschillende burger en milieuprojecten daarbinnen. Lerende evaluaties hebben veel minder een afrekenmentaliteit (“jullie hebben deze en die doelstelling niet gehaald”), maar vooral een leermentaliteit (“wat moet er gebeuren om zaken een bepaalde richting op te doen bewegen”). Een leerevaluatie is sterk aan verandering onderhevig. Indien een leerpunt door het Programmteam werd opgepakt, veranderde hiermee tegelijkertijd weer de focus en locus van de studie. Een voorbeeld hiervan is onze interventie dat de projecten meer vanuit de beleving van de burgers moesten worden opgezet en uitgevoerd werd opgepakt, waarna we gingen kijken hoe dat precies in de praktijk werd gebracht. Hiermee veranderde tegelijkertijd wat we verder gingen onderzoeken (locus) en op welke manier we dat zouden gaan volgen (focus).

Lerende evaluaties vertonen dan ook een sterk co-evolutionair karakter: de evaluatie verandert mee met de ontwikkelingen in het te bestuderen object. Leeronderzoek vertoont daarmee sterke dynamiek omdat zij ontwikkelingen in de beleidspraktijk

actief volgt. Constateringen en conclusies zijn daarmee wel beperkt houdbaar. Daarom zijn ook korte lijnen voor terugkoppeling noodzakelijk.

Waarneming, reflectie en (re)actie: veel kennis absorberen en te weinig feed back

Bij de start van de lerende evaluatie was het de bedoeling van het evaluatietraject om frequent terug te koppelen naar het programmteam om hen de gelegenheid te geven om op basis van het evaluatiemateriaal het Stimuleringsprogramma als totaal en de verschillende concrete projecten daarbinnen te kunnen bijsturen. Het programmteam gaf aan dat de terugkoppeling van de onderzoeksresultaten frequenter had gekund. Het programmteam had graag meer tussentijdse resultaten gezien om continu te kunnen waarnemen wat er gebeurde in de projecten.

De insteek van de lerende evaluatie was om korte en frequente cycli van waarneming (door de evaluatoren), gezamenlijke reflectie (inzicht krijgen en betekenis verlenen) en (re)actie te organiseren. In het begin hebben we veel tijd geïnvesteerd om grip te krijgen op de 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 datgene wat wij via onderzoek hadden waargenomen.

Toen wij uiteindelijk een ‘noodkreet’ lieten horen, kwam dat geluid tamelijk onverwacht. We hadden meer tijd moeten vrijmaken voor terugkoppeling van tussentijdse indrukken in de leersessies. Tenslotte was de evaluatietijd te gering om ook de bijgestuurde acties van het programmteam intensief te kunnen monitoren en hierover terug te kunnen rapporteren. Toen de rapportage moest worden afgerond was de implementatie van de leerpunten in volle gang.

Rol evaluatoren: spanning tussen betrokken en op afstand

Tijdens de evaluatie traden we als evaluator in verschillende hoedanigheden op. Soms waren we beoordelaars (daar vroeg de opdrachtgever vaak expliciet om) door positief en soms negatief uit te spreken over hoe het programma en de projecten daarbinnen verliepen. Sprekend voorbeeld hiervan is onze opstelling in een van de leersessies, waarin we aangaven dat het programma te weinig voortgang boekte en onvoldoende ver ging in het betrekken van burgers bij de ontwikkeling van het milieubeleid.

Vervolgens traden we op als expert en adviseur op het vlak van burgerbetrokkenheid door hen te wijzen waar men op moest letten als het ging over het betrekken van

burgers bij de vorming van milieubeleid. We werden ook op onze rol van expert/adviseur aangesproken toen expliciet aan ons gevraagd werd op welke andere wijze het Stimuleringsprogramma gestalte moest worden gegeven.

Dan weer namen we de rol aan van coach/facilitator door leden van het programmateam actief te begeleiden in de diverse leersessies in hun zoektocht naar ideevorming over en filosofie achter het Stimuleringsprogramma. Ook de doelstelling en toekomst van het programma werden in deze sessies onder begeleiding van de evaluatieteam besproken.

Op sommige momenten werden we bijna als 'medeambtenaren' gezien die een gelijk belang hadden in het tot een succes maken van het Stimuleringsprogramma. We werden als partners gezien die een gelijk belang leken te hebben, namelijk het tot een succes maken van het Stimuleringsprogramma.

Als we terugkijken kunnen we constateren dat het voor ons moeilijk is geweest om de goede balans te vinden tussen de verschillende rollen. 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. We beoordeelden toen in klassieke zin (zijn de doelstelling om milieubeleid meer van, voor en door de burger te maken gehaald) en in moderne zin (is er leervermogen aan de dag gelegd, is het Stimuleringsprogramma kwalitatief beter geworden). In die fase namen we ook meer afstand van het programmateam.

Dat leverde strijd op. Het Programmateam zag ons eerst als helpers, meedenkers, coaches en vervolgens trof zij ons in de rol van beoordelaar, examiner. Dat was ook voor het evaluatieteam lastig. De intensieve kennismaking met het project en de kennis over de ideeën erachter en de lastige praktijk, maken het lastig om voldoende afstand te nemen en zakelijk conclusies te trekken. Te makkelijk wordt de kool en de geit gespaard.

Een lerende evaluatie vereist van onderzoekers dat zij verschillende rollen kunnen uitoefenen; ze moeten beschikken over verschillende vaardigheden en competenties. Daarnaast worden onderzoekers aangesproken op hun 'schakelend vermogen': het tijdig en vloeiend kunnen overstappen naar verschillende rollen.

Die spanningen in roltypen zie je overigens ook terug in de literatuur over advisering (bijvoorbeeld: Feltmann, 1984; Van Twist e.a., 1998) en management van grote complexe projecten (Teisman, 1992; Otten, 2000) over de rol van de controller. De controller kan zijn rol opvatten als klassieke controleur, waarin toetsen van beleid,

kritische distantie en afrekenen op gestelde doelen centraal staat, of als adviseur met als dominante karakteristieken verbeteren van beleid, kwaliteitsgerichte betrokkenheid en gericht op leren van gemaakte fouten. De eerste rol staat sterk in het teken van afrekenen en beheersen, terwijl de tweede rol meer uitgaat van motiveren en ontwikkelen.

Doorwerking: het oppakken van de waarnemingen en lessen

Uiteindelijk heeft de opdrachtgever veel constatering en lessen opgepakt uit de evaluatiestudie. In eerste instantie werden deze op schrift gesteld in de zogenaamde 'intensiveringsnotitie'. Later werden ze ook in acties vertaald door projecten vanuit de agenda van de burger vorm te geven en de burger een meer actieve rol te geven in de ontwikkeling van beleid (via bijvoorbeeld de methode 'burgerforum'). Bovendien is de samenvatting van het evaluatierapport integraal naar de Tweede Kamer gestuurd. Ook onze telefonische enquête naar de bekendheid van het programma werkte door. Deze bekendheid bleek relatief laag. Het programmateam heeft, in reactie daarop, haar communicatie intern versterkt. Dit bleek te werken. Bij de tweede telefonische enquête was de bekendheid van het programma een stuk groter.

De moeilijkheid hierbij was dat wij als evaluatoren soms het idee hadden ons erg dienstbaar op te stellen naar de opdrachtgever. Met name aan het einde van het evaluatieproces, toen de balans moest worden opgemaakt, stelde de opdrachtgever zich defensief op. Wij hadden als evaluatoren veel moeite gedaan om in de lijn van het programma mee te denken. Dat maakte het in een later stadium lastig om weer enige afstand te nemen en 'neutraal' conclusies te formuleren over het functioneren van het programma. Op dat moment moesten we balanceren tussen enerzijds betrokkenheid en inlevingsvermogen (door in termen van het stimuleringsprogramma te denken en praktische en bruikbare tips te geven), en anderzijds voldoende distantie om onafhankelijk te kunnen blijven en onze eigen professionaliteit te bewaren.

5. Conclusie en discussie

In dit artikel stond een casusbeschrijving van een lerende evaluatie centraal. In de literatuur over evaluaties wordt veel gesproken over verschillende soorten evaluaties en ontwikkelingen in evaluatiestudies. Dit artikel wilde voorzien in een empirische

beschrijving van wat wij een lerende evaluatie hebben genoemd. In deze paragraaf proberen we enkele conclusies over deze vorm van evalueren te trekken. We pogen daarbij een koppeling te maken met verschillende soorten evaluaties die in de literatuur over evaluaties terug te vinden zijn.

Lerend evalueren: actieoriëntatie noodzakelijk

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). Lerend evalueren is dan een proces van proberen en aanpassen, van handelen en interpreteren.

Een lerende evaluatie is een evaluatievorm waarin gebruikers (de geëvalueerden) en uitvoerders van de evaluatie (de evaluatoren) in nauwe onderlinge interactie en samenspraak vorm geven en uitvoering geven aan evaluaties. Een belangrijk element is dat er frequente cycli van waarneming, constatering en (re)actie optreden. De waarneming en constatering zijn niet het sluitstuk van een evaluatie, maar vormen veelmeer het startschot voor de evaluatie. Dit komt overeen met de 'reflective practitioner' (Schön, 1983) en 'situated responsiveness' (Alkin, 1990). De evaluator staat in een voortdurende interactie met de actoren, op wie zijn evaluatie betrekking heeft. Zij reageren op zijn tussentijdse conclusies, waarna de evaluator weer op zoek gaat naar de effecten daarvan.

Een meerstemmige context

De evaluator dient bij de lerende evaluatie verschillende doelgroepen tegelijkertijd. In de eerste plaats de opdrachtgever (in deze casus: het programmateam), maar daarnaast ook andere partijen die er verder bij betrokken zijn (in deze casus: de ambtelijke projectleiders, betrokken maatschappelijke groepen, burgers die nauwer bij milieubeleid betrokken zouden moeten worden en ook de Tweede Kamer).

Deze meerstemmigheid geeft gemakkelijk aanleiding tot rolconflicten. De evaluator zal zelf moeten aangeven dat hij niet exclusief wenst geclaimd te worden door één

specifieke doelgroep. Zijn rol is mede gelegen in het overbruggen van tegenstellingen. Hij speelt daarbij de rol van onderhandelaar (Guba & Lincoln, 1989) waarbij niet vergeten mag worden dat ook zijn eigen probleemdefinitie voorwerp van onderhandeling moet zijn. Daarmee is een spanningsveld gecreëerd (zie Weiss, 1998). Enerzijds dient een evaluator zoveel mogelijk te streven naar een onderzoek waarin recht gedaan wordt aan de pluraliteit van beelden en ambities van betrokkenen. Hij moet energie investeren in het bereiken van overeenstemming over zijn evaluatieaanpak, zijn methoden en zijn uitkomsten. En anderzijds moet hij afstand bewaren, onafhankelijkheid garanderen om in staat te zijn met enig gezag aan politieke en ambtelijke bazen zijn conclusies te presenteren. Het vinden van de juiste balans tussen verschillende rollen (beoordelaar, facilitator, coach, onderhandelaar, etc.) wordt daarmee een kerncompetentie van de evaluator.

Meervoudige evaluatiearrangementen

In de literatuur over evaluaties worden vaak verschillende evaluaties naast elkaar gepresenteerd als onverenigbare polen (zie bijv. Guba en Lincoln, 1989; Stufflebeam, 2001). Zo wordt de klassieke rationele evaluatie tegenover de moderne constructivistische evaluatie geplaatst. In de evaluatiepraktijk wordt men welhaast gedwongen om een exclusieve keuze voor het ene of het andere te maken. Dit artikel biedt het inzicht dat men bij de uitvoering van een lerende evaluatie een dergelijke eenzijdige keuze niet wenselijk is. Een evaluator moet in de praktijk verschillende rollen vervullen; een evaluator meet, beoordeelt, spiegelt zijn beschrijvingen, onderhandelt, adviseert, faciliteert, maar heeft vooral als taak om mensen tot leren te verleiden. Deze verleidingskunst speelt in elke evaluatiestudie een rol, al is het maar om te bereiken dat de evaluatie-uitkomsten doorwerking krijgen. Maar het verleiden tot leren is wel in het bijzonder verbonden aan de lerende evaluatie, waarbij het leeraspect in het gehele evaluatietraject centraal staat.

Het is de kunst om elementen van verschillende soorten evaluaties werkbaar met elkaar te verbinden, om zo in elke specifieke context een werkbaar evaluatiearrangement te creëren. De specifieke situatie waarin de lerende evaluatie moet worden uitgevoerd bepaalt in sterke mate welke precieze vorm het aanneemt, welke elementen van de klassieke, constructivistische, responsieve en participatieve evaluatie het opneemt. In de praktijk van alledag ontstaan dus verschillende configuraties van lerend evalueren.

Een aantal elementen zijn echter cruciaal in elk lerend evaluatietraject. Dat is allereerst een open start waarbij evaluatoren en geëvalueerden samen de evaluatiedoelen formuleren en op zoek gaan naar de eigenlijke ambities achter het te evalueren project. Tijdens de rit is de frequente interactie tussen beide partijen van groot belang: niet alleen om tussentijdse resultaten terug te koppelen, maar ook om met elkaar ‘on speaking terms’ te blijven. Bij het implementeren van tussentijdse veranderingen spelen de evaluatoren een belangrijke coachende en adviserende rol. Na afronding van het evaluatieonderzoek verdient de verslaglegging bijzondere aandacht. Ook hier is interactie cruciaal, maar moet de opdrachtgever wel de souplesse hebben om de onafhankelijkheid van de evaluator te waarborgen.

Enkele reflexieve woorden tot slot

De lerende evaluatie kan opgevat worden als een hybride evaluatievorm die elementen van verschillende evaluatievormen in zich draagt. Het bevat elementen van een rationele evaluatie (Provus, 1971), omdat zij focust op de vraag of doelstellingen daadwerkelijk zijn gerealiseerd. Daarnaast heeft het overeenkomsten met een constructivistische evaluatie (Fischer en Forester, 1993), omdat zij het beleidsproces benadert als het resultaat van een proces van vormgeving en zingeving. Bovendien sluit zij aan bij 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. Tenslotte vertoont zij overeenkomsten met een op doorwerking georiënteerde evaluatie (Patton, 1997), omdat zij streeft naar het maximaliseren van de bruikbaarheid van haar conclusies, adviezen en aanbevelingen voor de beleidspraktijk.

Maar de lerende evaluatie heeft vooral meerwaarde, omdat zij het leervermogen en het adaptatievermogen van de betrokken actoren tracht te waarderen. Daarnaast wordt zij niet voorafgaand of na afloop van een beleidsprogramma uitgevoerd, maar tijdens de rit. Dit samenstel van uiteenlopende karakteristieken maakt de lerende evaluatie een bruikbare maar tevens spannende en spanningsvolle vorm van evalueren.

Referenties

- Abma, T. 1996. *Responsief evalueren. Discoursen, controversen en allianties in het postmoderne*. Delft: Eburon.
- Alkin, M.C. 1990. *Debates on Evaluation*, London: Sage.
- Edelenbos, J. en M. van Eeten. 2001. The Missing Link. Processing Variation in Dialogical Evaluation. *Evaluation* 7 (2):211-217.
- Edelenbos, J., M.W. van Buuren, M. van As, M. Duijn, A. Slob, en W. Vermeulen. 2003. *Milieubeleid meer van, voor en door de burger. Resultaten van de lerende evaluatie van het Stimuleringsprogramma Burger en Milieubeleid*. Erasmus Universiteit Rotterdam.
- Edelenbos, J. en M.W. van Buuren. 2005. Learning Evaluation. A theoretical and empirical exploration, in: *Evaluation Review (te verschijnen)*.
- Feltmann, C.E. 1984. *Adviseren bij organiseren: een studie over interventiekunde ten behoeve van organisatie-ontwikkeling en maatschappij-georiënteerd organiseren ("vermaatschappelijking")*. Amsterdam: De Perscombinatie.
- Fischer, F. en J. Forester, eds. 1993. *The Argumentative Turn in Policy Analysis and Planning*, Durham/London: Duke University Press.
- Guba, E. en Y. Lincoln. 1989. *Fourth Generation Evaluation*. Newbury Park: Sage.
- House, E.R. en K.R. Howe. 2000. Deliberative democratic evaluation. *New Directions for Evaluation* 85:3-12.
- Otten, M. 2000. *Verstrikt in grote projecten. Hoe de stadshuizen in Amsterdam en Apeldoorn totstandkwamen*. Den Haag: VNG Uitgeverij.
- Patton, M.Q. 1997. *Utilization-focused Evaluation. The next century text*. 3th edition. Thousand Oaks: Sage Publications.
- Provus, M.N. 1971. *Discrepancy evaluation*. Berkeley, CA: McCutcheon.
- Schön, D.A. 1983. *The reflective practitioner: how professionals think in action*, New York: Basic Books.
- Schwandt, T.A. 1984. *An examination of alternative models for socio-behavioral inquiry*. Ph.D. diss., Indiana University.
- Scriven, M. 1991. *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Stake, R.E. 1983. Program evaluation, particularly responsive evaluation. In: G.F. Madaus, M. Scriven en D.L. Stufflebeam (red.) *Evaluation models*. Boston: Kluwer-Nijhoff, pp. 287-310.
- Stufflebeam, D.L. 2001. Evaluation Models. *New Directions for Evaluation* 89: 8-98.

- Teisman, G.R. 1992. *Complexe besluitvorming: een pluricentrisch perspectief op besluitvorming over ruimtelijke investeringen*. Den Haag: VUGA.
- Teisman, G.R., F.B van der Meer, E.H. Klijn, J. Edelenbos, H.L. Klaassen en M.A. Reudink. 2002. *Evalueren om te leren. Naar een evaluatiearrangement voor de Vijfde Nota RO*. Erasmus Universiteit Rotterdam.
- Tweede Kamer, 2001-2002. Motie Feenstra, Klein Molenkamp en Augusteijn-Esser, 28 000 XI: 21.
- Twist, M.J.W. van, J. Edelenbos en M. van der Broek. 1998. In dilemma's durven denken, in: *M&O. Tijdschrift voor Management en Organisatie*, nr.5, pp.7-23.
- Wadsworth, Y. 2001. Becoming responsive – and some consequences for evaluation as dialogue across distance. *New Directions for Evaluation* 92:45-58.
- Weiss, C.H. 1998. *Evaluation*. 2^e editie. Upper Saddle River: Prentice Hall.