

Evalueren in interactie

De mogelijkheden van lerende evaluaties
voor het Milieu- en Natuurplanbureau

W. Kuindersma
F.G. Boonstra
S. De Boer
A.L. Gerritsen
M. Pleijte
T.A. Selnes

r a p p o r t e n

WOT
Wettelijke Onderzoekstaken Natuur & Milieu

Evalueren in interactie

De inhoudelijke kwaliteit van dit rapport is beoordeeld door Melchert Reudink, Milieu- en Natuurplanbureau.

Het rapport is geaccepteerd door Birgit Elands en Florence van den Bosch, opdrachtgevers namens de unit Wettelijke Onderzoekstaken Natuur & Milieu.

De reeks 'Rapporten' bevat onderzoeksresultaten van uitvoerende organisaties die voor de unit Wettelijke Onderzoekstaken Natuur & Milieu opdrachten hebben uitgevoerd.

WOT-rapport 19 is het resultaat van een onderzoeksopdracht van het Milieu- en Natuurplanbureau (MNP) en de WOT Natuur & Milieu aan Alterra en LEI. Dit onderzoeksrapport draagt bij aan de kennis die verwerkt wordt in meer beleidsgerichte publicaties zoals de Natuurbalans, (thematische) verkenningen en quick scans. Het rapport is geen MNP-product.

Evaluëren in interactie

De mogelijkheden van lerende evaluaties
voor het Milieu- en Natuurplanbureau

W. Kuindersma

F.G. Boonstra

S. De Boer

A.L. Gerritsen

M. Pleijte

T.A. Selnes

Rapport 19

Wettelijke Onderzoekstaken Natuur & Milieu

Wageningen, mei 2006

Referaat

Kuindersma, W., F.G. Boonstra, S. de Boer, A.L. Gerritsen, M. Pleijte & T.A. Selnes, 2006. *Evalueren in interactie. De mogelijkheden van lerende evaluaties voor het Milieu en Natuurplanbureau*, Wageningen, Wettelijke Onderzoekstaken Natuur & Milieu, WOt-rapport 19. 88 blz. 3 fig.; 1 tab. 52 ref.; 1 bijl.

Dit rapport bevat een verslag van een inventariserende studie naar de mogelijkheden van lerende evaluaties voor het Milieu- en Natuurplanbureau (MNP). Lerende evaluaties zijn evaluaties waarin zowel verantwoordens als leren centraal staan. We concluderen dat er vele verschijningsvormen van lerende evaluaties zijn. Met een omgevingsanalyse kunnen evaluatieonderzoekers een goede keuze daarbinnen maken. Ook in de evaluatiewetenschap en in vergelijkbare organisaties als het MNP is men bezig met het experimenteren met lerende vormen van evaluaties. Het MNP zelf gaat nog grotendeels uit van klassieke beleidsevaluatie en heeft nog relatief weinig lerende elementen in haar evaluaties ingebouwd. Deze studie beveelt aan om de mogelijkheden van lerende evaluaties binnen het MNP nader te onderzoeken en er in de praktijk daadwerkelijk mee te gaan experimenteren.

Trefwoorden: Beleidsevaluatie, evaluatie, bestuurskunde, beleidswetenschappen, governance.

Abstract

Kuindersma, W., F.G. Boonstra, S. de Boer, A.L. Gerritsen, M. Pleijte & T.A. Selnes, 2006. *Interactive evaluations: possibilities of learning-oriented evaluation at the Netherlands Environmental Assessment Agency*. Wageningen, Statutory Research Tasks Unit for Nature and the Environment. WOt-rapport 19. 88 p. 3 Figs.; 1 Tab. 52 Refs.; 1 annex

The report presents the findings of a survey study into the possibilities of learning-oriented evaluation at the Netherlands Environmental Assessment Agency (MNP). Learning-oriented evaluation is a form of evaluation focusing on both accountability and learning. We conclude that learning-oriented evaluation takes many forms, and that evaluation experts should analyse the situation to allow the right form of evaluation to be selected. Experts on the theory of evaluation and institutions comparable to MNP are also experimenting with learning-oriented evaluation. MNP itself is still largely using the traditional forms of policy evaluation and has so far included relatively few learning-oriented elements in its evaluations. We recommend further studies into the possibilities of learning-oriented evaluation at MNP and experiments with such options in actual evaluations.

Key words: policy evaluation, evaluation, public administration studies, policy studies, governance.

ISSN 1871-028X

©2006 **Alterra**

Postbus 47, 6700 AA Wageningen.

Tel: (0317) 47 47 00; fax: (0317) 41 90 00; e-mail: info.alterra@wur.nl

Landbouw-Economisch Instituut

Postbus 29703, 2502 LS Den Haag

Tel: (070) 335 83 30; fax: (070) 361 56 24; e-mail: informatie.lei@wur.nl

De reeks 'WOt-Rapporten' is een uitgave van de unit Wettelijke Onderzoekstaken Natuur & Milieu, onderdeel van Wageningen UR. Dit rapport is verkrijgbaar bij het secretariaat. Het rapport is ook te downloaden via www.wotnatuurenmilieu.wur.nl.

Wettelijke Onderzoekstaken Natuur & Milieu Postbus 47, 6700 AA Wageningen

Tel: (0317) 47 78 44; Fax: (0317) 42 49 88; e-mail: info.wnm@wur.nl; Internet: www.wotnatuurenmilieu.wur.nl

Alle rechten voorbehouden. Niets uit deze uitgave mag worden veeelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever. De uitgever aanvaardt geen aansprakelijkheid voor eventuele schade voortvloeiend uit het gebruik van de resultaten van dit onderzoek of de toepassing van de adviezen.

Inhoud

Samenvatting	7
Summary	9
1 Inleiding	11
1.1 Inleiding	11
1.2 Methoden van beleidsevaluatie onder de loep	11
1.3 Probleemstelling	15
1.4 Methode	16
1.5 Leeswijzer	17
2 Analytisch kader voor lerende evaluaties	19
2.1 Inleiding	19
2.2 Voorbereiding evaluatieonderzoek	20
2.3 Ontwikkelen aanpak voor beantwoording onderzoeksvragen	26
2.4 Uitvoering evaluatieonderzoek	29
2.5 Evaluatie van de evaluatie	30
2.6 Tot slot	31
3 Op zoek naar lerende evaluaties	33
3.1 Inleiding	33
3.2 Lerende vormen van evaluatie	33
3.2.1 Pestman - De Algemene Rekenkamer	33
3.2.2 Dammers - Ruimtelijk Planbureau	34
3.2.3 Edelenbos - Erasmus Universiteit	36
3.2.4 Abma – Universiteit van Maastricht	37
3.3 Reflectie	39
4 Reflecties op MNP evaluaties	41
4.1 Inleiding	41
4.2 Evaluatiestudie 'Schuiven op zand'	42
4.2.1 Evaluatieonderzoek	42
4.2.2 Analyse van de evaluatie	43
4.2.3 Conclusie en reflectie	47
4.3 Evaluatiestudie 'Van aankoop naar beheer'	48
4.3.1 Evaluatieonderzoek	48
4.3.2 Analyse van de evaluatie	49
4.3.3 Conclusie en reflectie	54
4.4 Evaluatiestudie 'Bouwtekening Transitie'	55
4.4.1 Evaluatieonderzoek	55
4.4.2 Analyse van de evaluatie	57
4.4.3 Conclusie en reflectie	60
4.5 Conclusies	61
5 Verhalen uit de praktijk van lerende evaluaties	63
5.1 Inleiding	63

5.2	Ontwikkelen met kwaliteit	64
5.3	Evaluatie reconstructie zandgebieden	67
5.4	Waardevolle Cultuurlandschappen	70
5.5	Reflecties op evaluatieverhalen	73
6	Conclusies en aanbevelingen	77
6.1	Inleiding	77
6.2	Conclusies	77
6.3	Aanbevelingen	78
	Literatuur	81
Bijlage 1	Geïnterviewden in dit onderzoek	85

Samenvatting

Lerende evaluaties zijn evaluatievormen die een of meer elementen van klassieke beleidsevaluatie en responsieve evaluatie combineren. Het gaat dus om een mengvorm waarin verschillende combinaties mogelijk zijn. Het doel van deze studie is om de mogelijkheden van deze lerende evaluaties voor het Milieu- en Natuurplanbureau (MNP) te verkennen. Hiermee willen we de discussie over evaluatiemethoden binnen en buiten het MNP faciliteren en stimuleren.

In deze studie staan de volgende onderzoeksvragen centraal:

1. Welke evaluatievormen en methoden passen in een governance-omgeving, zijn gericht op leren, maar doen tegelijkertijd ook recht doet aan de behoefte binnen het MNP aan afrekenbaarheid?
2. Op welke manieren geven evaluatiewetenschappers buiten het MNP vorm aan lerende vormen van evaluatie en welke lessen zijn hieruit te trekken voor het MNP?
3. Welke methodische keuzes liggen ten grondslag aan bestaande MNP-evaluaties en in welke mate bevatten deze elementen van lerende evaluaties?
4. Op welke problemen stuiten evaluatie-onderzoekers in de praktijk bij het combineren van verantwoorden en leren in evaluaties?

Het overzicht van lerende evaluaties is gebaseerd op literatuurstudie. Het vormgeven van een lerende beleidsevaluatie is geen invuloefening. Het gaat om een tussenvorm tussen een klassieke beleidsevaluatie en een evaluatie puur gericht op leren. Er zijn dus talrijke tussenvormen denkbaar met meer of minder accent op leren dan wel op verantwoorden. De evaluatieonderzoeker moet dus keuzes maken. Daarbij is het belangrijk om hieraan vooral veel aandacht te besteden in de voorbereidingsfase. Voordat voor een bepaalde vorm gekozen wordt, zal de onderzoeker zich eerst allerlei vragen moeten stellen en bij voorkeur ook het beleidsveld goed in beeld brengen. Op deze manier kan worden bepaald in hoeverre het beleidsveld wordt gekenmerkt door kenmerken van governance. Op dat moment kan het zinvol zijn om meer lerende elementen in de evaluatie (en de vraagstelling) in te brengen.

Op basis van interviews met vier evaluatiewetenschappers is (een beperkt) beeld geschetst van externe ervaringen met lerende evaluaties. Het blijkt dat de zoektocht naar lerende vormen van beleidsevaluatie niet exclusief een vraagstuk van het MNP is. Ook andere evaluatiewetenschappers en vergelijkbare instituties als de Algemene Rekenkamer en het Ruimtelijk Planbureau (RPB) zijn op zoek naar dergelijke evaluatievormen. Wel blijkt ook hier dat binnen dit spectrum heel veel verschillende keuzes zijn te maken. Zo blijft de Algemene Rekenkamer vanwege haar formele rol erg op verantwoorden zitten. Zij proberen voorzichtig meer lerende elementen in haar evaluaties te betrekken. Vooral de behoefte aan verklaringen en het (multi-level) governance-achtige karakter van beleid zijn stimulerende factoren. Iemand als Abma kiest helemaal voor leren. Verantwoorden speelt in haar evaluaties nauwelijks nog een rol. Daartussen bevinden zich vele mengvormen, zoals Edelenbos (EUR) en Dammers (RPB) laten zien.

Vervolgens is een drietal MNP-evaluaties geanalyseerd op aanwezige elementen en mogelijkheden voor lerende evaluaties. Ingegaan is op de MNP-evaluaties van de Reconstructie, de beleidsomslag van verwerven naar beheer en de evaluatie van transities. De evaluaties zijn geanalyseerd met behulp van het kader van deze studie en de projectleiders zijn geïnterviewd. Uit de analyse blijkt dat het MNP sterk de nadruk legt op verantwoorden en

uitgaat van een governmentcontext. Het MNP is huiverig om andere actoren dan de rijksoverheid een belangrijke rol te geven in het vormgeven van de evaluatie. Toch zijn er vaak wel contacten met andere actoren, maar deze hebben meestal een beperkte en hoogstens consulterende rol. Het gevolg is wel dat er een soort standaardaanpak van beleidsevaluatie lijkt te bestaan. Veel van de mogelijke vragen die gesteld kunnen worden voordat een evaluatie start, worden feitelijk niet gesteld. Het is de vraag in hoeverre er daadwerkelijk behoefte is aan nieuwe vormen van beleidsevaluatie binnen het MNP. Ook kun je je afvragen in hoeverre opdrachtgevers een dergelijke evaluatie van het MNP verwachten. De evaluatie van transities vormt een mogelijkheid voor het MNP om wel met lerende beleidsevaluatie te experimenteren. Tot nu toe gaat het echter om een 'droogzwem oefening'. Of het MNP bij de daadwerkelijke evaluatie van het transitiebeleid voor de voorgestelde lerende vorm van evaluatie gaat kiezen, is nog onzeker.

Vervolgens is aan de hand van verhalen van Alterra-onderzoekers een aantal ervaringen met lerende evaluaties in beeld gebracht. Ingegaan is op de ervaringen met de evaluaties van het beleid voor de Waardevolle Cultuurlandschappen (WCL's), de evaluatie Reconstructie en de evaluatie van het beleidsconcept Ontwikkelen met kwaliteit. De ervaringen geven aan dat het werken met dergelijke evaluaties nieuwe dilemma's oproept voor onderzoekers en ook nieuwe vaardigheden vragen van onderzoekers. In de verhalen komen de volgende punten naar voren:

- De onderzoeker kan door zijn interacties het beleid gaan beïnvloeden. In hoeverre spoort dit met het beeld van de onderzoeker als onpartijdige buitenstaander?
- Lerende evaluaties vragen om betrokkenheid van de belangrijke actoren in het beleidsveld. Dit vraagt om specifieke procesvaardigheden van de onderzoekers.
- Veel interactie in het onderzoek kan leiden tot draagvlak voor het onderzoek (en mogelijk tot beleidsverbetering). Toch zal de onderzoeker uiteindelijk kritisch moeten zijn. Dit kan tot spanning leiden omdat verantwoording kan leiden tot afrekenen. Hierdoor verdwijnt vaak de nuance uit het oordeel en bestaat het gevaar dat betrokken partijen zich niet meer in de conclusies herkennen.

Op grond hiervan bevelen we aan om:

1. De effecten van MNP-evaluaties op beleid en de kwaliteit van de aanbevelingen te onderzoeken.
2. Een strategisch debat binnen het MNP te voeren over de mogelijkheden van lerende beleidsevaluaties (samen met externe deskundigen).
3. Bij belangrijke beleidsevaluaties is het verstandig om vooraf het beleidsveld te onderzoeken op governancekenmerken, zodat beter beargumenteerd kan worden of een klassieke beleidsevaluatie of een lerende beleidsevaluatie nodig is.
4. Te experimenteren met lerende vormen van evaluatie in de MNP-praktijk (bijvoorbeeld bij de evaluatie van het transitiebeleid).

Summary

Learning-oriented evaluation is a form of evaluation that combines one or more elements of classical policy evaluation and responsive evaluation. It is a hybrid form of evaluation that may consist of various combinations. The present study explored the opportunities that learning-oriented evaluation offers for the Netherlands Environmental Assessment Agency (MNP), in order to facilitate and stimulate the debate on evaluation methods within and outside MNP.

The study focused on the following questions:

1. What forms and methods of evaluation are suitable for a governance situation, are learning-oriented and at the same time meet the MNP's accountability requirements?
2. What forms of learning-oriented evaluation have been developed by experts on evaluation theory outside MNP and how can MNP staff learn from these developments?
3. What are the methodological choices underlying the existing MNP evaluations and to what extent do these methods feature elements of learning-oriented evaluation?
4. What practical problems are evaluation experts confronted with when combining accountability and learning aspects in evaluations?

The report offers a survey of learning-oriented evaluations, based on a study of the available literature. Developing a learning-oriented evaluation study is not a matter of simply filling in an existing template. What needs to be developed is a hybrid between classical policy evaluation and a form of evaluation purely focused on learning. Such a hybrid can take many different forms, depending on whether the emphasis is mostly on learning or on accountability. Hence, developing evaluations means making choices, particularly during the preparatory stages. Before deciding on a particular form, a researcher will need to ask various questions, and should preferably explore the policy area in detail. The results of such an exploration will show to what extent the policy area at stake is characterised by features of governance. If that is the case, it may be useful to include more aspects of learning in the evaluation (and the problem definition).

We have used interviews with four experts on evaluation theory to produce a limited overview of experience gained with learning-oriented evaluation outside MNP. This has shown that the search for learning-oriented forms of policy evaluation is by no means limited to MNP's situation. Other experts on evaluation theory, and comparable institutions like the Netherlands Court of Audit and the Netherlands Institute for Spatial Research are also looking for such forms of evaluation. The interviews also showed, however, that many different choices can be made within the available spectrum of forms of evaluation. Due to its formal role, the Netherlands Court of Audit tends to emphasise accountability and is cautiously trying to introduce more learning-oriented elements into its evaluations. Factors stimulating the inclusion of learning-oriented elements are particularly the need for explanations and the (multi-level) governance nature of policy.

We have analysed three evaluations implemented at MNP in terms of their constituent elements and the opportunities they offer for learning-oriented evaluation. We analysed the evaluation of the Reconstruction programme¹, that of the recent Dutch policy shift from land acquisition for nature conservation towards countryside stewardship by existing landowners,

¹ This programme aims to restructure the sandy soil areas and is an example of integrated rural policy with special attention on the environmental impact and economical position of intensive livestock farming.

and the evaluation of transition policy² The evaluations were analysed using the framework developed for the present study, and project managers were interviewed. The analysis showed that MNP strongly emphasises accountability and bases its evaluations on the assumption of a government-type rather than a governance-type context. The MNP hesitates to give other actors than the national government a major say in the form of evaluation to be used. Although there are contacts with other actors, they usually play a limited and at best consultative role. As a result, there appears to be a more or less standardised approach to policy evaluation. Many of the questions that could be asked before the start of an evaluation are not being asked. It is unclear whether there is a real need for new forms of policy evaluation at MNP, or whether clients expect such new forms of evaluation. Although the evaluation of transition policy offers a good opportunity for MNP to experiment with learning-oriented evaluation, such experiments have so far been limited to 'dry runs'. It remains uncertain whether MNP will be applying the proposed form of learning-oriented evaluation in the actual evaluation of the transition policy.

We have also interviewed a number of researchers at the Alterra research institute to assess their experience with learning-oriented evaluation. This related to evaluations of (1) the policy on valuable man-made landscapes, (2) the Reconstruction programme and (3) the policy concept of high-quality landscape development. The interviews showed that working with such new forms of evaluation creates new dilemmas for researchers and requires them to learn new skills. The following elements emerged from their accounts:

- Learning-oriented forms of evaluation allows researchers to influence the policy-making process. To what extent does this conflict with the concept of researchers as independent experts?
- Learning-oriented evaluation implies involving the main actors of the relevant policy area in the evaluation process, which means that the researchers must possess specific process skills.
- A large interactive component in a policy evaluation can result in wider support for the evaluation (and potentially also in improved policymaking). At the same time, however, researchers will have to retain a critical attitude. This can cause tensions, as accountability can lead to judgements. This may remove certain nuances from the assessment and give the parties involved the feeling that the conclusions no longer reflects their views.

On the basis of our findings, we have formulated the following recommendations:

1. The effects of MNP's evaluations on the policy-making process and on the quality of the ensuing recommendations should be studied.
2. A strategic debate should be initiated within MNP (in consultation with external experts) on the possibilities of learning-oriented policy evaluations.
3. Major policy evaluations should start by an examination of the policy area involved to identify the presence of governance features, so as to inform the decision on whether to use classical policy evaluation or learning-oriented evaluation.
4. Forms of learning-oriented evaluation should be experimented with in MNP practice (for instance in the evaluation of transition policies).

² Transitions are structural, technological and social changes, which may often take considerable time to materialize. In the 4th National Environmental Policy Plan the management of transitions towards sustainable development is part of the governments' strategy to solve the major environmental problems.

1 Inleiding

1.1 Inleiding

Het Natuurplanbureau³ bestaat sinds 1998 en evalueert sindsdien het natuurbeleid van de rijksoverheid. Het nadenken over beleidsevaluatie binnen het natuurplanbureau gaat al terug naar deze beginjaren. In 1999 is er bijvoorbeeld een handleiding beleidsevaluatie voor medewerkers aan de Natuurbalans ontwikkeld (Kuindersma, 1999) en is er een overzicht gemaakt van mogelijke alternatieve evaluatiemethoden (Kuindersma en Pleijte, 1999). Mede onder invloed van de discussies binnen het Natuurplanbureau over governance in het natuurbeleid (zie bijv. Van der Zouwen en Van Tatenhove, 2002; Kuindersma (red), 2002), is de vraag ontstaan of de huidige methoden van beleidsevaluatie nog voldoen. In 2005 is een overzicht gegeven van een breed spectrum aan evaluatiemethoden (Kuindersma en Boonstra, 2005) Dit loopt van beleidsevaluaties die puur gericht zijn op de formele doelen van het beleid tot methoden waarbij het proces of het gemeenschappelijke belang van alle betrokken partijen centraal staat. Dit onderzoek bouwt voort op de resultaten van dit rapport uit 2005.

In dit hoofdstuk schetsen we eerst de hoofdlijnen uit de verkennende studie uit 2005 (par 1.2). Daarna gaan we in op de probleemstelling van dit rapport (par. 1.3). In paragraaf 1.4 schetsen we de onderzoeksmethodieken die in deze studie zijn toegepast en paragraaf 1.5 bevat een leeswijzer.

1.2 Methoden van beleidsevaluatie onder de loep

In de studie 'Methoden van beleidsevaluatie onder de loep' (Kuindersma en Boonstra, 2005) is een breed scala aan mogelijke evaluatievormen geschetst. Het doel van deze studie was om een bijdrage te leveren aan de discussie binnen het Milieu- en Natuurplanbureau (MNP) over nieuwe vormen van evaluatie. Deze nieuwe evaluatievormen moeten meer recht doen aan ontwikkelingen in beleid en bestuur, zoals de betrokkenheid van steeds meer (bestuurlijke) lagen (multi-level governance) en de betrokkenheid van steeds meer (maatschappelijke en bestuurlijke) actoren bij besluitvorming (multi-actor governance). De veronderstelling is dat de behoefte aan deze evaluatievormen toeneemt naarmate het beleid en de samenleving veranderd.

De bruikbaarheid van een evaluatievorm is bepaald aan de hand van twee (deels tegengestelde) bestuurlijke trends waar het MNP mee te maken heeft, namelijk de trend naar governance (meer partijen die betrokken zijn bij besluitvorming) en de trend naar afrekenbaarheid in beleid.

Governance als bestuurlijke trend gaat over een algemene trend in het openbaar bestuur die kan worden aangeduid met de ontwikkeling van 'government' naar 'governance' (Kuindersma red., 2002). Government duidt op de situatie dat er een centrale overheid is die stuurt. Governance is dan 'een op samenwerking gerichte stijl van regeren, waarin overheidsinstellingen en niet-overheidsinstellingen deelnemen aan gemengde openbare en

³ Het Natuurplanbureau maakt sinds kort deel uit van het Milieu- en Natuurplanbureau (MNP). Om deze reden spreken we in dit rapport over het Milieu- en Natuurplanbureau.

private netwerken' (Van der Zouwen en Tatenhove, 2002: 24). Deze trend duidt daarmee op de ontwikkeling dat de besluitvorming over beleid steeds meer plaatsvindt buiten de formele kaders van de natiestaat, en verschuift naar interacties tussen staat, civil society en bedrijfsleven.

Deze trend bestaat uit twee aspecten: multi-actor governance en multi-level governance. Multi-actor governance wijst in feite op de toenemende horizontalisering van het bestuur van de samenleving. Met andere woorden besluitvorming over publieke zaken vindt steeds meer plaats door overheden in samenwerking met burgers, maatschappelijke organisaties en bedrijven. Het impliceert niet alleen een andere manier van sturen vanuit de overheid, maar ook dat de vervlechting tussen overheid, civil society en de markt (of het bedrijfsleven) toeneemt. Of zoals Pestman en Van Tatenhove (1998: 256) het zeggen: "In hedendaagse samenlevingen kunnen de subsystemen markt, staat en civil society niet meer als gescheiden werelden worden beschouwd, met eigen logica's (concurrentie, hiërarchische controle en solidariteit), maar is er sprake van een wederzijds doordringen van deze subsystemen" (zie ook: Giddens, 1990).

Multi-level governance legt de nadruk op de ontwikkeling dat besluitvorming ook steeds meer op verschillende niveaus plaatsvindt. Daarbij gaat het niet alleen om de verschillende bestuurslagen op overheidsniveau, zoals gemeenten, provincies, nationale overheid, Europese Unie, Verenigde Naties. Nadrukkelijk gaat het ook op al deze niveaus weer om de interacties van overheden met maatschappelijke organisaties en bedrijven.

Naast de trend naar governance, is nog een andere trend te onderscheiden in het openbaar bestuur. Dat is de trend naar afrekenbaarheid. Deze ontwikkeling is te typeren als: "...a still increasing emphasis on transparency, measurable results and accountability. Policy documents should specify clear goals, the attainment of which should be measured by unequivocal (and if possible quantitative) indicators. Policymakers should be hold accountable for the result thus assessed" (Van der Meer en Edelenbos, 2002: 3). Op rijksniveau heeft deze ontwikkeling zich vertaald in de VBTB-systematiek van de rijksbegroting. VBTB staat hierbij voor het project 'Van Beleidsbegroting Tot Beleidsverantwoording' en is erop gericht om de te bereiken resultaten van rijksbeleid beter meetbaar te maken en te koppelen aan het bestede geld. Ook in de relatie tussen overheden is er een impuls tot het benoemen van afrekenbare resultaten, bijvoorbeeld in het kader van het ILG (Investeringsbudget Landelijk Gebied) waarin het rijk concrete afspraken wil maken met de provincies over te bereiken doelstellingen in het landelijk gebied.

Deze ontwikkeling lijkt op gespannen voet te staan met de eerder genoemde ontwikkeling van government naar governance. Zo noemen Van der Meer en Edelenbos (2002) een aantal problemen, dat ontstaat met deze 'accountability' in beleid waarin governanceprocessen steeds belangrijker worden:

- Het multi-actor en interactieve karakter van veel beleidsprocessen zorgt ervoor dat niet precies meer duidelijk is wie waarvoor verantwoordelijk is;
- Het aantonen van een verband tussen beleid en maatschappelijke processen is vanwege de complexiteit, ambiguïteit en interafhankelijkheid lastig te bepalen;
- Veel beleidsprocessen hebben geen duidelijke doelstellingen als uitgangspunt en zijn meer 'goal seeking' dan 'goal driven'.

Er bestaat dus geen eenduidigheid over de aanwezige bestuurlijke trends. De beweging van government naar governance is een duidelijke ontwikkeling. Hierin staat centraal dat meerdere partijen voortaan betrokken zijn bij beleidsvorming en beleidsuitvoering. Binnen de (rijks)overheid vindt tegelijkertijd ook een ontwikkeling plaats naar afrekenen op resultaten van

beleid. Tussen deze trends bestaan spanningen, maar vooralsnog gaan we ervan uit dat beide trends naast elkaar bestaan.

In de studie 'Methoden van beleidsevaluatie onder de loep' is verder een aantal evaluatiemethoden geschetst en is gezocht naar evaluatievormen vanuit verschillende wetenschappelijke discoursen. Onderscheiden zijn de volgende wetenschappelijke discoursen:

1. Systeemanalytisch discours
2. Kritisch-theoretisch discours
3. Sociaalconstructivistische discours

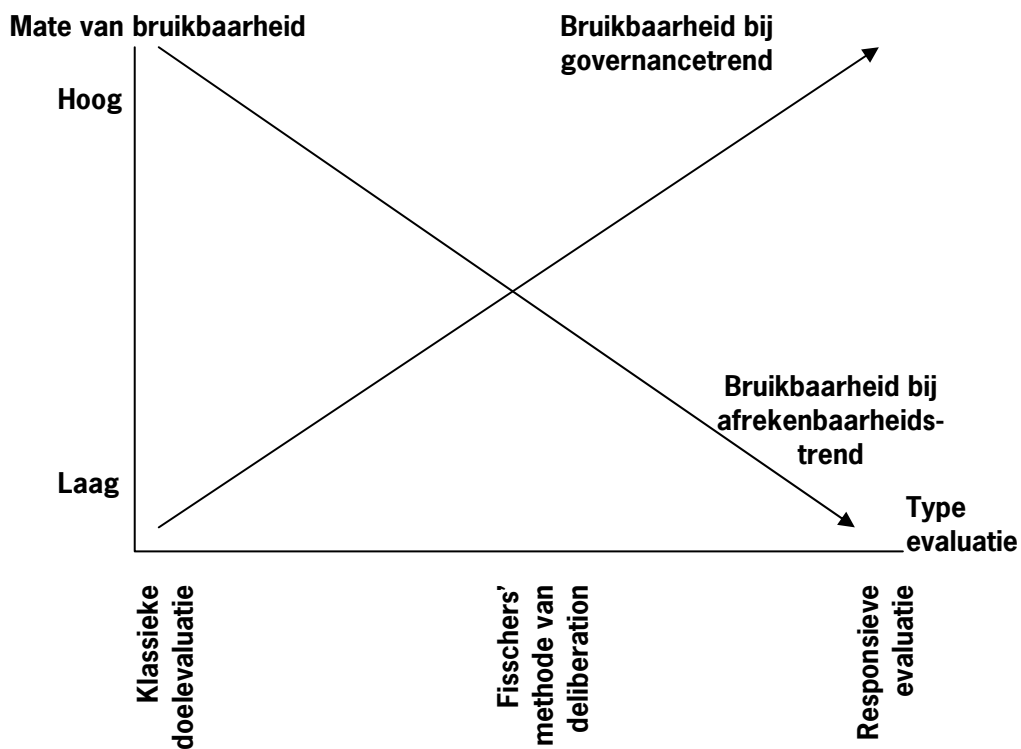
In het systeemanalytisch discours gaat de evaluator uit van de doelen van het beleid (zie bijv. Bressers en Hoogerwerf, 1991). Hij stelt deze doelen niet ter discussie, maar bepaalt in welke mate deze doelen gehaald zijn en (eventueel) wat de redenen zijn voor het (gedeeltelijk) niet behalen van de doelstellingen. Evaluatiemethoden binnen dit discours passen heel goed in de trend naar afrekenbaarheid en sluiten ook goed aan bij de wettelijke rol van het MNP als de evaluator van het nationale milieu- en natuurbeleid. Het is dan ook de belangrijkste evaluatiemethode van het MNP. Binnen deze methode zijn ook vormen van beleidsevaluatie waarin meer aandacht is voor de doelen van andere partijen of de beschrijving van het proces. Indien de nadruk in het beleidsproces en de achterliggende sturingsfilosofie ligt op het doelzoekende karakter van het beleid en de doelen van meerdere betrokken organisaties en bestuurslagen (multi-actor en multi-level governance), zijn deze evaluatiemethoden minder bruikbaar.

Aan de andere kant van het evaluatiespectrum vinden we het sociaalconstructivistische discours (zie bv. Abma, 1996; Guba en Lincoln, 1989; Fortuin, 1994). Dit discours vormt een breuk met de unieke dan wel speciale positie die aan de formele overheidsdoelen wordt toegekend. Deze benadering gaat uit van het gegeven dat er geen echte objectieve waarheden of doelstellingen zijn. Alles is een sociale constructie. Dit levert hoogstens een door een groep gedeelde of aangenomen intersubjectieve waarheid op. Het sociaalconstructivisme als wetenschappelijk discours omvat veel stromingen. Binnen enkele daarvan zijn methodische aanwijzingen voor beleidsevaluatie uitgewerkt waarvan wij er twee hebben besproken: de pluricentrische evaluatie en de responsieve evaluatie. Over het algemeen sluiten deze evaluatiemethoden goed aan op de governancetrend in beleid en bestuur. Binnen de trend naar afrekenbaarheid passen dergelijke benaderingen minder goed. Ook onderzoekers komen in deze benaderingen meer in een procesrol terecht. Een dergelijke methode kan afschrikken omdat onderzoekers in samenspraak met anderen tot een oordeel komen en dus niet meer als "neutrale" expert boven de partijen staan. Beleidsmakers krijgen met dergelijke evaluaties geen feiten maar interpretaties van feiten. Het is de vraag of beleidsmakers hierop zitten te wachten. Bovendien: accepteren ze een dergelijke benadering van het MNP dat volgens velen alleen geacht wordt met feiten te komen?

Het *kritisch-theoretische discours* heeft als belangrijkste kritiek op de systeemanalytische benadering dat de kritische houding ten opzichte van de formele beleidsdoelen ontbreekt (zie bv. Fischer, 1995). Evaluaties vanuit dit discours blijven vanuit de formele beleidsdoelen redeneren, maar voegen daar de normen, waarden en doelstellingen van andere betrokken partijen aan toe. Binnen dit discours hebben we de deliberatieve benadering van Fischer behandeld. Deze insteek is interessant omdat ze uitgaat van een analyse van de formele beleidsdoelen en daarmee bruikbaar is vanuit de afrekenbaarheidstrend. Daar voegt Fischer dan analyses aan toe waarin vragen worden gesteld over de bijdrage van het bereiken van deze doelen aan de oplossing van het maatschappelijke probleem, aan bredere maatschappelijke wensen en doelen en zelfs aan het maatschappelijk systeem. Dit laatste biedt mogelijkheden om ook doelen van andere actoren een rol te geven in de evaluatie.

Daarmee kan ook in zekere mate recht gedaan worden aan governance. Deze evaluatie biedt dus een soort midden tussen de sociaalconstructivistische benaderingen en de systeemanalytische methoden.

We constateren dat de omgeving van het MNP snel verandert. Trends naar meer afrekenbaarheid in beleid en governance doen zich in de praktijk beiden voor. Er zijn veel verschillende vormen van evaluatie met elk eigen wetenschappelijke uitgangspunten. Sommige methoden passen beter bij afrekenbaarheid en andere methoden beter bij governance. Een MNP dat in deze context wil evalueren zal meerdere beleidsevaluatiemethoden in haar gereedschapskist moeten hebben. Of het MNP zal moeten beseffen dat het een evaluatie van beleid in een governance omgeving niet zelf uit kan voeren en aan anderen moet overlaten. Maar ook indien het MNP geen evaluatiemethoden wil ontwikkelen die beter passen bij de governancetrend, zal men ook bij traditionele evaluaties met deze governance omgeving te maken krijgen. Hiertoe zullen in ieder geval evaluatiemethoden moeten worden ontwikkeld waarin combinaties worden gemaakt tussen methoden die meer geschikt zijn bij verantwoorden en bij governance. Om te kunnen bepalen welke evaluatiemethode in welk geval toegepast dient te worden, zou er voor uitvoering van een evaluatie een keuzemoment (evaluatietoets) moeten zijn. Op dit moment kan het MNP een analyse uitvoeren naar de vraag in hoeverre de evaluatie plaatsvindt in een context van afrekenbaarheid of van governance (of een mengvorm).



Uiteindelijk denken wij dat een MNP dat in staat is om meerdere evaluatiemethoden toe te passen, haar positie als onafhankelijke evaluator van het rijksnatuurbeleid kan versterken. Hiertoe dienen dan wel ook binnen het MNP de methodische kennis en vaardigheden in huis te zijn om vast te kunnen stellen in hoeverre andere evaluatiemethoden toegepast moeten of kunnen worden en om dergelijke evaluaties uit te voeren of aan te sturen.

1.3 Probleemstelling

Deze studie bouwt voort op de resultaten uit het genoemde rapport (Kuindersma en Boonstra, 2005). Het doel is echter nu niet zozeer het geven van overzichten, maar het daadwerkelijk operationaliseren van nieuwe evaluatiemethoden. Daarbij richten we ons op het evalueren van beleid met elementen van de klassieke evaluatie (verantwoorden/afrekenen) en de responsieve benadering (leren) (zie Tabel 1.1). Deze tussenvorm noemen we de lerende evaluatie. Lerende evaluaties zijn evaluatievormen die een of meerdere elementen van klassieke beleidsevaluatie en responsieve evaluatie combineren. Het gaat dus om een mengvorm waarin verschillende combinaties mogelijk zijn.

Tabel 1.1. Overzicht evaluatievormen

	Klassieke evaluatie	Lerende evaluatie	Responsieve evaluatie
Wetenschapsdiscours	Systeemanalytisch	↔	Sociaalconstructivisme
Doel	Afrekenen/verantwoording	↔	Leren (beleidsverbetering)
Visie op beleid	Government (monocentrisch)	↔	Governance (pluricentrisch)
Beoordelingskader	Formele beleidsdoelen van de (centrale) overheid	↔	Aanvankelijk geen, in interactie met stakeholders opgesteld
Rol evaluator	Afstandelijk, onafhankelijk, objectief	↔	Betrokken
Soort kennis/vaardigheden	Vakkennis, onderzoeksvaardigheden	↔	Procesvaardigheden/ procesfacilitator
Rol stakeholders	Passief, alleen informant	↔	Actief: leveren van input, opstellen beoordelingskader, gebruik resultaten

In dit project staan lerende vormen van beleidsevaluatie centraal. Daarmee ontkennen we niet het belang en/of de waarde van klassieke doelevaluaties of pure responsieve evaluaties. In pure governmentsituaties waarin verantwoording/afrekenen centraal staat, is een klassieke beleidsevaluatie nog steeds heel geschikt. En in situaties waarin het vooral gaat om leren en beleidsverbetering is de responsieve benadering heel geschikt. Wij beperken ons tot de lerende evaluatie omdat het MNP al veel ervaring heeft met de klassieke evaluaties en juist niet met meer lerende vormen van evaluatie. Puur responsieve evaluaties zijn wellicht minder geschikt vanwege de formele (afreken)rol van het MNP in het nationale natuur- en milieubeleid.

Het MNP heeft op dit moment nog nauwelijks ervaring met lerende evaluaties. Dit is een probleem omdat steeds meer beleid in een governance context wordt vorm gegeven en uitgewerkt. Indien het MNP in de evaluatie van dit soort beleid een rol wil spelen zal het andere vormen van beleidsevaluatie moeten ontwikkelen en ook zich meer bewust moeten worden in welke situatie, welke methode van beleidsevaluatie het meest bruikbaar is.

Het doel van het project is om het leren op het gebied van methoden van beleidsevaluatie en hun toepassing binnen het Milieu- en Natuurplanbureau te stimuleren. Dit is een nadrukkelijke

verbreding ten opzichte van de eerdere projecten die zich vooral op het Natuurplanbureau richtten (Kuindersma, 1999; Kuindersma en Pleijte, 1999 en Kuindersma en Boonstra, 2005). Het leren houdt voor ons in dat we een bijdrage leveren aan de uitbreiding van de methoden van beleidsevaluatie die beschikbaar zijn in het MNP (gereedchapskist) en de bewustwording binnen het MNP over de (on)mogelijkheden van deze nieuwe methoden van beleidsevaluatie.

In het project zijn we zowel op zoek gegaan naar de ervaringen met lerende evaluaties en hebben we ook de verschillende verschijningsvormen hiervan proberen te schetsen.

Hierbij hebben we de volgende vragen beantwoord:

1. Welke evaluatievormen en methoden passen in een governance-omgeving, zijn gericht op leren, maar doen tegelijkertijd ook recht doet aan de behoefte binnen het MNP aan afrekenbaarheid?
2. Op welke manieren geven evaluatiewetenschappers buiten het MNP vorm aan lerende vormen van evaluatie en welke lessen zijn hieruit te trekken voor het MNP?
3. Welke methodische keuzes liggen ten grondslag aan bestaande MNP-evaluaties en in welke mate bevatten deze elementen van lerende evaluaties?
4. Op welke problemen stuiten evaluatie-onderzoekers in de praktijk bij het combineren van verantwoorden en leren in evaluaties?

Hiermee willen we een bijdrage leveren aan de discussie binnen (en buiten) het MNP over beleidsevaluatie. Daartoe zijn ook tussenresultaten van dit project gepresenteerd tijdens het Platform Beleidsevaluatie van het MNP. Ook kan dit rapport een bijdrage leveren aan de (nieuwe) handreiking beleidsevaluatie van het MNP.

1.4 Methode

Dit onderzoeksrapport is opgebouwd uit verschillende onderdelen, die deels onafhankelijk van elkaar tot stand zijn gekomen. Elk hoofdstuk behandelt een van de onderzoeksvragen. In deze paragraaf geven we kort aan welke methode gebruikt is om de afzonderlijke onderzoeksvragen te beantwoorden.

Onderzoeksvraag 1 is beantwoord met behulp van literatuurstudie.

Bij onderzoeksvraag 2 hebben we gekozen voor het interviewen van een selectie van vier evaluatiewetenschappers, waarvan bekend is dat ze ervaring hebben met nieuwe vormen van beleidsevaluatie. Met hen zijn gesprekken gevoerd aan de hand van een beperkt aantal gespreksthema's. Tevens zijn enkele ondersteunende documenten en/of artikelen van hun hand bestudeerd.

Bij onderzoeksvraag 3 is gekozen voor een aantal cases in de vorm van bestaande MNP-evaluaties. Voor de selectie is het Platform Beleidsevaluatie binnen het MNP ingeschakeld. Zij hebben enkele bestaande evaluaties geselecteerd op basis van hun behoefte om reflectie te krijgen op deze MNP-evaluaties. Voorwaarde was wel dat de projectleider instemde met een kritische analyse en er ook zelf iets van wilde leren. Voor de analyse van deze studies is een vragenlijst opgesteld die is gebaseerd op het analytisch kader uit hoofdstuk 3. Er is steeds gebruik gemaakt van documentenstudie van op zijn minst de evaluatierapporten en een aanvullend gesprek met een betrokken MNP-onderzoeker.

Bij onderzoeksvraag 4 hebben we ervoor gekozen om gebruik te maken van de praktische ervaringen met lerende evaluatie binnen Alterra. Drie onderzoekers is gevraagd om te

reflecteren op hun eigen ervaring met lerende evaluatie. Voorwaarde was wel dat de evaluatie een voorbeeld moest zijn van een lerende evaluatie en dus een evaluatie met zowel verantwoordten als leren. De onderzoekers hebben een bepaald dilemma uit hun eigen evaluatiepraktijk geselecteerd en in een verhalende stijl uitgewerkt.

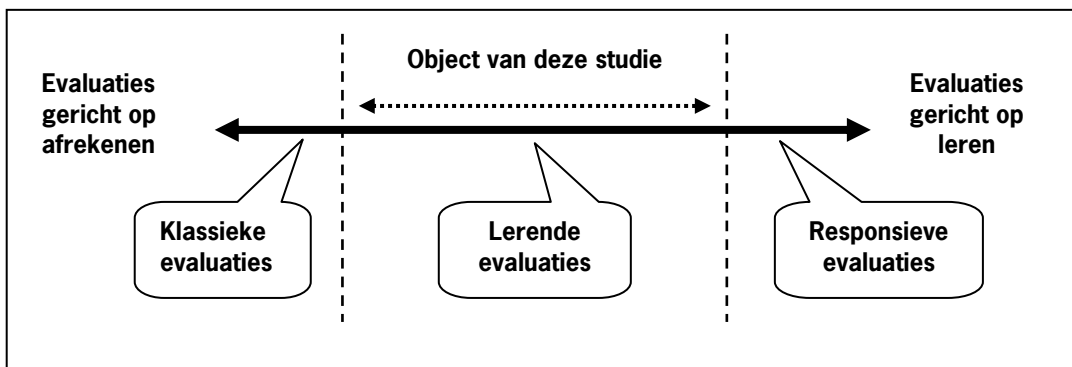
1.5 Leeswijzer

In elk van de volgende hoofdstukken behandelen we een van de vier genoemde vragen. Hoofdstuk 2 schetst een analytisch kader voor lerende evaluaties. Hoofdstuk 3 gaat in op ervaringen met lerende beleidsevaluatie buiten het MNP. In hoofdstuk 4 gebruiken we dit kader bij de reflectie op drie bestaande MNP-evaluaties. Daarbij gaan we vooral in op de gemaakte keuzes en de lerende elementen in deze evaluaties. In hoofdstuk 5 schetsen drie Alterra-onderzoekers in een verhalende vorm enkele ervaringen met beleidsevaluatie waarin zowel leren als verantwoordten centraal staan. In hoofdstuk 6 trekken we enkele algemene conclusies en doen we enkele aanbevelingen voor het MNP.

2 Analytisch kader voor lerende evaluaties

2.1 Inleiding

In het vorige hoofdstuk zijn we ingegaan op verschillende vormen van evaluatie, zoals deze zijn geschetst in een eerdere studie voor het MNP (Kuindersma en Boonstra, 2005). Er is onderscheid gemaakt tussen twee extremen: evaluaties gericht op verantwoorden (klassieke evaluaties) en evaluaties gericht op leren (responsieve evaluaties). Gesteld is dat er vele manieren zijn om deze ideaaltypische vormen van beleidsevaluatie te combineren en dat juist hierin een uitdaging ligt voor het MNP. Evaluaties die zowel gericht zijn op verantwoorden als op leren noemen we lerende evaluaties (Fig. 2.1).



Figuur 2.1: Positionering lerende evaluatie

Omdat er veel verschillende manieren zijn om leren en verantwoorden met elkaar te combineren, bestaat er niet één recept voor een lerende evaluatie. Elke specifieke evaluatievraag en –context vergt zijn eigen methodologische aanpak. Toch levert bestaand evaluatieonderzoek en literatuur over evaluatiemethodologieën wel rode draden op die benut kunnen worden bij het ontwerp van een evaluatieonderzoek. In dit hoofdstuk geven we per fase in het evaluatieproces aan welke vragen er voorliggen en werken we de opties die passen bij een lerende evaluatie verder uit. De centrale vraag luidt dan ook:

Welke evaluatievormen en methoden passen in een governance-omgeving, zijn gericht op leren, maar doen tegelijkertijd ook recht doet aan de behoefte binnen het MNP aan afrekenbaarheid?

Beantwoording van deze vraag beoogt een aanzet te leveren voor een analytisch kader voor lerende beleidsevaluatie. In concrete evaluatiestudies van het MNP kan het verder worden uitgewerkt en toegepast

In het navolgende gaan we achtereenvolgens in op de voorbereiding van evaluatieonderzoek (2.2), de ontwikkeling van een beoordelingskader en de daadwerkelijke beoordeling (2.3), de uitvoering van evaluatieonderzoek (2.4) en de kwaliteitscriteria die je op deze nieuwe vormen van evaluatieonderzoek kunt toepassen (2.5). De belangrijkste antwoorden op de centrale vraag vatten we samen in paragraaf 2.6.

2.2 Voorbereiding evaluatieonderzoek

Beleidsevaluatie is de beoordeling van waargenomen inhoud, processen of effecten van een beleid aan de hand van bepaalde criteria (Bressers en Hoogerwerf, 1991). Dit is een nogal algemene omschrijving. De exacte betekenis van het begrip “beleidsevaluatie” varieert met het antwoord op de vraag: Wie evalueert wat voor wie, wanneer, waarom en hoe? Zo zijn er externe en interne evaluaties, wetenschappelijke en gebruikgerichte evaluaties, ex post en ex ante evaluaties, lerende en afrekenende evaluaties en worden zowel kwantitatieve als kwalitatieve onderzoeksmethoden ingezet (Verhoef, 1992).

De wie, wat, wanneer, waarom en hoe-vraag is typisch een kwestie die in de voorbereiding van een evaluatieonderzoek aan bod komt, zij het dat de beantwoording niet definitief hoeft te zijn. Evalueren is een iteratief proces: een groeiend inzicht in het empirisch veld kan tot aanscherping van de methoden van onderzoek leiden, dit kan vervolgens de afbakening van de probleemstelling beïnvloeden enzovoorts enzoverder. Bovendien heeft evaluatieonderzoek in een governance-omgeving per definitie een explorerend karakter. De onderzoeksvraag en beleidsdoelen van de opdrachtgever zijn een belangrijk vertrekpunt, maar in de eerste fase van het onderzoek worden nadrukkelijk ook de doelen en verwachtingen van andere betrokkenen bij het te evalueren beleid betrokken. Dit kan leiden tot bijstellingen van de onderzoeksvraag, het voorwerp van evaluatie en de methode.

Voordat de wie, wat, wanneer, waarom en hoe-vraag in een samenspel tussen opdrachtgever en opdrachtnemer beantwoord kan worden moet de evaluator echter een globaal beeld hebben van inhoud en organisatie van het beleidsveld waarvan het te evalueren beleid deel uit maakt. Wat zijn de formele beleidsdoelen, hoe verhouden doelen en middelen zich tot elkaar, welke discussies en onzekerheden spelen er, wie zijn naast de opdrachtgever de belangrijkste stakeholders, op welk bestuurlijk niveau opereren zij, wat is hun onderlinge relatie, wat zijn hun standpunten ten aanzien van het te evalueren beleid, in hoeverre zijn zij potentiële gebruikers van de evaluatieresultaten enzovoort. Dit alles moet duidelijk maken in hoeverre sprake is van een governance-omgeving. Zoals gezegd in hoofdstuk 1 stelt een governance-omgeving namelijk specifieke eisen aan een evaluatieonderzoek en het te ontwikkelen evaluatiearrangement. Wanneer de evaluator zich hiervan in de beginfase niet bewust is loopt hij het risico zich teveel op de perceptie van de opdrachtgever te verlaten en kansen op bredere benutting van het evaluatieonderzoek mis te lopen.

In het navolgende gaan we in op de wijze van beantwoording van de wie, wat, waarom en wanneer-vraag in de voorbereidingsfase. De methodische hoe-vraag reserveren we voor paragraaf 2.3. Resultaat is een gedegen onderzoeksplan of evaluatiearrangement. Daarbij kan nog worden opgemerkt dat “voorbereiding” bij benuttinggericht evaluatieonderzoek in een governance-omgeving een enigszins misleidende aanduiding is voor deze fase. Met de verkenning van het beleidsveld, de onderlinge relaties en afhankelijkheden en de wensen en verwachtingen van de opdrachtgever en de stakeholders is het onderzoek eigenlijk al begonnen.

Wat is de beleidsomgeving?

Eerst moet de evaluator vaststellen of en in welke mate het beleid deel uit maakt van een governance-context. “Governance” is een op samenwerking gerichte vorm van beleidsvoering waarbij vertegenwoordigers van overheid, markt en civiele samenleving deelnemen aan gemengde openbare en private netwerken. Naast een nieuwe vorm van overheidssturing verwijst de term naar een toenemende verflechting tussen overheid, markt en civiele samenleving (Van der Zouwen en Van Tatenhove, 2002).

Daarnaast moet de evaluator zich bewust zijn van het feit dat hij/zij deel uit maakt van een politieke omgeving. Formulering van doel- en vraagstelling van evaluatieonderzoek is onderwerp van politieke strijd. Conclusies van beleidsevaluatoren wedijveren met andere standpunten in de politieke arena. Evaluaties zijn tenslotte inherent normatief (Fischer, 1995) en kennen winnaars en verliezers (Guba en Lincoln, 1989).

Voor het vaststellen van de aanwezigheid van 'governance' en de verschillende politieke belangen in een bepaald beleidsveld kunnen de volgende indicatoren worden benut (Boonstra et al., te verschijnen a):

- Multi-actor: beleidsontwikkeling en –uitvoering zijn een coproductie van overheden, marktpartijen, maatschappelijke partijen en/ of burgers.
- Multi-level: beleidsontwikkeling en –uitvoering zijn een coproductie van verschillende bestuurslagen (EU, nationaal, provinciaal, gemeentelijk).
- Multi-betekenis: beleidsontwikkeling- en uitvoering weerspiegelen verschillende probleemdefinities, oplossingsrichtingen, doelen en visies.

Wanneer sprake is van zowel multi-actor, multi-level beleidsvoering als multi-betekenisgeving spreken we van governance. Natuurlijk zal er in de praktijk altijd sprake zijn van gradaties en mengvormen.

Voor het vaststellen van governance zijn meer en minder geavanceerde methoden voor handen: krachtenveldanalyse, stakeholderanalyse, discoursanalyse, (beleids)netwerk analyse en beleidsarrangementen analyse. Afhankelijk van tijd en middelen kan voor een van deze methoden worden gekozen. Wel kennen de meer geavanceerde benaderingen als de beleidsnetwerk- en beleidsarrangementenbenadering duidelijke voordelen. Zo besteedt de beleidsnetwerkbenadering veel aandacht aan de onderlinge (afhankelijkheids)relaties tussen actoren en hun effect op het beleidsproces en probeert de beleidsarrangementenbenadering inhoud en organisatie van een beleidsveld met elkaar te verbinden. Voor de aanpak van een beleidsnetwerkanalyse kan worden verwezen naar Driessen en Glasbergen (2000). De methodische aanpak van een beleidsarrangementenbenadering staat onder meer beschreven in Boonstra (2004) en Van der Zouwen (2006). Vaak kan de evaluator in deze fase echter gebruik maken van bestaande analyses van het beleidsveld.

Wie evalueert voor wie?

Hierbij gaat het om de vraag wie de producenten en consumenten van de evaluatie zijn. Wat betreft de producenten is het gebruikelijk een onderscheid te maken tussen interne- en externe evaluaties. Bij interne evaluatie vervullen de geëvalueerden zelf een centrale rol in het bepalen van de vraagstelling en aanpak van de evaluatie en in het beoordelen van de resultaten ervan. Bij externe evaluatie voert een onafhankelijke of toezichhoudende instantie de evaluatie uit (Edelenbos en Van Buuren, 2005). Wanneer echter gepoogd wordt verantwoording met leren te combineren, dan vervaagt het onderscheid tussen intern en extern evalueren en wordt het evaluatieproces een coproductie tussen geëvalueerden en externe onderzoekers. Dat betekent overigens niet dat er geen verschillende rollen en verantwoordelijkheden zijn in de verschillende fasen van het proces. Zo kan worden afgesproken dat de geëvalueerden een rol spelen in het beoordelen van de resultaten van de evaluatie, maar dat de evaluatoren deze beoordeling aanvullen met externe observaties en reflecties. Hiermee wordt enerzijds voorkomen dat externe evaluatoren conclusies trekken die niet gerechtvaardigd zijn volgens de geëvalueerde organisatie en daardoor in de spreekwoordelijke la verdwijnen. Anderzijds wordt op deze manier het interne leerproces verrijkt door "onafhankelijke" observaties (Van der Meer en Edelenbos, 2002). Door verwachtingen van rollen goed met elkaar te bespreken en te komen tot spelregels kunnen

interessante, wederzijds bevruchtende, combinaties ontstaan van intern en extern evalueren. Hiervoor zijn geen algemene richtlijnen te geven, per situatie moeten verschillende rollen, verantwoordelijkheden en activiteiten op elkaar worden afgestemd.

Ook de beantwoording van de vraag wie de consumenten van de evaluatie zijn is minder eenduidig dan op het eerste gezicht lijkt. Naast de directe opdrachtgever(s) van de evaluatie zijn er tal van andere betrokkenen bij het beleidsproces die als gebruiker kunnen worden aangemerkt. Ook de notie van opdrachtgever verdient nuancering. Zo is er waarschijnlijk een verschillende behoefte aan evaluatie tussen de beleidsambtenaar die de evaluatieopdracht uitzet, zijn directeur of directeur-generaal en zijn politieke baas. De andere betrokkenen in het beleidsproces en hun onderlinge afhankelijkheidsrelaties zijn als het goed is tijdens de verkenning van het beleidsveld al geïnventariseerd. Guba en Lincoln (1989) maken in dit verband onderscheid tussen drie typen stakeholders: "agents", "beneficiaries" en "victims". Agents zijn de direct betrokkenen in het beleidsproces zoals de opdrachtgever van de evaluator, andere beleidsmakers, uitvoerende diensten, vertegenwoordigers van maatschappelijke partijen, bestuurders en politici. Beneficiaries zijn de mensen die direct of indirect profiteren van het te evalueren beleid. Victims zijn degenen die negatief worden beïnvloed door het beleid. Deze verschillende stakeholders zullen uiteenlopende behoeften hebben aan het evaluatieonderzoek. In een gebruikersgerichte evaluatie is het van belang ook deze behoeften in een vroeg stadium te onderzoeken. Op basis hiervan kan de probleemstelling van het onderzoek worden bijgesteld en kunnen opdrachtgever en evaluatoren beslissen welke rollen deze potentiële gebruikers in het evaluatieonderzoek kunnen vervullen.

Evaluatieonderzoek in een governance-omgeving dat een bijdrage wil leveren aan beleidsverbetering, impliceert namelijk dat sommige stakeholders meer moeten zijn dan passieve consumenten van het evaluatieonderzoek. Oorzaak is de gespreide verantwoordelijkheid voor de beleidsresultaten die het kenmerk is van een governance-omgeving (zie Fig. 2.2). Het volstaat dan niet dat een instantie leert dan wel zijn gedrag aanpast, de herkenning en acceptatie van evaluatieresultaten door andere sleutelactoren is cruciaal. Bovendien beschikken deze stakeholders over waardevolle informatie die evaluatoren nodig hebben. Dit betekent dat zij, net als de directe opdrachtgever, nauw bij het evaluatieproces betrokken moeten worden. Dit kan in verschillende gradaties. Mede-opdrachtgeverschap gaat het verst waarbij stakeholders gezamenlijk onder meer de probleemstelling van het onderzoek bepalen. Verder kunnen stakeholders het beoordelingskader mee-ontwikkelen, worden betrokken bij de interpretatie en beoordeling van de verzamelde gegevens of deze gegevens zelf verzamelen. Het minst ver gaat een rol als informant.

Natuurlijk zitten er grenzen in het betrekken van stakeholders bij het evaluatieproces. Het is in de context van opdrachtonderzoek meestal niet haalbaar om stakeholders te blijven betrekken tot het verzadigingspunt qua informatievoorziening is bereikt (zie bijv. Guba en Lincoln, 1989) Een geschikt criterium voor afbakening vormt wederzijdse afhankelijkheid. Van welke actoren is de opdrachtgevende instantie sterk afhankelijk bij het doorvoeren van eventuele beleidsveranderingen? Het is zaak in ieder geval deze een belangrijke rol te geven in het evaluatiearrangement.

Verder brengt het betrekken van stakeholders automatisch meerstemmigheid met zich mee. De evaluatoren moeten zich hierop bedacht zijn en de consequenties in de voorbereidingsfase doordenken. Stakeholders hebben immers meestal verschillende verwachtingen van de evaluatie, van het te evalueren beleid en komen tot verschillende beoordelingen. Hoewel het nuttig is op hoofdlijnen overeenstemming tot stand te brengen, betekent deze

meerstemmigheid niet dat alle stakeholders het met alle onderdelen van de evaluatieopzet eens hoeven te zijn. Wel moeten er genoeg elementen in zitten die het interessant maken om te participeren. Een mogelijkheid om recht te doen aan meerstemmigheid is deze te laten doorklinken in het beoordelingskader. Dit betekent dat dit criteria en indicatoren omvat die gebaseerd zijn op verschillende verwachtingen en doelen van stakeholders. Een andere optie is het combineren van verschillende evaluatieactiviteiten die in het belang zijn van verschillende stakeholders. Denk aan een combinatie van effectevaluaties met procesevaluaties of een evaluatie op centraal niveau met evaluaties op regionaal en lokaal niveau. Wel moet dan goed nagedacht worden over hoe deze evaluaties met elkaar kunnen worden verbonden. In het geval van een evaluatie op centraal en regionaal niveau kan bijvoorbeeld gedacht worden aan het organiseren van gezamenlijke reflectie met actoren van beide niveaus over gemeenschappelijke beoordelingscriteria. Dit heeft als positief neveneffect dat communicatie en samenwerking tussen de niveaus wordt bevorderd, waar coördinatie in de beleidspraktijk vaak een weerbarstig is (Van der Meer en Edelenbos, 2002). Ook heeft het toelaten van meerstemmigheid consequenties voor de verslaglegging. Hierover meer in paragraaf 2.3.

Waarom wordt geëvalueerd?

Er zijn verschillende motieven voor het uitvoeren van een evaluatie onderzoek. Naast beleidsinhoudelijke motieven zijn er andere motieven die betrekking hebben op tactisch of strategisch gebruik van het onderzoek met andere dan beleidsinhoudelijke bedoelingen. Zo kunnen onderzoeksopdrachten worden verstrekt om bepaalde beleidsbeslissingen te vertragen (de ijskastfunctie van onderzoek) of te vermijden. Verder kunnen onderzoeksresultaten worden aangewend om bijvoorbeeld de positie in interdepartementaal overleg te versterken (de vijgenbladfunctie) of die van anderen te verzwakken (Arentsen, 1991). In de praktijk zal het overigens vaak gaan om een mix van inhoudelijke en strategische motieven.

Er mag van worden uitgegaan dat aan een gebruikgericht onderzoek een inhoudelijke informatie- of leerbehoefte ten grondslag ligt. Het onderzoek dient met andere woorden antwoord te geven op bepaalde vragen met betrekking tot beleid. Zo kan men willen weten of het beleid werkt en tot welke effecten het beleid leidt. Ook kunnen er specifieke problemen of onzekerheden zijn ten aanzien van de werking of effecten van beleid. Beleidsevaluatie kan verder bedoeld zijn voor het ontwerpen of bijstellen van beleid. In de eerste plaats kan een evaluatie een argument voor of tegen het beleid leveren. De consequentie hiervan kan zijn dat het beleid wordt uitgebreid of ingekrompen. Deze vorm wordt wel summatieve evaluatie genoemd en past over het algemeen slecht bij een lerende benadering. Wanneer een beleid niet goed werkt, is daarmee nog geen alternatieve aanpak voor het beleidsprobleem voorhanden. Bovendien kunnen er vele andere, politieke, redenen zijn waarom het goed is zulk beleid voor te zetten dan geen enkel beleid meer te voeren. Formatieve evaluaties gaan daarom naast beoordeling ook in op de verklaring van de beleidsresultaten: hoe komt het dat de beleid goede dan wel slechte resultaten boekte? Juist dit streven naar verklaring kan informatie opleveren waarmee ook adviezen kunnen worden ontwikkeld voor verbetering van beleid (Bressers en Hoogerwerf, 1991).

De leerbehoefte kan ook minder instrumenteel zijn. Zo kunnen beleidsevaluaties bijdragen aan reflectie op de aansluiting van de beleidsdoelen op bredere maatschappelijke wensen, doelen en arrangementen en een rol spelen in de constructie van rationele ideologieën (zie Fischer, 1995; Verhoef, 1992). Ook kan een evaluatie bedoeld zijn om het lerend vermogen van de betrokkenen bij het beleidsproces te verhogen (zie Edelenbos en Van Buuren, 2005).

Het is zaak de verschillende motieven voor evaluatie zoveel mogelijk duidelijk te krijgen in de voorbereidingsfase van het onderzoek en je er van bewust te zijn dat ze ook nog eens kunnen

verschillen per actor. Maar weinig evaluatiestudies zullen zo breed van opzet kunnen zijn dat zij aan alle informatiebehoeften van alle relevante stakeholders kunnen voldoen. Verder is het van belang voor jezelf af te wegen met welke motieven je wel en niet uit de voeten kunt en samen met de opdrachtgever en eventuele andere stakeholders te komen tot een heldere doelstelling voor het onderzoek (zie Fig. 2.2).

Wanneer wordt geëvalueerd?

Uitgaande van verouderde ideeën over het volgtijdelijk verloop van de beleidscyclus wordt in de literatuur wel onderscheid gemaakt tussen ex ante, ex durante en ex post evaluaties (zie Fig. 2.2). Ex ante evaluatie beoogt dan om tijdens de beleidsvorming een min of meer complexe beslissituatie inzichtelijk te maken voor de beleidsmakers. Ex durante evaluatie betreft het beleid-in-uitvoering. De vergaarde en geïnterpreteerde gegevens wordt teruggekoppeld naar hetzij de beleidsmakers in de beleidsvorming hetzij de beleidsuitvoerders in het implementatieproces. Ex post evaluatie gaat om evaluatie op een moment dat het beleid uitgevoerd is, of tenminste gedurende langere tijd in uitvoering is en de effecten zichtbaar zijn (Verhoef, 1992).

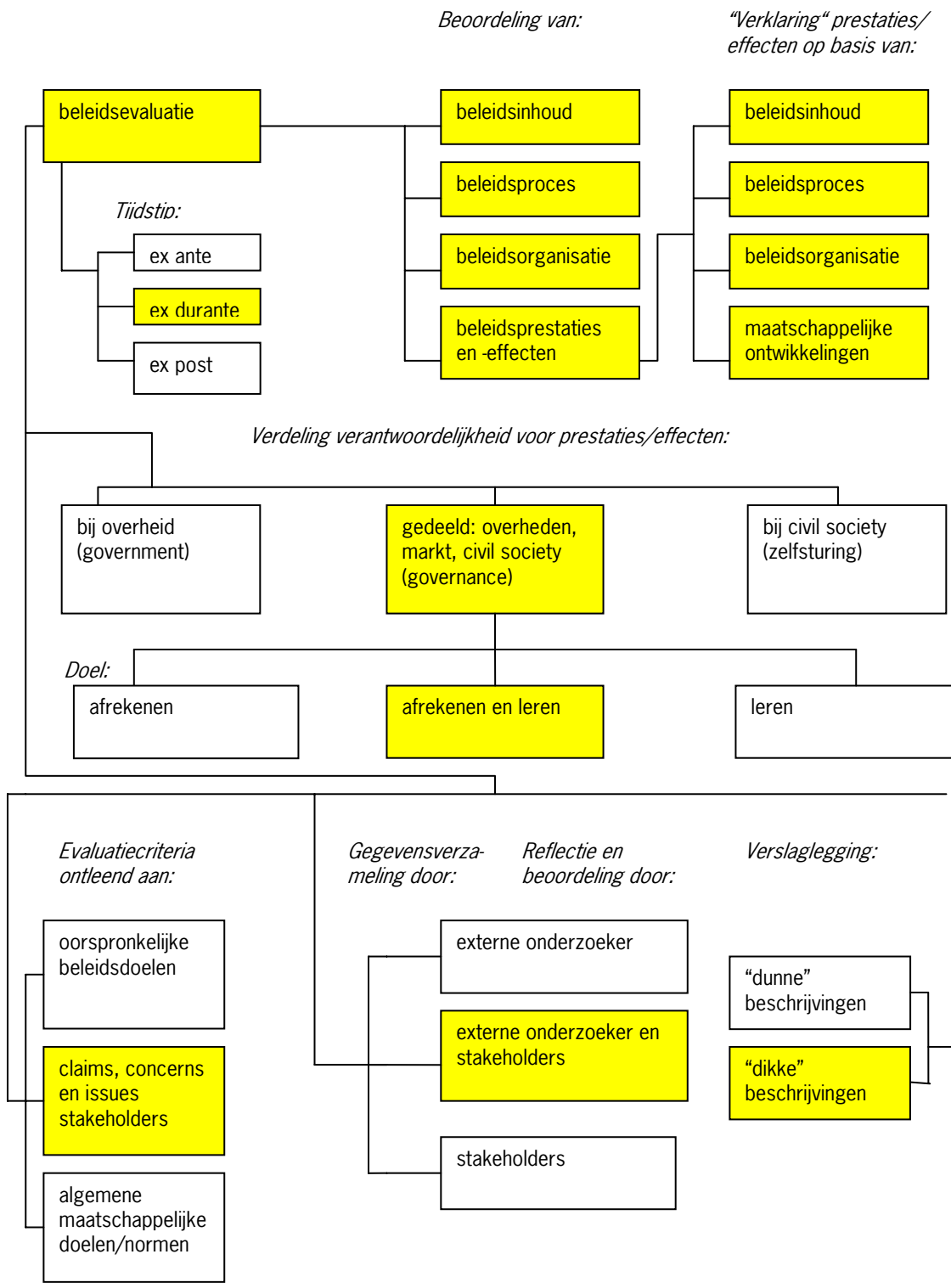
Als beleid wordt beschouwd als doorgaand proces en beleidsontwikkeling en –uitvoering steeds vaker gelijktijdig plaatsvinden, wordt het onderscheid tussen ex ante, ex durante en ex post evaluatie kunstmatig. Bovendien is het de vraag in hoeverre pure ex post evaluaties voldoen aan het doel van een benuttingsgerichte evaluatie. In beginsel is ex post evaluatie alleen interessant om in analoge toekomstige situaties van te leren (Verhoef, 1992). Het is de vraag of die bestaan.

In een lerende evaluatie is evaluatie tijdens de rit eigenlijk de enige optie omdat er sprake is van onmiddellijke, korte termijn terugkoppeling en reflectie over wat wordt beoogd en wat feitelijk wordt bereikt. Nog tijdens de rit kan de beleidsontwikkeling of – uitvoering zo worden aangepast. Om een goede timing mogelijk te maken is een gedetailleerd inzicht in bestuurlijke en politieke besluitvormingsproces en de ambtelijke voorbereiding een absolute vereiste. Bovendien kan de gewenste timing per gebruiker verschillen. Ook hiermee moet de evaluator zoveel mogelijk rekening houden.

Wat is het voorwerp van evaluatie?

Afhankelijk van het doel van de evaluatie kan het voorwerp van evaluatie worden bepaald. Gangbaar is het onderscheid tussen inhoud, proces, organisatie en effecten van beleid (zie Fig. 2.2). Bij een inhoudsevaluatie kan het gaan om de ontworpen doel-middel structuur of kan de complete beleidstheorie worden beoordeeld. In procesevaluaties kunnen in beginsel alle deel- of aspectprocessen voorwerp van evaluatie zijn. In evaluaties van de organisatie gaat het om het functioneren van organisaties die het beleid maken of uitvoeren. In effectevaluatie gaat het om het vaststellen van effecten van beleid of beleidsprogramma's. Hieronder vallen doelbereik-, doeltreffendheids- en doelmatigheidsonderzoek, maar ook onderzoek naar niet door het beleid beoogde neveneffecten (Arentsen, 1991; Verhoef, 1992). Soms is het niet mogelijk om beleidseffecten vaststellen en wordt volstaan met het meten van beleidsprestaties.

Bij al deze vormen van evaluatie hoort een specifiek beoordelingskader. Dit beoordelingskader kan worden ontleend aan het beleid zelf, algemene uitgangspunten of normen (zoals algemene beginselen van behoorlijk bestuur), theoretische inzichten of aan de percepties van stakeholders. Typisch voor een lerende evaluatie in een governance context is dat de beoordelingscriteria niet louter aan de formele beleidsdoelen worden ontleend. Hierover meer in paragraaf 2.3.



Figuur 2.2: Keuzepad lerende evaluatie in governance-context

Verder kan een analyse van de beleidsinhoud, het beleidsproces en of de beleidsorganisatie(s) deel uit maken van de verklaring van beleidseffecten of –prestaties. Ze zijn dan geen zelfstandig voorwerp van beoordeling, maar de analyse ervan draagt bij aan een beter begrip van de beleidsresultaten of het gebrek daaraan. Andere onderzoeksobjecten die hiervoor in aanmerking komen zijn de institutionele context waarin het beleid tot stand komt en onderliggende maatschappelijke processen.

2.3 Ontwikkelen aanpak voor beantwoording onderzoeksvragen

De voorbereidende fase resulteert in een goed omschreven doel- en vraagstelling voor het onderzoek. Ook heeft de voorbereidende fase waarschijnlijk al ideeën over de te hanteren onderzoeksmethode opgeleverd. Op de methodische keuzes gaan we in dit hoofdstuk verder in. We gaan er hierbij wederom van uit dat we onderzoek doen in een governance-context en naast verantwoording ook leren beogen.

Eerst gaan we in op het opstellen van een beoordelingskader. Naast de onderzoeksvragen is met name dit beoordelingskader richtinggevend voor het evaluatieonderzoek. Het geeft aan op basis van welke criteria en indicatoren we het beleid gaan beoordelen. Een lerende evaluatie stelt specifieke eisen aan de wijze waarop dit beoordelingskader tot stand komt.

Vervolgens behandelen we de fase van daadwerkelijke beoordeling. Welke methoden van gegevensverzameling, analyse en interpretatie lenen zich hiervoor? Anders dan in een traditionele evaluatie gaat het hierbij niet alleen om pure onderzoeksmethoden, maar ook om methoden om de interactie tussen en reflectie van stakeholders te faciliteren. Op dergelijke methoden gaan we in dit verband niet verder in (zie bijvoorbeeld Abma, 1996; De Bruijn en Ten Heuvelhof, 1995)

Ten slotte gaan we in op de keuzes die moeten worden gemaakt over de verslaglegging van het onderzoek. Ook hieraan stelt een lerende evaluatie specifieke eisen.

Opstellen beoordelingskader

In de meeste beleidsevaluaties worden de beoordelingscriteria en indicatoren afgeleid uit de formele, oorspronkelijke beleidsdoelen van het te evalueren beleid (zie Fig. 2.2). Als deze beleidsdoelen onvoldoende concreet zijn, worden ze op basis van de informatie van betrokken beleidsmakers door de onderzoekers gereconstrueerd (Bressers en Hoogerwerf, 1991). Op deze werkwijze valt het nodige af te dingen. Beleidsprocessen in een governance-omgeving hebben steeds vaker een doelzoekend karakter waardoor oorspronkelijke doelen van de initiator te globaal zijn om op te evalueren en/of gedurende het publieke onderhandelingsproces wat op de beleidsvaststelling volgt minder relevant worden. Hajer en Wagenaar (2003: 88) beschrijven dit laatste fenomeen als “policymaking as constitutive of politics”. Beleidsvoering is hierin niet uitsluitend het resultaat van politieke besluitvorming. Steeds vaker creëert beleid een publieke ruimte waarin mensen van diverse pluimage beraadslagen over hun toekomst, hun onderlinge relaties en hun verhouding tot overheden.

Bovendien doet de gangbare methode om tot beoordelingscriteria te komen onvoldoende recht aan het uitgangspunt van een meervoudige werkelijkheid. Dit uitgangspunt betekent dat mensen hun eigen betekenis aan de wereld geven en dat die betekenissen uiteen kunnen lopen. Door de werkelijkheidsdefinitie van de beleidsinitiator als enige input voor het beoordelingskader te accepteren worden kansen op leren tussen diverse betrokkenen misgelopen.

Meerstemmigheid in het beoordelingskader kan op verschillende wijzen worden georganiseerd. Guba en Lincoln (1989) introduceren een methodiek die als inspiratiebron kan dienen. Zij komen via de 'claims', 'concerns' en 'issues' (CCI's) van diverse belanghebbenden ten aanzien van het te evalueren beleid tot een zogenaamde 'agenda voor onderhandeling'. De term 'claim' verwijst naar positieve verwachtingen ten aanzien van het beleid of programma, 'concerns' naar zorgen en 'issues' naar discussiepunten (Abma, 2000). Als belanghebbenden onderscheidt hij: "agents", "beneficiaries" en "victims". Agents zijn de direct betrokkenen in het beleidsproces zoals de opdrachtgever van de evaluator, andere beleidsmakers, uitvoerende diensten, vertegenwoordigers van maatschappelijke partijen, bestuurders en politici. Beneficiaries zijn de mensen die direct of indirect profiteren van het te evalueren beleid. Victims zijn degenen die negatief worden beïnvloed door het beleid.

Deze CCI's zijn niet vooraf kenbaar maar moeten door onderzoek worden verkend bijvoorbeeld via documentenanalyse, diepte-interviews en gesprekken met homogeen samengestelde groepen stakeholders. Naast het beschrijven van de CCI's moet de evaluator achterliggende motiveringen en waardensystemen verhelderen met het oog op de lerende functie van de evaluatie. Speciale aandacht moet uitgaan naar de controversen. Nadat de CCI's zijn verkend, stellen Guba en Lincoln voor om het proces in heterogeen samengestelde groepen te vervolgen aan de hand van een door de evaluator opgestelde agenda voor onderhandeling. De evaluator treedt hierbij op als bemiddelaar en facilitator en moet zorgen voor een open onderhandelingsproces tussen gelijkwaardige gesprekspartners. De evaluator kan hierbij ook zelf issues of (theoretische) inzichten inbrengen, mits dat gebeurt op niet autoritaire wijze.

Boonstra et al. (2006, te verschijnen b) hebben Guba en Lincolns aanpak geamendeerd zodat die meer tegemoet komt aan het verantwoordingsdoel van evalueren. Zij gaan naast huidige CCI's ook op zoek te gaan naar vroegere CCI's van stakeholders. Op die wijze kan de ontwikkeling van het beleidsveld worden beschreven. Bovendien kunnen zo de oorspronkelijke doeleinden van de beleidsmaker/opdrachtgever worden betrokken in het beoordelingskader zodat ook aan de formele verantwoordingsfunctie van het evaluatieonderzoek wordt voldaan. In plaats van een agenda voor onderhandeling komen zij zo tot een concept-beoordelingskader. Dit wordt vervolgens ter verdere uitwerking voorgelegd aan de heterogeen samengestelde groep stakeholders. Doel is waar mogelijk te komen tot gedeelde evaluatiecriteria en indicatoren. Resultaat is een set van criteria en indicatoren waarmee, afhankelijk van de onderzoeksvragen, de beleidsprestaties of –effecten, het beleidsproces, de beleidsinhoud of –organisatie kunnen worden beoordeeld en verklaard. Omdat in ambigue beleidsprocessen niet op alle onderdelen consensus of een gedeelde constructie zal ontstaan kan dit beoordelingskader inconsistenties of tegenstrijdige criteria omvatten. Zo vinden stakeholders in de reconstructie zandgebieden structuurverbetering van de intensieve veehouderij een belangrijk doel. Wanneer zij dit doel echter operationaliseren blijken verschillende visies op de toekomst van de veehouderij te leiden tot uiteenlopende evaluatiecriteria. Kan structuurverbetering voor de een worden beoordeeld aan de hand van het aantal geclusterde bedrijfslocaties, voor een ander is bedrijfsontwikkeling op de bestaande, verspreid liggende locaties, ook een indicator voor structuurversterking.

Er zijn overigens ook andere interactieve methodieken om tot evaluatiecriteria te komen. Zo kan de evaluator ook aan de hand van cruciale gebeurtenissen in het beleidsproces die stakeholders benoemen, komen tot een reconstructie van doelen en criteria. Vervolgens kunnen deze doelen en evaluatiecriteria weer worden uitgewisseld en uitonderhandeld.

Komen tot een beoordeling

Als de onderzoeksvragen en het beoordelingskader zijn opgesteld breekt de fase van daadwerkelijke beoordeling plaats. Om ook in deze fase leren onder stakeholders te stimuleren is het zaak gedurende deze fase frequente cycli van waarnemen, gezamenlijke reflectie en (re)actie te organiseren (Edelenbos en Van Buuren, 2005).

Wat betreft het waarnemen zijn verschillende keuzes mogelijk: door evaluatoren, met hulp van stakeholders of door zelfevaluatie van de geëvalueerden (zie Fig. 2.2). Afhankelijk van de onderzoeksvragen en het beoordelingskader kunnen verschillende kwalitatieve en kwantitatieve methodieken van gegevensverzameling worden ingezet zoals interviews, documentenanalyse, participatieve observatie, enquêtes. Dit zijn gangbare methodieken die niet onderscheidend zijn voor een interactieve evaluatie.

De wijze van interpretatie of analyse van de gegevens is wel nadrukkelijk anders dan in een traditionele evaluatie omdat het geen exclusieve activiteit van de onderzoekers. Betekenis geven aan onderzoeksresultaten en komen tot een beoordeling middels gezamenlijke reflectie met stakeholders is een essentieel onderdeel van een interactieve evaluatie. Daarom moet er ruimte zijn voor herhaaldelijke terugkoppeling van (tussen)resultaten. Dit kan middels bilaterale contacten tussen onderzoekers en individuele stakeholders, waarin stakeholders reageren op tussenresultaten en onderzoekers hun resultaten verrijken met de interpretaties van stakeholders. Meer leereffecten worden verwacht van gezamenlijke leersessies met stakeholders waarin ook de verschillen in betekenisgeving onderwerp van gesprek zijn. Abma benadrukt dat het hierin niet zozeer gaat om meningen en overtuigingen van stakeholders, maar vooral om hun ervaringen en praktijken (Abma, 2000).

De tussentijdse terugkoppelingsmomenten kunnen leiden tot (re)actie van de stakeholders. Vervolgens kunnen de evaluatoren ervoor kiezen deze acties onderwerp van evaluatie te maken. Dit laat het object en aanpak van het evaluatieonderzoek niet ongemoeid. Zo adviseerden de onderzoekers in een lerende evaluatie van het Stimuleringsprogramma Burger en Milieubeleid het VROM-programmateam op basis van tussenresultaten projecten meer vanuit de beleving van burgers op te zetten en uit te voeren. Vervolgens keken ze hoe het programmeam dit advies uitvoerde. Hiermee veranderde de locus (onderzoeksobject) en de focus (onderzoeksaanpak) van het onderzoek (Edelenbos en Van Buuren, 2005). Deze gang van zaken is nadrukkelijk anders dan in puur op verantwoording gerichte evaluaties, waarin de tussentijdse invloed van de onderzoeker meestal tot een minimum wordt beperkt en onderzoekers zeker geen actieve adviseursrol op zich nemen.

Verslaglegging

We gaan er vanuit dat in een evaluatie gericht op leren én verantwoording de onderzoeksresultaten ook gerapporteerd moeten worden⁴. Een dergelijke lerende evaluatie stelt eisen aan de verslaglegging van de onderzoeksresultaten (zie Fig. 2.2). De lezer moet niet alleen een beeld krijgen van het slagen of falen van een beleidscategorie, hij/zij moet ook gevoel krijgen voor de complexiteit van het beleidsproces. Gangbare onderzoeksverslagen presenteren alle data in een alomvattend kader of metaverhaal en maken kwistig gebruik van tabellen en schema's (Abma, 1996: 213). Er is meestal sprake van een reductionistische benadering "*which extracts the general from the particular and then sets the particular aside as detail, background or qualification*" (Flyvbjerg, 2001: 133). Beschrijvingen zijn 'dun': zaken worden genoemd en verklaard, details en specifieke tijds- en plaatsaanduidingen ontbreken. De waargenomen

⁴ In een evaluatie met alleen een leerdoel kan het leerproces het product van de evaluatie zijn en wordt rapportage overbodig.

observaties worden beschreven vanuit een extern en afstandelijk perspectief. Het is een monoloog van auteurs. Het beleid en de evaluatie zijn in de tekst gescheiden.

Meer verhalende, “thick descriptions” geven beter de complexiteit, gelaagdheid en dubbelzinnigheid van beleefde ervaringen weer. Gekozen wordt voor natuurlijk taalgebruik in verhaalvorm, waarbij de wetenschappelijke en argumentatieve redeneertrant wordt losgelaten. Concrete gebeurtenissen staan centraal. De onderzoeker verschuilt zich niet als passieve, onschuldige onthuller van feiten, maar positioneert zich als actieve constructor van de tekst (Abma, 1996: 213-214). Andere stemmen en interpretaties komen veelvuldig aan bod. Deze zogenaamde meerstemmigheid creëert een zekere openheid van de tekst en laat daardoor ruimte aan lezers om eigen en mogelijk andere conclusies te trekken (Alvesson en Sköldberg, 2000). Met name casestudy’s lenen zich voor een dergelijke wijze van verslaglegging.

Tussenvormen zijn natuurlijk ook mogelijk. Denk bijvoorbeeld aan een metaverhaal over beleidsprestaties of -effecten, afgewisseld met stukjes “thick description” waarin de nadruk ligt op concrete gebeurtenissen die tot specifieke beleidseffecten hebben geleid, de institutionele context en de betekenisgeving van diverse betrokkenen.

2.4 Uitvoering evaluatieonderzoek

Het mag duidelijk zijn: een lerende evaluatie vergt nogal wat van de onderzoeker. Volgens Edelenbos en Van Buuren (2005) is lerend evalueren 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 beleidspraktijk. Lerend evalueren is bij uitstek iteratief: het is een proces van proberen en aanpassen, van handelen en interpreteren. Dit vergt een explorerende en flexibele houding van onderzoekers (Edelenbos en Van Buuren, 2005). Wel is het zaak in dit proces in overleg met de opdrachtgever duidelijke ijkmomenten af te spreken, waarin (tussen)producten worden opgeleverd en onderzoeksplannen voor een volgende fase worden vastgesteld (Boonstra et al, te verschijnen b).

De evaluator heeft in een lerende evaluatie verschillende rollen: hij onderzoekt, beoordeelt, spiegelt zijn beschrijvingen, adviseert en faciliteert. In de rol van onderzoeker wordt de evaluator gevraagd om op wetenschappelijk verantwoorde wijze de onderzoeksvragen te beantwoorden. In de rol van beoordelaar moet hij op basis van de onderzoeksuitkomsten komen tot een negatief dan wel positief oordeel over het betreffende beleidsprogramma, de wet of regeling. Als “spiegelaar” komt het er op aan dat de evaluator empirisch rijke beschrijvingen van de beleidspraktijk maakt, waarin stakeholders zich kunnen herkennen. Bovendien moet de evaluator in deze rol zo onbevooroordeeld mogelijk kunnen luisteren en observeren, om de betekenisgeving van stakeholders tot zijn recht te laten komen. Als expert of adviseur wordt hij gevraagd op basis van wetenschappelijke inzichten of praktijkervaringen te adviseren over bijvoorbeeld het proces van beleidsuitvoering, de vormgeving van nieuwe beleidsinstrumenten of de werkwijze van een betrokken overheid. In de rol van facilitator tenslotte moet de evaluator leerproces van betrokken stakeholders actief begeleiden.

Een lerende evaluatie vereist dat onderzoekers deze verschillende rollen beheersen; ze moeten beschikken over de benodigde vaardigheden en competenties. Daarnaast moeten ze kunnen “schakelen”; tijdig en vloeiend overstappen van de ene naar de andere rol (Edelenbos en Van Buuren, 2005)

Het verdient verder aanbeveling om, voor zover mogelijk, verschillende rollen over verschillende leden van het projectteam te verdelen, zodanig dat individuele competenties en vaardigheden tegemoetkomen aan de gevraagde rollen. Een dergelijke taakverdeling heeft als bijkomend voordeel dat dit de transparantie voor opdrachtgever en stakeholders vergroot: men leert van wie men wat kan verwachten. Zo beschrijven Edelenbos en Van Buuren (2005) dat in de evaluatie Stad en Milieu het voor het betreffende programmteam lastig te volgen was dat de evaluators eerst vooral een rol hadden als meedenker en helper en zich uiteindelijk ontpopten als kritische beoordelaar (zie ook paragraaf 2.3).

Ook van opdrachtgever en stakeholders vergt een lerende evaluatie andere rollen dan een traditionele evaluatie. Ze moeten het explorerende en deels onvoorspelbare karakter van de evaluatie als uitgangspunt accepteren. Een lerende evaluatie vraagt bovendien een actieve bijdrage van genoemde groepen. Zij moeten de verkregen inzichten willen en kunnen vertalen in bijsturing van hun handelen. Dit laatste wordt ook wel leervermogen genoemd (Edelenbos en Van Buuren, 2005). Complicerende factor hierbij is dat veel opdrachtgevers en stakeholders in de groene beleidspraktijk evaluaties vooral zien als afrekenmechanismen. Hoewel men de problemen met dit soort evaluaties aan den lijve ondervindt, kent men weinig alternatieven. Om in die context een lerende evaluatie te initiëren en tot een goed einde te brengen is continu verwachtingsmanagement door de evaluatoren op zijn plaats (Boonstra et al., te verschijnen b).

2.5 Evaluatie van de evaluatie

Voor de beoordeling van de kwaliteit van lerend evaluatieonderzoek kunnen we gebruik maken van verschillende kwaliteitscriteria. Het gaat hierbij op de eerste plaats om criteria voor de methodologische kwaliteit: de argumentatieve kwaliteit van de onderzoeksconclusies (Maso en Smaling, 1998). Op de tweede plaats gaat het om criteria waarmee de bijdrage van de evaluatie aan de participatie en het leren van stakeholders beoordeeld kan worden.

Bij de methodologische criteria is er discussie over de vraag of traditioneel-wetenschappelijke criteria als validiteit en betrouwbaarheid geschikt zijn voor dit soort onderzoek. Auteurs als Guba en Lincoln (1989) en Abma (1996) kiezen voor andere termen met als belangrijkste argument dat “de” werkelijkheid niet bestaat en interpretaties van de werkelijkheid veranderen, mede onder invloed van het onderzoeksproces. Hierdoor verliezen de gangbare operationalisatie van validiteit (de correspondentie tussen bevindingen en de werkelijkheid), en betrouwbaarheid (de onafhankelijkheid van bevindingen van onderzoeker, tijd en onderzoeksinstrumenten) hun waarde. Maso en Smaling (1998) kiezen er voor de begrippen betrouwbaarheid en validiteit te handhaven maar ze een andere invulling te geven dan in gangbaar, kwantitatief, onderzoek. Kijkend naar de concrete uitwerking dan zijn de verschillen tussen beide zienswijzen echter niet zo groot en is het dus meer een kwestie van smaak of strategie welke termen men gebruikt.

Geloofwaardigheid (of interne validiteit) verwijst naar de mate van correspondentie tussen het verslag van de onderzoeker en kijk van de participanten op de door hen geuite belevingen en gebeurtenissen (Abma, 1996). Naast een zorgvuldige omgang met en analyse van empirisch materiaal zijn er verschillende manieren om aan geloofwaardigheid te werken. Middels zogenaamde “member check” krijgen respondenten interpretaties voorgelegd met de vraag of zij zich erin kunnen herkennen (Abma, 1996; 2000). Triangulatie van methoden, bronnen, perspectieven en onderzoekers kan de reikwijdte en geschakeerdheid van gegevens vergroten. Het levert bovendien argumenten om één interpretatie te benadrukken boven andere, omdat die beter past bij de rest van het empirisch materiaal (Alvesson en Sköldberg, 2000).

Overdraagbaarheid (of externe validiteit) verwijst naar de mate waarin bevindingen naar andere personen, fenomenen, situaties en tijdstippen dan die van het onderzoek en is gebaseerd op analogieredeningen. Niet de onderzoeker zelf bepaalt of en welke analogieën bestaan tussen de onderzochte en de niet-onderzochte situaties, maar de lezer of gebruiker. Belangrijk is dat overdraagbaarheid lang niet altijd nodig is in evaluatieonderzoek. Methoden om aan overdraagbaarheid te werken zijn het maken van zogenaamd “thick descriptions”. Hierdoor kan de lezer bepalen of de bevindingen ook in andere situaties opgeld doen (Guba en Lincoln, 1989; Abma, 1996; Maso en Smaling, 1998).

Afhankelijkheid (of betrouwbaarheid) verwijst naar de aannemelijkheid van methodologische beslissingen en naar transparantie over de rol van de onderzoeker. Omdat feitelijke herhaling van een lerende evaluatie niet mogelijk is en door de actieve interventie van de onderzoeker ook nooit identieke resultaten zal opleveren, gaat het er bij dit criterium om het onderzoek zo uit te voeren dat andere onderzoekers de gang van het onderzoek kunnen nagaan en hun eigen oordeel kunnen vormen. Een manier om de afhankelijkheid te verkleinen is de vervaardigen van een zogenaamd ‘audit-trail’, een systematische boekhouding van alle documenten, tussenrapporten, briefwisselingen, notities, geluidsopnames etc. zodat een onafhankelijke onderzoeker als een soort accountant het geheel kritisch kan inspecteren (Abma, 1996; Maso en Smaling, 1998). Ook triangulatie kan de betrouwbaarheid vergroten.

Naast geloofwaardigheid, overdraagbaarheid en afhankelijkheid introduceren Guba en Lincoln (1989) zogenaamde authenticiteitscriteria. *Eerlijkheid* verwijst daarbij naar de mate waarin verschillende stakeholders indringend worden bevroegd en hun ervaringen worden belicht in het evaluatieproces. Alle stakeholders moeten volgens dit criterium gelijke kans hebben deel te nemen aan het proces. De *overige authenticiteitscriteria* hebben betrekking op het toegenomen individuele inzicht, het toegenomen wederzijdse begrip, de mate waarin het evaluatieproces handelen faciliteert en stimuleert en de mate waarin participanten zijn toegerust om te handelen (“empowerment”).

De opsomming van kwaliteitscriteria is hiermee niet uitputtend. Ook benuttingswaarde of implementatiewaarde zijn bijvoorbeeld waardevolle criteria voor gebruikersgerichte evaluaties.

2.6 Tot slot

Aan de keuze en het ontwerp van een lerend evaluatieonderzoek ligt een gedegen analyse van de omgeving van het te evalueren beleid ten grondslag. Op basis van deze omgevingsanalyse kan worden geconcludeerd dat sprake is van een governance-omgeving. Ook kunnen uiteenlopende wensen voor het onderzoek van potentiële gebruikers worden vastgesteld. In overleg met de opdrachtgever kunnen betrokken onderzoekers vervolgens kiezen voor een evaluatieonderzoek waarin zowel sprake is van verantwoord en als leren. Kenmerkend hiervoor zijn de interactie met stakeholders of gebruikers in verschillende fasen van het onderzoek, de tussentijdse beleidsinterventie en de beperkte voorspelbaarheid van het onderzoeksproces.

Maar er zijn ook overeenkomsten met meer klassieke vormen van beleidsevaluatie. Deze liggen vooral op het punt van het onderzoeksobject, de te beantwoorden onderzoeksvragen en methoden voor gegevensverzameling. Onderzoekers moeten bovendien de juiste onderzoekscompetenties hebben. Meer traditionele, afstandelijke onderzoeksrollen worden in een lerend onderzoek echter gecombineerd met betrokken posities als beleidsadviseur of facilitator.

Ten slotte is de verantwoording van de onderzoeksanpak bij een lerende evaluatieonderzoek bij uitstek van belang, omdat van gebaande paden wordt afgeweken. Wel komen hiervoor andere criteria zoals geloofwaardigheid en eerlijkheid in aanmerking.

3 Op zoek naar lerende evaluaties

3.1 Inleiding

Het Milieu- en Natuurplanbureau (MNP) is natuurlijk niet de enige instantie die kampt met de geschetste problemen met de klassieke evaluatievormen. In meer of minder mate zijn wetenschappers praktisch bezig met het experimenteren met lerende evaluaties. In dit hoofdstuk gaan we in op ervaringen die ook leerzaam kunnen zijn voor het MNP. De vraag die hier centraal staat is:

Op welke manieren geven evaluatiewetenschappers buiten het MNP vorm aan lerende vormen van evaluatie en welke lessen zijn hieruit te trekken voor het MNP?

Voor dit onderdeel hebben we een beperkt aantal respondenten geselecteerd op basis van hun (vermeende) vernieuwende visie op evaluatiemethoden. Ook is gezorgd voor een zekere spreiding over verschillende organisaties. Tineke Abma (Universiteit van Maastricht) en Jurian Edelenbos (Erasmusuniversiteit Rotterdam) zijn geselecteerd op basis van hun wetenschappelijke publicaties over responsieve en lerende evaluatiemethoden. Daarnaast hebben we gesproken met onderzoekers van traditionele evaluatieinstanties als de Algemene Rekenkamer en het Ruimtelijk Planbureau (RPB). De reden hiervoor is dat deze organisaties qua visie op formele beleidsevaluaties (Algemene Rekenkamer) of qua werkveld en formele rol (RPB) veel lijken op het MNP. Binnen deze organisaties is wel gesproken met onderzoekers die op het terrein van beleidsevaluatie vernieuwend werken (Paul Pestman en Ed Dammers). Aanvullend op deze gesprekken is gebruik gemaakt van literatuur die is aangedragen door deze deskundigen.

3.2 Lerende vormen van evaluatie

3.2.1 Pestman - De Algemene Rekenkamer

De Algemene Rekenkamer is een zogenaamd Hoog College van Staat en heeft een onafhankelijke positie in het staatsbestel. Vanuit die positie houdt de Rekenkamer zich bezig met het toetsen en verbeteren van het functioneren en externe presteren van het rijk. Het college zelf bestaat uit drie mensen, zij beslissen uiteindelijk over de strategie en activiteiten van de Rekenkamer. De Rekenkamer doet sporadisch opdrachtonderzoek, bijvoorbeeld in opdracht van de Tweede Kamer of van een Minister.

Het werk van de Rekenkamer kent twee pijlers. De eerste pijler gaat over verantwoording en toezicht. De hoofdactiviteit is het jaarlijkse rechtmatigheidsonderzoek: de goedkeuring van de rijksrekening en rapporten bij de jaarverslagen. De tweede pijler gaat over beleidsevaluatie en kent activiteiten gericht op het beleid en de uitvoeringspraktijk bij het rijk. Wij gaan hier in op de tweede pijler. Het evaluatieonderzoek van de Rekenkamer beperkte zich van oudsher tot doelmatigheids- en doeltreffendheidsonderzoek. Sinds 2004 is hierin iets veranderd (Algemene Rekenkamer, 2004a). De Rekenkamer heeft een nieuwe strategie vastgesteld. Belangrijk is dat de Rekenkamer niet alleen meer een oordeel wil vellen over het beleid (goed of slecht beleid), maar ook onderzoek wil doen naar de oorzaken van beleidssucces of beleidsfalen. De filosofie is dat hiermee de conclusies en aanbevelingen een hogere beleidsrelevantie kunnen krijgen.

In dit kader heeft de Rekenkamer in 2004 (Algemene Rekenkamer, 2004b) een intern project uitgevoerd rond de ontwikkeling van verklaringen. Het project kent drie onderdelen: (1) de ontwikkeling van een methodische gereedschapskist voor verklarende beleidsevaluatie, (2) de ontwikkeling van beleidsrelevante bestuurlijke boodschappen en (3) de organisatorische consequenties van 1 en 2. Naast een lijvig rapport heeft dit project een handzame brochure opgeleverd. Wat betreft de gereedschapskist maakt het project onderscheid tussen verklaringen voor achterblijvende prestaties en effecten binnen en buiten de directe beïnvloedingssfeer van het rijk. Voor verklaring van achterblijvende prestaties buiten de directe beïnvloedingssfeer van het rijk beveelt het project bijvoorbeeld bedrijfsvoeringsonderzoek bij derden aan of netwerk- en stakeholderanalyses. Voor onderdeel 2 hebben projectmedewerkers de kwaliteit van vroegere aanbevelingen geanalyseerd. Deze bleek mager. Veel aanbevelingen hebben het karakter van: dit is niet aanwezig, zorg dat het wel aanwezig is. Door te werken aan verklaringen wil de Rekenkamer komen tot betere aanbevelingen. In hoeverre dit gaat lukken, moet natuurlijk nog blijken.

Bij de Rekenkamer is meepraten van stakeholders over het onderzoeksontwerp en de onderzoeksvragen uit den boze. De onderzoeksagenda is afgeleid van de strategische visie van de Rekenkamer. Daarin zijn wel de ideeën en signaleringen van medewerkers verwerkt, maar om haar onafhankelijkheid te borgen wordt er geen beïnvloeding van buitenaf georganiseerd. Wel zijn er andere momenten waarop meepraten van stakeholders wordt gewenst. Zo wordt informatie over (verschuivende) beleidsdoelen van ambtenaren betrokken in de beoordelingscriteria. Uiteindelijk is het echter de Rekenkamer zelf die beslist over de beoordelingscriteria. Verder worden ze betrokken als informant middels interviews naar beleidsprestaties en verklaringen. De nota van bevindingen wordt ook voorgelegd aan het betreffende ministerie. Dit kan tot bijstellingen leiden, maar wederom is het de Rekenkamer zelf die daarover beslist. Tenslotte geeft het betreffende ministerie een bestuurlijke reactie op het rapport van de Rekenkamer.

Ook binnen de Rekenkamer zijn er mensen die het aantal momenten van interactiviteit proberen te verhogen. Zo zijn in een onderzoek naar decentralisatie tussentijdse bevindingen voorgelegd aan een brede groep van betrokkenen, niet alleen rijk, maar ook provincies en gemeenten. Ook is in dit onderzoek een groep experts gevraagd om een oordeel te geven over de kernboodschap van het project.

3.2.2 Dammers - Ruimtelijk Planbureau

Het Ruimtelijk Planbureau (RPB) is een van de vier planbureaus in Nederland. De andere planbureaus zijn het Milieu- en Natuurplanbureau, het Sociaal Cultureel Planbureau en het Centraal Planbureau.

Het RPB heeft onder meer als taken om ruimtelijk relevante maatschappelijke ontwikkelingen te verkennen en te signaleren en het monitoren van maatschappelijke en ruimtelijke ontwikkelingen. Evaluatie staat niet met zoveel woorden in de taakomschrijving. Wel is daarin het analyseren van ruimtelijk relevant rijksbeleid opgenomen en het ontwikkelen van varianten van ruimtelijk beleid en scenario's. Op het eerste gezicht lijkt het RPB dus wat minder gericht op evaluatie van beleid dan het MNP maar de analyse van ruimtelijk beleid raakt hier wel aan.

Net zoals bij het MNP en de Algemene Rekenkamer staat bij het RPB onafhankelijkheid centraal. Een verschil is wel dat het RPB explicieter aangeeft niet alleen gedegen kennis te willen aanreiken, maar ook het debat wil verdiepen door nieuwe onderwerpen op de agenda te plaatsen. Bovendien is men wel gericht op het Rijk (regering en parlement), maar werkt men ook voor andere overheden.

Dammers et al. (2004) gebruikten in een studie naar ontwikkelingsplanologie inzichten op het terrein van lerend evalueren. Inspiratiebronnen vormden vooral Pressman and Wildavsky (1973) en Guba and Lincoln (1989). Het bijzondere van de methodiek van Guba en Lincoln ("fourth generation evaluation" of responsieve evaluatie) is dat ze alle betrokken partijen aan het woord laten en oog hebben voor zogenaamde "winnaars" en "verliezers" bij het te evalueren beleid. Zo krijgen ze zicht op doelen, handelen, denken, conflicten en verschillen.

In de studie over ontwikkelingsplanologie wilde men naast de verhalen van betrokkenen ook waarnemingen van onderzoekers betrekken. Dammers wilde verder niet alleen toetsen op het criterium "als de betrokkenen maar tevreden zijn" maar de cases ook beoordelen op het fysieke resultaat.

Uiteindelijk zijn er voor de studie drie cases geselecteerd (Blauwe Stad, Regiodialoog Noord-Limburg en IBA Emscherpark). Doel was de sleutelkenmerken van innovatie te achterhalen. De onderzoekers gingen ook zelf kijken in de gebieden. Daarnaast deden ze interviews en maakten ze gebruik van eerdere interviewresultaten.

Ook is in de studie Ontwikkelingsplanologie gewerkt met een simulatiespel rond de ontwikkeling van het stedelijk netwerk Brabant. Hieraan namen naast de mensen die hier in de dagelijkse (werk)praktijk direct bij betrokken zijn, ook mensen deel van buiten dit netwerk. Participatieve methoden werken namelijk niet goed als alleen de insidergroep van bestuurlijk betrokkenen deelnemen. Bij dit simulatiespel hebben de onderzoekers dit proberen open te breken door maatschappelijke actoren als treinreizigersgroep ROVER, gehandicapten-organisaties en organisaties voor ouderen en jongeren aan het spel te laten deelnemen. Ook zijn er ontwerpers bij gehaald. Dat werd allemaal gedaan om zicht te krijgen op de mogelijkheden en beperkingen die zich in de praktijk voordoen bij ontwikkelingsplanologie, gericht op systeeminnovaties. Het was erg confronterend voor de deelnemers, vooral voor Bureau BrabantStad dat er niet goed in slaagde om alle initiatieven die ontstonden te coördineren. Daarnaast zorgde de inbreng van buitenstaanders ook voor meer creativiteit in het spel. Het simulatiespel is alleen bedoeld om hypothesen te formuleren over de mogelijkheden en beperkingen die zich in de praktijk bij ontwikkelingsplanologie kunnen voordoen. De hypothesen werden daarna in de cases onderzocht.

Het ging het RPB in deze studie niet zozeer om afrekenen maar veel meer om te leren van de bestaande ervaringen met ontwikkelingsplanologie. Een formeel beleidsdoel kan je in de visie van Dammers misschien nog wel als vertrekpunt gebruiken, maar vervolgens moet je meer met de bewegingen in het krachtenveld meegaan. Halsstarrig vasthouden aan bestaande doelen is niet altijd nuttig omdat je het beleidsproces daarmee niet verder helpt. Je moet als onderzoeker het beleid een spiegel voorhouden en als het beleid vastloopt, kan je technieken aandragen om verder te komen.

In het project is ook gewerkt met een soort 'roadshow' in de provincies. Dat was een interactieve manier om op verschillende plaatsen discussie te voeren. Dit was nuttig voor het RPB project én voor de projecten in de provincies. Dammers denkt dat het MNP ook op een soortgelijke, levendiger manier te werk zou kunnen gaan. Guba en Lincoln geven ook goede (methodische) voorbeelden van hoe evaluaties minder technisch en levendiger opgezet kunnen worden. Verder zou het MNP volgens Dammers zich wat minder moeten vereenzelvigen met het rijksbeleid en de rijksdoelen.

3.2.3 Edelenbos - Erasmus Universiteit

Edelenbos noemt alle evaluaties waar een mengvorm van leren en verantwoorden aan de orde is 'lerende evaluatie'. Hij onderscheidt acht tot tien hoofdvormen (zie ook: Edelenbos en Van Buuren 2005). De precieze verhouding en vorm kunnen in de praktijk verschillen al naar gelang de wensen van de opdrachtgever, de beleidscontext, de vraagstelling etc. Daarbij wordt pragmatisch geput uit de bestaande evaluatieliteratuur. De stelling dat leren en verantwoorden niet samengaan van Guba en Lincoln (1989) of Abma (2000) deelt hij dus niet. Uitgangspunt van zijn zoektocht naar lerende evaluatie is dat de klassieke aanpak niet meer werkt. Dat wil zeggen dat de klassieke aanpak onvoldoende aangrijpingspunten voor leren en beleidsverbetering oplevert en bovendien ongewenst neveneffecten als ontwijkgedrag uitlokt. Mogelijk dat de klassieke aanpak nog wel geschikt is voor evaluaties die puur een afrekendoel hebben. Edelenbos vindt de klassieke evaluatie verder niet interessant omdat er weinig wetenschappelijke vernieuwing in zit.

Het onderzoek naar lerende evaluatie begon met een conceptuele verkenning in opdracht van VROM-DGR. Ze konden het concept daarna empirisch toepassen in een vervolgonderzoek voor VROM-DGM over het Stimuleringsprogramma Burger en Milieubeleid. Hieraan lag zowel een vraag van de Tweede Kamer naar de effectiviteit van het programma als een behoefte van het programmateam om te leren ten grondslag. Nu is men bezig met een vervolg dat gaat over de vraag: "hoe leert de organisatie"?

Centraal in deze onderzoeken staat dat de evaluatie is verweven met het beleidstraject. Je staat als onderzoeker niet aan de zijlijn van het beleidsproces, maar maakt daar deel van uit. Edelenbos spreekt van korte, frequente cycli van waarneming, reflectie en advisering door de onderzoeker en interpretatie en actie door de opdrachtgever. Hierdoor verspringen de focus en locus van de evaluator steeds. Het te evalueren beleid verandert immers 'tijdens de rit' door de interventies van de evaluator en deze veranderingen roepen nieuwe onderzoeksvragen op. Ook verschillen de rollen van de evaluator in de tijd of over personen. De onderzoeker is soms beoordelaar, dan weer coach en in een ander geval weer adviseur. Soms lijkt het wel of de onderzoeker op de stoel van de beleidsambtenaar gaat zitten.

Het combineren van deze rollen is overigens geen sinecure. Het programmateam van het burger en milieuonderzoek schrok er bijvoorbeeld van dat de onderzoekers ineens een beoordelende rol innamen en met een kritisch eindoordeel kwamen terwijl ze daarvoor zo leuk hadden meegedacht. Ook vergen de verschillende rollen uiteenlopende competenties van de onderzoekers. Het is maar de vraag of iedereen die heeft. Het is in ieder geval zaak je projectteam zo samen te stellen dat de verschillende competenties aanwezig zijn. Zo zegt Edelenbos dat hij zelf de neiging heeft om zich wat terug te trekken als er een te groot beroep op hem in de rol als medebeleidsambtenaar wordt gedaan. Andere projectteamleden hadden daar minder problemen mee.

Andere ervaringen tot nu toe zijn dat je vaak niet de kans hebt verschillende cycli van waarneming, interpretatie, advies en actie mee te maken. Twee cycli zijn echter minimaal nodig. Wel kunnen de cycli worden verkort door waarneming en reflectie tijdens een bijeenkomst plaats te laten vinden. Vraag is wel of je dan voldoende rust hebt voor zorgvuldige interpretatie. Wat ook nog niet gelukt is tot nu toe, is het organiseren van gezamenlijke reflectie van programmateam en projectteam op de rol van de evaluatoren. Dit gaan ze wel doen in het huidige project. Verder hebben de andere stakeholders, zoals de burgers die aan de projecten meededen, niet meegepraat over de evaluatieaanpak. Deze groepen waren overigens wel als informant/respondent betrokken bij het onderzoek.

De kritiek dat je als evaluator in een lerende evaluatie niet meer onafhankelijk bent en dus ook niet meer kritisch kan zijn, deelt Edelenbos niet. Hij geeft aan dat de samenvatting van het eindrapport dat naar de Tweede Kamer is gestuurd, best kritisch was. Er stond keihard in dat de originele doelstelling van het programma niet was gehaald. Daaraan voegde het rapport echter toe dat doelstellingen die in de loop van de tijd waren gerezen, zoals een meer burgergerichte houding bij de beleidsambtenaren, wel gehaald waren, mede als gevolg van de interventie van de evaluatoren. Zo hadden die eerder in het proces een zogenaamde intensiveringnotitie geschreven met veranderpunten. Hetzelfde geldt voor het lerend vermogen, ook dat was volgens betrokkenen toegenomen in de loop van het proces. Les hieruit is dat het belangrijk is zowel originele beleidsdoelstellingen als nieuwe doelen in het beoordelingskader te betrekken. Hierdoor gaat een evaluatie ook verder dan louter een tevredenheidstoets. De indruk bestaat dat de Tweede Kamer er niets mee gedaan heeft, bij de bestuurlijke top is het wel aangekomen. Overigens heeft de onderzoeker wel er sterk op aangedrongen dat er een samenvatting van het rapport naar de Kamer ging. De opdrachtgever vond een brief genoeg.

In het huidige DGM-project waarbij het lerend vermogen van de organisatie centraal staat wil Edelenbos de opdrachtgever nog beter betrekken bij onderzoeksmatige keuzes. Verder wil hij vaker terugkoppelen: twee cycli van waarneming en leren moeten haalbaar zijn in de zeven maanden die voor het project staan. Verder wil hij de opdrachtgever en evaluatoren samen laten reflecteren, niet alleen op de rol van de opdrachtgever maar ook op die van de evaluatoren. Overigens werkt hij in de lerende evaluatie samen met mensen van TNO. Zij begeleiden de leersessies. Hij ervaart dit als heel positief: zij hebben die expertise, terwijl zijn eigen expertise meer analytisch is.

3.2.4 Abma – Universiteit van Maastricht

Het werk van Abma hebben we eerder besproken in een project voor het MNP (Kuindersma en Boonstra, 2005). Daarbij baseren we ons met name op haar promotie-onderzoek waarin ze de responsieve evaluatiemethode van Guba en Lincoln verder heeft uitgewerkt (Abma, 1996). Deze subparagraaf steunt voornamelijk op een interview over haar ervaringen met de toepassing van de methode in de laatste jaren.

De responsieve benadering heeft als uitgangspunt dat de issues van de stakeholders centraal staan en niet de beleidsdoelen van de overheid. Wanneer onderzoekers op zoek gaan naar deze issues dan kijken ze meestal naar opvattingen, meningen of cognitieve overtuigingen. Abma vindt ervaringen interessanter dan opvattingen en meningen. Ervaringen krijgen betekenis in verhalen. Ze heeft de responsieve evaluatiemethode daarom uitgebreid met de concepten “verhalen” en “dialogoog”. De issues worden achterhaald via verhalen en gedeeld via dialogoog. Voor het faciliteren van deze dialogoog is het volgens Abma niet zinvol om het te hebben over opvattingen en meningen want dan kom je in een debat terecht met argumenten voor of tegen. Je krijgt daardoor oppositie en strijd. Wie wint het spel? Wie is het meest gewiekst? Wie kan het beste zijn tegenstander pareren? In een echte dialogoog gaat het daarentegen om het uitwisselen van ervaringen, het openstaan voor elkaars ervaringen ofwel het leren van elkaar.

Abma kiest in tegenstelling tot de meer pragmatische insteek van Edelenbos heel duidelijk voor haar eigen benadering. In het interview zegt ze het zo:

Ik zelf doe geen doe-evaluaties. Ik zal altijd werken vanuit een sociaalconstructivistische of postmodernistische perspectief. Ik kan niet anders. Ik kan niet twee overtuigingen gelijk aanhangen. Ik kan niet doen alsof de werkelijkheid een gegeven is en doen alsof de werkelijkheid multi-interpretabel is. Maar er vanuitgaande

dat wij de wereld construeren kan ik best werken met mensen die kwantitatief onderzoek doen vanuit een logisch-positivistische opstelling. Alleen zal ik die niet voor waar aannemen, maar als "een waarheid" zien en er mee in gesprek gaan'.

Abma heeft vooral veel projecten binnen de gezondheidszorg gedaan: de psychiatrie, revalidatie, chronische zorg en paletieve zorg. Daarnaast was ze actief in de sociale vernieuwing, onderwijs, de kunst en het (eerder genoemde) Schiphol project. Ze kiest daarbij uitsluitend voor op leren gerichte evaluaties. Opdrachtgevers weten dat ook en zullen haar niet vragen voor een op verantwoording gerichte evaluatie. Illustratief voor de aanpak van Abma is het volgende citaat uit het interview:

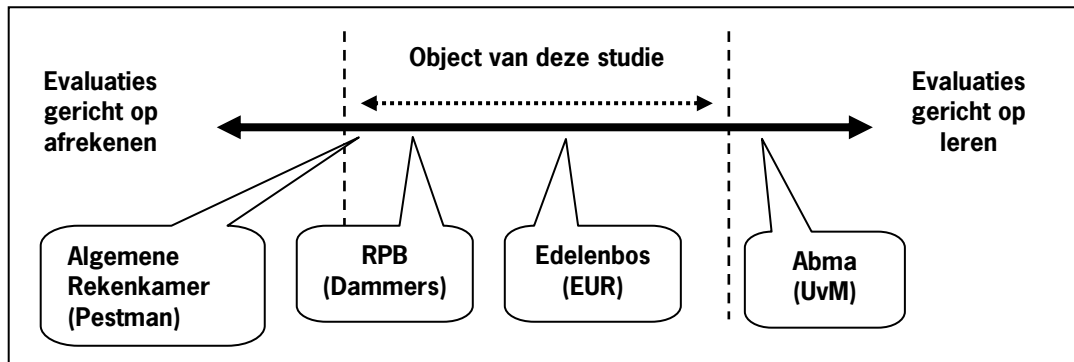
'Mijn insteek is altijd het op gang brengen van een dialoog tussen stakeholders als voertuig voor praktijkverbetering. Dus ik probeer een proces te organiseren. Dat was ook zo grappig in het psychiatrieproject. Ik was helemaal vergeten het eindrapport te schrijven. Ik dacht: het proces is toch het product. Ik had wel een boek gemaakt, maar dat had ook een procesfunctie ter bestendinging en verspreiding van informatie. Maar de participanten in dat project hadden toch het idee: er moet ook een rapport komen. Nou prima dan maak ik ook nog een rapport. Klaar'.

Veel klassieke doelevaluaties concluderen volgens haar dat de doelen niet gehaald zijn. Dat is heel voorspelbaar. Wat je dan vaak niet weet zijn de neveneffecten van het beleid en je weet niet waardoor het komt. Vaak is de reactie vanuit het beleid op zo'n conclusie ook nog eens: stop de handel maar, het was geen goede beleidsveronderstelling. Maar je kunt ook zeggen de onderzoeksmethodologie is niet goed. Als je een andere methodologie had gehanteerd, had je een andere uitkomst gehad. Nu werkt de uitkomst ook nog eens tegen je in de zin dat de betrokkenen bij het beleid zich dan ook nog eens tegen je keren. De responsieve benadering gaat er meer vanuit hoe je het beleid beter invulling kan geven. Waar lopen de stakeholders tegenaan? Wat is er moeilijk aan de uitvoering van het beleid? Wat vinden zij belangrijk?

In een recent evaluatieproject in de psychiatrie kwam overigens wel naar voren dat de responsieve en klassieke benadering elkaar kunnen aanvullen. Het project ging over het toepassen van dwang en drang in de psychiatrie in Nederland. Een aantal jaren terug is het besef ontstaan dat het raar is dat wij zoveel mensen opsluiten. Toen zijn, in samenwerking met het veld, professionals en oud-cliënten en familie, er kwaliteitscriteria ontwikkeld voor dwang en drang. Die criteria zijn twee jaar geleden geïmplementeerd en deze implementatie heeft Abma geëvalueerd. De implementatie in 12 instellingen werd gevolgd. De instellingen hadden veel vrijheid om het nieuw beleid zelf in te vullen. In de evaluatie is dit proces gevolgd en vooral een leerproces tussen die instellingen gefaciliteerd. De stakeholders, verpleegkundigen, cliënten etc., zijn bij elkaar gebracht om hun issues op tafel te krijgen. Hoe kunnen zij met elkaar in gesprek gaan en van elkaar leren? Daarnaast vond ook een kwantitatieve evaluatie plaats. Daarin werd gekeken of het aantal separaties (= dwang) afnam in die instellingen. Van sommige afdelingen hadden de onderzoekers op basis van de dialoogsessies het gevoel dat zij het ontzettend goed deden. Toch bleek uit het kwantitatieve onderzoek dat het aantal separaties daar omhoog ging. Op dat moment moet je volgens Abma toch in gesprek en is de ervaringsdeskundigheid van de lokale afdeling nodig om de gegevens te kunnen duiden. Kwantitatieve onderzoekers terug kunnen wel het effect vaststellen, maar hebben geen inzicht in de waarom-vraag. Op dit punt kunnen methoden elkaar aanvullen, zonder dat het helemaal van tevoren is geregisseerd.

3.3 Reflectie

In dit hoofdstuk hebben we een aantal evaluatiedeskundigen van buiten het MNP aan het woord gelaten over hun ervaringen met lerende vormen van evaluatie. Samenvattend kunnen we de evaluatiedeskundigen op de as van evaluatievormen als volgt rangschikken (zie Fig. 3.1). Dit is een relatief beeld en gaat vooral uit van de evaluatiedeskundigen die we in deze studie hebben gesproken.



Figuur 3.1: Evaluatiedeskundigen op de 'meetlat'

Met deze inventarisatie is natuurlijk geen compleet beeld ontstaan van de praktijk van lerende beleidsevaluatie in Nederland. De gesprekken leveren wel een aantal opvallende constatering op en ook enkele mogelijke lessen voor beleidsevaluatie binnen het MNP.

We noemen de belangrijkste conclusies en lessen:

1. Een belangrijke reden voor het opstarten van een lerende evaluatie is de behoefte om te verklaren waarom beleid niet werkt. Deze vraag kan vaak niet voldoende beantwoord worden in klassiek en kwantitatief evaluatieonderzoek. Op dit punt kunnen klassieke en lerende beleidsevaluaties elkaar aanvullen. Zelfs de puur op leren gerichte responsieve benadering van Abma, blijkt in combinatie met klassieke effectmetingen nog voor meerwaarde te kunnen zorgen.
2. Een tweede belangrijke reden voor lerende beleidsevaluaties is de behoefte om handvatten voor beleidsverbetering aan te reiken. Klassieke beleidsevaluaties zouden hierin tekort schieten. Zelfs een vrij traditioneel ingesteld instituut als de Algemene Rekenkamer experimenteert met elementen van lerende evaluaties. Daartoe heeft het zelfs een eigen onderzoek uitgevoerd naar de kwaliteit en het gebruik van de eigen beleidsaanbevelingen. Wellicht is een dergelijke zelfreflectie op de kwaliteit en het gebruik van de aanbevelingen (en conclusies) van het MNP ook een goede suggestie.
3. Binnen de lerende beleidsevaluaties zijn verschillen in de selectie van actoren die bij de evaluatie betrokken worden. Sommige lerende evaluaties betrekken alleen de beleidsactoren en andere beleidsinsiders, andere lerende evaluaties gaan verder en hebben ook de ambitie om burgers, slachtoffers en doelgroepen van het beleid actief bij de evaluatie te betrekken.
4. Combinaties van verantwoord en leren in de vorm van lerende evaluaties zijn mogelijk, maar zorgen in de praktijk voor soms lastige rolconflicten voor de onderzoekers. Drie van de vier onderzoekers in onze inventarisatie zoeken naar deze combinaties. Abma kiest duidelijk voor de puur op leren gerichte responsieve evaluatie.
5. Mogelijk kan het MNP leren van de ervaringen van evaluatieonderzoekers al dan niet binnen organisaties als de Algemene Rekenkamer en het RPB met lerende evaluaties.

Ook binnen een organisaties als de Algemene Rekenkamer worstelt men met het dilemma tussen de traditionele rol van onafhankelijke en afstandelijke beleidsevaluator en de nieuwe vormen van lerende en interactieve evaluaties. Voor het MNP is het aan te bevelen om te werken aan de uitwisseling van ervaringen met deze of andere instanties en onderzoekers op dit vlak.

4 Reflecties op MNP evaluaties

4.1 Inleiding

Dit rapport gaat over de mogelijkheden van lerende evaluaties voor het Milieu- en Natuurplanbureau (MNP). Lerende evaluaties zijn zowel gericht op leren als op verantwoorden. Het MNP voert elk jaar verschillende evaluaties uit. Het leek ons daarom verstandig om eens vanuit het kader van lerende evaluaties naar een aantal bestaande evaluatie-onderzoeken van het MNP te gaan kijken. Hiertoe hebben we in overleg met het evaluatieplatform van het MNP drie cases geselecteerd van afgeronde of nog lopende beleidsevaluaties. Daarbij gold als randvoorwaarde dat de projectleider van deze evaluatie ook zelf open stond voor een kritische reflectie op zijn/haar aanpak van de evaluatie en hier mogelijk van wilde leren.

De vraag die in dit hoofdstuk centraal staat is:

Welke methodische keuzes liggen ten grondslag aan bestaande MNP-evaluaties en in welke mate bevatten deze elementen van lerende evaluaties?

Bij de beantwoording van deze vragen is het kader uit hoofdstuk twee leidend geweest. Dat wil dus ook zeggen dat we de evaluaties vooral hebben beoordeeld op de mate waarin aandacht is geweest voor elementen van een lerende evaluatie. Indien we de evaluatie vanuit het gezichtspunt van een klassieke evaluatie hadden beoordeeld was de kritiek en de daarbij behorende aanbevelingen natuurlijk heel anders geweest. Het gaat dus nadrukkelijk om een beoordeling vanuit het perspectief van de lerende evaluatie met dus een combinatie van verantwoorden en leren.

Elke studie is kritisch 'beoordeeld' op de verschillende aspecten van lerende evaluaties. Hiermee willen we de keuzes die zijn gemaakt in deze evaluaties aan de orde stellen. Zoals is aangegeven in hoofdstuk 2, is het in een evaluatie belangrijk om aan te geven waarom voor een bepaalde evaluatievorm is gekozen. In deze reflecties zetten we vraagtekens bij bepaalde keuzes om daarmee de discussie over de mogelijkheden en onmogelijkheden van lerende evaluaties voor het MNP te voeden. Het gaat dus nadrukkelijk niet om een integraal kwaliteitsoordeel over de MNP-evaluaties.

Uiteindelijk zijn de volgende MNP-evaluaties geselecteerd:

1. Evaluatie 'Schuiven op zand' (over de reconstructiewet zandgebieden)
2. Evaluatie 'beleidsomslag van verwerving naar beheer'
3. Evaluatiemethodiek transitie

Hiertoe is telkens het onderzoeksrapport (en of andere projectstukken) bestudeerd en is voor elke studie een interview afgenomen met de projectleider van de evaluatie. De kritische kanttekeningen zijn voorgelegd aan deze onderzoekers, maar zijn uiteindelijk voor rekening van de auteurs van dit rapport.

4.2 Evaluatiestudie 'Schuiven op zand'

4.2.1 Evaluatieonderzoek

Het rapport 'Schuiven op zand' (Van Wezel et al., 2004) behelst een ex ante evaluatie van de twaalf (ontwerp-) reconstructieplannen die zijn opgesteld in het kader van de Reconstructiewet concentratiegebieden. Het rapport analyseert de reconstructieplannen inclusief de uitvoeringsprogramma's die daar onderdeel van uitmaken, waarbij is uitgegaan van de 'stand van de planvorming' in september 2004. De reconstructieplannen zijn getoetst naar de te verwachten realisatie van de doelen zoals aangegeven in de Reconstructiewet en in navolgende brieven van de minister van LNV aan de Tweede Kamer. Daarbij wordt opgemerkt dat deze doelen in kwalitatieve termen zijn gesteld, terwijl in aanpalende nationale en internationale regelgeving op het gebied van natuur, water, landbouw en milieu wel kwantitatieve doelen zijn vastgelegd. Om die reden wordt in de evaluatiestudie ook onderzocht in hoeverre uitvoering van de reconstructieplannen bijdraagt aan het behalen van deze sectorale beleidsdoelen.

Het rapport geeft allereerst een overzicht van de huidige situatie en autonome ontwikkeling op het gebied van milieu, water, natuur, landschap en landbouw in de twaalf reconstructiegebieden en beschrijft de ontwikkelingen in relevante regelgeving. Dan volgt een beschrijving van de twaalf reconstructieplannen: hoe de zoneringsplannen eruit zien voor het totale reconstructiegebied, hoe de uitgaven verdeeld zijn over de diverse thema's en hoe het beschikbaar juridisch instrumentarium gebruikt wordt. Vervolgens worden de verwachte effecten van de reconstructieplannen op verduurzaming van de landbouw, kwaliteitsverbetering van landschap en natuur, waterbeheer en milieu beschreven.

De conclusies van het rapport hebben betrekking op de thema's bestuurlijke aspecten, zoneringsplannen, natuur en landschap, waterbeheer, duurzame landbouw, veterinaire kwetsbaarheid en milieu. Allereerst wordt geconcludeerd dat er draagvlak is voor de reconstructie bij de partijen die betrokken zijn geweest bij de planvorming. Dit wordt als een belangrijke succesfactor voor de uitvoering van de reconstructieplannen gezien. Daarnaast stelt het rapport dat de reconstructieplannen tot een sterkere ruimtelijke scheiding van de functies landbouw, wonen en natuur leiden. Daarbij wordt opgemerkt dat nabij natuurgebieden in veel provincies ruimte blijft voor de uitbreiding van zittende bedrijven. Een andere kanttekening die wordt gemaakt is dat de zoneringsplannen veelal betrekking heeft op kleine gebieden, terwijl zoneringsplannen op grotere gebieden volgens het rapport effectiever kan zijn.

Op de derde plaats concludeert het rapport dat de twaalf reconstructieplannen gezamenlijk een investering van €7 miljard vragen voor de periode tot 2016, waarbij vooral geïnvesteerd wordt in natuurverwerving, waterbeheer en landbouw. Wat betreft de investering in natuurverwerving en waterbeheer concludeert het rapport dat uitvoering van de reconstructieplannen bij zal dragen aan de realisatie van de Ecologische Hoofdstructuur (EHS) en dat de beoogde investeringen in waterberging, beekherstel en verdrogingsbestrijding bij kunnen dragen aan het doelbereik voor kwantitatief waterbeheer. De investeringen in de landbouw zijn volgens de onderzoekers vooral gericht op verbetering van het toekomstperspectief van de intensieve veehouderij, terwijl voor het voorkomen van dierziekten nauwelijks concrete maatregelen zijn voorzien.

Bovendien stelt het rapport dat ondanks de grote milieupoging van de reconstructieplannen zich evenmin richten op vermindering van de nutriëntenstromen. Een laatste conclusie van het rapport luidt dan ook dat de reconstructieplannen geen substantiële bijdrage leveren aan de verbetering van milieu- en natuurkwaliteit op gebiedsniveau. In de plannen wordt op

deze punten verwezen naar generiek beleid en autonome ontwikkelingen. De onderzoekers vinden dit een risicovolle keuze, omdat onzeker is of het generieke milieubeleid (met name op het gebied van mest) wel effectief zal zijn.

4.2.2 Analyse van de evaluatie

Vorbereiding evaluatieonderzoek

Het beleidsveld rond de reconstructie is in de voorbereiding van het evaluatieonderzoek niet geanalyseerd. Dit is wel gebeurd in het kader van een deelproject dat uitbesteed is aan de Universiteit van Utrecht en een aparte rapportage heeft opgeleverd, getiteld 'Platteland in beweging' (Driessen en De Gier, 2004). Dit deelonderzoek beschrijft het proces van de totstandkoming van de reconstructieplannen, onder meer op basis van interviews met betrokkenen. Ook is een analyse gemaakt van de plaats die de Reconstructiewet inneemt binnen het omgevingsrecht en is onderzocht welke belemmeringen deze wetgeving opleveren voor de totstandkoming van de reconstructieplannen (zie Driessen en De Gier 2004). De resultaten van dit onderzoek zijn in samengevatte vorm opgenomen in het rapport 'Schuiven op Zand' (onder het kopje 'bestuurlijke aspecten', p. 60).

Het rapport is geschreven in opdracht van de directie van het MNP, mede naar aanleiding van een verzoek van de ministeries van VROM en LNV. Het MNP constateert dat er veel energie gestoken is in het opstellen van reconstructieplannen en dat er ook veel geld mee gemoeid is. Het MNP heeft dan ook de behoefte om te evalueren wat de bijdrage van de reconstructieplannen is aan de realisatie van de doelen uit de Reconstructiewet en van sectorale doelen die voortvloeien uit aanpalend beleid. Daarbij speelt een rol dat op het moment van evaluatie de reconstructieplannen op het punt staan goedgekeurd te worden door provinciale bestuurders en betrokken ministers, als gevolg waarvan aanzienlijke budgetten beschikbaar komen, terwijl de betrokkenheid van de Tweede Kamer op dat moment matig is. Daarnaast bestaat ook de behoefte om een overzicht te hebben van de inhoud van de verschillende reconstructieplannen.

Dezelfde kennisbehoefte bestaat ook bij het ministerie van VROM (Directie Bodem, Water, Landelijk Gebied). De betrokken medewerkers van VROM zoeken contact met de directies Platteland en Regionale Zaken van LNV, die sterk betrokken zijn bij de reconstructieplannen. Vervolgens wordt een formele adviesaanvraag opgesteld door de DG van het ministerie van LNV, mede namens het ministerie van VROM, waardoor MNP aan de slag kan met het onderzoek. De adviesaanvraag bestaat uit een verzoek de reconstructieplannen te evalueren. De directie van het MNP heeft het onderzoek uit eigen middelen gefinancierd. Het MNP wordt 'lump sum' gefinancierd en beslist dus zelfstandig over het onderzoek dat zij uitvoert. Daarbij wordt wel rekening gehouden met de bestaande kennisbehoefte, in dit geval van de ministeries van VROM en LNV. MNP wil met het onderzoek echter niet voldoen aan de behoefte van VROM om de standpunten van het ministerie te ondersteunen in de belangstrijd met de lagere overheden.

De primaire doelgroep van het onderzoek zijn de bewindslieden van VROM en LNV en Tweede Kamerleden. De secundaire doelgroep bestaat uit de medewerkers van de twee ministeries en andere bij de reconstructie betrokken partijen (op rijks- en provinciaal niveau).

De onderzoekers kiezen voor een ex ante evaluatie van de twaalf (ontwerp-) reconstructieplannen. Het gaat om een klassieke beleidsevaluatie waarin het verwachte doelbereik van de reconstructieplannen getoetst wordt. De onderzoekers zijn daartoe bewust teruggeslagen naar de oorspronkelijke bedoeling van de Reconstructiewet. Uitgangspunt van de evaluatie zijn dus de formele doelen zoals die gesteld zijn in de wettekst en in navolgende

brieven van de minister van LNV aan de Tweede Kamer. Daar zijn doelen aan toegevoegd uit aanpalende nationale en internationale regelgeving, al dan niet in voorbereiding. Vooral het ministerie van VROM heeft behoefte aan een dergelijke evaluatie, waarin ook doelen uit aanpalend beleid meegenomen worden. Voor het ministerie van LNV geldt dat niet. Wat daarnaast een rol speelt bij deze keuze is dat de doelen zoals geformuleerd in de wettekst wel niet gekwantificeerd zijn en daarmee dus moeilijk operationaliseerbaar. Daarentegen zijn veel doelen zoals geformuleerd in (de grote hoeveelheid aan) nationaal en internationaal beleid en regelgeving op het gebied van natuur, water, landbouw en milieu wel in kwantitatieve termen geformuleerd. Dit helpt de onderzoekers bij het operationaliseren van de evaluatiecriteria.

De doelen van de lagere overheden die betrokken zijn bij de reconstructie (provincies, gemeenten en waterschappen) worden niet meegenomen in de evaluatie. De focus van het onderzoek ligt op de bijdrage van de reconstructieplannen aan het rijksbeleid, mede vanuit de overweging dat er veel rijks gelden gemoeid zijn met de reconstructie. Daarnaast is slechts in beperkte mate gekeken naar de doelen van de reconstructie ten aanzien van recreatie, wonen, werken en leefklimaat en economische structuur. De belangrijkste reden daarvoor is dat de uitvoeringsprogramma's nauwelijks in maatregelen op deze terreinen blijken te voorzien. Er wordt gekozen voor een focus op de thema's waar het grootste budget voor gereserveerd is. Een andere reden voor deze keuze is dat het MNP/RIVM niet over de benodigde expertise beschikt. Tenslotte behelst het onderzoek uitsluitend een evaluatie van de reconstructieplannen zelf. In het ontwikkelingsgericht werken na de planvorming en de bestuursovereenkomsten had het MNP/RIVM op het moment van het onderzoek nog geen inzicht.

Het voorwerp van de evaluatie zijn de verwachte resultaten van de twaalf reconstructieplannen, inclusief de uitvoeringsprogramma's die daar onderdeel van uitmaken. Daarbij wordt niet gekeken naar mogelijke verklaringen voor deze effecten. Wel worden de verwachte resultaten van de reconstructieplannen afgezet tegen de huidige situatie en de autonome ontwikkeling in de reconstructiegebieden ten aanzien van de verschillende thema's waarop geëvalueerd wordt: milieu en water, natuur en landschap en landbouw.

Methodiek evaluatieonderzoek

Zoals gezegd is het beoordelingskader van het onderzoek gebaseerd op de formele beleidsdoelen van de Reconstructiewet, aangevuld met sectorale beleidsdoelen uit aanpalende regelgeving. De reconstructiedoelen zijn als volgt geoperationaliseerd (p. 11): 1. duurzame landbouw (onder andere verbeteren ruimtelijke structuur en inrichting); 2. het verminderen van de veterinaire kwetsbaarheid; 3. duurzaam waterbeheer (verbetering kwaliteit, berging, verdrogingsbestrijding en beekherstel); en 4. de realisering van de EHS (reductie ammoniakdepositie, verdrogingsbestrijding en ontsnippering). De sectorale doelen uit aanpalend beleid worden in het rapport niet expliciet genoemd.

De gegevensverzameling bestaat uit het analyseren van de reconstructieplannen en aanvullende literatuur. Het analyseren van alle plannen, kaarten, financiële overzichten is een zeer omvangrijke klus gebleken. Daarnaast is een aantal GIS analyses uitgevoerd. Daarbij is gekeken naar hoe de plannen er in ruimtelijk opzicht uitzien: waar komen de landbouwontwikkelings-, verwevings- en extensiveringsgebieden te liggen, wat voor schaalniveau is voor deze gebieden gekozen en wat zijn de regionale verschillen daarin? Vervolgens zijn de uitvoeringspakketten geanalyseerd: hoe zijn de middelen verdeeld over de reconstructiegebieden en wat zijn de belangrijkste kostenposten? Dat laatste is een lastig onderdeel van de analyse, omdat sommige onderdelen van de plannen onder verschillende categorieën geschaard kunnen worden. Zo kan de verplaatsing van landbouwbedrijven zowel onder de categorie natuur als onder de categorie landbouw gerekend worden. Uiteindelijk is

het bij de categorie landbouw ingedeeld omdat de bijdrage van bedrijfsverplaatsing voor natuur op dat moment gering is. Hier is na het verschijnen van het rapport kritiek op gekomen (zie Lamers, 2005; Franken en Van Wezel, 2005). Vervolgens is ook de juridische kant van de reconstructieplannen onderzocht, waarbij onderzocht is in hoeverre bestaande instrumenten ingezet worden, waaronder de bestaande regelgeving op het gebied van emissie door de intensieve veehouderij, de ruimte-voor-ruimte regeling en de doorwerking naar streek- en bestemmingsplannen.

Daarop volgt de beoordeling van de reconstructieplannen. Daarbij staat centraal: wat draagt de zonering bij en in hoeverre dragen de plannen bij aan verduurzaming van de landbouw, de kwaliteit van natuur en landschap, de realisatie van de waterdoelen en de milieukwaliteit in de reconstructiegebieden. Op basis van de voor de reconstructieplannen opgestelde milieueffectrapportages worden ten slotte alternatieve beleidsmaatregelen geanalyseerd, waarbij vooral gekeken is naar de meerkosten van extra milieumaatregelen.

Behalve de reconstructieplannen, bijbehorende uitvoeringsprogramma's en achtergrond-documentatie zijn in het kader van het onderzoek geen andere bronnen gebruikt. In het kader van het deelrapport 'Platteland in beweging' zijn wel interviews gehouden met betrokkenen, waaronder vertegenwoordigers van provincies, LTO en milieubeweging (zie Driessen en De Gier, 2004). Uit dit deelonderzoek blijkt onder andere dat de betrokkenen tevreden zijn met het proces en dat er in de verschillende gebieden draagvlak is voor de reconstructieplannen, maar dat onduidelijk is wat de reconstructie moet bijdragen aan het realiseren van de verschillende sectorale beleidsdoelen en wat verwacht moet worden van ander ingezet beleid. Deze conclusies komen in het rapport 'Schuiven op zand' terug bij de eindconclusies (p. 60).

De onderzoekers concluderen dat de reconstructieplannen vooral bijdragen aan het realiseren van de doelen ten aanzien van kwantitatief waterbeheer en aan de realisatie van de EHS. Aan de realisatie van een aantal andere doelen is de bijdrage echter gering. Dit geldt voor de doelen op het gebied van milieu, maar ook om de doelen ten aanzien van de veterinaire kwetsbaarheid, terwijl dat laatste dé aanleiding was voor de reconstructie. In de visie van de onderzoekers vormt de reconstructie in de praktijk dan ook niet de integratie van al het bestaand beleid voor het landelijk gebied. Deze gedachte komt onder meer tot uitdrukking in het schema op p. 57.

Bij de analyse van het onderzoeksmateriaal is geen sprake geweest van grote verschillen in inzicht tussen de verschillende onderzoekers in het projectteam. De analyses zijn tussentijds niet voorgelegd aan de bij de gebiedsprocessen betrokken partijen, maar wel aan de ministeries van VROM en LNV, inclusief de LNV Directie Regionale Zaken. Laatstgenoemden zijn niet blij met de resultaten van het onderzoek. Voor hen vormen de onderzoeksresultaten een verstoring van het reconstructieproces.

Uitvoering evaluatieonderzoek

Bij het projectteam zijn vijf personen betrokken. De projectleider verzorgt de coördinatie binnen het projectteam en stuurt (met de vervangend projectleider) het uitbestede deelproject aan. Daarnaast voert de projectleider de analyses van de uitvoeringsprogramma's uit en schrijft ze, samen met de vervangend projectleider, het eindrapport. De andere onderzoekers hebben zich respectievelijk bezig gehouden met het uitvoeren van economische analyses, het uitvoeren van GIS-analyses, het maken van kaarten en figuren en het schrijven aan het eindrapport. Daarnaast hebben mensen van buiten het projectteam gereflecteerd op de resultaten van het onderzoek, waaronder de directeur van MNP en de leden van de klankbordgroep (zie hieronder).

Voorafgaand aan het evaluatieonderzoek is een klankbordgroep in het leven geroepen om het projectteam van advies te voorzien. In deze klankbordgroep zijn de betrokken ministeries vertegenwoordigd, inclusief de Directie Regionale Zaken van LNV. Tijdens het onderzoek is de klankbordgroep een aantal malen bij elkaar geweest. Zo zijn presentaties gehouden over het onderzoek en zijn conceptteksten voorgelegd aan de klankbordgroep. De leden van de klankbordgroep hebben constructieve vragen gesteld en opmerkingen gemaakt die door de onderzoekers zijn verwerkt in het onderzoek. De vragen en opmerkingen zijn hoofdzakelijk inhoudelijk van aard (en niet methodisch). Daarbij gaat het onder meer om de selectie van onderwerpen die aan de orde moeten komen en de interpretatie van het onderzoeksmateriaal.

Andere partijen die betrokken zijn bij de reconstructie (lagere overheden, maatschappelijke actoren) zijn niet in het onderzoeksproces betrokken. Wel hebben de vijf betrokken provincies medewerking verleend en inzage gegeven in de conceptplannen en andere relevante informatie zoals kaarten en achtergrondstudies. Bovendien zijn de onderzoeksresultaten voorafgaand aan de publicatie van het eindrapport voorgelegd en toegelicht aan de vijf betrokken provincies.

Gebruik evaluatieonderzoek

In reactie op het rapport heeft minister Veerman van LNV namens het kabinet een brief geschreven aan de Tweede Kamer. Daarin kondigt hij een vervolgonderzoek aan, uit te voeren door Alterra (zie ook paragraaf 5.3), omdat hij van mening is dat aan een aantal aspecten, waaronder recreatie en leefklimaat, maar ook aan de proceskant van (uitvoering van) de reconstructie, te weinig aandacht is besteed in 'Schuiven op zand'.

Er bestaat grote onvrede over het rapport bij de Directie Regionale Zaken van LNV. Dat heeft te maken met de conclusies van het rapport, waarin van een aantal onderdelen van de reconstructie (milieu en veterinaire kwetsbaarheid) wordt gezegd dat de reconstructieplannen daar slechts een geringe bijdrage aan leveren. De regionale directies van LNV hebben met bloed, zweet en tranen gewerkt aan de totstandkoming van de reconstructieplannen en zijn dus ongelukkig met de conclusies en de timing van het rapport. De projectleider van het onderzoek vindt dit begrijpelijk, maar benadrukt dat hun rapport ook de positieve effecten van de reconstructieplannen belicht.

In een aantal provincies is het rapport bij de behandeling van de reconstructieplannen in de Provinciale Staten bediscussieerd. Daarbij is onder meer verwezen naar de extra milieumaatregelen die in het rapport geopperd worden. Ook Groen Links heeft het rapport gebruikt om de extra milieumaatregelen onder het voetlicht te brengen.

Het onderzoek is, nog voordat het eindrapport officieel verschenen was, gepresenteerd op een Ruimteconferentie van het Ruimtelijk Planbureau. Als gevolg daarvan zijn de conclusies van het rapport uitgelekt, omdat er journalisten aanwezig waren bij de conferentie. Vermoedelijk was dit op uitnodiging van de milieubeweging, die aandacht wil voor de conclusie van het onderzoek dat de reconstructieplannen niet bijdragen aan de realisatie van milieudoelen op gebiedsniveau. Dit leverde een vervelende situatie op voor de onderzoekers. Het is namelijk gebruikelijk dat de minister eerst reageert op een rapport voordat het in de openbaarheid verschijnt. Over deze gang van zaken is de minister dan ook *'not amused'*, ook al ligt op dat moment het conceptrapport al bij LNV en is de minister dus al bekend met de conclusies. Hier komt bij dat de kop van één van de krantenartikelen luidt: 'RIVM kraakt reconstructieplannen af', terwijl de kop van het MNP persbericht benadrukt dat de reconstructie goed is voor landbouw en natuur, maar niet goed voor milieu en veterinaire kwetsbaarheid. In de media worden alleen de negatieve conclusies belicht.

Achteraf concludeert de projectleider van het onderzoek dat bepaalde keuzen die gemaakt zijn in het onderzoek beter beargumenteerd hadden moeten worden in het eindrapport. Ook is in de loop van het onderzoek het contact met LNV niet altijd even intensief geweest, mede als gevolg van personele wisselingen bij het ministerie. Om de betrokkenheid van de opdrachtgever bij het onderzoek te vergroten zou het wenselijk geweest zijn intensiever contact te hebben. Tegelijkertijd benadrukt de projectleider van het onderzoek de rol van MNP als onafhankelijk planbureau en de noodzaak om in die rol sociaal ongewenste conclusies te trekken. Het rapport zelf bevat volgens de onderzoekers een heldere boodschap en is bovendien vernieuwend, aangezien het als eerste een algemeen overzicht geeft van de inhoud van de twaalf reconstructieplannen.

Wat beleidsevaluatie in een governance context betreft benadrukt de projectleider het belang van het toetsen op doelbereiking: wat is het oorspronkelijke doel van de reconstructie en wat wordt daarvan wel en niet gerealiseerd middels de reconstructieplannen? Het feit dat alle betrokken partijen tevreden zijn met de plannen zegt niet dat het goede plannen zijn waarin alle verschillende belangen vertegenwoordigd zijn. Niet alle belangen in het landelijk gebied zijn goed vertegenwoordigd geweest in de processen waarin de reconstructieplannen tot stand zijn gekomen. Alleen toetsen op de tevredenheid van de bij de planvorming betrokken partijen levert in de visie van de projectleider dus geen volledig beeld op van het maatschappelijk nut van de reconstructieplannen.

De rol van 'onafhankelijk onderzoeker op afstand' is daarbij niet altijd gemakkelijk. Er moet een balans gevonden worden tussen enerzijds waardering hebben voor de complexe processen waarin plannen tot stand komen en anderzijds kritiek leveren op de punten waar dat nodig is. In de visie van de projectleiding is een goede communicatie over het onderzoek daarbij van groot belang, zodat niet alleen de negatieve punten uit het eindrapport opgepikt worden.

4.2.3 Conclusie en reflectie

Uit de voorafgaande analyse trekken we een aantal conclusies met betrekking tot de voorbereiding, de methodiek, de uitvoering en het gebruik van het evaluatieonderzoek 'Schuiven op zand'. Op de eerste plaats heeft de keuze voor een klassieke beleidsevaluatie er toe geleid dat in het eindrapport weinig aandacht is besteed aan de processen waarin de reconstructieplannen tot stand zijn gekomen. Het is zeer aannemelijk dat het verloop van deze processen een goede verklaring biedt voor de uiteindelijke inhoud van de reconstructieplannen. Doordat deze processen niet in de evaluatie zijn betrokken, is dan ook een kans gemist om inzicht te verkrijgen in de afwegingen die ten grondslag hebben gelegen aan de inhoud van de reconstructieplannen.

Daarnaast is met de keuze voor een klassieke beleidsevaluatie, waarbij teruggegrepen wordt op de oorspronkelijke doelen van de reconstructie, voorbijgegaan aan het feit dat de verschillende betrokken partijen verschillende doelen nastreven met de reconstructie en dat deze doelen bovendien in de loop van de jaren zijn verschoven. Ten slotte is de visie die de onderzoekers hebben op de reconstructie opvallend te noemen. Terwijl de reconstructie voor veel betrokkenen een middel is om bestaand sectoraal beleid integraal aan te pakken, zien de onderzoekers de reconstructie als één van de vele beleidssporen die relevant zijn voor het landelijk gebied. Als direct gevolg van deze visie op de reconstructie, besteedt de evaluatie geen aandacht aan de meerwaarde die de reconstructieplannen hebben vanwege het feit dat er integrale belangenafwegingen aan ten grondslag liggen waarin de verschillende sectorale belangen (natuur en landschap, milieu, water, landbouw, recreatie en leefklimaat) zijn meegenomen.

De gehanteerde onderzoeksmethode heeft bijgedragen aan een enigszins eenzijdige en statische benadering van de reconstructieplannen. Door alleen de tekst van de reconstructieplannen en bijbehorende uitvoeringsprogramma's te analyseren zonder door middel van interviews de visie van de verschillende betrokkenen op (de totstandkoming van) deze plannen en programma's in beeld te brengen is het sociaal-politieke karakter van de reconstructie geheel buiten beschouwing gelaten. Ook zijn de methodische keuzes die gemaakt zijn tijdens het onderzoek, in het eindrapport niet altijd duidelijk verantwoord. Veel methodische keuzes zijn wel benoemd maar zijn niet of nauwelijks beargumenteerd. Dat geldt bijvoorbeeld voor de afbakening van het onderzoek, maar ook voor de keuzes die gemaakt zijn ten aanzien van de gebruikte onderzoeksmethoden. Bovendien zijn de evaluatiecriteria die gehanteerd zijn niet gemakkelijk te achterhalen in de tekst van het eindrapport, met name de criteria die gebaseerd zijn op de kwantitatieve doelen uit aanpalend beleid.

Bij de uitvoering van het onderzoek valt op dat geen link is gelegd met de beleidspraktijk in de reconstructiegebieden. Daarmee zijn de onderzoeksresultaten ver van de beleidsprocessen op gebiedsniveau af komen staan. Voor het uiteindelijke gebruik van het eindrapport heeft dit bovendien als gevolg gehad dat de conclusies van het eindrapport door de bij de reconstructie betrokken partijen (met uitzondering van de milieubeweging) negatief zijn ontvangen. Wanneer de verschillende belanghebbenden bij de reconstructie intensiever bij het onderzoek betrokken waren geweest, en ze bijvoorbeeld de ruimte hadden gekregen om de onderzochte thema's mede te bepalen en mee te denken over de gehanteerde evaluatiecriteria, had dit wellicht geleid tot een proces van gezamenlijk leren en een groter draagvlak voor de conclusies van het onderzoek.

4.3 Evaluatiestudie 'Van aankoop naar beheer'

4.3.1 Evaluatieonderzoek

Het evaluatieonderzoek 'Van aankoop naar beheer' (Van Egmond en De Koeijer, 2005) is een evaluatie van de omslag in het natuurbeleid van grondaankopen naar beheer door particulieren en agrariërs. Deze omslag moet nog grotendeels zijn beslag krijgen. Sinds 2002 wil het kabinet en de rijksoverheid agrariërs en andere particulieren sterker betrekken bij het natuurbeheer. In plaats van het aankopen van grond door grote terreinbeherende natuurorganisaties zoals Vereniging Natuurmonumenten, Staatsbosbeheer en de Landschappen wil het kabinet en de rijksoverheid particulieren en agrariërs aanzetten tot natuurbeheer in de hoop hiermee het draagvlak voor het natuurbeleid te vergroten en de kosten voor de realisering van de Ecologische Hoofdstructuur (EHS) te drukken. Het Milieu- en Natuurplanbureau is in opdracht van het Ministerie van LNV, Directie Natuur nagegaan of die verwachting realistisch is. Het MNP heeft hiertoe drie verschillende organisatievormen van natuurbeheer (terreinbeheerders, particulieren en agrariërs) met elkaar vergeleken op hun consequenties voor ecologie, kosten en deelnamebereidheid aan natuurbeheer.

Het MNP concludeert dat de omslag onder de huidige condities niet het gewenste effect zal hebben. Voor particulier natuurbeheer is vooralsnog weinig animo en agrarisch natuurbeheer zal met de huidige regelingen en beschikbare financiën zelden de natuurdoelen kunnen halen die terreinbeheerders wel kunnen halen.

Kosten

De jaarlijkse kosten van *particulier natuurbeheer* zijn voor het ministerie van Landbouw, Natuurbeheer en Voedselkwaliteit (LNV) gemiddeld twintig procent lager dan beheer via

terreinbeherende natuurorganisaties. Dit komt doordat LNV naast de beheerkosten alleen de waardedaling hoeft te vergoeden van de grond die de bestemming natuur krijgt.

De kosten bij *agrarisch natuurbeheer* zijn zestig procent lager dan bij uitvoering door terreinbeherende organisaties, omdat de overheid naast de kosten voor het gevoerde beheer ook de opbrengstderiving door natuurbeheer compenseert. Wel wordt met de omslag in het natuurbeleid een deel van de uitgaven naar een latere periode verschoven. Tot 2018 treedt een besparing op van circa 650 miljoen euro. Na realisatie van de EHS komen er extra uitgaven van ongeveer 400 miljoen euro, omdat het beheer blijvend duurder is.

Deelnamebereidheid

Uit marktonderzoek van het Nationaal Groenfonds blijkt dat de deelnamebereidheid onder *particulieren* op termijn niet voldoende zal zijn om de areaaltaakstelling van ruim 40.000 hectare particulier natuurbeheer in 2018 te realiseren. De knelpunten zitten vooral in het aanvraagtraject: financiële onzekerheden, eenzijdige financiële risico's bij het mislukken van het aanvraagtraject en een gebrek aan kennis bij potentiële deelnemers en betrokken instanties zoals gemeenten.

De deelnamebereidheid onder agrariërs is veel rooskleuriger. Zij zijn over het algemeen tevreden over de regeling voor agrarisch natuurbeheer. Tachtig procent wil het agrarisch natuurbeheer continueren en zelfs uitbreiden na afloop van het nu lopende contract. Bedrijfseconomische analyses geven aan dat er een grens ligt bij circa 50% van de bedrijfsoppervlakte die onder agrarisch natuurbeheer kan worden gebracht en dat daarboven de bedrijfseconomische kosten hoger zijn dan de te ontvangen vergoeding. Agrariërs kiezen overwegend voor de lichtere vormen van agrarisch natuurbeheer.

Ecologische doelen

Om dezelfde natuurdoelen te verwezenlijken als met terreinbeheer zou de regeling voor agrarisch natuurbeheer moeten worden aangepast en zou de continuïteit moeten worden gewaarborgd, bijvoorbeeld via een oplopende vergoeding. Daarnaast zijn vaak inrichtingsmaatregelen nodig betreffende de waterhuishouding om natuurdoelen te realiseren.

4.3.2 Analyse van de evaluatie

Vorbereiding evaluatieonderzoek

Aan het onderzoek is geen beleidsanalyse vooraf gegaan. Wel is een aantal actoren onderscheiden. Het gaat dan met name om de (rijks)overheid en de beheerders. De beheerders zijn weer in drie typen opgedeeld: grote terreinbeherende natuurorganisaties, particulieren en agrariërs. In het onderzoek is vooral geëvalueerd vanuit het perspectief van de rijksoverheid en de formele beleidsdoelen, en niet vanuit het perspectief van de beheerders. De doelen van de beheerders heeft men wel in het achterhoofd gehouden, omdat het bepaalt wat een beheerder gaat doen. Maar in het onderzoek stond het overheidsdoel centraal. *'Als we naar de doelen van de beheerder keken, dan was het van: wat gaat de beheerder doen en worden daarmee de doelen van de overheid gerealiseerd?'* Van een beleidsanalyse vanuit een multi-actor perspectief was dus geen sprake. Naar de rol van bijvoorbeeld agrarische natuurverenigingen, LTO-afdelingen of verenigingen van kleine landgoedeigenaren is *niet* gekeken. Als reden hiervoor is aangevoerd dat het onderzoek in een beperkte tijdsspanne moest worden uitgevoerd. Het ging in het onderzoek om de ecologische kant, om de economische kosten voor de overheid en om de deelnamebereidheid onder beheerders. De eerste twee aspecten zijn meetbaar en daarvoor was, in de visie van de onderzoekers, de betrokkenheid van andere actoren niet nodig. Bij deelnamebereidheid ligt dit iets anders, maar ook daar is dit niet gebeurd.

De Directie Natuur van LNV is de opdrachtgever van het onderzoek. De vraag voor wie het MNP evalueert is verder niet zo expliciet aan de orde geweest. De beheerders zijn bijvoorbeeld niet vooraf gevraagd naar hun behoefte aan evaluatieonderzoek. Wel gingen de onderzoekers er vanuit dat de beheerders mogelijke gebruikers waren van het onderzoek. Zij kunnen hun voordeel doen met de onderzoeksresultaten. Na het onderzoek is aan de beheerders wel gevraagd wat ze vonden van de onderzoeksresultaten. De reden voor de nadruk in het onderzoek op LNV en de LNV-doelen en het beperkt meenemen daarin van het perspectief van de beheerders, lag volgens het MNP besloten in de onderzoeksvragen. Deze luiden: Wat is het effect van de omslag voor het natuurbeleid? En: Kunnen we de doelen van het beleid toch nog halen, nadat het beleid is aangepast? Het ging de opdrachtgever dus niet om de doelen van de beheerder en daarom zijn deze ook niet expliciet aan de orde gekomen.

De onderzoekers hebben ook nooit overwogen om de drie verschillende beheerdergroepen medeopdrachtgever te maken dan wel op te nemen in een begeleidingscommissie. Indien deze groepen opgenomen zouden worden in een begeleidingscommissie hebben zij ook echt iets te zeggen over de resultaten. Dat is iets wat volgens de onderzoekers echt niet kan binnen het MNP. Er is wel overwogen om hen in een platform op te nemen, maar daar is uiteindelijk ook van afgezien. De reden hiervoor was dat er nooit echt een start is geweest van dit project. De onderzoekers waren al jaren bezig met agrarisch natuurbeheer in verschillende projecten. Daar is op den duur wel een rapport van gemaakt. Het was dus een proces van vele jaren, zonder een duidelijk startpunt. Omdat er geen duidelijk startpunt was, kon men ook niet bij aanvang bij deze groepen langs gaan. Pas aan het eind was het project echt duidelijk en kon men ook met de resultaten bij deze partijen langs gaan.

De belangrijkste reden voor de evaluatie is volgens de onderzoekers dat de Minister van LNV aan de Tweede Kamer heeft beloofd om aan te geven wat de effecten zijn van de omslag in het natuurbeleid. In hoeverre er (andere) politieke motieven zijn of dat (de ambtenaren van) LNV met het onderzoek een omslag terug zou willen maken, dat is door de onderzoekers niet nagegaan. De vragen van de opdrachtgever zijn voor kennisgeving aangenomen. Wel hebben zij aan hun opdrachtgever voorgelegd hoe zij de vragen wilden beantwoorden, waarbij zij instemming kregen. De onderzoekers hebben niet doorgevraagd naar de redenen achter de vragen. Het voornemen is om dit bij deel II van dit onderzoek wel te doen. Daarbij vragen ze zich wel af of LNV als opdrachtgever überhaupt de ruimte biedt om die vragen te stellen. De praktijk is dat ze gewoon een kort mailtje of belletje krijgen van LNV met de boodschap: "dit zijn in het kort mijn vragen".

De evaluatie is gericht op verantwoording aan de Tweede Kamer en niet op leren (voor beheerders bijvoorbeeld). De onderzoekers hebben gekozen voor een ex durante evaluatie waarbij de nadruk lag op ex post (in deel 1). Voor de aansluiting op het beleidsproces is er op tijd geëvalueerd: de evaluatie was gereed voordat de rapportage aan de Tweede Kamer af moest zijn. Momenteel is het MNP bezig met deel 2 van een rapportage over de omslag in het natuurbeleid.

De evaluatie was zowel gericht op organisatie, proces, inhoud en effecten. Ook in de organisatie van Programma Beheer zijn er factoren waardoor de deelname aan agrarisch natuurbeheer zo laag is, en worden de ecologische doelen van het natuurbeleid niet gerealiseerd via SAN door de vormgeving van de pakketten. De onderzoekers hebben wel geprobeerd de resultaten te verklaren en hebben hier verschillende bronnen voor gebruikt. Ook is rond het probleem van de deelnamebereidheid door de onderzoekers een workshop georganiseerd om zicht te krijgen op de factoren die dit bepalen. Het idee bestaat dat LNV met deze uitkomsten hiervan wel aan de gang is gegaan.

Methodiek evaluatieonderzoek

Een van de belangrijke onderdelen in een evaluatieonderzoek is het opstellen van een evaluatiekader. In dit geval is dit kader gebaseerd op de formele beleidsdoelen van het rijk. Er is nooit overwogen om andere actoren (bijvoorbeeld de beheerders) een rol te geven bij het opstellen van het beoordelingskader, die verder ging dan die van informant. In lerende evaluatie worden soms de informanten ook betrokken als (mede)opsteller van het evaluatiekader of wordt dit kader mede gebaseerd op hun doelen. Dit is vanwege de vraagstelling van LNV in dit onderzoek niet aan de orde geweest. Bovendien zijn deze groepen dicht betrokken bij het Programma Beheer. De onderzoekers hebben juist geprobeerd daar zo ver mogelijk van af te staan. De band met deze partijen was juist minimaal om op een onafhankelijke manier te kunnen bepalen wat er geëvalueerd moest worden. Het MNP heeft de band met de beheerders zelf zo zakelijk mogelijk willen houden om objectief te kunnen vaststellen in hoeverre de natuurdoelen zijn gerealiseerd. Zij zijn niet betrokken om te bekijken wat het MNP moest evalueren: dat stond vast op basis van de vraag van de Tweede Kamer (doelen natuurbeleid).

De operationalisering van de formele doelen is volledig door de onderzoekers zelf uitgevoerd. Zo hebben de onderzoekers er voor gekozen om het vergroten van draagvlak op te vatten als deelnamebereidheid van agrariërs en is binnen agrarisch natuurbeheer vooral gefocust op weidevogelbeheer. Daarbij zijn degenen die agrarisch natuurbeheer daadwerkelijk dienen uit te voeren niet betrokken. De opdrachtgever was wel op de hoogte van de onderzoeksvragen en de aanpak. Binnen het onderzoek is zowel kwantitatief onderzoek als kwalitatief onderzoek uitgevoerd. Daar waar kwantitatief onderzoek mogelijk was genoot dit de voorkeur van de onderzoekers. Maar er zijn ook interviews uitgevoerd bij gemeenten, provincies, intermediaire organisaties als SBNL en beheerders.

Bij de formele doelen ging het om: de ecologische doelen, de kosten en de participatiebereidheid. Voor het onderdeel ecologische doelen waren de formele doelen ook niet meteen duidelijk. De formele natuurdoelen verschillen van de precieze doelen van het Programma Beheer. Het werd duidelijk dat met Programma Beheer de natuurdoelen niet altijd worden verwezenlijkt. Voor het verzamelen van informatie over de realisatie van de ecologische doelen was men afhankelijk van beschikbare gegevens. Landelijk bleken er nauwelijks data voorhanden om de effecten van beheer te meten. Vanuit Alterra werd een optie aangereikt om de effecten van beheer theoretisch te evalueren. Het MNP wilde echter toch graag ook wat echte praktijkgegevens opnemen. Men heeft daarom gekozen voor twee casestudies. De generalisatie van deze cases is echter wel een probleem. Ook gezien het feit dat men dus niet zelf kon kiezen en de cases voornamelijk op basis van de beschikbaarheid van data zijn gekozen. De onderzoekers hebben hiermee vooral willen laten overkomen, dat de beschikbaarheid van data een probleem is. LNV heeft dit vooralsnog niet opgepikt. Voor een vervolgonderzoek hebben ze het MNP weer de opdracht gegeven om een (landsdekkende) ecologische evaluatie uit te voeren.

Voor het verzamelen van gegevens over bijvoorbeeld de contracten die zijn afgesloten, was het MNP aangewezen op LNV en andere instanties (bijv. beheerders) voor informatie. Dat bleek nog niet zo eenvoudig. Het was bijvoorbeeld belangrijk voor LNV dat de gebruikte gegevens klopten met de gegevens die de Minister (via de gangbare verantwoordingssporen) aan de Tweede Kamer had gemeld. De interpretatie van deze gegevens is door de onderzoekers zelf uitgevoerd. Toch zijn de interpretaties van de cijfers wel voorgelegd aan de verstrekkers ervan. De onderzoekers hebben echter zelf bepaald of ze het commentaar hierop al dan niet verwerkten.

Een ander doel van de omslag is om het draagvlak voor het natuurbeleid te vergroten. In het onderzoek is dit geoperationaliseerd door de participatiebereidheid (binnen het Programma Beheer) van agrariërs en particuliere beheerders te meten. Het vergroten van draagvlak is daarmee teruggebracht tot participatiebereidheid. Daarnaast wordt dit alleen bij individuele particuliere en agrarische beheerders nagegaan zonder te bezien wie het natuurbeheer bij deze doelgroepen zou kunnen stimuleren. Belemmeringen worden vooral gevonden in regels en kennisleemten bij instanties en financiële onaantrekkelijkheden. Belemmeringen in educatie, voorlichting, intermediairs van de doelgroepen zijn wel als mogelijke verklaringen genoemd, maar niet erg uitgewerkt. Terwijl toch vrij aanneembaar is dat ook dit soort aspecten wel degelijk de deelnamebereidheid beïnvloedt. In de onderliggende deelonderzoeken is wel gebruik gemaakt van enquêtegegevens en van interviews, maar de data hiervan worden teruggebracht tot abstracte argumenten die in het eindrapport contextloos zijn gepresenteerd als feiten. Hierdoor ontstaat niet het gevoel in welk geval iets wel of niet goed kan werken, ofwel voor de casuïstiek, maar worden bevindingen gegeneraliseerd tot algemeenheden.

Toch hebben ook verschillende partijen een rol gespeeld bij de interpretatie van de gegevens. Zo is aan het eind van het onderzoek een workshop gehouden om de resultaten voor te leggen, in de trend van: "Dit willen wij als conclusies trekken of als resultaten brengen, kunnen jullie daar al dan niet in meegaan?" De workshop was vooral bedoeld om draagvlak voor de onderzoeksresultaten te verkrijgen en om tot aanscherping van de bevindingen te komen. Het gaat dus vooral om een toetsende rol, nadat de onderzoekers de onderzoeksgegevens al hadden geïnterpreteerd. De betrokkenen mochten de onderzoeksresultaten interpreteren om te bezien of er foute conclusies waren getrokken. Om ze echt medebeoordelaar te maken, dat ging de onderzoekers te ver. Men wilde komen tot een onafhankelijk oordeel en een onafhankelijk advies, daar past een dergelijke rol voor betrokken partijen niet in. De respondenten mochten dus alleen aan het begin van het project input leveren en later zeggen wat zij van de resultaten vonden. De onderzoekers bepaalden vervolgens zelf of zij hier al dan niet wat mee deden.

Het evaluatierapport bevat geen uitgebreide methodologische verantwoording van de gemaakte keuzes. De gegevens en conclusies zijn gebaseerd op vele onderzoeksrapporten. Deze onderzoeksrapporten dienen zelf bekeken te worden om te bezien welke onderzoeksmethoden en -technieken zijn toegepast.

Voor meerstemmigheid is in het onderzoeksrapport geen ruimte geboden. Hier zijn door het MNP keuzes gemaakt, want de gegevens waren wel degelijk op bepaalde punten verschillend te interpreteren. Aan de ene kant bleek dat agrarisch natuurbeheer relatief goedkoop is. Aan de andere kant vielen de resultaten in het veld en de belangstelling voor particulier beheer nog wat tegen. Je kunt dan kiezen voor de sombere boodschap ('het gaat slecht'), maar je had net zo goed de volgende boodschap kunnen afgeven: de omslag is een goed idee, maar het instrumentarium deugt nog niet en daar moet je wat aan doen. Het MNP heeft uiteindelijk voor de eerste, kritische, benadering gekozen. Daar is echter wel discussie over geweest binnen het projectteam. Er was dus wel sprake van meerstemmigheid tussen de onderzoekers binnen het MNP, maar deze meerstemmigheid is niet in de tekst tot uitdrukking gebracht. Het gevolg was wel dat in de pers alleen de boodschap 'agrariërs en particuliere natuurbeheerders kunnen er niets van' werd opgepakt. De rest van het verhaal, waarin mogelijkheden voor verbetering staan, daar was weinig aandacht voor. De onderzoekers hadden ook niet voorzien dat dit zo zou lopen. Voor een volgende keer wil men hier lering uit trekken.

Uitvoering evaluatieonderzoek

De projectleider en het projectteam waren vooral beoordelaars en beleidsadviseurs en in die rollen voelden zij zich ook het meest thuis: op basis van eigen expertise adviezen geven. Er

waren allemaal deelonderzoeken die elkaar vaak opvolgden in de tijd. Er was niet echt sprake van een projectteam. Voor het hoofdrapport waren er twee auteurs. Zij hadden een verdeling gemaakt om de deelonderzoeken beide binnenshuis aan te sturen, bij WUR en MNP-Bilthoven.

De onderzoekers hadden de gegevens al in december 2004 terwijl het project pas in september 2005 gereed is gekomen. De interactie met de Directie Natuur hield in dat regelmatig stukken konden worden voorgelegd. In het voortraject waren veel deelonderzoeken uitgezet. Daar waren nog geen resultaten van. Daardoor gebeurde er lange tijd niets. *We hadden anderhalf jaar lang niets te bespreken, dat is best lang*. Gaandeweg kon tijdens het onderzoek wel worden ingespeeld op extra vragen omdat het budget nog niet volledig was uitgeput.

Toch waren de contacten van de onderzoekers met LNV beperkt tot de contactpersoon bij de Directie Natuur en zijn medewerkers die zich met het Programma Beheer bezig hielden. Binnen die Afdeling werd er wel verschillend gedacht. Opvallend was dat deze beleidsmedewerkers sterk focusten op weidevogelbeheer omdat dat in deze periode politiek actueel was.

Voor onderzoekers is het lastig om steeds met andere contactpersonen bij LNV te maken te hebben. Het gevolg was ook dat LNV in het vervolgonderzoek gewoon weer vragen stelt die al in het eerste rapport zijn beantwoord. Dat betekent dus dat het onderzoek dus nog niet geheel is aangekomen.

Naast LNV heeft vooral Natuurmonumenten zich tijdens het onderzoek met de aanpak bemoeid. Medewerkers van Natuurmonumenten hebben de onderzoekers tijdens het onderzoek bestookt met e-mails van "dat zou je eigenlijk ook nog mee moeten nemen" en "heb je dit ook al uitgezocht?". De onderzoekers hebben wel geprobeerd zo goed mogelijk die extra vragen mee te nemen, voor zover zij ze relevant achtten voor het onderzoek. Overigens ging het niet allemaal puur om kennisvragen. Het resultaat van het onderzoek was niet op alle punten wat Natuurmonumenten beoogde, maar daar heeft men niet over geklaagd.

Gebruik onderzoek

Het onderzoek is door de opdrachtgever gebruikt om het beleid aan te passen al is niet met zekerheid te zeggen dat dit een gevolg was van het onderzoek. Ondanks dat de beleidsmedewerkers het rapport niet altijd goed hebben gelezen (zie hiervoor), is het rapport op andere punten wel gebruikt in het beleid. Zo worden er nu alleen nog zware beheerpakketten gehonoreerd en geen lichte meer. Dit zou toe te schrijven zijn aan het onderzoek, al is dat niet hard aantoonbaar. De uitkomsten van het rapport zijn overal neergelegd en dat is gaan rondzingen. Dat werkt blijkbaar op die manier.

Het onderzoek is door het MNP zelf gebruikt in de Natuurbalans 2005 (MNP, 2005). De onderzoekers hebben een negatieve reactie van SBNL (een organisatie voor particulier natuurbeheer) gehad over het onderzoek, naar aanleiding van de boodschap in de Natuurbalans. Overigens ging de kritiek van SBNL in de visie van het MNP vooral over een andere groep particulieren: de particuliere boseigenaren. In de evaluatiestudie ging het duidelijk om de particulieren die te maken kregen met functiewijziging van landbouwgrond in natuur. In de visie van de auteurs is dit in feite een andere groep particulieren dan de boseigenaren. SBNL heeft uiteindelijk een boze brief gestuurd aan de directeur van het MNP en aan de Tweede Kamer. Men vond dat de particulieren in de Natuurbalans onnodig negatief werden neergezet. De reacties van Natuurmonumenten en Staatsbosbeheer waren meer informatief. Men had vragen over de manier waarop de kosten waren berekend. Daar wilde men zelf wat van leren.

Zoals aangegeven heeft de pers een eenduidige boodschap uit het onderzoek gehaald: 'agrariërs en particuliere natuurbeheerders kunnen er niets van'. In feite is het een vrij genuanceerd verhaal, maar dat pikt de pers niet op. De pers haalt er koppen uit. Groen Links concludeerde uit persberichten dat er grond aangekocht moest worden. Het MNP heeft de conclusies uit het rapport vervolgens toegelicht bij deze fractie. De ZLTO wilde met het rapport laten zien dat boeren het beheer zelf kunnen doen en goedkoper. Ook WLTO wil het gebruiken als steuntje in de rug voor agrarisch natuurbeheer. Misschien komt dit doordat de conceptversie die richting LTO-organisaties was gegaan positiever van toonzetting was over agrarisch natuurbeheer. Na deze versie is de meer kritische toonzetting er pas ingekomen. Hierdoor was de definitieve versie van het rapport veel kritischer over agrarisch natuurbeheer. Dit had te maken met de eerder genoemde discussie binnen het MNP over de toonzetting. Over deze wijzigingen ten opzichte van het toegestuurd concept hebben de onderzoekers geen commentaar gehad van bijvoorbeeld de LTO.

Terugkijkend zou de onderzoeker het onderzoek niet anders hebben aangepakt. Zij heeft wel het gevoel dat zij als onderzoekers meer draagvlak hebben gekregen door veel tijd uit te trekken voor het weergeven en presenteren van de onderzoeksresultaten in het uiteindelijke rapport. Het grootste punt waar zij mee zit is de moeite om LNV te betrekken. De belangrijkste kwaliteitscriteria zijn volgens de MNP-onderzoeker het stellen van een heldere vraag en het geven van heldere antwoorden. Op deze punten is zij wel tevreden over het onderzoek. Ook het feit dat relevante actoren het onderzoek ondersteunen vindt zij een belangrijk criterium.

4.3.3 Conclusie en reflectie

Tot besluit van deze analyse trekken we een aantal conclusies over de voorbereiding, de methodiek, de uitvoering en het gebruik van het evaluatieonderzoek 'Omslag van verwerven naar beheer'.

In de voorbereiding van het onderzoek is geen expliciete aandacht geweest voor de politieke of strategische achtergronden van de evaluatievraag. De vragen van de opdrachtgever zijn als gegeven aangenomen en de onderzoekers hebben zich beperkt georiënteerd op de context waarbinnen particulier of agrarisch natuurbeheer al dan niet tot stand komt.

In het onderzoek is aangenomen dat het beleid zich bevindt in een government context. Of misschien zou je kunnen zeggen dat het bestaan van een governance context gewoon is genegeerd. De onderzoekers erkennen namelijk wel dat verschillende actoren een rol spelen rond het Programma Beheer (multi-actor) en dat de verschillende actoren op een andere manier tegen de problematiek aankijken (multibetekenis), maar de onderzoekers hebben ervoor gekozen om het Rijk als centrale actor te nemen. Ze kiezen dus voor het governmentperspectief. Het stond ook in de vraagstelling van LNV centraal en het MNP heeft dit niet ter discussie gesteld. Of basis van de hiervoor genoemde overwegingen zou ook best een governance context centraal hebben kunnen staan. Dan had ook de evaluatie er echter heel anders uitgezien.

Uiteindelijk is in dit onderzoek gekozen voor een klassieke (doel)evaluatie. Deze keuze heeft nooit ter discussie gestaan. De operationalisering van de beleidsvraag in onderzoeksvragen en de operationalisering van concepten is voor een groot deel door de onderzoekers zelf uitgevoerd aan de hand van de relevante rijksdoelen. Daarbij hebben de onderzoekers zich niet willen wagen aan het betrekken van de belangrijkste actoren rond agrarisch en particulier natuurbeheer bij het opstellen van het beoordelingskader. Volgens de onderzoekers stond het beoordelingskader in dit geval vast: worden de beleidsdoelen gehaald door verandering in

beleid. Indien je deze partijen betreft bij de evaluatie, krijgen ze ook iets te vertellen over de evaluatie en dat tast in het beeld van het MNP de onafhankelijkheid aan. Een dergelijke betrokkenheid van actoren gaat dus gegeven het eigen onafhankelijkheidsbeeld van het MNP dus te ver.

Wel zijn deze actoren betrokken in de evaluatie als respondent, als informatiebron. Daarmee werden gegevens verzameld en ook een bijdrage geleverd aan verklaringen van beleidsprestaties. Ook zijn de conclusies uiteindelijk voorgelegd aan de betrokken partijen. Men wilde hiermee draagvlak bereiken voor de resultaten. Wel hebben de MNP-onderzoekers daarbij altijd aangegeven dat de onderzoekers uiteindelijk bepaalden of kritiek zou worden overgenomen of niet.

Opvallend is dat er uiteindelijk wel binnen het MNP sprake was van meerstemmigheid over de manier waarop de resultaten van dit onderzoek zouden worden opgeschreven. Aanvankelijk kwam men met meer genuanceerdere conclusies over agrarisch natuurbeheer, maar tegen het eind van het proces zijn deze conclusies kritischer geworden. Dat was voor een deel nog na de consultatie van organisaties als de LTO. Ook in de Natuurbalans zijn de resultaten van het onderzoek vrij kritisch/negatief neergezet. Daar is ook kritiek op gekomen vanuit SBNL, een organisatie voor particulier natuurbeheer.

Overigens is in het algemeen de gehanteerde onderzoeksmethode niet goed uit het evaluatierapport op te maken. Hiervoor zouden eerst de vele deelrapporten moeten worden geraadpleegd. Het rapport bevat alleen een schets op hoofdlijnen. Het is begrijpelijk dat beleidsmakers niet in de tekst worden vermoeid met de methoden en direct ter zake wordt gekomen met de resultaten, maar voor de eis van wetenschappelijke reproduceerbaarheid en navolgbaarheid is het toch aan te bevelen om dan in een bijlage aan te geven van welke methoden en technieken gebruik is gemaakt. De gehanteerde evaluatiecriteria zijn impliciet uit het rapport op te maken. Zij zijn niet vooraf als beoordelingskader gepresenteerd. Daarnaast is niet altijd beargumenteerd waarom nu juist deze criteria zijn gekozen.

Een evaluatie vanuit een governance-setting had ertoe bijgedragen dat veel meer oog was geweest voor de betrokken belanghebbenden. Hierdoor had wellicht voorkomen kunnen worden dat bepaalde organisaties (zoals SBNL) nu vooral hun energie stopten in het leveren van kritiek op het onderzoek, terwijl ze er ook wat van hadden kunnen leren. Focussen op leren had waarschijnlijk veel meer positieve energie opgeleverd en waarschijnlijk ook aanbevelingen die minder instrumenteel van aard zijn. Het lijkt alleen wel alsof dit überhaupt geen optie was vanwege de gangbare manier waarop het MNP evaluaties uitvoert. Afrekenen staat daarin centraal en het MNP wil onafhankelijk blijven. Overigens betekent onafhankelijkheid binnen een governance-setting veel meer dan niet bewust de stellingname van een van de betrokken partijen meer aandacht krijgt dan de andere. In deze studie kiest het MNP er voor om veel meer de eigen stellingname centraal te stellen, binnen de context van de vraagstelling van LNV en de formeel vastgestelde beleidsdoelen. De interne discussie over de formulering van de conclusies toont wel aan dat ook hier uiteindelijk keuzes zijn gemaakt.

4.4 Evaluatiestudie ‘Bouwtekening Transitie’

4.4.1 Evaluatieonderzoek

Het rapport ‘Bouwtekening Transitie’ (Ros e.a., 2005) is een relatief lijvig document van 117 pagina’s met een aanpak voor de evaluatie van transitieprocessen. Het gaat dus vooral om een methodisch kader om uiteindelijk het transitiebeleid uit het Nationaal Milieubeleidsplan 4

(VROM, 2001) te kunnen evalueren. Het kader is wel als een soort oefening globaal toegepast op vijf transitieonderwerpen. Deze voorbeelden zijn niet opgenomen in de uiteindelijke rapportage, ze waren vooral bedoeld als een vingeroefening voor de medewerkers. Ondanks dat het hier niet gaat om een definitieve evaluatie, maar om een soort tussenstap, hebben we ervoor gekozen om deze studie mee te nemen in deze analyse. De reden hiervan is dat het transitiebeleid veel kenmerken vertoont van governance. Zo zijn er geen duidelijke einddoelen. Transities vormen veel meer een doelzoekend proces. Meer lerende evaluaties zouden hier beter bij passen dan klassieke beleidsevaluaties, zo is de veronderstelling uit hoofdstuk 1 en 2. In de analyse van deze MNP-studie kijken we in hoeverre hier (tot nu toe) vorm aan is gegeven door het MNP.

Het doel van het project is een evaluatie van de transities uit het Nationaal Milieubeleidsplan 4 (NMP4). Transities zijn ingrijpende maatschappelijke transformatieprocessen met technologische, ecologische en maatschappelijke veranderingen. Transities zijn relatief autonoom van aard en ontstaan doordat de innovatieve activiteiten van allerlei actoren in elkaar grijpen en elkaar versterken. De rijksoverheid wil bepaalde transities bevorderen (zie: Selnes en Aalders, 2005). Het probleem met de evaluatie van transities is dat transities breken met klassieke beleidsprocessen: er zijn geen einddoelen omdat het om doelzoekende processen gaat. Ook ontbreken er (inhoudelijke) doelen waarop getoetst kan worden en ook (voorlopig) de meetbare resultaten. Soms zijn er wel inhoudelijke doelen, maar deze hebben veel meer betrekking op de korte termijn en zeggen niets over de systeemverandering. Dit maakt een klassieke doelevaluatie al vrijwel onmogelijk. Transitieprocessen kenmerken zich door tal van onzekerheden. Het zijn meerzijdige processen die gekenmerkt worden door complexiteit. Het gaat dan ook om hardnekkige maatschappelijke vraagstukken. Toch moet het MNP iets met transities. Ze staan immers centraal in het vierde Nationaal Milieubeleidsplan (VROM, 2001), het belangrijkste milieuplan van dit moment.

Om in de evaluatie van transities toch enig houvast te hebben, kiest het rapport voor de zogenaamde systeemoptie. Systeemopties zijn mogelijke onderdelen van het toekomstig maatschappelijk systeem. In het rapport worden vier criteria (of randvoorwaarden) voor systeemopties beschreven:

- het moet een doorbraak betekenen en niet slechts een aanvulling op bestaande systemen;
- het moet het niveau van een productieproces overstijgen;
- het moet relevant zijn voor milieu, natuur en landschap in Nederland en/of daarbuiten;
- het moet aansluiten bij de aandacht/keuzes in de maatschappij/ministeries.

Een systeemoptie kan je omschrijven en dan wordt het gaande weg mogelijk om vervolgstappen daaraan te verbinden. Je kunt de verschillende fasen van ontwikkeling van een systeem proberen te beschrijven. Onderscheiden worden de voorontwikkelingsfase, de take-off fase en de versnellingsfase, waarin de motivatie bij betrokkenen toeneemt. Op deze manier zijn systeemopties op te bouwen. Met de systeemoptie komt het object van onderzoek uit de wereld van de abstracte concepten en in de wereld van concrete vervolgstappen. Dit is ook nodig voor een daadwerkelijke evaluatie. Hierbij schrijft de methode niet een keuze voor tussen systeemopties. Het kunnen ook concurrerende systeemopties zijn.

In de evaluatie van transities zou volgens de 'Bouwtekening' sprake moeten zijn van de volgende typen activiteiten, die in de voorontwikkelingsfase van een systeemverandering in de praktijk herkenbaar zijn:

1. het identificeren van probleempercepties
2. het ontwikkelen van een toekomstvisie
3. het onderzoeken en beoordelen van de inzet van Research & Development (R&D)
4. de experimenten in de praktijk met transities.

Daarnaast bevat de methodiek een beoordelingskader van de systeemopties. Het gaat om het inschatten van de effecten op basis van de volgende twee dimensies:

1. effecten op de drie P's (people, profit, planet);
2. effecten op Nederland en/of elders in de wereld.

De 'bouwtekening' benadrukt dat de evaluatie van het transitieproces niet gericht moet zijn op het beoordelen van effecten, maar juist inzicht in effecten moet genereren. Hiermee is de mogelijkheid geopend voor zowel kwantitatieve berekeningen als procesevaluaties. Bij het inschatten van de effecten kan het MNP gebruik maken van de eigen (kwantitatieve) modellen. Daarvoor is het dan wel eerst nodig om de vorm van de transitie te schetsen met behulp van meer kwalitatief onderzoek.

De bouwtekening presenteert vervolgens een long-list met mogelijke systeemopties die relevant kunnen zijn. In het werkproces is gebruik gemaakt van deelonderzoeken die zijn uitbesteed aan derden. Voor het beschrijven van de systeemoptie biobrandstoffen is bijvoorbeeld gebruik gemaakt van een deelstudie uitgevoerd door de universiteit van Utrecht (Suurs en Hekkert, 2005). Deze studie was ook vooral bedoeld om het MNP te helpen met de methodische aanpak aan de hand van een concreet voorbeeld. Elementen van hun aanpak zijn overgenomen in de bouwtekening. De 'bouwtekening' mondt uit in een plan van aanpak voor een tweede fase, waarin de echte toepassing van de methode op het transitiebeleid centraal zal staan. Anno 2006 worden zes systeemopties uitgewerkt: biobrandstoffen (vrijwel af), groene diensten, groene grondstoffen voor de chemie, thermische zonnecentrales met brandstofcelauto's, visvoer voor viskweek en micro-wkk. Het is de bedoeling dat deze tweede fase in 2006 wordt afgerond.

4.4.2 Analyse van de evaluatie

Vorbereiding evaluatieonderzoek

Het project is tot stand gekomen na een verzoek van de Tweede Kamer die inzicht wenst te hebben in het transitiebeleid: ook voor de Tweede Kamer een nieuw type beleid. Inhoudelijk heeft het MNP het beleidsveld in de voorbereiding van het onderzoek niet expliciet geanalyseerd. Er is wel goed gekeken naar uitgevoerde studies en het debat over transities, zowel op inhoud als op proces. Ook heeft het MNP zelf een workshop hierover georganiseerd.

Het project kent een lange voorgeschiedenis en is in feite één lange aanloop naar het daadwerkelijk evaluatieonderzoek. Er lag al twee jaar een stuk waarin het MNP een eerste aanzet had gedaan voor het ontwikkelen van een methodiek voor een transitieevaluatie (Ros e.a., 2003). Hiermee is zowel intern als extern een verwachtingspatroon gecreëerd over de evaluatieaanpak van het MNP. Dit maakte het lastig om later nieuwe inzichten en ervaringen met het oefenen met de pilot-systeemopties in de reeds bestaande evaluatiemethodiek in te passen. Toch is de bouwtekening wel op een aantal punten gewijzigd ten opzichte van de aanpak uit 2003.

Er is sprake van een duidelijke worsteling bij het MNP. Aan de ene kant ligt er een verzoek tot evaluatie die op zichzelf interessant is voor het MNP. Het betreft een belangrijk thema dat politiek gemotiveerd is en vele uitdagingen bevat. Aan de andere kant is het ook riskant voor het MNP omdat het niet op traditionele wijze 'hard' gemaakt kan worden. Hier is dus sprake van een spanningsveld. Het MNP had een paar jaar eerder al te maken gehad met een discussie over de hardheid van de modellen en bevindingen. De kritiek toen kwam hard aan bij een instituut dat zich voorstaat op de hardheid van haar bevindingen.

Op dat moment duikt de transitieopdracht op, waarin deze hardheid juist ver te zoeken lijkt. De transitie-evaluatie lijkt van start te gaan onder een slecht gesternte. Aanvankelijk verlopen interne MNP-discussies over transities dan ook weinig succesvol. Tegelijkertijd erkent het MNP dat transities deel uitmaken van de beleidsstrategie in het NMP-4 en heeft men wel het gevoel er 'iets mee te moeten'. Het is dus (in evaluatietermen) de vraag of het MNP het aandurft om in de evaluatie van transities een koppeling te maken tussen leren en verantwoorden. De keuze is op dit moment⁵ nog niet gemaakt.

Het project is tot stand gekomen na overleg met alle betrokken ministeries. In deze gesprekken is de Tweede Kamer als doelgroep van de evaluatie genoemd. Het project is deels gefinancierd door het ministerie van VROM. VROM heeft wel de ontwikkeling van de bouwtekening mede gefinancierd, maar niet de uitvoering van de evaluatie. VROM wil graag dat het MNP laat zien dat het transitiebeleid iets oplevert en het MNP heeft zelf ook het gevoel dat het iets met transities moet. Er waren echter weinig impulsen om snel met resultaten te komen. Wel was al in 2004 afgesproken dat september 2006 een goed moment zou zijn voor de eerste evaluatie. Het risico van het gebrek aan impulsen om snel met concrete resultaten te komen, is dat het project uiteindelijk weinig oplevert. Een voordeel is wel dat het MNP dus hier ook de ruimte heeft om te experimenteren met nieuwe vormen van evaluatie. Nader overleg met de opdrachtgever is nog nodig voor een concreet vervolg. Daarnaast moet het MNP zelf beslissen hoe het project verder ingevuld zal worden.

Methodiek

Het is duidelijk dat MNP worstelt met de evaluatie van transities. De kans bestaat natuurlijk dat de evaluatie sterk de nadruk gaat leggen op het berekenen van de effecten en dus op afrekenen. Hier ligt ook de kernexpertise van het MNP. Dat kan gebeuren door over de ontwikkeling van een transitie allerlei aannames te formuleren en op basis daarvan te gaan rekenen aan de effecten. De 'Bouwtekening' stelt wel voor om de (mogelijke) ontwikkeling van de transitie te baseren op kwalitatief onderzoek en mogelijk ook workshops. Er is dus ook binnen het MNP wel een verlangen naar vernieuwing, als het gaat om de evaluatie van transities. Het is alleen de vraag hoe diep en hoe (formeel) gedragen deze lijn is binnen het MNP.

In potentie is de geschetste methode geschikt voor een evaluatie met zowel afrekenen als leren (beleidsverbetering). De opgave nu is vooral het maken van de keuzes voor het vervolg. Het lijkt zinvol dat het MNP zich afvraagt wat hun competenties kunnen betekenen binnen een onderwerp als transitie.

Het is niet helder of MNP dit project een status aparte wenst te geven, bijvoorbeeld voor methodische ontwikkeling. Het was in ieder geval een SOR-project (strategisch onderzoek RIVM), tot twee jaar geleden. Voor de opdrachtgever is dit niet het centrale punt, die wil graag inzicht in de voortgang van het transitiebeleid. Op dit punt komt de methode in lastig vaarwater. Transities zijn vooral sociale leerprocessen. Uiteraard spelen zaken als regulering, technologie, machtsspel etc ook rollen in dit geheel, maar in het omgaan met dit soort 'wicked problems' staat het sociale leerproces rond het bouwen van een transitieagenda centraal. Het gaat om het stimuleren van vernieuwing door in te spelen op vertrouwensrelaties en capaciteitsverhoging. Loeber (2003) ziet transitie als een inbraak in het gangbare. Het gaat erom nieuwe wegen in te slaan. Toch lijkt de vraagstelling, inzicht in de voortgang, te neigen naar een klassieke evaluatie. De bouwtekening houdt hier rekening mee door interactie hoog te houden, maar heeft niet echt een oplossing voor de neiging naar een klassieke 'afrekening'. De methode is zowel deskresearch (documentenstudies, gegevens presenteren, interviews)

⁵ December 2005

als analyse via workshops/discussies over de systeemopties. Dit laatste punt is overigens in de bouwtekening niet zo uitgewerkt.

Uitvoering

Zoals eerder aangegeven gaat het hier om een evaluatieaanpak die nog niet daadwerkelijk is toegepast. Het gaat hier dus niet om een bespreking van de uitvoering van het onderzoek, maar om de aanpak zoals deze is geschetst in de bouwtekening (Ros e.a., 2005).

De auteur heeft duidelijk getracht om uit de conventies van disciplines en bestaande concepten te ontsnappen door zelf op zoek te gaan naar een eigen aanpak, verankerd in concepten uit de transitiebenadering. Dit spreekt aan omdat het juist de geest van het transitiedenken raakt: het ontsnappen uit conventies. Aan de andere kant vraagt een dergelijke aanpak een sterke betrokkenheid van de mensen die in de uitvoering van de evaluatie bezig zullen zijn. Ze moeten in staat zijn om onder de huid van de 'concepten' te kruipen en het gedachtegoed eigen maken en uit te dragen. In de tekst is dit voor een deel opgelost door relatief praktische handvaten te bieden. Die zijn er in de vorm van een eerste uitwerking van de systeemopties, met o.a. aandacht voor probleempercepties, toekomstvisies, inzicht in R&D en experimenten. Op dit punt is het een spannend verhaal.

Toch gaat het in de evaluatie van transities voor het MNP vooral om de effectiviteit van beleidsinspanningen. Voor de uitvoering daarvan is men vooral aangewezen op interviews en documentenstudies en spelen kwantitatieve gegevens en modellen pas in tweede instantie een belangrijke rol (bij het inschatten van de effecten). Het gaat er in eerste instantie om de toekomstvisies van de belangrijkste spelers verder te inventariseren. Welke rol speelt de systeemoptie voor de betrokkenen? Is er sprake van gezamenlijke visies of alliantievorming? Op zich lijkt dit nuttig, want er zijn enorm veel beschikbare strategische nota's, verslagen van workshops, verkenningen en opiniestukken, maar er is weinig overzicht. Een overzicht hiervan zal al toegevoegde waarde hebben. En gezien de voorgeschreven uitvoeringstermijn tot september 2006 is dat misschien de best haalbare optie.

De bouwtekening stelt ook voor om gegevens over budgetten voor R&D per systeemoptie inzichtelijk te maken, misschien zelfs internationaal te vergelijken, of de relatie tussen transitie en 'bestaande systeem' inzichtelijk te maken, of de relatie tussen publiek en privaat. De bouwtekening stelt ook voor om onderzoeksprojecten te turven, of patentanalyses uit te voeren. Hiermee wenst men het Nederlands onderzoek en de rol van de overheid in beeld te brengen. De vrees hier, aldus de bouwtekening, is dat de gegevens niet volledig zijn, of dat de gewenste informatie niet voorhanden is. Het project wordt op dit punt afhankelijk van externe partijen. Voor inzicht in de institutionele vernieuwing kan het project gericht worden op het inzichtelijk maken van welke studies er zijn geweest, en welke andere activiteiten en vernieuwingsprocessen het beleid heeft opgeleverd.

Voor de experimenten in de praktijk wil de evaluatie in beeld brengen met welke onderdelen van een systeemoptie er is geëxperimenteerd en in welke mate er een leerproces heeft plaats gevonden. Daarnaast is het de bedoeling te kijken of het een succes of een mislukking is geweest. Voor deze beoordeling is men aangewezen op interviews. Een belangrijke kwestie betreft de bedoelingen achter de experimenten. De bouwtekening stelt dat de analyse hiervan een belangrijk onderdeel van de evaluatie kan zijn. Daarmee belandt de evaluatie in een zoektocht naar de motivaties, actoren en de netwerken die een transitie moeten dragen. Het rapport suggereert dat een krachtenveldanalyse hier nuttig kan zijn, al is dit punt niet in detail uitgewerkt. Het zal een analyse moeten zijn waar een eerste overzicht geboden wordt over de aanwezige motivatie om de transitie verder te brengen. Getracht wordt dan om aan te geven welke invloed is uitgegaan van het (transitie)beleid op de ontwikkeling van de transitie.

Vervolgens stelt de bouwtekening vast dat als blijkt dat de daadwerkelijke motivatie gering is, een verdere evaluatie naar motivaties rond een take-off fase weinig zin heeft. Overigens is niet helder wie dan bepaalt wanneer de motivaties te gering zijn.

De evaluatie van de transitie, zoals voorgesteld in de 'Bouwtekening', kan in de termen van deze studie worden beschouwd als een lerende evaluatie. De evaluatie beoogt namelijk zowel leren als afrekenen. Als het gaat om het daadwerkelijk invloed uit laten oefenen op het beoordelingskader en het bepalen van de effecten is de evaluatie nog wat klassiek. Met andere woorden men is terughoudend om hier de verschillende actoren rond de transitie meer invloed te geven. Zo zou men de thema's in de evaluatie kunnen laten selecteren door de betrokken actoren. Zover lijkt het MNP echter niet te willen gaan. De thema's zullen waarschijnlijk door het MNP zelf worden geselecteerd. Dat zou waarschijnlijk ook een stap te ver zijn voor het MNP dat gewend is aan de rol van afstandelijk en onafhankelijk onderzoeksinstituut. Toch zou overwogen kunnen worden om (op termijn) bij de evaluatie van transitie meer te werken met dialoog en interactieve sessies.

Overigens is de manier waarop de bouwtekening is opgesteld ook niet al te interactief geweest. Dat hoeft ook niet, want het is een MNP project, en MNP maakt zelf de keuzes. Er zijn wel gesprekken gevoerd met beleidsmakers, maar het conceptuele werk is vooral deskresearch geweest en hetzelfde geldt voor het definiëren van de systeemopties. Er zijn wel workshops, discussies en presentaties geweest met transitiewetenschappers en andere betrokkenen.

In het algemeen heeft de bouwtekening gefunctioneerd als een facilitator van de meningsvorming binnen MNP over transitie-evaluatie. Er zijn ook voldoende handvaten ontwikkeld waar het project mee verder zouden kunnen gaan. Indien de uitvoering van de evaluatie zal gaan op de manier waarop de bouwtekening dit voorstelt, kan de evaluatie een leertraject vormen voor het MNP. Het gaat dan in feite om een experiment met een lerende manier van evalueren. In dat geval is het verstandig om deze leerervaring ook zelf te evalueren, zodat het MNP kan bepalen of (en hoe) men verder wil met dergelijke evaluaties.

Gebruik evaluatieonderzoek

Transities zijn ontwikkelingen waarin veel verschillende actoren een rol spelen. Een evaluatie van transitie die (ook) gebruikt wil worden voor beleidsverbetering van deze transitie kan zich dan ook beter niet beperken tot alleen het rijk als actor. Het voorlopige plan van aanpak van de evaluatie richt zich wel voornamelijk op het rijk. De bedoeling was dat er terugkoppeling plaats zou vinden met de ministeries via het 5-DG overleg, het ambtelijk transitieoverleg en bilateraal naar de ministeries. Inmiddels bestaat het 5-DG overleg niet meer, en de terugkoppeling zal nu via een nieuwe interdepartementale directie bij VROM lopen. Het gevaar hier is dat het vooral een papieren evaluatie blijft voor de ministeries, met weinig effect op de transitie die zich grotendeels buiten de ministeries afspelen.

4.4.3 Conclusie en reflectie

Uit ander evaluatieonderzoek blijkt dat er behoefte is aan evaluaties rond verbanden tussen het conceptuele gedachtegoed uit de transitiebenadering en de wereld van de praktijk (Selnes en Aalders, 2005). Wat dat betreft is de bouwtekening zeker van waarde.

Het evalueren van beleidsontwikkelingen (rond natuur- en milieubeleid) is bovendien de kerncompetentie van het MNP. In het geval van 'transitie' ontmoet de organisatie echter een interessant spanningsveld. Aan de ene kant moet het MNP hier aandacht aan schenken omdat het een centraal concept is in het nationale milieubeleid en het MNP is verantwoordelijk voor

de evaluatie van dat beleid. Aan de andere kant vragen transities om meer lerende vormen van evaluatie. Het MNP is juist gespecialiseerd in klassiek evaluatieonderzoek en heeft nog weinig ervaring met lerende beleidsevaluatie.

Dit is interessant mede omdat we uit de organisatiesociologie weten dat gedrag binnen organisaties 'gepast' dient te zijn binnen een bepaalde setting en arena. Althans, om iets te bereiken. Tegelijkertijd wordt het gedrag ook sterk gestuurd door deze 'gepastheid'. March en Olsen (1989) noemen dit *'the logic of appropriateness'* en hieruit volgt de verwachting dat MNP zelf niet veel met de transitie zal doen, want het breekt met de oriëntatie en gedragscultuur binnen de organisatie. March en Olsen (1989) zagen echter ook dat actoren meerdere identiteiten en meerdere soorten 'regels' volgen. Gedrag is daardoor ook veranderbaar, door uiteenlopende impulsen. Het is daarom niet ondenkbaar dat de transities toch een rol zullen spelen voor het MNP, mits er impulsen komen die transitie-evaluaties stimuleren. De vraag is of en eventueel hoe transities een rol kunnen spelen voor MNP. En onder welke condities is het meer en minder waarschijnlijk dat het een rol zal spelen. Het lijkt waarschijnlijk dat MNP zoekt naar manieren om de aanwezige kerncompetenties in te zetten voor transitie-evaluaties zonder afbreuk te doen aan de dominante oriëntatie op klassieke doelevaluaties.

Ook voor het effect van de evaluatie is het van groot belang dat de resultaten niet alleen 'gepresenteerd' worden aan ministeries en eventueel kamerleden. De manier waarop er gecommuniceerd wordt over de evaluatie is van groot belang. Anders dreigt de evaluatie in een 'ladekast te verdwijnen'. Het verdient aanbeveling om in de evaluatie lerende elementen in te bouwen. De bouwtekening doet hiervoor al een aantal aanbevelingen. Wellicht dat het aantal momenten waarop interacties plaatsvinden met de belangrijke actoren rond de transities nog wel iets omhoog kan. Zo kan men ook aansluiting zoeken bij de bestaande netwerken over transities.

Wellicht dient het project een experimentele status aparte te krijgen binnen het MNP, om het project ruimte voor creativiteit te geven. MNP kan hiermee ervaring opdoen met lerende evaluatie en bovendien beter inzicht krijgen in de nodige capaciteiten en competenties in verband met dit soort projecten. Daarvoor is wel commitment vanuit de MNP-top noodzakelijk. Het project heeft risico's voor het MNP, maar indien beleid in de toekomst steeds meer het karakter krijgt van governance, zoals in het transitiedenken het geval is, dan is het ook van strategisch belang voor het MNP dat men hiermee al ervaring heeft opgedaan.

4.5 Conclusies

In dit hoofdstuk zijn drie MNP-evaluaties onder de loep genomen. Daarbij is vooral gekeken naar de mate waarin de evaluaties op leren en interactie gericht zijn of uitgaan van een governance-context.

Het eerste dat opvalt is dat het MNP sterk hangt aan klassieke doelevaluaties. Daarbij hoort een government-context en een onafhankelijke en afstandelijke evaluator. Daarbij is men weinig kritisch op de rijksdoelen. Deze dienen als basis voor de evaluatie en het beoordelingskader. Zeer huiverig is men voor het betrekken van de doelen van andere actoren in het beoordelingskader. In de reconstructie-evaluatie en de evaluatie van de beleidsomslag is hiervoor in ieder geval geen (aantoonbare) ruimte gegeven. Wel moeten de formele beleidsdoelen vaak nog geoperationaliseerd worden, voordat ze onderzoekbaar zijn. Hier bepalen de onderzoekers hoe deze vertaalslag gemaakt wordt.

In de evaluatie naar de transitie blijkt echter dat het soms heel moeilijk is om zo'n klassieke benadering te blijven hanteren. Hier is het nog niet precies duidelijk of en zo ja hoe deze evaluatie wordt ingevuld. Het lijkt in ieder geval een evaluatie, waarin het experimenteren met lerende vormen van beleidsevaluatie, zeer passend is. Toch blijkt ook uit de (tussentijdse) reflectie op deze evaluatie dat dergelijke nieuwe evaluatiebenaderingen botsen met de formele en informele 'regels' (dus de cultuur) binnen het MNP. Het is de vraag in hoeverre deze transitie-evaluatie ook tot een transitie binnen de MNP-aanpak van beleidsevaluatie zal leiden.

Het gevolg van het toepassen van klassieke (doel)evaluaties lijkt wel te zijn dat in de uiteindelijke rapporten vrij weinig aandacht is voor de verklaring van de beleidsprestaties. Een nadere beschrijving van het proces en de rol van andere actoren daarin, had wellicht betere verklaringen kunnen geven voor achterblijvende beleidsprestaties (of van beleidssuccessen). In de MNP-evaluaties naar de Reconstructie en de beleidsomslag van verwerven naar beheer is wel aandacht geweest voor het beschrijven van dit proces, ook met interviews ed, maar is dit als zeer feitelijke en abstracte beschrijvingen terecht gekomen in de hoofdrapportages.

Een tweede opvallend punt is het gebrek aan verantwoording van methodologische keuzes in de evaluaties. Dit lijkt voort te komen uit de vanzelfsprekendheid waarmee voor een klassieke benadering wordt gekozen. Klassieke doelevaluaties zijn dus de norm binnen het MNP. Het is gebruikelijk dat evaluaties op een dergelijke manier worden uitgevoerd. Daarbij lijken de evaluatoren zich niet sterk bewust van het feit dat er ook andere mogelijke vormen van evaluatie zijn. Indien men zich daar meer van bewust is, zal ook het maken van keuzes voor een bepaalde evaluatievorm (gegeven een bepaalde beleidscontext) meer nadruk (moeten) krijgen.

Een derde opvallend punt is dat er in de evaluaties overal wel sprake is van interactie met belangrijke actoren in het beleidsveld. Daarbij heeft de opdrachtgever (het Rijk) natuurlijk altijd een speciale positie. Maar ook andere actoren worden gebruikt als informatiebron via interviews of als aanleveraar van gegevens. Ook worden concepten meestal voorgelegd aan deze partijen. Commentaar wordt dan al of niet verwerkt door de onderzoekers. Het is niet precies duidelijk hoe dat gebeurt. In ieder geval bepaalt het MNP welke kritiek men wel of niet wil verwerken in het rapport.

Ten slotte lijkt de klassieke evaluatiebenadering en de gerichtheid op de rijksoverheid er wel toe te leiden dat er veel kritiek is op de uitgevoerde evaluaties van de (andere) betrokken partijen. Provincies (reconstructie) en particuliere natuurbeheerder (beleidsomslag) herkennen zich niet (of beperkt) in het beeld dat in de evaluatie wordt neergezet. In hoeverre dit het gebruik van de resultaten van deze evaluaties beïnvloedt, is in deze studie niet onderzocht. Wellicht zou het goed zijn om in een studie naar het gebruik van de MNP-evaluaties (zie ook wat de Algemene Rekenkamer op dit gebied doet) dit aspect ook mee te nemen.

In dit hoofdstuk hebben we gekeken naar een aantal MNP-evaluaties. Op de transitie-evaluatie na, waren deze te kenmerken als vrij klassieke doelevaluaties. Het MNP zou er ook voor kunnen kiezen om deze evaluaties meer het karakter van lerende evaluaties te geven. Gegeven de governancecontext van het beleid dat geëvalueerd wordt, zou dit op zijn minst overwogen dienen te worden. Lerende evaluaties vormen dus een wenkend perspectief, maar zijn ook niet zonder problemen en dilemma's. Hier gaan we in het volgende hoofdstuk verder op in.

5 Verhalen uit de praktijk van lerende evaluaties

5.1 Inleiding

In dit hoofdstuk reflecteren drie Alterra-onderzoekers op eigen ervaringen met lerende evaluaties. De reden hiervan is dat van deze ervaringen in de MNP-praktijk mogelijk weer geleerd kan worden. Daarbij pretenderen we niet de wijsheid in pacht te hebben. In de reflecties leggen de onderzoekers ook hun eigen twijfels en onzekerheden bloot.

In dit hoofdstuk is gekozen voor een bijzondere vorm: de verhalen van onderzoekers. Voordat we ingaan op de selectie van de cases geven we eerste een toelichting op deze keuze.

Waarom verhalen?

Verhalen spelen een steeds grotere rol in de sociale wetenschappen. Zogenaamde interpretatieve, argumentatieve en discoursbenaderingen beschouwen sociale interactie grotendeels als 'verhalen vertellen'. Om deze sociale interactie te begrijpen en/of te beïnvloeden verzamelt en analyseert de onderzoeker deze verhalen ("tales of the field" Van Maanen, 1988). Het verzamelen van verhalen kan door middel van ongestructureerde interviews, observatie van discussies en documentenanalyse. Analyse is vervolgens niets meer of minder dan het confronteren van de verzamelde verhalen met andere verhalen (theorie, eerder onderzoek, persoonlijke ervaringen enz.).

De keuze voor verhalen als onderzoeksobject (input) heeft een epistemologische/ontologische achtergrond. Kortweg komt het er op neer dat de werkelijkheid niet eensluidend is en niet begrepen kan worden in termen van causale relaties en wetmatigheden. Mensen wisselen daarentegen in interactie individuele werkelijkheden uit en trachten zo tot een gemeenschappelijke werkelijkheid te komen. Om dit proces te begrijpen is een focus op de verhalen en argumenten die mensen gebruiken zinnig.

Ook nemen steeds meer onderzoekers verhalende elementen op in hun onderzoeksrapportages (zie bijvoorbeeld: Boonstra, 2004) of presenteren al hun empirische en theoretische bevindingen in samenhang als goedlopend verhaal (zie voor prachtig voorbeeld: Flyvbjerg 1998). Door Czarniawski (1998) wordt deze praktijk "tales from the field" genoemd: het zijn geen letterlijke weergaven van praktijkverhalen ("tales of the field") maar door de onderzoeker geconstrueerde verhalen op basis van gedetailleerd veldonderzoek.

Verhalen zijn in dit hoofdstuk gebruikt als onderzoeksobject en wijze van verslaglegging. Extra reden voor dit laatste is dat verhalen beter dan abstracte verhandelingen leiden tot herkenning bij de lezer. Ze nodigen uit tot het maken van eigen interpretaties van de beschreven gebeurtenissen en bieden daardoor aanknopingspunten voor leren en verdere actie (zie ook: Flyvbjerg, 2001)

Selectie van verhalen

Onze verhalen gaan over de rollen, activiteiten en intenties van bestuurskundige onderzoekers in evaluatieprocessen. De betreffende onderzoekers schrijven deze verhalen zelf. Doel is zicht te krijgen op de diversiteit aan evaluatiemethoden en praktische keuzes of dilemma's die je tegenkomt bij het opzetten van een interactieve evaluatiemethode.

Hiervoor hebben we vooraf drie evaluaties geselecteerd die een zeker lerend element in zich hebben. Daarmee zijn ze te onderscheiden van meer klassieke (doel)evaluaties (Kuindersma en Boonstra, 2005).

We hebben gekozen voor de volgende evaluaties:

- Evaluatie Ontwikkelen met kwaliteit
- Evaluatie Reconstructie zandgebieden
- Evaluatie van het beleid voor Waardevolle Cultuurlandschappen

In elk van de verhalen staat een bepaald dilemma of probleem centraal. Het verhaal begint met een korte algemene inleiding over de evaluatie. Daarna vertelt de betrokken onderzoeker over een specifieke gebeurtenis of dilemma bij het uitvoeren van de evaluatie. Het is dus geen integrale beschrijving van alle problemen en dilemma's in het onderzoeksproces!

Dit hoofdstuk heeft tot doel om de discussie over lerende evaluaties te stimuleren. We sluiten daarom af met enkele algemene lijnen, zonder de pretentie te hebben dat dit algemeen geldende conclusies zijn.

5.2 Ontwikkelen met kwaliteit

Met de introductie van het concept *Ontwikkelen met kwaliteit* in 2004 wil het Rijk stimuleren en faciliteren dat decentrale overheden en particulieren ontwikkelingen aangrijpen om een kwaliteitsslag in het landschap te bereiken. Hoe mooi ook, deze doelstelling is tamelijk vaag. Het Milieu- en Natuurplanbureau (MNP) vraagt zich dan ook af hoe het deze nieuwe beleidsstrategie moet evalueren. Evaluatiecriteria zijn immers niet rechtstreeks af te leiden uit de beleidsdoelstelling. Het MNP start daarop een onderzoeksproject om door interviews met belanghebbenden te komen tot beoordelingscriteria. Het project vormt tevens een mooie leerschool voor het op interactieve wijze opstellen van evaluatiecriteria. Dit is een nieuwe, gebruikelijke werkwijze voor het MNP (zie verder: Gerritsen et al., 2005).

Een projectteam van Alterra pakt de uitdagende klus op. De onderzoekers spreken met het MNP af eerst open interviews te houden op rijksniveau en geïnterviewden hierin uit te dagen om zelf met criteria en voorbeelden te komen. Op basis van deze verkenning selecteren de onderzoekers praktijkvoorbeelden in de provincies Gelderland, Overijssel en Limburg om aan de hand daarvan meer grip te krijgen op het concept Ontwikkelen met kwaliteit.

Al snel blijkt dat belanghebbenden bij overheden nog niet zo ver zijn met de uitwerking van de beleidsstrategie. Bovendien is de afbakening ten opzichte van nauw verwante beleidsstrategieën problematisch. Dit brengt de betrokken onderzoeker in een lastig parket. Hij krijgt het gevoel zelf structuur aan te moeten brengen in het weerbarstige beleidsdossier en voelt zich soms meer beleidsontwikkelaar dan onderzoeker. In het navolgende reflecteert hij op dit rollenconflict en de uitdagingen die het met zich meebrengt.

Beleidsonderzoek of beleidsontwikkeling?

A.L. Gerritsen

Op een mooie zomermorgen ben ik op weg naar kasteel Neerijnen, in het rivierengebied. Ik ga praten met een gemeentelijke beleidsmedewerker over landschap, kwaliteit en ontwikkelen. Vanwege het moment van de dag heeft het geheel nog iets mysterieus. De plek had niet beter kunnen passen bij het doel van het gesprek. Op deze locatie is ontwikkelen met kwaliteit

overduidelijk gelukt. Het is een historische plek, met een mooi kasteel en een voorname oprijlaan. Het geheel straalt uit dat verschillende generaties hier iets moois van hebben willen maken. Ik ken het gebied goed en zou graag eens in het kasteel zelf komen, maar het gesprek vindt helaas plaats in een bijgebouw.

Ik steek meteen van wal met de vraag wat de gemeente Neerijnen verstaat onder het begrip "Ontwikkelen met kwaliteit". De gemeenteambtenaar antwoordt: *"het is voor de gemeente een nieuwe term, die vooral door het Ministerie van LNV gebruikt wordt. Er is toevallig deze week in Neerijnen een bijeenkomst georganiseerd. Door deze bijeenkomst werd de term bekend bij mij. Ik vermoed dat het bedoeld is om de ontwikkelingsgerichte landschapsstrategie een nieuwe impuls te geven."* Hoewel het niet de eerste keer is dat ik onbekendheid met het begrip Ontwikkelen met kwaliteit ontmoet, verbaast zijn antwoord me. Eerdere geïnterviewden van rijk en provincie vonden de gemeente Neerijnen namelijk een typisch voorbeeld waar Ontwikkelen met kwaliteit voortvarend wordt aangepakt. En nu is de gemeenteambtenaar waarmee ik spreek, zich niet eens bewust van zijn voorloperstatus binnen deze nieuwe beleidsstrategie!

Ik vervolg met de vraag hoe de gemeente bezig is met de relatie tussen ruimtelijke ordening en landschap. Nu komt de ambtenaar los. Landschapskwaliteit is wel degelijk een belangrijk beleidsdoel van de gemeente. Dit geldt ook voor ontwikkelingsgericht werken. De ambtenaar noemt twee succesvolle voorbeelden in zijn gemeente: het herstellen van een zichtlijn door het verwijderen van kassen en het bouwen van woningen en de realisatie van een kleine woonwijk in streekeigen architectuur met een recreatieve verbinding richting een landgoed. Maar deze ontwikkelingen hebben hun oorsprong in de reconstructie en worden daarom niet verbonden met Ontwikkelen met kwaliteit. Ik vraag me ongerust af of Ontwikkelen met kwaliteit wel bestaat op het lokale niveau. Ben ik niet bezig een nieuw beleidsveld te creëren? Als dit zo is, heeft dit ook consequenties voor het projectdoel om belanghebbenden zelf criteria te laten benoemen, waarmee Ontwikkelen met kwaliteit geëvalueerd kan worden.

Problemen met de afbakening

Ook voor het Rijk bleek het overigens lastig om Ontwikkelen met kwaliteit te scheiden van ander beleid. Er bestaat bijvoorbeeld ook nog zoets als ontwikkelingsplanologie, waarvoor het Ministerie van VROM verantwoordelijk is. Door het voeren van gesprekken bij het Ministerie van VROM en het Ruimtelijke Planbureau probeerde ik om dit duidelijk te scheiden van ontwikkelen met kwaliteit, maar dat is eigenlijk niet goed gelukt. Wat betreft de benadering van ruimtelijke ontwikkelingen zijn het op zijn minst nauw verwante beleidsvelden. Het lijkt zo te zijn dat ontwikkelingsplanologie meer over regionale ontwikkelingen gaat en Ontwikkelen met kwaliteit over kleinschalige ontwikkelingen, zoals het aanplanten van houtwallen en bomenrijen. Maar dat is mijn interpretatie en ik ben er niet zo zeker van dat anderen dit ook zo zien. Afbakening ten opzichte van ander beleid is in ieder geval moeilijk. Het is zelfs de vraag of dit wel wenselijk is. Misschien moet Ontwikkelen met kwaliteit gewoon als een op landschapskwaliteit gerichte vorm van ontwikkelingsplanologie gezien worden. Feit is in ieder geval dat de mensen waarmee ik sprak, daar verschillend over lijken te denken.

Vanaf het moment dat ik gesprekken ging voeren bij provincies en gemeenten werd het nog ingewikkelder. Vooral op het gemeentelijk niveau is het landschapsbeleid eigenlijk niet te scheiden van ruimtelijke ordening en plattelandsontwikkeling in het algemeen. Zo was bijvoorbeeld bij mijn gesprek bij de gemeente Tubbergen een medewerker aanwezig die zich bezig houdt met economische ontwikkeling. Ruimte bieden aan boeren om hun bedrijf duurzaam voort te zetten en het mooier maken van het landschap hadden in de ogen van deze beleidsmedewerkers alles met elkaar te maken. Misschien is dit ook een positief gegeven. Landschapselementen die een inkomen opleveren, zullen eerder behouden of ontwikkeld

worden dan landschapselementen die dit niet doen. Het maakt het daadwerkelijk evalueren van deze beleidsstrategie er echter niet makkelijker op.

De onderzoeker als beleidsontwikkelaar

De afbakeningsproblematiek had tot gevolg dat ik er zelf maar aan begon. Doordat ik in interviews steeds vroeg naar criteria en ook wel eens wat inbracht uit eerdere gesprekken voelde ik me soms een bemiddelaar of nieuwsbrenger. Geïnterviewden gingen wellicht opnieuw nadenken over het thema, de nota's weer eens inkijken of zelfs contact opnemen met het Ministerie van LNV. Ik weet niet of het ook echt zo gewerkt heeft, maar zou het erg zijn? Ik kan me zelfs voorstellen dat het een impliciet doel is geweest van onze opdrachtgever van het MNP, hoewel hij dit zelf nooit zo heeft gesteld. Anderzijds bestaat er ook een kans dat ik een storende functie had in het operationeel maken van het nieuwe beleid, een ongericht projectiel die van alles zegt wat LNV net niet zo bedoeld heeft.

Dergelijke ervaringen leiden tot beschouwingen over mijn rol als onderzoeker. Ben ik nu een onpartijdige buitenstaander of een deelnemer aan het beleid? Dit is geen nieuwe discussie, maar ik word er in dit project direct mee geconfronteerd. Ook ik ben 'opgevoed' met de gedachte dat onderzoekers buitenstaanders moeten zijn; vakmensen die met een vogelvluchtperspectief objectief analyseren wat zij observeren en kunnen analyseren hoe dingen beter kunnen. In de praktijk werkt het dus niet altijd zo. Ik merk bijvoorbeeld dat dit onderzoek mij mede boeit omdat ik me betrokken voel bij de ontwikkeling van het landschap. Of Nederland mooier of lelijker wordt raakt ook mij. Dit geldt zeker voor plekken waar ik me mee verbonden voel, zoals het kasteel Neerijnen en de Betuwe.

Ik vind het daarentegen wel belangrijk om deze persoonlijke motivaties op de achtergrond te houden en slaag daar aardig in. Maar persoonlijke motivaties zijn misschien ook tactisch in te zetten om meer te weten te komen. Door je interesse in hun gebied en werk te tonen, kan een gesprek in ieder geval prettiger verlopen. Dit geldt ook voor betrokkenheid bij het landschap, maar dan alleen als die betrokkenheid ook aanwezig is bij de persoon waarmee je spreekt. De gemeente Neerijnen kende ik goed. Hierdoor kon ik betere vragen stellen.

Ik gebruik mijn eigen achtergrond alleen als ik dit relevant vindt voor het onderzoek; om informatie boven tafel te krijgen. Lastiger is het om de consequenties van het doen van je onderzoek te overzien. Dit geldt in het bijzonder voor verkennende studies. Puur doordat ik criteria ga beschrijven, ben ik eigenlijk beleid aan het maken. Je ontkomt in ieder geval niet aan het interpreteren van de antwoorden die je krijgt. Ik voel me daar ongemakkelijk bij omdat ik vind dat ik niet aan beleidsontwikkeling moet doen. Dat moeten de mensen doen die hier in de praktijk mee aan de slag moeten. Ik kan daar wel in bemiddelen en om de discussie scherp te krijgen kan ik ook wel af en toe opmerkingen maken en lastige vragen stellen, maar met het benoemen van wat de doelen van beleid zouden moeten zijn, moet ik me niet bezig houden. Dat is politiek bedrijven. Maar misschien kun je onderzoek en politiek wel nooit helemaal scheiden. Bij Alterra doe je onderzoek voor beleidsmedewerkers en bestuurders en is alleen het bestaan van een onderzoek in wezen al een politieke daad. De opdrachtgever wil immers iets met je onderzoek. Hiervan moet je op zijn minst bewust zijn.

Slot

Terwijl ik het terrein van het kasteel afloop, kijk ik achterom en vraag ik me af of dit nou met kwaliteit ontwikkeld is. Was deze kwaliteit het doel of slechts een afgeleide van het bouwen van een prestigieus onderkomen? Ook moet ik eraan denken dat het in de rest van de gemeente waarschijnlijk veel moeilijker is om te bepalen wat kwaliteit heeft. Of iets mooi of waardevol is om te behouden voor toekomstige generaties is meestal niet zo duidelijk als met

een eeuwenoud landgoed. Bovendien spelen ook vaak belangen een rol, zoals werkgelegenheid of het willen uitbreiden van een bedrijf. Het gaat kortom om politieke stellingnames.

Maakt het verder uit dat op dit lokale niveau niet helemaal duidelijk is waar het ene beleid begint en het andere ophoudt? Landschap is per definitie een samenspel van veranderingen in grondgebruik en het aanzien van de activiteiten die dit met zich meebrengt. Invloed van overheden heeft vanzelf meerdere invalshoeken. Juist op dit lokale niveau komt alles bij elkaar. Op één gebied kunnen meerdere beleidsstrategieën van toepassing zijn. Bij de gemeente Tubbergen werd dit ook benadrukt: *“niemand weet nog wat het beleid is”*, zei een geïnterviewde. Hoe moet je dan de effectiviteit van een beleidsstrategie vaststellen?

Hoe zou het hier over 20 jaar uitzien? Ik weet dat er in vergelijking met 20 jaar geleden al veel veranderd is, dat zal naar de toekomst toe ook wel zo zijn. Hopelijk heeft mijn onderzoek er toe bijgedragen dat er iets nuttigs gezegd kan worden over de pogingen van overheden om deze veranderingen te sturen.

5.3 Evaluatie reconstructie zandgebieden

De evaluatie reconstructie zandgebieden komt voort uit een bepaling in de reconstructiewet. Hierin staat dat de wet vier jaar na inwerkingtreding geëvalueerd moet worden. Het Ministerie van LNV is opdrachtgever en belast Alterra met de klus. De onderzoekers krijgen twee jaar om de beleidseffecten van de reconstructie in kaart te brengen, want begin 2007 wil Minister Veerman hierover verantwoording afleggen aan de Tweede Kamer.

Dit is echter gemakkelijker gezegd dan gedaan. De reconstructie zandgebieden heeft tal van typische governance kenmerken. Hoewel het rijk, getuige de wet, initiatiefnemer is van de ruimtelijke herinrichting van de zandgebieden, is het bij verdere ontwikkeling en uitvoering van dit beleid afhankelijk van lagere overheden, maatschappelijke ondernemers, boeren, burgers en buitenlui. Deze actoren en bestuurslagen hebben ieder zo hun eigen opvattingen over de reconstructie. Bovendien raakt de reconstructie aan een groot aantal andere beleidsdossiers in het omgevingsbeleid.

De onderzoekers besluiten daarom in overleg met de opdrachtgever LNV tot opdeling van het onderzoeksproces in twee fasen: een verkennende fase (2005) en een uitvoerende fase (2006). In de verkennende fase willen zij zich oriënteren op het complexe beleidsdossier en bijbehorende beleidsnetwerk en op interactieve wijze een evaluatiekader voor het onderzoek ontwikkelen. Tijdens de uitvoerende fase staat de daadwerkelijke uitvoering van het onderzoek op basis van het evaluatiekader centraal.

Het navolgende verhaal speelt zich af tijdens de verkennende fase. De projectleider van het onderzoek vertelt over haar worsteling tussen betrokkenheid en distantie en trekt daaruit lessen voor de daadwerkelijke uitvoering van de evaluatie.

Van moeten naar willen

F.G. Boonstra

“Waarom doet het rijk dit? Als provincie zitten we niet te wachten op een evaluatie, we willen nu eindelijk met de uitvoering van de reconstructie aan de gang”. Aan het woord is de programmaleider uitvoering reconstructie van de provincie Overijssel. Vertegenwoordigers van andere provincies knikken instemmend. Argwaan is voelbaar in het volle zaaltje van LNV DRZ-West in Utrecht. Het is 13 januari 2005. Ambtelijke betrokkenen bij de reconstructie van LNV,

VROM, VNG, Brabant, Overijssel en Gelderland zijn een dag bijeen rond de start van de evaluatie reconstructie. Zelf ben ik aanwezig als projectleider van het evaluatieonderzoek samen met enkele van mijn teamleden.

Wat voor veel aanwezigen in Utrecht de eerste confrontatie is met de evaluatie, begint voor mij al in de zomer van 2004. In de zogenaamde kaderbrief, waarmee het Ministerie LNV zijn onderzoeksvragen bij de DLO-instituten uitzet, staat ook de vraag naar een evaluatie van de reconstructiewet. Achtergrond vormt een verplichting uit de reconstructiewet tot evaluatie van de beleidseffecten, vier jaar na inwerkingtreding van de wet.

Je hoeft geen insider te zijn om het problematische gehalte van deze bepaling in te zien. De inkt van de verplichte provinciale reconstructieplannen is nog maar amper droog en de uitvoering moet nog beginnen. Wat valt er in deze fase dan te evalueren? Hoe zullen partners van LNV in het reconstructieproces hierop reageren, druk als ze zijn om hun plannen goedgekeurd te krijgen door datzelfde LNV?

Voor mij is zonneklaar dat wil de evaluatie impact hebben op de beleidspraktijk, sleutelfiguren uit provincies en gebieden er over mee moeten denken. In mijn promotieonderzoek naar besluitvormingsprocessen in gebiedsgericht beleid ben ik nadrukkelijk geconfronteerd met de afstand tussen Den Haag en de gebiedspraktijk. Er is veel miscommunicatie, wrijving en soms zelfs strijd. Een evaluatie die uitsluitend uitgaat van Haagse behoeften, doelen en wensen, kan de broze vertrouwensrelatie wel eens te zwaar op de proef stellen. Bovendien zijn nu de provincies en gebieden aan zet in de uitvoering van de reconstructie en ander gebiedsgericht beleid. Zij moeten dus iets kunnen met eventuele aanbevelingen.

Doorleefde geschiedenissen

Ik tref in mijn opdrachtgever van LNV een medestander. Wat hem betreft moet de evaluatie geen verplichte oefening worden. Ze moet naast verantwoording ook een aanzet geven tot beleidsverbetering en leren. Dat daarvoor betrokkenheid van in ieder geval provincies nodig is, staat voor hem als een paal boven water. Op de bewuste startbijeenkomst zegt hij: "Provinciemensen kunnen heel goed concreet maken wat er gebeurt met de reconstructie. De uitdaging voor deze evaluatie is om doorleefde geschiedenissen te schetsen, die dat ook op landelijk niveau duidelijk maken." Ook ik probeer in mijn inleiding tijdens het interactieve middagedeelte onrust weg te nemen en zeg: "Het zou jammer zijn als alleen LNV straks iets aan de resultaten van de evaluatie heeft. Daarom willen we wensen en ideeën voor de evaluatie van stakeholders actief peilen en zoveel mogelijk meenemen in de projectdefinitie. Vanmiddag maken we daarmee een start." Hoewel de aanwezigen constructief aan de slag gaan in kleine groepjes en voorstellen formuleren voor het doel van de evaluatie, het gewenste resultaat en de rol van stakeholders, blijft het voor sommigen moeilijk de evaluatie te zien als kans. Keer op keer komt de onbeantwoorde vraag bovendrijven wat LNV nu eigenlijk wil met de evaluatie.

Terugkijkend op de dag analyseren we als projectteam dat de argwaan van sommige provinciemensen voortkomt uit angst dat het rijk de evaluatie zal aangrijpen om aanvullende eisen te stellen aan bijvoorbeeld reconstructieplannen ('van maatwerk naar uniformiteit'). Sommige provincies willen bovendien niet terugblikken op het eigen proces, omdat ze bang zijn daarmee oude discussies op te rakelen. Achtergrond van deze zorgen is dat ze volop bezig zijn met de start van de uitvoering en niet willen dat een te kritische evaluatie die uitvoering verstoort. Ondanks de weerstand zijn we niet ontevreden over de dag: we weten nu beter wat we kunnen verwachten en uit andere processen weten we dat voor het bijstellen van opvattingen meer interactie nodig is. Dit is slechts het begin van een proces dat minimaal twee jaar gaat duren.

Politiek-strategisch gehalte

Welgemoed beginnen we daarom met een eerste ronde diepte-interviews met bestuurlijke betrokkenen en managers van verschillende partijen. Deze interviews moeten, voortbouwend op de startbijeenkomst, wensen voor de doel- en vraagstelling van de evaluatie opleveren. Bovendien willen we de interviews, à la Guba en Lincoln, benutten als input voor beoordelingscriteria. Hiervoor laten we mensen praten over hun verwachtingen van de reconstructie, hun zorgen en discussievragen. Ook vragen we ze naar cruciale gebeurtenissen in het afgelopen reconstructieproces.

De gesprekken verlopen voorspoedig. Over de reconstructie raakt men niet uitgepraat: ondanks de complexiteit roept het proces veel enthousiasme op. Wel is het politiek-strategisch gehalte van sommige gesprekken hoog en praten veel geïnterviewden in abstracte termen. Er is verder waardering voor de interactieve wijze waarop we deze verkenning uitvoeren, maar er is ook kritiek. Een gedeputeerde is ronduit tegen de evaluatie. Hij is bang dat het zijn ambtenaren teveel tijd gaat kosten, waarop de interviewer in de ongemakkelijke positie terecht komt dat hij het evaluatieonderzoek moet gaan verdedigen.

Gelukkig heeft de opdrachtgever voorzien in agendering van het evaluatieonderzoek op bestuurlijk niveau. Voor het bestuurlijk overleg reconstructie van mei 2005, een overleg tussen Minister Veerman, de staatssecretarissen Van Geel en Schultz en de verantwoordelijke gedeputeerden van de reconstructieprovincies, schrijf ik een korte notitie met daarin een voorstel voor de evaluatie op hoofdlijnen op basis van de interviews. Ik vermijd zoveel mogelijk wetenschappelijk jargon en probeer me in te leven in de belevingswereld van bestuurders. Wat minister Veerman precies zegt over de evaluatie krijg ik niet te horen en evenmin of de genoemde gedeputeerde de discussie is aangegaan. Wel hoor ik achteraf van mijn opdrachtgever, dat we verder kunnen volgens de geschetste hoofdlijn.

Vuurproef

Voor de zomervakantie starten we een tweede ronde interviews. Nu zijn de “praktijkmensen” in de reconstructie aan de beurt: de gebiedscoördinatoren, programmaleiders, voorzitters, reconstructiecommissieleden enzovoort. Onze verwachting dat deze mensen de abstracties uit de eerste ronde kunnen concretiseren en aanvullen, komt grotendeels uit. Men is ook opener en bereid de hand in eigen boezem te steken over zaken die niet goed lopen. De houding over de evaluatie is verder een stuk positiever. Is het de interventie van de minister in het Bestuurlijk Overleg Reconstructie, onze interactieve insteek of is de reconstructie in een rustiger vaarwater beland doordat de reconstructieplannen inmiddels zijn goedgekeurd? Wie zal het zeggen.

De vuurproef moet echter nog komen. Op 16 november staat een bijeenkomst gepland in Kasteel Groeneveld in Baarn waarin we de aanzet voor het evaluatiekader dat we op basis van de interviews hebben opgesteld, met geïnterviewden willen bespreken. De aanloop naar de bijeenkomst voelt goed. Het projectteamlid dat de workshop voorbereidt, steekt er veel tijd en energie in. Samen met een ander projectteamlid analyseer ik de interviews. De opdrachtgever toont zich uitermate betrokken. Beoogde deelnemers willen graag komen. Even dreigt een potentieel conflict als de opdrachtgever zich bemoeit met de deelnemerslijst. Begaan met de bestuurlijke inbedding van de evaluatie wil hij vooral mensen uitnodigen die er op politiek-bestuurlijk niveau toe doen. Als projectleider wil ik echter het evaluatiekader een stapje verder helpen. Daarvoor zijn mensen nodig met praktijkervaring én analytisch vermogen. Bovendien wil ik putten uit de ruim 40 geïnterviewden (incl. deelnemers aan de startbijeenkomst), anders ben ik bang dat de dag onttaardt in een herhaling van zetten. Uit de literatuur over lerende evaluaties weet ik bovendien dat het voor “leren” essentieel is een zogenaamde leergemeenschap te creëren, die meerdere malen bij elkaar komt zodat cycli van

gezamenlijke reflectie en actie mogelijk worden. Hoewel deze ambitie misschien te hoog gegrepen is, wil ik dat de 16 november groep ook een rol krijgt bij de interpretatie van tussenresultaten volgend jaar en dat we samen iets opbouwen. Of mijn argumenten werkelijk aankomen bij de opdrachtgever weet ik niet. Als het puntje bij het paaltje komt laten “zijn favorieten” het stuk voor stuk afweten zodat ik “mijn favorieten” er probleemloos in kan fietsen. Ik neem me voor dit een volgende keer niet weer te laten gebeuren en van te voren een duidelijke rolverdeling af te spreken. Ik ben als projectleider verantwoordelijk voor methodische keuzes, zoals de deelnemerslijst voor de workshop. Hij is opdrachtgever van het onderzoek en mag suggesties doen voor afzonderlijke onderdelen, maar is niet verantwoordelijk voor de inhoud en organisatie daarvan.

Taaie kost

De dag zelf verloopt als een trein. Na een korte introductie vraagt de procesbegeleider de deelnemers meteen in de benen te komen. Aan de hand van verschillende voorgestructureerde antwoorden op de vraag “Ik kan de evaluatie gebruiken om ...” moeten de deelnemers bij “ja” of “nee” gaan staan en vervolgens zichzelf introduceren en hun antwoord toelichten. De positieve vraagstelling pakt goed uit, de antwoorden die we op basis van de interviews hebben geformuleerd zijn goed getroffen en de ervaring van de procesbegeleider doet de rest. De sfeer voor de rest van de dag is gemaakt. 's Ochtends bijt men zich blijmoedig vast in de taaie kost van het evaluatiekader. 's Middags gaat het gesprek over door deelnemers zelf aangedragen brandende kwesties in de uitvoering en vindt kennisuitwisseling plaats. Ook wij hebben er wat aan: tal van slaag- en faalfactoren die we volgend jaar nader kunnen onderzoeken passeren de revue.

Tijdens de gebruikelijke borrel na afloop concludeer ik voor mezelf dat de evaluatie al lang niet meer eng is. Over “te vroeg” heb ik ook niemand meer gehoord. Zijn we daarmee geslaagd in onze opzet? Ik ben te moe om hierop een goed antwoord te formuleren. Later kom ik met aanwezige wetenschappelijke klankbordgroepleden en projectteamleden terug op deze vraag. De interviews, workshop en bestuurlijke inbedding tijdens de verkennende fase hebben bijgedragen een bredere acceptatie van de evaluatie en de evaluatoren, daarover is iedereen het eens. Ook hebben workshopdeelnemers kennis kunnen uitwisselen: er is geleerd. Maar aan het interactieve proces kleven ook risico's. Als we volgend jaar kritische uitspraken doen over de reconstructie kan dit de goede sfeer bederven. Wat als de huidige medestanders zich dan ontpoppen als criticasters en de evaluatie-uitkomsten naast zich neerleggen? Tot nu toe hebben we voornamelijk geluisterd en een proces gefaciliteerd, hiermee zijn mogelijk verwachtingen gecreëerd die niet stroken met de rol van externe beoordelaar die we onherroepelijk gaan innemen. Ik neem me voor in het vervolg van deze evaluatie meer aandacht te besteden aan het “management” van deze verwachtingen en dit aspect ook met onze opdrachtgever te bespreken.

5.4 Waardevolle Cultuurlandschappen

Het Structuurschema Groene Ruimte van 1993 (Ministerie van LNV en VROM, 1993) introduceert beleid voor Waardevolle Cultuurlandschappen (WCL's). Het zijn gebieden met belangrijke natuur- en landschapswaarden, grote cultuurhistorische en aardkundige betekenis en recreatief-toeristische aantrekkelijkheid. Landbouw (en soms bosbouw) is in belangrijke mate de drager van het landschap. Het WCL-beleid heeft drie speerpunten:

- het verminderen van de spanning tussen landbouw, recreatie en natuur en landschap;
- het zoeken van nieuwe economische impulsen;
- herstellen of behouden van milieu-, natuur- en landschapskwaliteit.

Het Structuurschema belooft bovendien een evaluatie van het WCL-beleid in 2000 voor de Tweede Kamer. Alterra voerde het monitorings- en evaluatieonderzoek uit in opdracht van het Ministerie van LNV. Hierbij stonden twee invalshoeken centraal: monitoring en evaluatie van de toepassing van het WCL-beleid (1) en evaluatie van het WCL-beleid als instrument (2). Interactie met beleids- en gebiedsactoren vond gedurende het hele traject plaats middels interviews, groepsgesprekken, workshops en expertjudgements. In elf gebieden ontstond zo een netwerk van circa 200 personen die regelmatig werden geconsulteerd.

In het navolgende doet de projectleider van het onderzoek verslag van de discussie die ontstaat over de manier van rapporteren. Hoe doe je in de eindrapportage recht aan de diversiteit aan gebiedservaringen en meervoudige betekenisgeving die gedurende het proces naar boven komen? (zie verder: Pleijte et al., 2000).

Kansen en bedreigingen rond een interactieve evaluatie

M. Pleijte

Het is 9 februari 2000. Enigszins gespannen reis ik richting Kasteel Groeneveld te Baarn waar we de resultaten van de WCL-evaluatie gaan presenteren aan betrokkenen bij het onderzoek. Ik koester warme herinneringen aan deze locatie. Gedurende het vijfjarige onderzoeksproces hebben we hier al twee maal eerder tussenresultaten van het onderzoek met stakeholders besproken. Het waren levendige bijeenkomsten. Toch groeit mijn onzekerheid bij het betreden van het terrein. Een kasteelheer voel ik mij allerminst, eerder een onzekere ridder met een broos harnas die dadelijk als projectleider en hoofdverantwoordelijke de steken mag gaan opvangen van degenen die de uitkomsten van het onderzoek niet bevallen.

Totaalbeeld

De achtergrond van deze onzekerheid vormt een gebeurtenis van zo'n twaalf maanden eerder. Tijdens een overleg met de opdrachtgevers van het Ministerie van LNV staat de eindrapportage op de agenda. Voorafgaand hieraan zijn per gebied al uitgebreide gebiedsrapporten verschenen. Nu is de aan de orde hoe de eindrapportage er uit moet zien. De opdrachtgevers willen een krachtig totaalbeeld, waaraan de lezers bij wijze van spreken in oogopslag kunnen aflezen hoe het met het WCL-beleid is gesteld.

Ik voel me ongemakkelijk over deze vraag. Achter de eenvoud van een totaalbeeld gaan immers zeer intensieve en complexe interactieve processen en rijke data schuil. Ik vrees dat het de meerstemmigheid in het beschikbare empirisch materiaal platslaat. Bovendien vind ik het belangrijk de context van de resultaten van het WCL-beleid te schetsen. Wat voor het ene gebied een goed resultaat is hoeft dat voor het andere immers niet te zijn. Zo sprak een geïnterviewde boer me aan op het feit dat boeren in Waterland, onder de rook van Amsterdam, veel makkelijker verbredingsopties kunnen oppakken dan elders.

Toch ben ik wel gevoelig voor de opmerking van de opdrachtgevers dat veel beleidsmakers en kamerleden geen tijd hebben om alle gebiedsrapporten te lezen en zelfs het eindrapport al veel van hen vraagt. Hun eerdere tip om gelaagdheid in de informatie aan te brengen door gebiedsrapporten en een daarop gebaseerd eindrapport op te stellen, sprak me wel aan. Het dwingt me aan te tonen dat ik het hele onderzoeksveld goed overzie en de hoofdlijnen uit het onderzoeksmateriaal kan destilleren. Wel vond ik het jammer dat de gebiedsrapporten niet een bijlage van het hoofdrapport zijn. En nu komt daar dus de vraag om een totaalbeeld overheen.

Stoplicht

Blijkbaar trek ik een vertwijfeld gezicht. Een projectteamlid, die te boek staat als “die hard bèta”, kijkt mij verbaasd aan. Ook uit zijn houding maak ik op dat hij niet snapt waarom ik niet enthousiast reageer. Hij ‘overruled’ mij vervolgens in mijn rol als projectleider, door mee te buigen met de opdrachtgevers en zegt: “Natuurlijk kunnen wij een simpel en helder overzicht maken van hoe de gebieden op verschillende criteria scoren.” Hij heeft bovendien al ideeën over hoe hij dat wil doen en ontvouwt enthousiast zijn plan. Hij wil het totaalbeeld geven via stoplichtkleuren in een tabel, want die laten zich gemakkelijk communiceren. Rood staat dan natuurlijk voor onvoldoende, oranje voor redelijk en groen voor goed. Een legenda is overbodig! De tabel maakt het bovendien mogelijk de gebieden onderling te vergelijken op procesmatige en inhoudelijke resultaten. Wij als onderzoekers geven de scores. Al dat interactieve gedoe rond het project vindt hij toch al niks. Je moet volgens hem als onderzoeker zelf de deskundigheid in huis hebben om criteria op te stellen en data te beoordelen. Daar ga je toch niet over overleggen met Jan en alleman. Het wordt al gauw polderen, waarbij je door de dikke stroop heen moet varen omdat iedereen zo nodig zijn eigen geurvlag moet plaatsen. Weinig efficiënt volgens hem.

Gelukkig denken de opdrachtgevers daar heel anders over. Een van hen refereert aan de oorspronkelijke opdracht om rondom het monitorings- en evaluatieproces een communicatietraject op te zetten, waarin de opzet, aanpak en resultaten zich geleidelijk ontwikkelde. Ze voelt haarfijn aan dat het Ministerie afhankelijk is van het enthousiasme van het veld om het beleid uit te voeren en de evaluatie te gebruiken om de eigen prestaties te verbeteren. De opdrachtgever houdt daarom vast aan de eerder ingezette lijn om draagvlak te krijgen voor de onderzoeksresultaten door gebiedsactoren te betrekken. Besloten wordt de stoplichtkleuren zoveel mogelijk via gezamenlijke ‘expert-judgements’ vast te stellen.

Betrokkenheid

Hoewel ik achteraf tevreden ben over de volmondige steun van de opdrachtgever voor het interactieve traject, moet ik als ik eerlijk ben toegeven dat er ook nadelen aan kleven. Voor de verklaring van de gebiedsprestaties waren we grotendeels aangewezen op de gebiedsbetrokkenen. Zelf ontbeerden we inzicht om geconstateerde effecten of prestaties al dan niet aan het WCL-beleid toe te schrijven. Hierdoor werden we afhankelijk van wat de gebiedsbetrokkenen aan verklaringen opperden. Als onderzoekers waren we geneigd om plausibel klinkende verklaringen over te nemen. Maar is dat wel terecht? Respondenten beoordelen natuurlijk voor een groot deel ook hun eigen werk en zijn geneigd strategisch te antwoorden. Het raadplegen van onafhankelijke gebiedsdeskundigen had hiervoor uitkomst kunnen bieden, evenals het laten beoordelen van elkaars WCL-gebieden. Deze extra slagen voor wederhoor konden we omwille van de tijdsdruk nauwelijks toepassen. Hierdoor kan ik het herhaaldelijke verwijt van mijn projectteamlid dat ik teveel bij de gebiedsbetrokkenen op schoot ben gekropen ook minder goed pareren.

Een ander probleem betreft het streven naar consensus. De opdrachtgever wilde dat er onder betrokken stakeholders overeenstemming zou ontstaan over de onderzoeksresultaten maar eigenlijk was voor dit proces te weinig tijd ingeruimd. Het dwong ons bovendien in de rol van bemiddelaar of procesbegeleider. Dit versterkte het beeld van te grote betrokkenheid nog eens.

Rode lap

Dit soort bespiegelingen maken me onzeker over de presentatie van het eindbeeld tijdens de slotmanifestatie met stakeholders. Ik vraag mijn directeur de presentatie te verzorgen. Helaas is hij ziek en wordt hij vervangen door een collega die voldoende van het onderwerp af weet en niet als een betrokken onderzoeker wordt gezien. Als geen ander kan hij immers

presentaties geven waarbij hij mensen met flair en charisma voor zich weet te winnen. Ook nu stelt hij de mensen uit de slecht scorende gebieden snel gerust: “Het schijnt dat de rijksoverheden hun nieuw gebiedsgericht beleid willen richten op de gebieden die nog slecht scoren en die zichzelf nog niet financieel kunnen bedruipen. Ik hoop dat een rood vakje daarom niet op uw werkt als een rode lap op een stier, maar als een mogelijkheid om te voorkomen om rood te blijven staan.” En zo gingen ook mensen met een slechte score voldaan naar huis!

Achteraf leidden de stoplichtkleuren tot het formuleren van nieuwe doelen en de aanpassing van strategieën. Waar wij vreesden voor negatieve reacties vanuit het veld op slechte scores, is dit door het veld als een kans ervaren om extra financiering en beleidsaandacht te krijgen. En zo kon het gebeuren dat betrokkenen uit vrees geen geld meer te krijgen probeerden om hun goede score af te zwakken. Het gras is immers altijd groener bij de burens.

5.5 Reflecties op evaluatieverhalen

In dit hoofdstuk stonden drie verhalen van Alterra-onderzoekers over praktijkervaringen met lerende evaluaties centraal. We gaan in op de belangrijkste elementen uit deze verhalen en bespreken enkele leerpunten voor dergelijke vormen van evaluatie.

In het verhaal over de evaluatie ‘Ontwikkelen met kwaliteit’ gaat de onderzoeker in op zijn dilemma’s in de evaluatiestudie naar het beleidsthema Ontwikkelen met kwaliteit. Het gaat vooral om een explorerende benadering waarin de onderzoeker op zoek gaat naar de verschillende doelen en beelden die betrokken actoren hebben van dit beleidsveld. Hij komt terecht in een nog nauwelijks ontwikkeld beleidsveld dat nog niet is afgebakend en waar duidelijke doelen, ook bij de rijksoverheid, ontbreken. De onderzoeker worstelt vooral met zijn rol in het beleidsveld. Soms lijkt het wel alsof hij het beleid zelf beïnvloedt. Dat spoot niet met zijn beeld van de onderzoeker als onpartijdige buitenstaander.

Dit laatste beeld komen we ook tegen bij de reflecties op de MNP-evaluaties (in hoofdstuk 4). Ook de MNP-evaluator wil graag zijn werk doen als onpartijdige buitenstaander. Toch heeft men ook uiteindelijk tot doel dat de opdrachtgever (indien nodig) het beleid bijstuurt. Alleen kiest het MNP over het algemeen voor de formele weg. Men publiceert een rapport, stuurt dit aan de Tweede Kamer en verspreid een persbericht. De doorwerking van de resultaten kan beginnen. Met meer lerende vormen van evaluatie begint de doorwerking (idealiter) al eerder. In een responsieve evaluatie die alleen maar op leren is gericht, zoals de voorbeelden van Abma (hoofdstuk 3), is het leren in feite afgelopen na afsluiting van de evaluatie. De leereffecten zijn bereikt in de interacties van de evaluatie zelf en een rapportage heeft dan vooral een symbolische waarde of een betekenis voor de buitenstaanders.

Meer leren in beleidsevaluatie betekent dus dat je als onderzoeker het beleid meer direct gaat beïnvloeden. Dat levert nieuwe dilemma’s op voor onderzoekers die gewend zijn om juist als afstandelijke buitenstaander het beleid te beoordelen. Indien een MNP zich meer op leren in beleidsevaluatie gaat richten zou het goed zijn, om zich hiervan bewust te zijn.

Het tweede verhaal over de evaluatie reconstructie schetst de start van een evaluatie die duidelijk op leren en afrekenen is gericht. De achtergrond hiervan is dat de onderzoekers de beleidscontext definiëren als een governance omgeving. Er zijn zeer veel bestuurlijke en andere partijen betrokken bij de reconstructie en de inschatting is dat het niet verstandig is om alleen de rijksdoelen als uitgangspunt te nemen. Bovendien zouden andere partijen ook iets kunnen leren van de evaluatie, maar dan moet de evaluatie hun doelen en belangen wel

serieus nemen. Bovendien is het Rijk zelf als opdrachtgever een voorstander van een dergelijke evaluatie. Ook omdat reconstructie (ook formeel) bij uitstek een vorm van (op zijn minst) multi-level governance is.

In de evaluatie verzamelen de onderzoekers eerst de wensen en verwachtingen van de belangrijkste actoren in open interviews. Vervolgens maken de onderzoekers hiervan een beoordelingskader en wordt dit kader verrijkt en getoetst in een interactieve bijeenkomst met de belangrijkste actoren in de reconstructie. Dit vraagt om heel andere vaardigheden dan gebruikelijk bij (evaluatie)onderzoekers. Het komt nu ook aan op procesmethoden en procesvaardigheden.

Het resultaat van de bijeenkomst lijkt een groter draagvlak voor de onderzoeksaanpak te zijn. Aanvankelijk wantrouwen van de betrokken partijen tegen een evaluatie lijkt te zijn verdwenen. Maar uiteindelijk zullen er (later) ook conclusies worden getrokken die niet iedereen even leuk zal vinden. Het lijkt een risico voor de evaluatie dat er uiteindelijk wel 'beoordeeld' zal worden. In hoeverre accepteren de betrokken partijen dit van onderzoekers die in een eerste fase vooral begripvol hebben geluisterd naar de zorgen en verwachtingen van de beleidsactoren? Hier eindigt dit verhaal, omdat deze evaluatiestudie nog niet is afgerond.

Het verhaal laat zien dat mengvormen tussen leren en verantwoorden naast analytische vaardigheden ook om heel andere vaardigheden vragen. Procesvaardigheden en communicatieve vaardigheden zijn belangrijk. Daarbij is verwachtingenmanagement van belang. Het moet duidelijk zijn voor de betrokkenen dat er uiteindelijk beoordeeld zal worden. Hoe genuanceerd ook, er zal uiteindelijk een rapport worden geschreven, dat ook kritisch kan zijn over het beleid of de rol die bepaalde actoren hierin spelen. De veronderstelling is dat het gemeenschappelijk draagvlak dat voor het beoordelingskader is ontstaan, een positieve invloed zal hebben op de acceptatie en het gebruik van de uiteindelijke resultaten van de evaluatie. Maar in hoeverre dit daadwerkelijk het geval zal zijn, moet nog blijken in 2006/2007.

Er is in ieder geval een opvallend contrast met de MNP-evaluatie van de reconstructie (hoofdstuk 4) waarin de onderzoekers juist expliciet ervoor kiezen om de rijksdoelen centraal te stellen en andere actoren geen invloed te laten uitoefenen op het beoordelingskader. Dit werd vooral beargumenteerd vanuit het perspectief van de onafhankelijke onderzoeker.

De worsteling tussen de rollen van onafhankelijk en afstandelijk beoordelaar van beleid en betrokken onderzoekers gericht op leren komt ook naar voren in het derde verhaal. In dit verhaal over de evaluatie van het WCL-beleid, vertelt de onderzoeker over de verschillende stromingen binnen het projectteam (afrekenen versus leren). De spanning hiertussen komt ook juist naar boven in de eindfase van het project. Daarvoor konden beide sporen blijkbaar nog naast elkaar lopen. Ook omdat de opdrachtgever zowel leren als afrekenen wilde.

Aan het einde van het project wilde men graag een totaalbeeld schetsen van het beleid. Het stoplichtmodel (rood, oranje of groen) impliceerde al een sterke nadruk op afrekenen. Hierbij is het nadeel dat alle nuance die in de verschillende gebiedsstudies is geschetst grotendeels verloren gaat. LNV krijgt het totaalbeeld, maar de gebiedsactoren herkennen zich niet meer in de simpele stempels van rood, groen of oranje.

Dit illustreert het dilemma van dit soort evaluaties. Aan de ene kant willen de onderzoekers het verloop van het beleidsproces kunnen verklaren en begrijpen. Aan de andere kant wil men ook een oordeel vellen. Het gevaar is dan, dat deze twee onderdelen weinig meer met elkaar te maken hebben. Met andere woorden lijkt het zo te zijn: meer lerende vormen van evaluatie

leiden tot genuanceerde verhalen en minder duidelijke oordelen over het beleid in termen van goed of slecht. Indien uiteindelijk dat simpele eindoordeel toch geveld moet worden, bestaat het gevaar dat de informatie die dit oordeel nuanceert niet gebruikt wordt. Het gevolg kan zijn dat dit juist meer weerstand oproept bij de betrokkenen bij de evaluatie. Het lijkt aan te bevelen om ook in het eindoordeel van dergelijke evaluaties meerstemmigheid toe te laten. Het vinden van de juiste verhouding tussen simpele eindoordelen en genuanceerde verhalen is maatwerk en vraagt veel van de manoeuvreerkunsten van de betrokken onderzoekers.

6 Conclusies en aanbevelingen

6.1 Inleiding

In deze studie stonden lerende evaluaties centraal. Lerende evaluaties zijn evaluaties waarin zowel leren als verantwoordden centraal staan. Het doel van deze studie was op de mogelijkheden voor lerende evaluaties van het Milieu- en Natuurplanbureau te onderzoeken. In dit hoofdstuk geven we nog eens kort de belangrijkste conclusies van het onderzoek (par. 6.2). Daarna volgt een aantal aanbevelingen voor het vervolgdebat over beleidsevaluaties binnen het Milieu- en Natuurplanbureau (par. 6.3).

6.2 Conclusies

De belangrijkste conclusies van dit onderzoek zijn:

1. Het vormgeven van een lerende beleidsevaluatie is geen invuloefening. Het gaat om een tussenvorm tussen een klassieke beleidsevaluatie en een evaluatie puur gericht op leren, dus er zijn talrijke tussenvormen denkbaar met meer of minder accent op leren dan wel verantwoordden. De evaluatieonderzoeker staat dus voor tal van keuzen. Daarbij is het belangrijk om hieraan vooral veel aandacht te besteden in de voorbereidingsfase. Voordat voor een bepaalde vorm gekozen wordt, zal de onderzoeker zich eerst allerlei vragen moeten stellen en bij voorkeur ook het beleidsveld goed in beeld brengen. Op deze manier kan worden bepaald in hoeverre het beleidsveld wordt gekenmerkt door governancekenmerken. Indien dat het geval is, kan het zinvol zijn om meer lerende elementen in de evaluatie (en de vraagstelling) in te brengen.
2. De zoektocht naar nieuwe vormen van beleidsevaluatie en combinaties van leren en verantwoordden is niet exclusief een vraagstuk van het MNP. Ook andere evaluatiewetenschappers en vergelijkbare instituties als de Algemene Rekenkamer en het RPB zijn op zoek naar dergelijke evaluatievormen. Wel blijkt ook hier dat binnen dit spectrum heel veel verschillende keuzes zijn te maken. Zo blijft de Algemene Rekenkamer vanwege haar formele rol erg op verantwoordden zitten en probeert deze organisatie voorzichtig meer lerende elementen in haar evaluatie te betrekken. Vooral de behoefte aan verklaringen en het (multi-level) governance-achtige karakter van beleid zijn stimulerende factoren. Iemand als Abma kiest helemaal voor leren. Verantwoordden spelen in haar evaluaties nauwelijks nog een rol. Daartussen bevinden zich vele mengvormen, zoals Edelenbos (EUR) en Dammers (van het RPB) laten zien.
3. Uit de analyse van bestaande MNP-evaluaties blijkt dat men sterk de nadruk legt op verantwoordden en uitgaat van een governmentcontext. Het MNP is er huiverig voor om andere actoren dan de rijksoverheid een belangrijke rol te geven in het vormgeven van de evaluatie. Toch zijn er vaak wel contacten met andere actoren, maar deze hebben meestal een beperkte en hoogstens consulterende rol. Het gevolg is wel dat er een soort standaardaanpak van beleidsevaluatie lijkt te bestaan. Veel van de mogelijke vragen die gesteld kunnen worden voor dat een evaluatie start, worden feitelijk niet gesteld. Het is de vraag in hoeverre er daadwerkelijk behoefte is aan nieuwe vormen van beleidsevaluatie binnen het MNP en in hoeverre dit wel van het MNP wordt verwacht. De evaluatie van transities kan gezien worden als een mogelijkheid om wel met lerende

beleidsevaluatie te experimenteren. De onderzochte studie is echter nog een 'droogzwemoefening'. Of het MNP bij de daadwerkelijke evaluatie van het transitiebeleid voor de voorgestelde lerende vorm van evaluatie gaat kiezen, is nog onzeker.

4. De ervaringen met mengvormen tussen leren en verantwoorden in evaluatie geven wel aan dat deze vormen van evaluaties nieuwe dilemma's oproepen voor onderzoekers en ook nieuwe vaardigheden vraagt van onderzoekers. Indien het MNP met dergelijke evaluatievormen gaat experimenteren, is het goed dat men zich hier ook bewust van is

6.3 Aanbevelingen

Uit deze studie ontstaat de indruk dat het MNP zeer behoudend is als het gaat om evaluatiemethoden. De klassieke beleidsevaluatie is de standaardaanpak. Toch zijn er goede redenen vanuit het beleid en de wetenschap om het toepassen van lerende beleidsevaluaties in sommige gevallen te overwegen. Op grond van deze studie bevelen we aan om:

1. *De effecten van MNP-evaluaties op beleid en de kwaliteit van de aanbevelingen te onderzoeken.*
Het zou goed zijn om (aansluitend aan initiatieven van organisaties als de Algemene Rekenkamer) eens te onderzoek te doen naar het gebruik en de argumentatieve kwaliteit van de bestaande MNP-evaluaties. Daarbij zou niet alleen het gebruik bij de ministeries en de Tweede Kamer, maar ook het gebruik bij andere bestuurlijke en maatschappelijke partijen centraal kunnen staan. Een dergelijke studie kan het startpunt zijn van een bredere reflectie van het MNP op de rol en de gewenste vorm van MNP-evaluaties.
2. *Een strategisch debat binnen het MNP te voeren over de mogelijkheden van lerende beleidsevaluaties (samen met externe deskundigen).*
Zoals blijkt uit deze studie is de klassieke beleidsevaluatie diep verankerd in de manier van werken en zelfs de cultuur van het MNP. Lerende evaluaties vormen een breuk met deze aanpak (of op zijn minst een belangrijke aanvulling). Het zou goed zijn om op strategisch niveau binnen het MNP hierover een debat te organiseren. Daarbij kan het zinvol zijn om daarbij ook externe (evaluatie)deskundigen te betrekken. Het debat zou moeten leiden tot (in ieder geval) een eerste uitspraak op het punt of het MNP iets wil met lerende evaluaties (of niet natuurlijk). Dat kan het startpunt zijn van de verdere ontwikkeling van dergelijke evaluaties binnen het MNP.
3. *Bij belangrijke beleidsevaluaties vooraf het beleidsveld te onderzoeken op governancekenmerken, zodat beter beargumenteerd kan worden of een klassieke beleidsevaluatie of een lerende beleidsevaluatie nodig is.*
Uit deze studie blijkt dat het MNP in beleidsevaluaties vaak niet echt de vraagstelling en probleemanalyse van de opdrachtgever ter discussie stelt. Het zou goed zijn dat voortaan het beleidsveld en het beleidsprobleem vooraf beter in beeld gebracht worden. Indien governancekenmerken het beleid, het beleidsveld en/of het beleidsprobleem domineren kan vervolgens worden overwogen om meer lerende elementen in de evaluatie in te brengen. Een onafhankelijk planbureau zou een dergelijke onafhankelijke positie ten opzichte van de opdrachtgever en de gegeven probleemstelling moeten kunnen innemen.

4. *Te experimenteren met lerende vormen van evaluatie in de MNP-praktijk (bijvoorbeeld bij de evaluatie van het transitiebeleid).*

De meest verregaande aanbeveling is om daadwerkelijk aan lerende beleidsevaluaties te gaan doen. Het opdoen van ervaringen is immers de beste manier om te ontdekken wat de mogelijkheden ervan zijn en is ook een goede manier om competenties op dit terrein te ontwikkelen. Hierbij kan externe kennis worden ingeschakeld, maar voor de leerervaring is het wel belangrijk dat er ook MNP-ers direct bij de uitvoering van dergelijke evaluaties betrokken zijn. De evaluatie van het transitiebeleid uit het NMP4 biedt hiervoor mogelijkheden en biedt een kans om een dergelijk experiment binnen het MNP op te zetten.

Literatuur

- Abma, T.A. Responsief evalueren, Discoursen, controversen en allianties in het postmoderne, 1996.
- Abma, T.A., Onderhandelend evalueren. 'Fourth Generation Evaluation' van Egon Guba en Yvonna Lincoln, In: Bestuurskunde, jaargang 9, nummer 8, december 2000: 393-403.
- Algemene Rekenkamer, Slagvaardig en transparant. Presteren en functioneren van het openbaar bestuur. Strategie Algemene Rekenkamer 2004-2009, Den Haag, 2004a.
- Algemene Rekenkamer, Verklaringen voor niet (goed) uitgevoerd beleid (eindverslag), Den Haag, 2004b.
- Alvesson, M. en K. Sköldbberg, Reflexive Methodology. New Vistas for Qualitative Research. London, Sage Publications, 2000.
- Arentsen, M.J. De voorbereiding van evaluatieonderzoek. In: J.Th.A. Bressers en A. Hoogerwerf, Beleidsevaluatie, Alphen aan den Rijn, 1991.
- Bergen, J. van, Zorgt reconstructie voor een duurzame herinrichting?, in: Landwerk, 1-2005, p. 23-24.
- Boonstra, F.G. Laveren tussen regio's en regels. Verankering van beleidsarrangementen rond plattelandsontwikkeling in Noordwest-Friesland, de Graafschap en Zuidwest Salland, Assen, 2004.
- Boonstra, F.G., W.W. Buunk en M. Pleijte. Governance of Nature. Milieu- en Natuurplanbureau (2006, in voorbereiding)
- Boonstra, F.G. en W. Kuindersma. Methodologisch kader evaluatie reconstructie zandgebieden, Alterra, Wageningen (2006, in voorbereiding).
- Bressers, J.Th.A., en A. Hoogerwerf, Beleidsevaluatie, Alphen aan den Rijn, 1991.
- Bruijn, J.A. en E.F. ten Heuvelhof, Netwerkmanagement: strategieën, instrumenten en normen, Utrecht, 1995.
- Czarniawska, B., A narrative approach to organization studies, Thousand Oaks, 1998.
- Dammers, Ed, Leren van de toekomst. Over de rol van scenario's bij strategische beleidsvorming, Delft, 2000.
- Dammers, E., F. Verwest, B. Staffhorst en W. Verschoor, Ontwikkelingsplanologie. Lessen uit en voor de praktijk, Ruimtelijk Planbureau, Rotterdam: NAI Uitgevers, 2004.
- Driessen, P.P.J. en A.A.J. de Gier, Platteland in beweging? Enkele inhoudelijke, bestuurlijke en juridische aspecten van de voortgang reconstructie concentratiegebieden intensieve veehouderij, RIVM rapport 500025001/2004, MNP/RIVM, Bilthoven, 2004.
- Driessen, P.P.J. en P. Glasbergen (red.). Milieu, samenleving en beleid. Den Haag, 2000.
- Edelenbos, J. en A. van Buuren. "Evalueren als leerproces. Een nadere kennismaking met de 'lerende evaluatie'." In: Bestuurskunde, 14, 2005/6: pp. 2-12.
- Egmond, P.M. en T.J. de Koeijer, Van aankoop naar beheer- Verkenning kansrijkheid omslag Natuurbeleid I. Bilthoven: Milieu- en Natuurplanbureau. Rapportnummer 408767001, 2005.

- Fischer, F., *Evaluating Public Policy*. Nelson-Hall Publishers: Chicago, 1995.
- Flyvbjerg, B., *Rationality and power. Democracy in practice*, Chicago: The University of Chicago Press, 1998.
- Flyvbjerg, B. *Making Social Science Matter. Why social inquiry fails and how it can succeed again*. Cambridge, University Press, 2001.
- Fortuin, K., *Evaluatieonderzoek in een paradoxale beleidscontext*, In: Francke, A. en R. Richardson (red.), *Evaluatieonderzoek, Kansen voor een kwalitatieve benadering*, Coutinho, Bussum, 1994: 155-176.
- Franken, R. en A. van Wezel, *Arcadis' reactie op ex ante evaluatie is juist 'te kort door de bocht'*. *Reactie RVM*, in: *Landwerk*, 1-2005, p. 22.
- Gerritsen, A.L., J. Kruijt en W. Kuindersma, *Ontwikkelen met kwaliteit. Een verkenning van evaluatiecriteria*, *Wettelijke Onderzoekstaken Natuur & Milieu, Werkdocumenten 18*, Wageningen, 2005,
- Giddens, A., *The consequences of modernity*, Cambridge, 1990.
- Guba, E.G. en Y.S. Lincoln, *Fourth generation evaluation*, Beverly Hills, Sage Publ., 1989.
- Hajer, M.A. en H. Wagenaar (red.) *Introduction*, In: M.A. Hajer en H. Wagenaar, *Deliberative Policy Analysis. Understanding Governance in the Network Society*. Cambridge, University Press, 2003.
- Kuindersma, W., *Beleidsevaluatie voor de Natuurbalans. Een handleiding voor medewerkers aan de Natuurbalans*, DLO. Werkdocument 1999/08. Natuurplanbureau, vestiging Wageningen, 1999.
- Kuindersma, W. en M. Pleijte, *Naar nieuwe vormen van beleidsevaluatie voor het Natuurplanbureau? Een overzicht van evaluatiemethoden en de toepasbaarheid daarvan*, DLO. Werkdocument 1999/13, Natuurplanbureau, vestiging Wageningen, 1999.
- Kuindersma, W. (red.), *Bestuurlijke trends en het natuurbeleid*, *Planbureaustudies nr 3*. Natuurplanbureau - vestiging Wageningen, Wageningen, 2002
- Kuindersma, W. en F.G. Boonstra, *Methoden van beleidsevaluatie onder de loep. Een zoektocht naar nieuwe evaluatievormen voor het Milieu- en Natuurplanbureau*, *Planbureauroport nr. 26*. Natuurplanbureau, vestiging Wageningen, 2005.
- Lamers, L., *Ex ante evaluatie reconstructieplannen gaat te kort door de bocht*, in: *Landwerk*, 1-2005, p. 16-21.
- Loeber, A., *Inbreken in het gangbare. Transitiemanagement in de praktijk: de NIDO benadering*, NIDO, Leeuwarden, 2003.
- Maanen, J. Van, *Tales of the field: on writing ethnography*, Chicago, 1988
- March, J.G. & J.P. Olsen, *Rediscovering Institutions*, New York, Free Press, 1989.
- Maso, I. en A. Smaling, *Kwalitatief onderzoek: praktijk en theorie*, Boom, Amsterdam, 1998.
- Meer, F-B. van der en J. Edelenbos, *Evaluation in multi-actor policy processes: accountability, learning and cooperation*, Paper presented for the 2002 EES Conference, Seville, October 2002.
- Milieu- en Natuurplanbureau, *Natuurbalans 2005*, Bilthoven, 2005
- Ministerie van LNV en Ministerie van VROM, *Structuurschema Groene Ruimte*, Den Haag, 1993.
- Ministerie van VROM, *Nationaal Milieubeleidsplan 4 (NMP4)*, Den Haag, 2001.

- Pestman, Paul en Jan van Tatenhove, "Reflexieve beleidsvoering voor milieu, ruimtelijke ordening en infrastructuur: Nieuwe initiatieven nader beschouwd." In: *Beleidswetenschap*, 1998/3: 254 - 272.
- Pleijte, M., R.P. Kranendonk, F. Langers en Y. Hoogeveen, *WCL's ingekleurd. Monitoring en evaluatie van het beleid voor Waardevolle Cultuurlandschappen*, Alterra, Wageningen, 2000.
- Pressman, J.L. en A. Wildavsky, *Implementation*, Berkely, University of California Press, 1973.
- Ros, J, e.a., *Methodiek voor de evaluatie van een transitie*, MNP, Bilthoven, 2003.
- Ros, J., e.a., *Geen titel ("De Bouwtekening")*, MNP, Bilthoven, 2005.
- Selnes, Trond en Marius Aalders, *Een kantelende overheid. Debat intensieve veehouderij en de transitie naar duurzame landbouw*, LEI, Den Haag, 2005.
- Suurs, R. en M. Hekkert, *Naar een methode voor het evalueren van Transitieprojecten. Functies van innovatiesystemen toegepast op biobrandstoffen*, Innovatiewetenschappen, Universiteit van Utrecht, 2005.
- Verhoef, R., *Evaluatie in de beleidscyclus*. In: H. van de Graaf en R. Hoppe, *Beleid en politiek: een inleiding tot de beleidswetenschap en de beleidskunde*, Coutinho, Muiderberg, 1992.
- Wezel, A.P. van, R.O.G. Franken, J.D. van Dam, W. Loonen en P. Cleij, *Schuiven op zand. Ex ante evaluatie van de reconstructieplannen*, RIVM rapport 718401002/2004, MNP/RIVM, Bilthoven, 2004.
- Zouwen, M. van der, *Nature Policy between Trends and Traditions. Dynamics in nature policy arrangements in the Yorkshire Dales, Doñana and the Veluwe*, Eburon, 2006.
- Zouwen, M. van der, en J.P.M. van Tatenhove, *Implementatie van Europees natuurbeleid in Nederland. Planbureaustudies nr. 1. Natuurplanbureau, vestiging Wageningen*, 2002.

Bijlage 1 Geïnterviewden in dit onderzoek

Voor dit onderzoek zijn in 2005 de volgende personen geïnterviewd:

Mw. T.A. Abma (Universiteit Maastricht)

Dhr. E. Dammers (Ruimtelijk Planbureau)

Dhr. J. Edelenbos (Erasmus Universiteit Rotterdam)

mw. T.J. De Koeijer (Milieu- en Natuurplanbureau)

Dhr. P. Pestman (Algemene Rekenkamer)

Dhr. M. Reudink (Milieu- en Natuurplanbureau)

Mw. A. van Wezel (Milieu- en Natuurplanbureau)

WOt-onderzoek

Verschenen documenten in de reeks Rapporten van de Wettelijke Onderzoekstaken Natuur & Milieu – vanaf september 2005

WOt-rapporten zijn verkrijgbaar bij het secretariaat van Unit Wettelijke Onderzoekstaken Natuur & Milieu (voorheen Natuurplanbureau), Lumengebouw, te Wageningen.

T 0317 – 47 78 44
F 0317 – 42 49 88
E info.wnm@wur.nl

WOt-rapporten zijn ook te downloaden via de WOt-website www.wotnatuurenmilieu.wur.nl

- 1 *Wamelink, G.W.W., J.G.M. van der Gref-van Rossum & R. Jochem* (2005)
Gevoeligheid van LARCH op vegetatieverandering gesimuleerd door SUMO
- 2 *Broek, J.A. van den* (2005)
Sturing van stikstof- en fosforverliezen in de Nederlandse landbouw: een nieuw mestbeleid voor 2030
- 3 *Schrijver, R.A.M., R.A. Groeneveld, T.J. de Koeijer & P.B.M. Berentsen* (2005)
Potenties bij melkveebedrijven voor deelname aan de Subsidieregeling Agrarisch Natuurbeheer
- 4 *Henkens, R.J.H.G., S. de Vries, R. Jochem, R. Pouwels & M.J.S.M. Reijnen*, (2005)
Effect van recreatie op broedvogels op landelijk niveau; Ontwikkeling van het recreatiemodel FORVISITS 2.0 en koppeling met LARCH 4.1
- 5 *Ehlert, P.A.I.* (2005)
Toepassing van de basisvrachtbenadering op fosfaat van compost; Advies
- 6 *Veeneklaas, F.R., J.L.M. Donders & I.E. Salverda* (2006)
Verrommeling in Nederland
- 7 *Kistenkas, F.H. & W. Kuindersma* (2005)
Soorten en gebieden; Het groene milieurecht in 2005
- 8 *Wamelink, G.W.W. & J.J. de Jong* (2005)
Kansen voor natuur in het veenweidegebied; Een modeltoepassing van SMART2-SUMO2, MOVE3 en BIODIV
- 9 *Runhaar, J., J. Clement, P.C. Jansen, S.M. Hennekens, E.J. Weeda, W. Wamelink, E.P.A.G. Schouwenberg* (2005)
Hotspots floristische biodiversiteit
- 10 *Cate, B. ten, H.Houweling, J. Tersteeg & I. Verstegen (Samenstelling)* (2005)
Krijgt het landschap de ruimte? – Over ontwikkelen en identiteit
- 11 *Selnes, T.A., F.G. Boonstra & M.J. Bogaardt* (2005)
Congruentie van natuurbeleid tussen bestuurslagen
- 12 *Leneman, H., J. Vader, E. J. Bos en M.A.H.J. van Bavel* (2005)
Groene initiatieven in de aanbidding. Kansen en knelpunten van publieke en private financiering

- 13** *Kros, J. P. Groenendijk, J.P. Mol-Dijkstra, H.P. Oosterom, G.W.W. Wamelink* (2005)
Vergelijking van SMART2SUMO en STONE in relatie tot de modellering van de effecten van landgebruikverandering op de nutriëntenbeschikbaarheid
- 14** *Brouwer, F.M, H. Leneman & R.G. Groeneveld* (2005)
The international policy dimension of sustainability in Dutch agriculture
- 15** *Vreke, J., R.I. van Dam & F.H. Kistenkas* (2005)
Provinciaal instrumentarium voor groenrealisatie
- 16** *Dobben, H.F. van, G.W.W. Wamelink & R.M.A. Wegman* (2005)
Schatting van de beschikbaarheid van nutriënten uit de productie en soortensamenstelling van de vegetatie. Een verkennende studie
- 17** *Groeneveld, R.A. & D.A.E. Dirks* (2006)
Bedrijfseconomische effecten van agrarisch natuurbeheer op melkveebedrijven; Perceptie van deelnemers aan de Subsidieregeling Agrarisch Natuurbeheer
- 18** *Hubeeck, F.B., F.A. Geerling-Eiff, S.M.A. van der Kroon, J. Vader & A.E.J. Wals* (2006)
Van adoptiekip tot duurzame stadswijk; Natuur- en milieueducatie in de praktijk
- 19** *Kuindersma, W., F.G. Boonstra, S. de Boer, A.L. Gerritsen, M. Pleijte & T.A. Selnes* (2006)
Evalueren in interactie. De mogelijkheden van lerende evaluaties voor het Milieu- en Natuurplanbureau



W O t

Wettelijke Onderzoekstaken Natuur & Milieu

