

Evaluëren om te leren

naar een evaluatiearrangement voor de Vijfde Nota RO

Opdrachtgever:

Ministerie van VROM
(RijksPlanologische Dienst')

Auteurs:

prof.dr. Geert R. Teisman
dr. Frans-Bauke van der Meer

dr. Erik-Hans Klijn
dr. Jurian Edelenbos
dr. Henk L. Klaassen
drs. Melchert A. Reudink

prof.dr. A. B. Ringeling (referent)

ErasmusUniversiteit
Rotterdam

maart 2002

Inhoudsopgave

Voorwoord	3
Overzicht	5
Deel 1: Evaluatie van ruimtelijk beleid	7
1. Inleiding en vraagstelling	7
1.1 Vraagstelling.....	7
1.2 Evaluatie, verantwoording en de sturingsfilosofie van de 5N.....	7
1.3 Aanpak van het onderzoek en opzet van het rapport.....	10
2. Theoretische verkenning	12
2.1 Inleiding.....	12
2.2 Paradoxen van evalueren	12
2.3 Visies op evaluatie.....	14
2.4 Gebruik van evaluatie.....	15
2.5 Vormen van evaluatie.....	16
3. Denken over en ervaringen met beleidsevaluatie in Nederland	18
4. Conclusies verkenningen.....	21
Deel 2: Handleiding voor evaluatief verantwoord en leren	23
1. Inleiding.....	23
2. Regels voor het ontwikkelen van een kwalitatief hoogwaardige evaluatie	24
Regel 1: Zet elke evaluatie zorgvuldig op conform breed aanvaarde regels	24
Regel 2: Zorg voor goede timing van evaluaties in relatie tot beleidsprocessen op verschillende niveaus.....	25
Regel 3: Onderscheid en verbind statische en dynamische evaluatie.....	26
3. Koppelen van evaluaties versterkt het lerend effect	28
Regel 4: Koppel interne (zelf)evaluaties aan externe	28
Regel 5: Verbind beoordeling en reflectie.....	28
Regel 6: Combineer kwantitatieve en kwalitatieve evaluaties gekoppeld met trendanalyse	29
4. Een evaluatiearrangement voor de ruimtelijke beleidsvorming in Nederland.	31
Regel 7: Onderscheid en verbind product- en procesevaluatie.....	31
Regel 8: Koppel evaluaties op verschillende niveaus.....	31
Regel 9: Organiseer multi-actorevaluatie	33
Deel 3: Gebruiksaanwijzing	35
1. Inleiding.....	35
2. Gebruiksaanwijzing vanuit het perspectief van het parlement	36
2.1 Perspectief	36
Investeren in evaluatie.....	37
2.2 Verantwoording als ambitie van evaluatie	38
2.3 Evaluatie om ervan te leren	39
2.4 Evaluatie om samenhang in overheidsbeleid te versterken	39
3. Gebruiksaanwijzing vanuit het perspectief van VROM	41
3.1 Perspectief	41
3.2 Investeren in evaluatie.....	41
3.3 Verantwoording	42
3.4 Leren.....	43
3.5 Interactie en samenwerking	43
4. Gebruiksaanwijzing vanuit het perspectief van Financiën	45
4.1. Perspectief	45
4.2. Investeren in evaluatie	45
4.3. Verantwoording.....	45

4.4.	Leren.....	46
4.5.	Samenhang en bevordering samenwerking	46
5.	Gebruiksaanwijzing vanuit het perspectief van andere departementen.....	47
5.1	Perspectief	47
5.2	Investeren in evaluatie.....	47
5.3	Verantwoording.....	47
5.4	Leren.....	48
5.5	Interactie en samenwerking.....	48
6.	Gebruiksaanwijzing vanuit het perspectief van regionaal bestuur	49
6.1	Perspectief	49
6.2	Investeren in evaluatie.....	49
6.3	Verantwoording.....	50
6.4	Leren.....	51
6.5	Samenwerking	51
7.	Gebruiksaanwijzing vanuit het perspectief van het lokale niveau.....	52
7.1	Perspectief	52
7.2	Investeren in evaluatie.....	52
7.3	Verantwoording.....	53
7.4	Leren.....	53
7.5	Interactie en samenwerking.....	53
	Deel 4: Achtergronden	54
1.	Inleiding.....	54
2.	Literatuurverkenning	55
2.1	Inleiding.....	55
2.2	Twee visies op evaluatie.....	55
2.3	Leren van evalueren	56
2.4	Multi-actorevaluatie	59
2.5	Joint fact-finding	60
2.6	Verklarende evaluatie.....	62
2.7	Kwaliteit van beleid.....	62
3.	Ervaringen met kengetallen en VBTB.....	65
3.1	Inleiding.....	65
3.2	Ervaringen van gemeenten: kengetallen en productbegrotingen.....	65
3.3	Ervaringen bij het Rijk: lessen uit het verleden.....	69
3.4	Ervaringen bij het Rijk: het VBTB-proces	70
3.5	Lessen voor de toekomst	72
3.6	VBTB en de 5N.....	72
4.	Doorwerking onderzoek Algemene Rekenkamer.....	74
5.	Ervaringen met evaluaties RO.....	76
5.1	Evaluaties van de Vierde Nota RO.....	76
5.2	Werkgroep 5N	79
5.3	Ex ante evaluatie 5N.....	82
5.4	Koepelplan Noord-Brabant.....	85
5.5	Waardevolle cultuurlandschappen.....	90
6.	Ervaringen met evaluatie en onderzoek in andere domeinen	92
6.1	Gasboringen in de Waddenzee	92
6.2	Kennisimpasse: onderzoek als middel in de strijd.....	96
6.3	ISV.....	99
	Bijlage 1: Leden Begeleidingscommissie.....	110
	Bijlage 2: Geïnterviewde personen	111
	Bijlage 3: Geraadpleegde literatuur.....	112

Voorwoord

Dit rapport presenteert voorstellen tot een evaluatiearrangement voor de Vijfde Nota Ruimtelijke Ordening. Uitgangspunt is een arrangement te ontwikkelen dat voor alle betrokken actoren toegevoegde waarde heeft. Daarbij is enerzijds rekening gehouden met het karakter van de Vijfde Nota en het daarin vervatte sturingsmodel en anderzijds gebruik gemaakt van recente wetenschappelijke inzichten en praktijkervaringen met (nieuwe) vormen van evaluatie gericht op verantwoording, leerprocessen en andere ondersteuning van beleidsprocessen.

De auteurs hopen dat met dit rapport een bijdrage levert aan het denken over en de praktijk van evaluatie van ruimtelijk beleid, op een zodanige wijze dat deze praktijken daadwerkelijk impulsen geven tot kwalitatief hoogwaardiger beleid, betere samenwerking en effectieve beleidsrealisatie. Dit rapport vormt richtinggevende input voor het organiseren van de beleidsevaluatie.

De auteurs danken de leden van de begeleidingscommissie¹ voor hun stimulerende commentaren en de mensen die we mochten interviewen² voor hun inzichten en wijsheid. We hebben er dankbaar gebruik van gemaakt.

Geert Teisman
Frans-Bauke van der Meer

¹ zie bijlage 1

² zie bijlage 2

Overzicht

De lezer treft hier een voorstel aan om tot leerzame evaluatie van ruimtelijk beleid te komen. Het voorstel werd geschreven in opdracht van de RPD (Ministerie van VROM). Het bouwt voort op ervaringen die zijn opgedaan met beleidsevaluaties, onder meer door de Rekenkamer, bij medeoverheden, van het VINEX-beleid, en van specifieke casus. Recente wetenschappelijke inzichten over evaluatie en de organisatie van evaluatie, leiden tot het voorstel om evaluatie niet alleen te benutten als een manier om verantwoording af te leggen voor het gevoerde beleid, maar ook als een methode voor beleidsmakers en organisaties om leerprocessen op gang te brengen en zelfs als een hulpmiddel om samenwerking tussen betrokken actoren te verbeteren. Door zo de functie van evaluatie aanzienlijk te versterken, krijgen alle betrokken meer belang bij het evalueren van beleid.

Het was onze opdracht om voorstellen te doen tot evaluatie van de uitvoering van de Vijfde Nota Ruimtelijke Ordening. We plaatsen deze opdracht in een breder wetenschappelijk debat over de functionaliteit van evaluaties. Evaluatie leidt wel vaak tot kritische uitspraken, maar vaak niet tot verandering van gedrag. Er wordt dan met andere woorden niet van geleerd. Het debat richt zich nu op de vraag of, wanneer en hoe organisaties leren van evaluaties. Of geleerd wordt van evaluaties blijkt af te hangen van de plaats die evaluaties hebben in beleidsprocessen (ze staan er vaak naast, of sluiten processen af) en de wijze waarop beleidsmakers betrokken zijn bij het voorbereiden, uitvoeren en presenteren van evaluaties.

Vaak beschouwen beleidsmakers evaluatie als mosterd na de maaltijd. Ons voorstel pleit ervoor evaluatie juist meer in beleidsprocessen te verankeren en over de mogelijkheden tot verankering, juist op momenten dat beleidsmakers openstaan voor leerervaringen, steeds vooraf aan beleidsprocessen gemeenschappelijke afspraken te maken. Dit richt de processen en biedt mogelijkheden om onderweg te leren.

Vaak ook beschouwen beleidsmakers evaluatie als iets van anderen. Die anderen zijn zowel uitvoerders van evaluaties (onderzoekers, controllers) als de ontvangers (controleurs en toezichthouders). Zonder afbreuk te doen aan de verantwoordingsfunctie van evaluaties, pleiten wij voor verankering van de leerfunctie. Deze vereist veranderingen in de rol van beleidsmakers bij evaluatie.

Ten slotte bevat ons voorstel de uitdrukkelijke aanbeveling om evaluatie meer te richten op de samenhang tussen het handelen van verschillende overheden en daarmee ook in te zetten ter versterking van de samenwerking tussen beleidsmakers. Juist publieke samenwerking is voor de effectiviteit van de Vijfde Nota van groot belang. .

Het voorstel dat we doen bestaat uit een handleiding en gebruiksaanwijzingen. De *handleiding* bestaat uit negen regels, aan de hand waarvan tot adequate evaluatie is te komen. We achten deze regels voor alle bij ruimtelijk beleid betrokken partijen van belang. De regels betreffen afzonderlijke evaluaties, de samenhang tussen evaluaties (bijv. een interne op leren gerichte evaluatie gekoppeld aan een externe op verantwoording gerichte evaluatie) en het geheel aan evaluaties dat de komende decennia wordt uitgevoerd op het beleidsterrein van ruimtelijke ordening. We spreken hier van een evaluatiearrangement. Deze drie-eenheid benadrukt dat een effectieve uitvoering van ruimtelijke beleid drie vereisten in zich draagt:

- partijen dienen aan hun politieke controleurs verantwoording af te leggen over hun individuele effectiviteit;
- partijen dienen te leren effectief te handelen door evaluaties uit te (laten) voeren over het door hen gevoerde beleid;
- partijen dienen te leren samen effectief te handelen door met regelmaat gezamenlijke evaluaties uit te (laten) voeren over de gezamenlijke resultaten van hun handelen.

Deze kijk op effectief beleid is in lijn met de Vijfde Nota. De Nota roept op om de ruimtelijke kwaliteit de komende jaren voortdurend te verhogen en laat ruimte om daaraan invulling te geven. Onderkend wordt dat overheden gezamenlijk verantwoordelijk zijn voor het ruimtelijke beleid en dat het van belang is om via evaluaties zowel de effectiviteit van het eigen handelen, als de onderlinge samenhang in handelen te vergroten. Om tot een goed evaluatiearrangement te komen moeten overheden hun evaluaties op elkaar afstemmen.

Om daarbij behulpzaam te zijn bevat ons voorstel ook zes *gebruiksaanwijzingen*. Voor de Tweede Kamer, het ministerie van VROM, het ministerie van Financiën, andere departementen die een relatie met ruimtelijk beleid hebben, Provincies en Gemeenten geven we aan hoe evaluaties, samenhang in evaluaties en het te ontwikkelen evaluatiearrangement deze organisatie kan helpen in hun zoektocht naar beleidsverbetering en ruimtelijke kwaliteit.

Deel 1: Evaluatie van ruimtelijk beleid

1. Inleiding en vraagstelling

1.1 Vraagstelling

Dit rapport doet verslag van een verkenning van de evaluatiemomenten en mogelijkheden in het ruimtelijke beleid. Aanleiding daartoe vormt de Vijfde Nota Ruimtelijke Ordening (verder aangeduid als 5N). De verkenning mondt uit in voorstellen tot een robuust evaluatiearrangement. Daarmee bedoelen we een arrangement dat stevig staat – in de zin dat het geaccepteerd wordt door betrokken actoren en verweven is in hun bestuurlijk en beleidsmatig handelen – en krachtig is – in de zin dat de evaluatieprocessen en resultaten daadwerkelijk worden benut voor verantwoording en leren. Het voorgestelde arrangement moet alle betrokkenen bij ruimtelijke beleid de komende jaren helpen in hun zoektocht naar ruimtelijke kwaliteit. We beschouwen evaluatie daarmee als een integraal onderdeel van beleidsprocessen. Wil evaluatie effectief zijn, dan moet er vanaf het allereerste moment dat overheden spreken over de uitvoering van de ideeën zoals deze staan verwoord in de 5N, aandacht zijn voor evaluatie. Evaluatie beschouwen we als een actief hulpmiddel voor partijen om verantwoording af te leggen aan hun toezichthouders, om zelf te leren en om te reflecteren op hun samenwerkend vermogen.

De verkenning is uitgevoerd in opdracht van het ministerie van VROM en uitdrukkelijk bedoeld om de evaluatie van de 5N krachtig vorm te geven. De verkenning is uitgevoerd vanuit de volgende tweedelige vraagstelling:

1. *Welke implicaties heeft de 5N en het daarin vervatte sturingsmodel voor de evaluatie van het beleid en beleidsprocessen?*
2. *Hoe kan een concrete aanpak van evaluatie van de beleidsdoelen en beleidsprocessen in het kader van de 5N er uitzien?*

In par. 1.2 wijden we een eerste beschouwing aan de eerste vraag. Die mondt uit in een aantal dilemma's die van belang zijn bij de vormgeving van evaluatieprocessen rondom de 5N. In par. 1.3 zetten we kort uiteen hoe tewerk zijn gegaan om manieren te vinden om met deze dilemma's om te gaan en een antwoord te geven op de tweede vraag. Daarbij wordt verwezen naar de verschillende delen waaruit dit rapport is opgebouwd.

1.2 Evaluatie, verantwoording en de sturingsfilosofie van de 5N

De 5N is een strategische nota. Naast een inhoudelijke naar landsdelen uitgewerkte visie op hoofdlijnen, bevat de 5N een uitvoeringsstrategie die niet is gebaseerd op “een ‘van bovenaf’ gedicteerde ruimtelijke inrichting” (VROM 2001c, p. 257). De realisatie van beleid ligt voor een belangrijk deel bij medeoverheden en private initiatiefnemers. Het Rijk toetst de uitwerking en uitvoering van het beleid aan de uitgangspunten en ‘richtinggevende uitspraken’ in de PKB en aan zeven algemene criteria voor ruimtelijke kwaliteit (VROM 2001c, p. 8). Naast eigen beleid op punten van nationaal belang (zie bijv. VROM 2001c, p. 267-271) ziet het Rijk voor zichzelf een stimulerende en een faciliterende rol weggelegd (VROM 2001c, p. 271-272). Lokale en regionale overheden worden geacht hun beleid op elkaar en op de rijksdoelen

aangaande beheer en ontwikkeling af te stemmen. Aan provincies wordt een belangrijke re-gisserende rol toegekend.

De 5N en de uitwerking in PKB deel 3³ geven richtingen aan en vormen een uitdaging aan mede-overheden, samenwerkingsverbanden van overheden, alsmede private investeerders en maatschappelijke organisaties om met voorstellen te komen. De te realiseren voorstellen moeten in streek- en bestemmingsplannen terechtkomen. Ze kunnen in het kader van stimulering en facilitering financieel door het Rijk worden ondersteund. Daarnaast draagt het Rijk zelf directe verantwoording voor projecten van (inter-)nationaal belang. De bijdrage van het Rijk aan zowel het realiseren als aan het stimuleren en faciliteren van ruimtelijke beleid wordt uit-gewerkt in tweejaarlijkse uitvoeringsprogramma's⁴

De 5N stelt vast dat er een aanzienlijke ruimtebehoefte is, mogelijk met de omvang van de provincie Zuid-Holland en dat er in Nederland geen vrije ruimte beschikbaar is. Een deel van de ruimte kan gevonden worden in de ruimtelijke krimp in de landbouw. Het grootste deel van de behoeften moeten echter bevredigd worden door functies toe te voegen aan bestaande gebieden.

Het beleid kent daarbij twee lijnen. De eerste is die van de bescherming van natuurwaarden. Via het aanbrengen van groene contouren wil het ministerie een krachtige bescherming be-werkstelligen van de ecologische en natuurlijke kwaliteiten in de gebieden binnen de contouren. De nadruk ligt hier op het behouden van bestaande kwaliteiten en het beschermen van deze kwaliteiten tegen nieuwe activiteiten (denk aan de Veluwe of de Waddenzee). Verbod, behoud, beheer zijn kernoriëntaties in dit beleid. De vraag luidt dan ook of de overheid goede toezicht- en beheerarrangementen heeft ontwikkeld om tot bescherming te komen.

De tweede lijn is die van ontwikkeling van stedelijke gebieden. In de nota wordt vooral aandacht besteed aan het trekken van rode contouren. Daarbinnen zouden nieuwe ruimteclaims moeten worden gerealiseerd. Maar met of zonder contour blijft de vraag vervolgens hoe deze claims daadwerkelijk worden toegewezen en hoeveel ruimtelijke kwaliteit daarmee wordt toegevoegd respectievelijk vernietigd. Het gaat hier om het realiseren van kwaliteit die vooraf lastig is te benoemen. Het vermogen tot samenwerken zal een belangrijk succescriterium zijn voor implementatie. Dit geldt voor de landsdelen en de stedelijke netwerken en andere functi-onele samenwerkingsverbanden. ***Bij evaluatie zal in ieder geval aandacht moeten zijn voor de bestuurlijke vermogens tot beheer enerzijds en hoogwaardige ontwikkeling anderzijds.*** Bij beheersing gaat het dan om het voorkomen van ongewenste ontwikkelingen en deels ook om het realiseren van vooraf geformuleerde concrete doelen. Bij ontwikkeling gaat het meer om een doelzoekend en creatief proces, waarbij de criteria pas in de loop van het ontwikke-lingsproces duidelijker worden. Het is de vraag of ze verschillende evaluatiemodellen verei-sen, of dat ze combineerbaar zijn? Welke gegevens zijn er vanuit elk van de perspectieven nodig en hoe kunnen die worden verzameld en geïnterpreteerd?

De 5N presenteert een programma van eisen. Deze zijn te beschouwen als een uitdaging aan de medeoverheden, de nieuwe samenwerkingsverbanden in stedelijke netwerken en op lands-delig niveau. Ze vormen ook een uitdaging voor private investeerders en maatschappelijke groeperingen om tot passende beheer- en inrichtingsvoorstellen te komen. Een belangrijke vraag voor de evaluatie is dan in hoeverre de gestelde eisen in de beleidsstrategieën van al deze actoren tot uiting komen.

Het strategische karakter van de 5N roept een aantal additionele en deels nieuwe vragen op over beleidsverantwoording en evaluatie. We laten hier een zevental de revue passeren:

³ De PKB deel 3 is in november 2001 vastgesteld (VROM 2001a)

⁴ Het eerste uitvoeringsprogramma Nationaal Ruimtelijk Beleid (VROM 2001d) is in december 2001 vastgesteld.

- Welke overheidsorganisatie is in welke mate en in welke opzichten verantwoordelijk voor en aanspreekbaar op concrete ruimtelijke ontwikkelingen, of op het achterwege blijven van gewenste ruimtelijke ontwikkeling? In 5N is de behoefte voelbaar om de verschillende verantwoordelijkheden ‘helder’ te definiëren. In het eerste uitvoeringsplan is een en ander nader uitgewerkt. Toch stelt het gekozen sturingsmodel op dit punt grenzen. Wie bepaalt bijvoorbeeld wat ‘ruimtelijke diversiteit’ in een specifieke lokale situatie betekent? In lijn met het sturingsmodel zal dat doorgaans de uitkomst van een interactieproces tussen vele actoren zijn. De 5N staat daarbij toe dat er regionale verschillen ontstaan. De 5 N pleit zelfs voor ruimtelijke diversiteit naar landsdelen. Tegelijkertijd kan de Minister door het Parlement op deze variëteit in uitvoering worden aangesproken. Hier ligt een spanning die aandacht vereist in de vormgeving van evaluaties en verantwoording.
- Hoe is vast te stellen in welke mate het ruimtelijke beleid is gerealiseerd, doel treft en doelmatigheid is, daar waar het beleid meer richting geeft dan einddoelen vastlegt? Vanuit het perspectief van verantwoording is er behoefte aan om te weten in hoeverre het beleid direct en autonoom ontwikkelingen bewerkstelligt en of de verhouding tussen kosten en baten acceptabel is. Een evaluatie kan alleen dergelijk inzicht opleveren wanneer de kosten en baten meetbaar zijn. Dat is in de 5N maar beperkt het geval en in de PKB en het uitvoeringsplan slechts in zekere mate. Op veel punten kan de vereiste duidelijkheid en meetbaarheid er in de loop van de tijd wel komen, maar dan opnieuw als resultaat van interactieprocessen tussen vele actoren. Dat heeft consequenties voor een passend evaluatiearrangement. De beoordelingscriteria zijn niet alleen deels het resultaat van de uitvoering zelf, ook is het zo dat diverse van de criteria het resultaat zijn van interactie en niet van een eenzijdige wilsbeschikking van het ministerie en/of de Tweede Kamer. De vraag dient zich aan hoe hiermee om te gaan?
- Daarop aansluitend dient zich de vraag aan of inhoudelijke dan wel procesaspecten bij de evaluatie centraal moeten staan? 5N is het meest concreet over het proces. Op zichzelf zijn daaruit toetsbare indicatoren af te leiden. (Zijn de rode contouren getrokken? Zijn vastgelegde tijdspaden gerealiseerd? Heeft toetsing plaatsgevonden? Heeft een provincie een krachtig beheerregime?) Maar er blijft ook behoefte om het beleid op zijn ‘uiteindelijke’ effecten te beoordelen. (Is de ruimtelijke kwaliteit toegenomen? Is Nederland mooier geworden? Komt dat voort uit het beleid of is date bereikt ondanks het beleid?) De vraag doet zich dan voor hoe dat geoperationaliseerd kan worden in het licht van het sturingsmodel van 5N. En is dat allemaal wel in direct meetbare grootheden te vangen?
- Hoe moet er bij evaluatie worden omgegaan met contextuele veranderingen, ontwikkelingen en verschuivende waardeoriëntaties? Deze vraag is niet specifiek voor het gekozen sturingsmodel, maar krijgt er wel een extra dimensie door. Doordat het beleid in samenwerking tussen actoren totstandkomt, zijn contextuele veranderingen, ontwikkelingen en maatschappelijke waardepatronen niet alleen een extern gegeven, maar deels ook het resultaat van het beleidsproces zelf. Leren in te spelen op deze veranderende context, vormt een belangrijke voorwaarde voor succesvol beleid. De vraag is hoe flexibel overheden mogen inspelen op de contextuele veranderingen die zich voordoen in relatie tot de noodzaak wel verantwoording af te leggen. Is het bijvoorbeeld mogelijk verantwoording af te leggen over het lerend vermogen in plaats van alleen over het bereiken van doelen die geformuleerd zijn in het verleden? Zeker in een dynamische samenleving is dit een gewichtige vraag.
- Daar waar er behoefte aan is om evaluatie niet alleen te beschouwen als een methode om achteraf verantwoording af te leggen, maar ook als een manier om gaandeweg de uitvoering van beleid effectiever gedrag aan te leren, hoe moeten dan evaluaties er uitzien?

Leerprocessen bij betrokken actoren zullen onder meer afhangen van de gebruikte evaluatiecriteria, wie die formuleert, wie de evaluatie uitvoert, welke stem men zelf heeft in vraagstelling, in de gegevensverzameling en bij het trekken van conclusies en op welke wijze er met contextuele ontwikkeling rekening wordt gehouden. Van belang zijn hierbij ook de gepercipieerde belangen van evaluerende instanties, de impulsen tot defensieve of immuniserende (re)acties en de (on)overzichtelijkheid van (over elkaar heen buitende) evaluatieactiviteiten. Is het met andere woorden voldoende om alleen naar aparte evaluaties te kijken, of verdient het aanbeveling om de samenhang tussen evaluaties te versterken?

- Daarmee is ook de vraag naar participatie in en onafhankelijkheid van de evaluatie gesteld. Participatie is in brede zin een issue dat al een aantal jaren speelt. Zo wordt gesproken over toenemende afhankelijkheid, over netwerkvorming ook in het publieke domein en de erosie van hiërarchie in het publieke domein. Eveneens wordt wel betoogd dat met name burgers zelf een rol kunnen spelen in de beleidsvorming. Het lijkt te verwachten dat de vraag zich ook bij evaluaties zal voordoen of het oordeel van burgers niet als onderdeel van evaluatie kan worden meegenomen? Met name bij een ingewikkeld en meervoudig begrip als ruimtelijke kwaliteit dient de vraag zich extra aan.
- Ten slotte kan de onderzoeker zelf niet buiten beeld blijven. Daar waar de feiten onomstreden zijn, is de onderzoeker als subject niet problematisch. Daar waar effecten en beoogde (ruimtelijke) kwaliteiten verbonden zijn met de opvattingen van actoren wordt de onderzoeker zelf actor in de beoordeling. Enerzijds zien we bij onderzoekers de nadruk leggen op geconstateerde tekortkomingen. Anderen daarentegen zien we neigen tot begrip de beleidsvoerende instantie en vooral excuses aanvoeren voor niet gerealiseerde beleidsambities. Beide invullingen van evaluaties dragen niet bij aan het lerend vermogen van organisaties.

Deze punten representeren dilemma's verbonden aan de evaluatie van de 5N. We kunnen de onderliggende spanningen niet opheffen. Wel is het mogelijk te zoeken naar evaluatievormen die recht doen aan verschillende behoeften (zoals verantwoording en leren), die inspelen op ontwikkelingen in context en waardeoriëntaties en die daadwerkelijk beleidsprocessen ondersteunen.. We geven nu de weg weer die we hebben gevolgd om tot een arrangement te komen dat aan deze eisen voldoet.

1.3 Aanpak van het onderzoek en opzet van het rapport

Om te beginnen hebben we ons georiënteerd op bestaande inzichten en ervaringen. Daartoe hebben we vier soorten verkenningen verricht. Onze bevindingen worden in hoofdlijnen in dit eerste deel van het rapport weergegeven. Meer gedetailleerde overzichten zijn te vinden in deel 4 ('Achtergronden').

1. Verkenning van wetenschappelijke literatuur over evaluatie van complexe multi-actor beleidsprocessen en over gebruik van evaluatie en organisationele leerprocessen (samengevat en toegelicht in deel 4, hoofdstuk 2); de verkenning vormt de basis van de theoretische beschouwing in hoofdstuk 2 hieronder;
2. Verkenning van ervaringen met het voorstel van het ministerie van Financiën om tot een betere verantwoording van beleid te komen (het zogenaamde VBTB) en met de toepassing van kengetallen als 'hard' evaluatie- en sturingsinstrument (zie hoofdstuk 3 en deel 4, hoofdstuk 3);

3. Inzicht in recente ontwikkelingen in het denken over evaluatie bij verschillende instanties in Nederland. Het gaat hier zowel om evaluatoren en toezichthouders, als om beleidsmakers in het veld. Het is gebaseerd op documentstudie en interviews (zie hoofdstuk 3);
4. bestudering van een aantal recente Nederlandse voorbeelden van evaluatie op het gebied van de RO alsmede enkele ervaringen met evaluaties van beleid waarbij het sturingsmodel verwantschap vertoont met dat van de 5N (hoofdstuk 3, deel 4., hoofdstukken 5 en 6)

Deze verkenningen vormen de basis van de beschouwingen in deel 1. Zij monden uit in aandachtspunten bij het ontwikkelen van een evaluatiearrangement voor de 5N. Deze zijn in hoofdstuk 4 samengevat. In deel 2 presenteren we een handleiding voor de ontwikkeling van een evaluatiearrangement rondom de 5N, gericht op verantwoording, leren en samenwerken. Zij is geconcretiseerd in een negental *regels*. Een van de inzichten die uit onze verkenningen naar voren komt, is dat de perspectieven en belangen van verschillende betrokkenen bij ruimtelijk beleid zodanig verschillen dat gezamenlijk gedragen en benutte evaluatiearrangementen geen vanzelfsprekendheid zijn. We hebben daarom gepoogd de handleiding zodanig vorm te geven dat het resulterende arrangement vanuit verschillende perspectieven toegevoegde waarde heeft. In deel 3 expliciteren we de achterliggende gedachtegang door vanuit een aantal verschillende perspectieven weer te geven hoe het bepleite arrangement bij het betreffende perspectief kan aansluiten, wat de toegevoegde waarde ervan kan zijn, en welke ‘investeringen’ dat vergt.

Deel 4 omvat, zoals gezegd uitvoerige informatie over de verrichte achtergrondstudies.

2. Theoretische verkenning

2.1 Inleiding

De gesignaleerde dilemma's bij de evaluatie van de 5N zien we ook in andere beleidsterreinen en andere landen terugkomen. Daarom hebben we een verkenning van de literatuur ondernomen omtrent de vraag hoe over deze problemen wordt gedacht en welke manieren er zijn ontwikkeld om ermee om te gaan. Onze voorstellen voor een werkbaar, robuust en effectief evaluatiearrangement zijn mede op deze literatuurstudie gebaseerd.

2.2 Paradoxen van evalueren

Evaluatie is een spanningsvolle aangelegenheid. In zekere zin is evaluatie ook juist bedoeld om spanning teweeg te brengen. Enerzijds gaat het hier om een beoordeling, waar de beoordeelde met spanning wacht op de uitkomst. Anderzijds gaat het om reflectie op de spanning tussen beleidsintenties en resultaten achteraf en de verklaringen voor deze spanning. Deze reflectie kan bijdragen aan leren. Beide opbrengsten van evaluatie, enerzijds het oordeel en anderzijds de les, hebben hun waarde. Beide hebben evenwel ook de neiging om aan waarde in te boeten.

Zo kan geconstateerd worden dat beoordelende evaluaties door de geëvalueerde(n) soms worden beschouwd als tamelijk vijandige activiteiten, waartegen men zich dient te wapenen en te immuniseren. Immunisering tegen beoordelende evaluaties vindt plaats door de situatie zo gunstig mogelijk voor te stellen, door veranderingen in het beleid aan te brengen vooraleer het evaluatierapport verschijnt en door tekortkomingen niet expliciet te maken.

Bij reflexieve evaluaties, waarbij de organisatie zichzelf evalueert (ongeacht of de organisatie de evaluatie zelf uitvoert of doorvoor een onderzoeksbureau inhuurt), is er juist de neiging om vooral te kijken naar positieve resultaten en wordt daarmee een rooskleurig beeld geschetst. Ook is er dan een neiging om vooral conclusies te trekken ten gunste van de eigen positie en het eigen bestaande instrumentarium. Er is dan minder sprake van leren en meer van zelfbevestiging.

Kort door de bocht gezegd, lijken negatieve uitspraken door de evaluator of de gebruikers van een beoordelende evaluatie in de richting van de geëvalueerde, vaak te leiden tot anticiperende of immuniserende reacties..

En even kort door de bocht gezegd leiden reflexieve evaluaties nogal eens tot bemoedigende uitspraken over wat is goed gegaan en bestaat er een neiging om tekortkomingen vooral te wijten aan het gedrag van anderen. Om daarop in te spelen wordt regelmatig voorgesteld het sturend vermogen van de onderzochte organisatie te versterken. Schema 1.1 brengt een en ander in beeld.

Aggregatieproblemen

Een ander probleem waar evaluaties mee worstelen is dat van de aggregatie. Met name externe beoordelingen zijn vaak bedoeld voor gebruikers die niet deskundig zijn aangaande het geëvalueerde vraagstuk en die vooral geïnteresseerd zijn in een aantal globale kengetallen, om zo vergelijkend uitspraken kunnen doen. Dit is met name het geval in de beleidsverslagen die op de tweede woensdag in juni in de Tweede Kamer worden gesproken. Daarin wordt met behulp van de VBTB-regels per departement een indruk gegeven van de bereikte resultaten in relatie tot de (financiële) inzet.

Schema 1.1: Twee soorten evaluaties, hun kracht en hun ontkrachtende werking

	<i>Beoordelende Evaluaties</i>	<i>Reflexieve evaluaties</i>
<i>Kracht</i>	<p>Uit de wetenschap dat er een extern oordeel komt, letten overheden beter op wat ze doen.</p> <p>Door externe oordelen bestaat de mogelijkheid organisaties 'te dwingen' zich aan te passen.</p>	<p>Door wetenschap dat eigen criteria worden gehanteerd is geëvalueerde eerder bereid zich bloot te geven.</p> <p>Door interne oordelen is het mogelijk om via zelfreflectie te leren.</p>
<i>Ontkrachtende werking</i>	<p>Omdat externe beoordelaars als 'vijanden' worden beschouwd, is de beschikbaar gestelde informatie sterk gekleurd.</p> <p>Externe beoordelaars neigen ertoe vooral zwakke punten te benadrukken, hetgeen in media en bij politiek wel eens leidt tot veroordeling in plaats van beoordeling; dit zet beoordeelde weer aan tot inzet immuniseringstechnieken.</p>	<p>Interne beoordelaars hebben als 'peers' meer begrip voor de beleidsproblemen en neigen tot toedekken van zwaktes in beleidsstrategie.</p> <p>Interne beoordelaars neigen ertoe de waarde van de gekozen aanpak te benadrukken, tekortkomingen vaak extern te zoeken (onwillige uitvoerders b.v.) en de oplossing te zoeken in versterking van het eigen beleidsinstrumentarium.</p>

Het blijkt echter moeilijk om betekenis te geven aan deze vergelijkende overzichten. Anders geformuleerd: de vergaande aggregatie van kwalitatieve gegevens op de werkvloer naar meer kwantitatieve samenballende indices maakt deze indices in zeker opzicht ook minder betekenisvol. Dit geldt voor de beoordeelde, die de complexe werkelijkheid waarin hij handelt gereduceerd ziet tot enkele indices, maar ook voor de beoordelaar (bijv. Tweede Kamer) die het gevoel heeft niet echt 'greep' te krijgen op de werkelijkheid van de beleidspraktijk. De gesignaleerde dilemma's zijn in schema 1.2 in beeld gebracht.

Wat leert de literatuur ons nu over deze dilemma's? Waar hangt de mate en de aard van het gebruik van evaluatie van af? Welke vormen van evaluatie zijn er te onderscheiden? Welke filosofie zit daarachter? Welke ervaringen zijn ermee opgedaan? Welke aandachtspunten levert dat op voor de evaluatie van de 5N?

We gaan hieronder in op verschillende manieren van denken over evaluatie die we in de literatuur aantreffen, op inzichten omtrent het gebruik van evaluatie en op de ontwikkeling van en ervaringen met nieuwe vormen van evaluatie die gericht zijn op een grotere effectiviteit en op adequaat omgaan met de gesignaleerde paradoxen en spanningen⁵.

⁵ Voor een uitvoeriger weergave van de literatuurverkenning zij verwezen naar deel 4, hoofdstuk 2.

Schema 1.2: Kracht en Zwakte van geaggregeerde en niet geaggregeerde evaluatiemodellen

	<i>Geaggregeerde modellen Perspectief van verantwoording</i>	<i>Niet geaggregeerde modellen Perspectief van lering</i>
Kracht	<p>Aggregatie via een vast stramien maakt onderlinge vergelijking mogelijk.</p> <p>Modellen vereenvoudigen het oordelen.</p>	<p>Door per situatie een passend evaluatiestramien te ontwikkelen, wordt recht gedaan aan de betreffende beleidspraktijk en is het mogelijk steun te krijgen bij de geëvalueerde voor het evaluatiestramien.</p> <p>Modellen bevorderen de kans op leren.</p>
Zwakte	<p>Aggregatie van variëteiten aan gegevens tot gestandaardiseerde kengetallen leidt tot uitkomsten die weinig recht doen aan de geëvalueerde praktijk en biedt weinig aangrijpingspunten voor verbetering.</p> <p>De oordelen leiden niet tot verbetering.</p>	<p>Door een passende evaluatiestramien kan een specifieke beleidssector zichzelf onvergelijkbaar maken met anderen en daarmee het lerend vermogen juist verminderen.</p> <p>De evaluaties leiden niet tot een verbetering van de performance.</p>

2.3 Visies op evaluatie⁶

Evaluaties vervullen een belangrijke rol in beleidsprocessen, of zouden dat althans moeten doen⁷. Wat die rol is en hoe gerichte evaluaties vorm kan krijgen, daarover verschillende de meningen in de literatuur⁸, afhankelijk van hoe men naar beleidsprocessen kijkt.

Een eerste benadering beschouwt beleid als een rationele activiteit die gericht is op het realiseren van vooraf geformuleerde doelen. In deze visie is het van belang beleidsdoelen zo concreet en meetbaar mogelijk te formuleren. Evaluatie is er dan op gericht om achteraf vast te stellen of en in hoeverre de doelen zijn gerealiseerd, idealiter door onafhankelijke onderzoekers. De resultaten van een dergelijke evaluatie zijn dan bruikbaar om verantwoording over het gevoerde beleid af te leggen en dat beleid te beoordelen.

Een tweede benadering van evaluatie⁹, die bij hetzelfde beeld van beleidsprocessen kan aansluiten legt minder de nadruk op beoordelen en verantwoorden, en meer op leren. Deze benadering wijst erop dat optredende 'effecten' mogelijkwerwijs geen effect van het beleid zijn, en dat uitblijven van bedoelde effecten misschien niet aan slecht beleid te wijten is. De vraag of zulks al of niet het geval is, kan slechts worden beantwoord door onderzoek te doen naar de

⁶ Zie deel 4, met name par. 2.2.

⁷ Zie par. 2.4 over (niet-)gebruik van evaluatie.

⁸ De drie in de tekst onderscheiden vormen kunnen worden gekarakteriseerd met de in de literatuur gangbare onderscheiding tussen summatieve en formatieve evaluatie en die tussen individuele en collectieve. Bij de eerste visie gaat het dan om een summatieve evaluatie vanuit het perspectief van een actor, bij de tweede om een formatieve evaluatie uit het perspectief van een actor. De derde visie oriënteert zich op beleid dat in interactie of samenwerking tussen meerdere actoren totstandkomt en heeft uit de aard der zaak vooral een formatief karakter.

⁹ Zie hierover deel 4, par. 2.6.

mechanismen die in beeld gebrachte ontwikkelingen verklaren. Op basis daarvan kan worden geconcludeerd of beleid werkt en waarom. Dat levert aanknopingspunten voor aanpassing van het beleid op.

Een derde benadering vertrekt vanuit een geheel andere optiek op beleidsprocessen. Zowel het beleid als de resultaten daarvan komen in deze visie tot stand in wisselwerking tussen verschillende actoren. Die actoren hebben vaak wel doelstellingen, maar die komen niet altijd met elkaar overeen. Bovendien verschillende de visies op hoe de doelstellingen het beste kunnen worden gerealiseerd. Zowel de doelstellingen als de visies op realiseren ervan kunnen bovendien in de tijd veranderen, onder invloed van nieuwe inzichten, ervaringen, of als extern ervaren ontwikkelingen. Dat betekent niet dat geen verantwoording kan worden afgelegd, of niet van evaluatie zou kunnen worden geleerd, maar wel dat de perspectieven van verschillende actoren in evaluaties moeten worden betrokken¹⁰ en dat recht gedaan moet worden aan multipliciteit en verschuiving van doelen. Een exclusieve oriëntatie op vooraf dooréén van de partijen geformuleerde oogmerken is vanuit deze optiek weinig zinvol. Daarmee is overigens niet gesteld dat partijen zich niet (ook) afzonderlijk moeten (kunnen) verantwoorden.

In de 5N zijn elementen van beide aangeduide visies op beleidsprocessen waar te nemen. In onze verkenning van de Nederlandse situatie¹¹ blijken sommige actoren meer op de ene en andere meer op de andere visie georiënteerd zijn. Een belangrijke vraag is dus of en hoe evaluatievormen zijn te ontwikkelen die aan de verschillende hier aangeduide oriëntaties recht doen. Dat is niet alleen van belang in het licht van de sturingsstrategie van de 5N (par. 1.2), maar ook voor de effectiviteit van evaluatie. Daarover gaat de volgende paragraaf.

2.4 Gebruik van evaluatie¹²

Er is veel onderzoek gedaan naar het gebruik van evaluatieresultaten. Sommige factoren helpen om het gebruik ervan te stimuleren. Genoemd worden de technisch-wetenschappelijke kwaliteit, de reputatie van evaluatoren, het bestaan van formele procedures terzake, de timing van evaluatierapporten, de ervaren relevantie en de concreetheid van de aanbevelingen. Geen van deze factoren is echter altijd noodzakelijk en geen ervan is op zichzelf of in combinatie doorslaggevend. Een belangrijke conclusie uit onderzoek is dat het niet zozeer ‘objectieve’ karakteristieken van evaluaties of evaluatieprocessen zijn die de doorwerking ervan bepalen, maar vooral de percepties die actoren hebben van de inhoud, de bedoeling en de relevantie van de evaluatie.

Anders gezegd, het antwoord op de vraag in hoeverre actoren zich herkennen in vraagstellingen, methodes en uitkomsten van de evaluatie, maar ook hun perceptie van wat anderen met de evaluatie (kunnen) doen, zal in hoge mate bepalen of en vooral ook wat zij met evaluatieresultaten zullen doen. Soms kan dat vooral tot defensief of immuniserend gedrag leiden, soms tot tevreden achterover leunen, soms tot nadere reflectie en leren. Inzicht in de perspectieven van de verschillende actoren, de mogelijkheden die zij zien en de randcondities en bedreiging die zij ervaren, kunnen helpen evaluatievormen te ontwikkelen waarbij actief gebruik en gezamenlijk leren van evaluatie wordt gefaciliteerd.

¹⁰ Zie par. 2.5 en deel 4, par. 2.4 en 2.5.

¹¹ Zie hoofdstuk 3 en deel 4, hoofdstuk 3 e.v.

¹² Zie ook deel 4, par. 2.3 en hoofdstuk 3 e.v.

2.5 Vormen van evaluatie¹³

De literatuur laat een bonte verzameling van vormen van evaluatie en hun toepassingen zien. We geven daar geen compleet overzicht van, maar noemen hier enkele vormen en onderscheidingen die voor de evaluatie van de 5N van belang lijken te zijn.

Om te beginnen het onderscheid tussen interne en externe evaluatie. De eerste vorm wordt door de geëvalueerde zelf geïnitieerd¹⁴ en is doorgaans gericht op de vraag of, hoe en waarom bepaalde (beleids)maatregelen werken en of en waarom er onbedoelde effecten optreden. De verzamelde gegevens en analyses vormen basis voor reflectie en organisatorische leerprocessen. De geëvalueerden vervullen zelf een centrale rol in het bepalen van de vraagstelling en aanpak van de evaluatie en in het beoordelen van de resultaten ervan en het daaraan verbinden van consequenties voor bijstelling van het beleid. Bij externe evaluatie voert in het algemeen een onafhankelijke of toezichhoudende instantie de evaluatie uit. Hierbij gaat het vaak om beoordeling van het gevoerde beleid, wat bijv. financiële consequenties kan hebben. Vraagstelling en methoden worden hier veelal door de evaluator of zijn opdrachtgever bepaald.

De aanduidingen 'intern' en 'extern' betreffen overigens de uiteinden van een continuüm. Er zijn tussenvormen denkbaar, bijv. als een afdeling van een overheidsinstantie een evaluatie initieert van beleid dat onder regie staat van een andere afdeling van diezelfde instantie.

Het onderscheid tussen beoordelende en reflexieve evaluatie dat in par. 2.2 is geïntroduceerd is nauw verwant aan het onderscheid tussen intern en extern. In principe kan een externe evaluatie tot interne reflectie aanleiding geven, maar of dat gebeurt, lijkt vooral afhankelijk van of en hoe de bevindingen aansluiten bij gedachten die binnen de geëvalueerde instantie al leven¹⁵. Omgekeerd kan interne evaluatie worden gebruikt voor verantwoording naar buiten, maar veelal zal de buitenwereld behoefte hebben aan een onafhankelijke second opinion.

Een tweede onderscheid betreft dat tussen kwantitatieve en kwalitatieve evaluatie. De eerste vorm wordt vaak geassocieerd met objectiviteit en van daaruit geschikt geacht voor verantwoording en beoordeling (mits de doelstellingen voldoende scherp zijn geformuleerd), maar ook om beter inzicht in beleidseffecten te krijgen en op grond daarvan te leren. De kritiek op deze benadering is dat lang niet alle beleidsdoelstellingen en -effecten te kwantificeren zijn. En als dat wel het geval is, blijkt de relatie tussen beleid en effect niet ondubbelzinnig vast te stellen (Pawson and Tilley 1997). Bovendien wordt betoogd dat in sociale en maatschappelijke processen interpretaties en zingeving door actoren in hoge mate het verloop en de uitkomsten bepalen - of zelfs de uitkomsten zijn (Guba and Lincoln 1989). Kwalitatieve beelden, waaronder nadrukkelijk perceptie en beleving van actoren, zijn daarom van belang om beleid te beoordelen en erover te leren.

Op basis van inzichten met betrekking tot het gebruik van evaluatie en met betrekking tot de aard van complexe beleidsprocessen zijn verschillende nieuwe vormen van evaluatie ontwikkeld die tezamen als multi-actorevaluatie worden aangeduid (deel 4, par. 2.4 en 2.5). Aan deze evaluaties ligt de gedachte ten grondslag dat er in complexe beleidsprocessen een pluraliteit aan waarden bestaat en dat de uiteindelijke effecten van beleid resulteren uit de interactie tussen de actoren. Er zijn daarom twee redenen waarom percepties van actoren een rol moeten spelen in evaluatie. De eerste is dat zonder rekenschap van percepties het verloop van beleidsprocessen niet is te begrijpen. De tweede is dat evaluaties waarin actoren zich niet op

¹³ Zie ook deel 4, par. 2.4 e.v.

¹⁴ We spreken hier ook van 'interne' evaluatie als de geëvalueerde zelf het initiatief neemt om een externe instantie een evaluatie te laten verrichten.

¹⁵ Zie hierover verder deel 4, par. 2.3 en hoofdstuk 4.

enigerlei wijze herkennen, minder effectief zullen zijn. In multi-actorevaluaties wordt daarom door verschillende actoren geparticipeerd, zowel bij het formuleren van de probleemstelling(en), het verzamelen van gegevens als het trekken van conclusies.

De betrokkenheid van een groot aantal publieke en private actoren bij ruimtelijk (relevant) beleid, lijkt het belang van multi-actorevaluatie te onderstrepen. Een belangrijke vraag voor te ontwikkelen evaluatiearrangementen is hoe daarbij aan de behoefte aan verantwoording en reflectie kan worden voldaan en hoe daarbij verschillende soorten kwantitatieve en kwalitatieve gegevens kunnen worden benut. In deel 2 (handleiding) worden hieromtrent suggesties gedaan.

3. Denken over en ervaringen met beleidsevaluatie in Nederland

De gedachte dat beleid geëvalueerd moet worden, ligt voor de hand en is al oud. Zo ontstond in 1477 al een voorloper van de huidige Algemene Rekenkamer (De Vries 2000, p. 27 e.v.). Lange tijd heeft de nadruk gelegen op financiële controle en verantwoording: zijn de uitgaven van het Rijk rechtmatig? Ook de aandacht voor evaluatie van de doelmatigheid en doeltreffendheid van beleid is niet nieuw. In de Comptabiliteitswet van 1927 wordt reeds van doelmatigheidscontrole gesproken (De Vries 2000, p. 29). De afgelopen decennia is de Algemene Rekenkamer daarop een groter accent gaan leggen. En vanaf 1970 wordt ook door de Rijksoverheid, met name vanuit het Ministerie van Financiën gestreefd naar prestatiebegrotingen, waarin de doelen van beleid worden geanalyseerd en zo mogelijk vertaald in meetbare prestatie-indicatoren (zie deel 4, par. 3.3). Dit streven is beperkt succesvol geweest. Ook pogingen op gemeentelijk niveau om door middel van productbegrotingen en kengetallen het beleidsproces en de resultaten daarvan transparanter te maken, verliepen moeizaam (zie deel 4, par. 3.2). Het eerste probleem is dat veel beleid nauwelijks meetbare output en outcomes heeft. Een ander probleem is dat de productie van beleidsprestaties geen homogeen en transparant proces is, waarin effecten makkelijk zijn toe te delen.

Toch blijft de behoefte aan resultaatgerichte beleidsvoering aanwezig en daaraan gekoppeld de behoefte om resultaten zichtbaar te maken. Het door het ministerie geëntameerde proces van beleidsverantwoording afleggen (VBTB) is daarvan een uitdrukking (zie deel 4, par. 3.4). Hoewel meningen over de mate waarin beleid en beleidseffecten in meetbare indicatoren zijn te vangen, soms scherp tegenover elkaar blijven staan, valt een zeker convergentie waar te nemen. Enerzijds wordt erkend dat outcomes niet altijd meetbaar en aan het beleid toeschrijfbaar zijn, en dat output soms beter in proces termen kunnen worden geschreven dan in termen van substantiële resultaten. Anderzijds wordt het beantwoorden van de vraag waar beleid toe leidt en in hoeverre er maatschappelijke problemen mee worden opgelost of gewenste maatschappelijke veranderingen worden bevorderd, steeds meer als noodzakelijk element van verantwoording gezien. Daarnaast is dat ook voor ontwikkelen van nieuw beleid van groot belang. Evaluatie en onderzoek over beleidseffecten mogen zich dan ook in een grote belangstelling verheugen. Hierin is een aantal lijnen te ontdekken.

Om te beginnen wordt door instanties als het Ministerie van Financiën en de Algemene Rekenkamer nagedacht en met anderen meegedacht over hoe VBTB nader ingevuld kan worden. De door de Minister van Financiën vastgestelde Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid (2001) is daar een voorbeeld van. Hierin worden richtlijnen gegeven voor het zo veel mogelijk definiëren van 'betrouwbare en controleerbare prestatiegegevens waarmee de doeltreffendheid en doelmatigheid van beleid en de doelmatigheid van de bedrijfsvoering kunnen worden beoordeeld'. Ook worden methodisch-technische criteria geformuleerd waaraan evaluaties moeten voldoen. Ten slotte wordt voorgeschreven dat alle beleidsdoelen in beginsel ten minste eens per vijf jaar geëvalueerd moeten worden.

De evaluatie van de VINEX door de Werkgroep 5N van de Tweede Kamer (zie deel 4, par. 5.2) weerspiegelt ook de behoefte aan herkenbare en meetbare beleidsdoelen (die de werkgroep achteraf in de Vierde Nota te weinig aanwezig vindt).

Onze verkenningen bevestigen de indruk dat velen het nut van dit streven onderschrijven, maar dat er tegelijkertijd vaak twijfel bestaat (bijvoorbeeld op decentraal niveau) over de mate waarin beleidsprestaties in kwantitatieve indicatoren zijn te vangen. Vanuit deze positie wordt benadrukt dat het weinig zin heeft alléén effecten in beeld te brengen. Er moet ook onderzoek worden gedaan naar oorzaken van die effecten: in hoeverre zijn ze aan beleid toe te schrijven en in hoeverre aan andere al dan niet beïnvloedbare factoren?

Een tweede lijn is erop gericht om reeds voor de implementatie van beleid meer inzicht te krijgen in de (mogelijke) resultaten en (onbedoelde) gevolgen ervan. Onderzoek en ex ante evaluatie komen dan in beeld (deel 4, par. 5.3, 6.1 en 6.2). Net als bij meer verklarende evaluatie achteraf gaat het hier niet primair om verantwoording, maar om reflectie en het entameren van leerprocessen. Een probleem bij dergelijke evaluaties blijkt te liggen in de gehanteerde veronderstellingen. Veelal zullen die door sommige actoren meer geaccepteerd worden dan door andere, met directe gevolgen voor hoe overtuigend men de resultaten van het betreffende onderzoek acht. Soms zijn modellen waarmee verkenningen worden uitgevoerd zo ingewikkeld dat de veronderstellingen die erin gemaakt worden, onzichtbaar zijn voor betrokkenen. Dat kan zowel tot ongefundeerd vertrouwen als tot ongefundeerd wantrouwen leiden.

Een derde lijn in Nederlandse evaluatiepraktijken, die steeds meer wordt aangetroffen, betreft evaluaties waarbij meerdere stakeholders betrokken zijn. ‘Stakeholders’ zijn betrokkenen of belanghebbenden. Daaronder vallen bij beleidsvorming en -uitvoering betrokken overheden, particuliere organisaties die ruimtelijk beleid mede realiseren, maatschappelijke organisaties die ruimtelijke, mobiliteits- en milieubelangen representeren en (groepen) burgers¹⁶. Er kunnen verschillende vormen van evaluatie met stakeholders worden onderscheiden. Eén daarvan is dat de evaluatie zodanig wordt opgezet dat die voor meerdere betrokkenen relevante informatie oplevert, zoals bij de Monitor Stedelijke Vernieuwing (zie deel 4, par. 6.3). Een andere vorm betreft gezamenlijk opdrachtgeverschap (een voorbeeld in deel 4, par. 6.2). Gekoppeld aan het NVVP bestaat er een evaluatieplatform waarin in betrokken overheden vertegenwoordigd zijn. Het platform bepaalt vooraf wat en hoe er geëvalueerd wordt en accordeert de resultaten. Het platform is daarmee eindverantwoordelijk voor de evaluatie. Combinatie van zelfevaluatie en een daarop volgende externe visitatie is weer een andere vorm. Het interessante van al deze vormen is dat probleemstellingen van verschillende actoren in evaluatietrajecten aan elkaar worden gekoppeld. Dat komt de herkenbaarheid en gepercipieerde bruikbaarheid van de evaluaties ten goede. Daarnaast bieden dergelijke evaluatietrajecten, veel meer dan bijv. het VBTP-proces, mogelijkheden om gewijzigde inzichten en verschoven doelstellingen tot uitgangspunt te nemen.

De genoemde lijnen zijn niet zozeer elkaar uitsluitende vormen van evaluatie, maar te onderscheiden denkmodellen, die wellicht in concrete evaluatiepraktijken aan elkaar gekoppeld kunnen worden. We zullen daar in deel 2 en 3 ook voorstellen voor doen. De derde lijn vestigt echter meer dan de beide andere de aandacht op het belang van ‘aansluiting’ van evaluatieprocessen en -bevindingen op de perspectieven van betrokken actoren, als ‘doorwerking’ en ‘leren’ belangrijke criteria voor zinvolle evaluatie zijn. Het begrip ‘aansluiting’ verwijst hier naar herkenbaarheid als relevant voor de betreffende actor, naar de perceptie dat voldoende rekening is gehouden met ervaren randcondities en dergelijke. Aansluiting zoeken is iets anders dan zeggen wat de betreffende actor wil horen. Het gaat er veeleer om dat pijnlijke of ongewenste informatie eerder tot leereffecten zal leiden als er óók herkenningpunten in zitten.

Ook voor verantwoording kan multi-actorevaluatie extra inzichten opleveren, bijvoorbeeld doordat een grotere verzameling van relevante condities en mechanismen in beeld komt.

We kunnen concluderen dat de drie visies op evaluatie die we in de literatuur aantreffen, ook in de Nederlandse ontwikkelingen aanwezig zijn. Waar het beleid, zoals dat op het gebied van de RO het geval is, in wisselwerking tussen actoren totstandkomt (en men daar ook bewust voor kiest), lijkt het derde perspectief aan belang toe te nemen. Naast evaluatie gericht op verantwoorden, beoordelen en evaluatie gericht op meer inzicht voor de evaluerende of geëvalueerde instantie, is het daarom in een robuust evaluatiearrangement voor de 5N nodig te voorzien in multi-actorevaluatie (bijvoorbeeld gekoppeld aan bestaande beleidsarena’s) en

¹⁶ Bij de uitwerkingen in dit rapport beperken we ons overigens tot publieke actoren.

koppeling van evaluaties die op verschillende niveaus plaatsvinden. De kunst zal zijn om vormen te vinden die het voor verschillende actoren aantrekkelijk maken op relevante punten in een dergelijk arrangement te investeren.

4. Conclusies verkenningen

Op grond van onze verkenning van literatuur, bestudering van recente evaluatiepraktijken en gesprekken met betrokkenen en deskundigen, is een aantal conclusies te formuleren. We beperken ons tot aspecten die bij de evaluatie van de 5N van belang zijn en aanknopingspunten bieden voor de te presenteren handleiding en gebruiksaanwijzingen.

Ten eerste valt te constateren dat er verschillende opvattingen bestaan over plaats, doel en functie van evaluaties. In relatie tot de 5N is wellicht het onderscheid tussen evaluatie als instrument voor verantwoording over en beoordeling van beleid en evaluatie als hulpmiddel voor reflectie en leerprocessen het belangrijkste. In het eerste geval wordt vaak de koppeling gemaakt met externe evaluatie gericht op vaststellen van de mate van bereiken van vooraf gestelde concrete doelen, de laatste met interne interactieprocessen gericht op de vraag hoe het beter zou kunnen.

Naar onze mening is een dergelijk scherp onderscheid niet noodzakelijk. De praktische vraag is of en hoe beide invalshoeken bij de evaluatie van de 5N in één traject combineerbaar zijn. Een aandachtspunt daarbij is wat het aggregatieniveau van de evaluatie moet zijn. Andere onderscheidingen betreffen product- versus procesevaluatie, en evaluatie vanuit a-priori doelen versus evaluatie vanuit zich ontwikkelende inzichten. Waar in de 5N ruimtelijke kwaliteit het centrale oogmerk is en het sturingsmodel in hoge mate mikt op interactie binnen en tussen verschillende bestuurlijke en publiek-private constellaties, en op een combinatie van ontwikkeling en beheersing, lijken beide invalshoeken van deze onderscheidingen relevant. De vraag hoe een combinatie kan worden gerealiseerd, stelt zich hier dan ook nadrukkelijk

Een tweede hoofdpunt is dat gebruik of doorwerking van evaluatie allerminst vanzelfsprekend is. Ongeacht wat het primaire oogmerk van evaluatie is, is dit een belangrijk punt. Een goede timing is een eerste voorwaarde voor effectieve evaluatie. Op welke momenten in beleidstrajecten kan er zinvol iets worden gedaan met evaluatieuitkomsten? Belangrijk is ook de vraag hoe de evaluatie aansluit op de 'wereld' van actoren, bij de handelingsmogelijkheden die zij zien en bij zaken die voor hun van belang zijn. Bij ruimtelijk beleid zijn veel actoren betrokken met uiteenlopende visies en perspectieven. De vraag is hoe daarmee in evaluatietrajecten kan worden omgegaan.

In de derde plaats is duidelijk geworden dat positionering en organisatie van evaluatieactiviteiten een belangrijk aandachtspunt is. Zo lijken externe evaluaties vaker kritische oordelen op te leveren en een dwingender karakter te hebben, maar zij roepen ook vaker defensieve en immuniserende reacties op bij geëvalueerden. Evaluaties die organisaties zelf initiëren kunnen wellicht op een opener ontvangst rekenen, maar zijn haast per definitie door de eigen perspectief gekleurd. Ook hier is de vraag of de voordelen kunnen worden bevorderd en de nadelen afgeremd en of combinatie mogelijk en zinvol is.

Een ander 'organisatie'-aspect is wie er op welke wijze bij de evaluatie betrokken zijn. Verschillende vormen van multi-actorevaluatie zijn met het oog op draagvlak, aansluiting en doorwerking in relatie tot de 5N interessante opties. Daarbij is dan wel meteen de vraag aan de orde onder welke condities evaluatietrajecten voor (alle) betrokken actoren voldoende interessant zijn om in te investeren.

Een laatste organisatieaspect betreft de relatie van evaluatie met het beleidsproces zelf. Aan de ene kant is er de opvatting, veelal in relatie tot verantwoording, dat evaluatie onafhankelijk en gescheiden van het beleidsproces vorm moet krijgen. Aan het andere uiterste is evaluatie een integraal onderdeel van beleidsprocessen en nauwelijks te onderscheiden van bijv. kwaliteitszorg. Keuzes die op deze dimensie worden gemaakt, hangen samen met visies op sturing en op het belang van het vinden van aansluiting bij actoren.

Ten slotte is er reden apart aandacht aan meetbaarheid te besteden. Aan de ene kant bestaat de opvatting dat duidelijke doelen waarvan de realisatie is te meten een voorwaarde is voor zinvolle evaluatie ('anders kun je er niets van leren'). Daartegenover staat de gedachte dat complexe beleidsprocessen doelzoekend zijn en dat gekwantificeerde indicatoren weinig recht doen aan de complexiteit en bovendien weinig aanknopingspunten voor sturing bieden. De vraag is hoe beide opvattingen recht te doen. Een zekere convergentie van opvattingen tekent zich af. Zo zal niemand ontkennen dat gekwantificeerde bevindingen soms reflectie en leerprocessen kunnen ontketenen of beïnvloeden. Omgekeerd wordt door vrijwel iedereen erkend dat niet alle beleid in meetbare doelen is te vertalen. In het kader van VBTB worden ook andere vormen om doelen en resultaten zichtbaar te maken en te vergelijken acceptabel geacht. Daarbij kunnen bijvoorbeeld ook tevredenheidsindicatoren of procescriteria een rol spelen.

Deel 2: Handleiding voor evaluatief verantwoord en leren

1. Inleiding

Deze handleiding bevat voorstellen voor de inrichting van een evaluatiearrangement dat passend is voor het in 5N uitgezette beleid. Het begrip ‘evaluatiearrangement’ verwijst naar het geheel aan evaluaties dat wordt uitgevoerd en naar de wijze waarop die evaluaties in elkaar grijpen. We oriënteren ons in onze voorstellen op de eerstkomende jaren.

Een concrete evaluatie vormt in onze optiek de bouwsteen van het hier voorgestelde evaluatiearrangement. Toch is het arrangement meer dan alleen een optelsom van de tientallen evaluaties die de komende jaren zullen worden uitgevoerd. Waar wij in de handleiding vooral naar zoeken is de lijn die nodig is om van de veelheid aan bouwstenen meer te maken dan een hoop losse stenen. We proberen met andere woorden samenhang aan te brengen in de evaluaties.

Dit gebeurt niet met de bedoeling een mooie geordende werkelijkheid te scheppen. In de praktijk zullen diverse evaluaties langs elkaar heen gaan schuren en zullen actoren elk hun eigen insteek kiezen. Soms ook zullen evaluaties ingezet worden in de strijd die verschillende overheidsorganisaties denken te moeten voeren. Het is naïef te veronderstellen dat het hier voorgestelde evaluatiearrangement betrokkenen direct overhaalt de strijdbijl te begraven. Wat wij wel aanbieden, is een handleiding van samenhangend evalueren en een stel gebruiksaanwijzingen, die betrokkenen – *als ze dat willen* – de mogelijkheid bieden op drie punten de kwaliteit en effectiviteit van evaluatie te vergroten:

- de kracht van *eigenstandige evaluaties* optimaal te benutten in de zoektocht naar ruimtelijke kwaliteit;
- de *verbindingen tussen* de verschillende *evaluaties* zo te leggen dat ze elkaar versterken;
- *evaluaties zo in te bouwen in het interorganisatorische beleidsproces* dat ze daadwerkelijk een rol van betekenis kunnen spelen als hulpmiddel om te *verantwoorden*, te *leren* en *samen te werken*.

In deze handleiding geven we een aantal *regels* die wijzen op mogelijkheden die er zijn om evaluaties zo op te zetten dat ze elkaar kunnen versterken. We onderscheiden daarbij de drie niveaus die samenhangen met bovenstaande drie doelen:

- een *evaluatie* moet zo goed in elkaar zitten dat deze combineerbaar is;
- een *samenstel van enkele evaluaties* die elkaar kunnen versterken;
- het *evaluatiearrangement* moet betrokken aanzetten tot adequate samenwerking

Op elk niveau stellen we een drietal regels voor, die we in enkele schema's verbeelden. Het gaat hierbij als het ware om een verbeelding van de implementatie van de 5N, waarbij uitsluitende de evaluatiemomenten zichtbaar worden gemaakt. Per voorstel geven we een korte toelichting.

Het arrangement, maar ook de regels op het niveau van een evaluatie en het samenstel van evaluaties zijn bruikbaar voor alle actoren. De gebruiksaanwijzingen in deel 3 geven voor verschillende actoren aan wat vanuit hun belang een zinvol evaluatieproces kan zijn.

2. Regels voor het ontwikkelen van een kwalitatief hoogwaardige evaluatie

Om de kracht van evaluaties optimaal te benutten moeten deze allereerst een hoge interne kwaliteit hebben. Op dit punt kunnen eerdere voorstellen tot evaluatie van de Vierde Nota RO (Herweijer et al. 1990) al het nodige meegeven aan de actoren, maar we menen dat er op een aantal punten aanpassingen gewenst zijn. Herwijer c.s. stelden de volgende stappen in een concreet evaluatietraject voor:

- stap 1: formulering doelstelling
- stap 2: inhoudsanalyse van het beleidsdocument om daaruit af te leiden wie wat wanneer zou moeten doen
- stap 3: formulering evaluatievragen: op basis van stap 2 criteria vaststellen voor doorwerking, conformiteit, doelbereiking en effectiviteit
- stap 4: empirisch onderzoek

Vooraf in de eerste stappen wordt nauw aangesloten bij de (oorspronkelijke) oogmerken van de centrale beleidsmaker. In de dynamische multi-actorcontext waarin ruimtelijk beleid plaatsvindt, is heroverweging van deze enkelvoudige oriëntatie nodig. Een dergelijke heroverweging is een van de gronden waarop de eerste drie regels die we hieronder presenteren zijn gebaseerd..

Regel 1: Zet elke evaluatie zorgvuldig op conform breed aanvaarde regels

Als eerste een regel die veel lijkt op een open deur. Wie is er nu niet voor een zorgvuldige opzet, zal de geroutineerde lezer van onderzoeksrapporten verzuchten? De regel krijgt evenwel betekenis als we deze nader specificeren. Zorgvuldigheid heeft namelijk gelijktijdig een aantal betekenissen.

Ten eerste is er de methodische zorgvuldigheid. Met name wetenschappers zullen op deze vorm van zorgvuldigheid wijzen. Wij pleiten er ook voor om deze kwaliteit van evaluaties ook zorgvuldig te blijven koesteren. Een goede reflectie op de opzet van een evaluatie, de procedures, de explicitering van begrippen en dergelijk draagt bij aan verantwoord evalueren. Op dit punt kan ons inziens volop geput worden uit vele eerdere publicaties over evaluaties. Voor de opzet van een concrete evaluatie in de ruimtelijke ordening kan bijvoorbeeld goed worden voortgebouwd op het stramien zoals dat is voorgesteld door Herweijer e.a (1990). Dat stramien is gemaakt voor de Vierde Nota Ruimtelijke Ordening en de daarin gepresenteerde methodiek lijkt ons nog steeds hanteerbaar.

Onze toevoegingen hebben met name betrekking op de inbedding van evaluaties. Om deze inbedding mogelijk te maken, voegen we wel een belangrijke additionele voorwaarde toe aan het stramien van Herweijer e.a. In hun voorstellen vloeien de tijdens de evaluatie te stellen vragen, de gehanteerde beoordelingscriteria en andere aandachtspunten bij evaluaties voornamelijk voort uit de ‘inhoud’ van de betreffende beleidsnota. Op dat punt heeft het bestuurskundig denken zich evenwel verder ontwikkeld en wel in de richting van beleidsnetwerken en beleidsdynamiek.

De consequentie daarvan is dat *methodische zorgvuldigheid zich ook uitstrekt tot het adequaat in kaart brengen van de perspectieven* (en dus evaluatievragen, beoordelingspunten en aandachtspunten) *van diverse actoren* in relevante beleidsnetwerken. Methodische zorgvuldigheid strekt zich verder ook uit tot de voortschrijdende inzichten van betrokken actoren. *De dynamiek van de evaluatievragen, beoordelingspunten en aandachtspunten wordt daarmee zelf ook onderwerp van methodische reflectie.* Deze verbreding sluit niet de mogelijkheid uit

om specifieke conclusies te trekken aan de hand van een bepaald criterium, of vanuit een specifieke actor, maar houdt wel de mogelijkheid op om tegelijkertijd ook andere conclusies te trekken.

Een tweede invulling van zorgvuldigheid betreft de rol en positie van de evaluerende instantie of persoon zelf. Breed wordt de opvatting verwoord dat deze een onafhankelijke positie moet kunnen innemen. Wij vragen extra aandacht voor deze positie. Onafhankelijkheid wordt vaak ingevuld als ‘een evaluerende instantie van buiten’, die geen relatie heeft met de geëvalueerde¹⁷. Ook wij zijn de mening toegedaan dat extern onderzoek in een zinvolle en zelfs onmisbare rol in evaluatietrajecten vervullen. Dit geldt met name daar waar het zelfkritische vermogen van de geëvalueerde onder druk staat. Een nadeel is echter dat evaluaties van buitenaf vaak alleen gebaseerd kan zijn op schriftelijke stukken en interviews. Dat leidt niet steeds tot een afgewogen oordeel. We pleiten er daarom voor onafhankelijkheid niet alleen in te vullen in termen van losstaand van de geëvalueerde, maar ook in te vullen als ‘zodanig goed op de hoogte van het krachtenveld in beleidsnetwerken dat een rijke, afgewogen beoordeling mogelijk wordt’. Met name daaruit zullen actoren lessen trekken. Bij regel 4 bouwen we voort op deze gedachte door te stellen dat combinaties van externe evaluaties en zelfevaluatie de effectiviteit van evaluatiearrangementen aanmerkelijk lijkt te vergroten (vgl. hoofdstuk 3).

Een derde invulling van zorgvuldigheid heeft ten slotte te maken met de inbedding van de evaluatie in het relevante beleidsnetwerk zelf. Anders geformuleerd: een opzet waarin veel van de betrokken beleidsmakers zich herkennen, zal niet alleen leiden tot een grotere bereidheid om de benodigde informatie te leveren, maar ook de aandacht voor de uitkomsten versterken. Herkenning geldt voor zowel de evaluatievragen, de beoordelingscriteria als de keuze van de evaluator en de vaststellingsprocedure van de conclusies.

Regel 2: Zorg voor goede timing van evaluaties in relatie tot beleidsprocessen op verschillende niveaus

Timing is voor de doorwerking van evaluaties van groot belang. We wijzen erop dat beleidsmakers sowieso al meer aandacht hebben voor beleidsontwikkeling dan voor beleidsevaluatie. Het is dus een goed teken van ‘klantgerichtheid’, wanneer de evaluatie op een moment plaats vindt en gepresenteerd wordt, dat de kans op aandacht het grootst is. Vaak vinden evaluaties te laat plaats, waardoor de uitkomsten nauwelijks nog een rol spelen bij het aanpassen van beleid en het formuleren van nieuw beleid. Wij pleiten er daarom voor evaluaties vroegtijdig in te plannen in beleidsprocessen en met name beschikbaar te hebben op de momenten dat beleidsmakers open staan voor beleidsaanpassing.

Timing is in een beleidsnetwerk nog gewichtiger. Omdat het bij de 5N niet alleen gaat om doorwerking bij centrale overheden, is het nodig aandacht te besteden aan het tijdspad van beleidstrajecten op decentraal niveau. Zo moeten medeoverheden conform de 5N vóór eind 2005 allerlei zaken (rode en groene contouren, regionale landschapsparken, waterbeleid) in streekplannen neerleggen. Vanuit het perspectief van de centrale overheid ligt het wellicht voor de hand om *na 2006* na te gaan in hoeverre een en ander daadwerkelijk is gerealiseerd en aan de in de PKB gestelde eisen voldoet. Provincies evenwel zullen evaluaties die *vooraf* gaan aan de vaststelling van streekplannen vermoedelijk meer gebruiken.

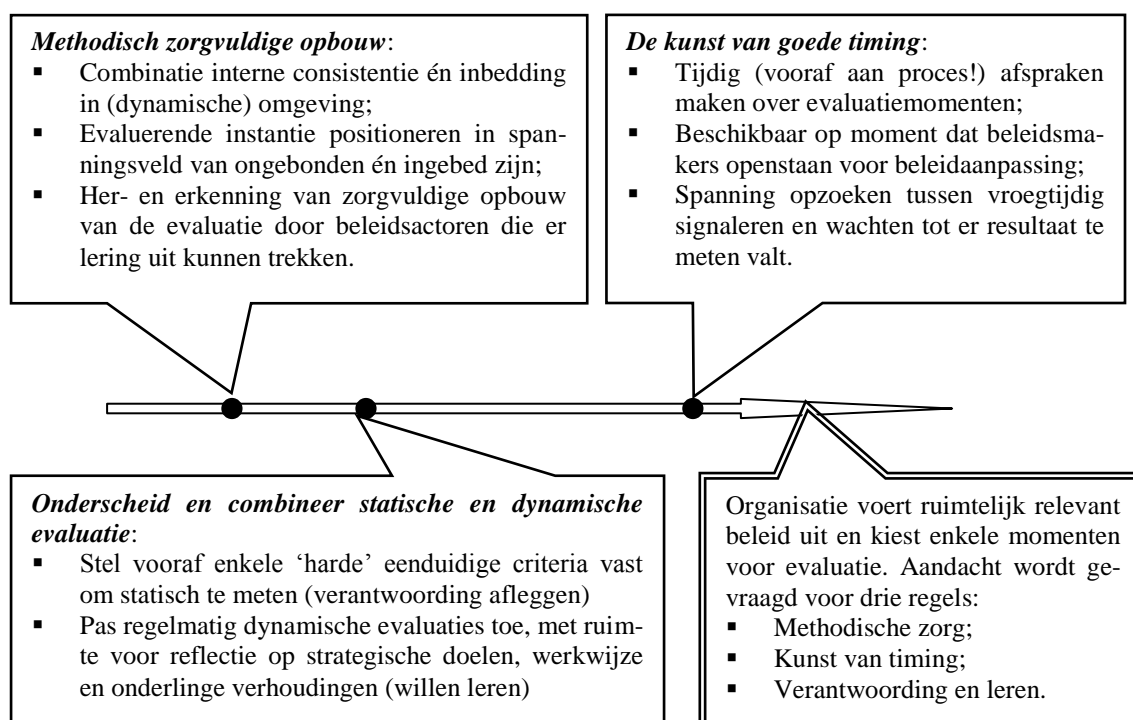
¹⁷ Zoals eerder gesteld, zijn er allerlei tussenvormen tussen zuiver interne en zuiver externe evaluatie mogelijk. Het betoog is er hier evenwel op gericht duidelijk te maken dat het voor effectieve evaluatie nodig is vormen van ‘onafhankelijkheid’ te kiezen die tegelijkertijd betrokkenheid op en deskundigheid over de onderhavige materie inhouden.

Keuze voor het ene moment, hoeft keuze voor de ander niet uit te sluiten, maar het stellen en beantwoorden van de vraag ‘*wat wanneer voor wie interessant en wie dus wanneer in evaluatie zal willen investeren*’, lijkt van groot belang voor de kwaliteit en de effectiviteit van die evaluatie. Behalve afstemming tussen beleidstrajecten van verschillende bestuurslagen, is ook afstemming tussen departementen van belang. Hoe en wanneer kunnen evaluaties in het kader van de 5N van nut zijn voor beleidstrajecten rondom het 2e Structuurschema Groene Ruimte of het 3e NVVP en omgekeerd?

Regel 3: Onderscheid en verbind statische en dynamische evaluatie

De 5N is een strategische nota. De oogmerken van het (centrale) beleid worden er in algemene termen in aangeduid. Doelen worden in de loop van het proces geconcretiseerd en kunnen regionaal verschillende worden ingevuld en door de tijd heen verschuiven. Hetzelfde geldt voor instrumenten, oplossingen voor ruimtelijke problemen etc. Het ligt vanuit het sturingsmodel van de 5N voor de hand in evaluatie recht te doen aan een variëteit aan concretiseringen en aan voortschrijdend inzicht. Het inspelen op variëteit en dynamiek maakt de oorspronkelijke doelen uit de 5N (of debat over de betekenis ervan) niet (voor iedereen) irrelevant. We achten het ook zeer wel mogelijk dat het ministerie van VROM bij het vaststellen van de convenanten met medeoverheden in het kader van het uitvoeringsprogramma van de 5N aangeeft op een aantal punten hard te willen evalueren. We spreken dan van een statische evaluatie. Zeker ook politici zoals de Tweede Kamer zullen behoefte hebben aan statische evaluaties.

Schema 2.1. Regels voor individuele evaluaties



Het is echter wel zaak daarin bijzonder gericht en terughoudend te zijn. Vervolgens is het nodig om af te spreken dat op andere punten sprake is van een dynamische evaluatie, waarbij steeds expliciete reflectie nodig is op de te stellen vragen en de invulling van beoordelingscriteria. Statische evaluaties zullen daarbij vaak door een actor worden uitgevoerd met de bedoe-

ling verantwoording af te leggen, terwijl dynamische evaluaties eerder gericht zijn op meer actoren en op leren en leren samenwerken. Het is van belang in in de samenhang van evaluaties zo helder mogelijk te zijn over wat de focus is van elke evaluatie en hoe ze zich onderling verhouden. Juist de combinatie is aantrekkelijk, omdat dat toegevoegde waarde kan hebben voor bezinning op oude en nieuwe visies en criteria.

In schema 2.1 worden de hierboven besproken regels visueel voorgesteld. De horizontale pijl geeft een afzonderlijk evaluatietraject weer, De zwarte bolletjes representeren de vormgeving van het evaluatietraject op basis van de geformuleerde regels. Daarbij is niet een bepaalde tijdsvolgorde geïmpliceerd.

3. Koppelen van evaluaties versterkt het lerend effect

Eerder al hebben we aangegeven dat verschillende soorten evaluaties ieder hun eigen effect hebben. Vaak ook zit er aan elke soort evaluatie een tekortkoming. In het onderstaande ontwikkelen we een aantal koppelingsregels die het lerend effect van evaluaties mogelijk kunnen versterken.

Regel 4: Koppel interne (zelf)evaluaties aan externe

Er is geconstateerd dat interne evaluaties de neiging hebben te veel begrip te tonen voor de situatie waarin de organisatie zich bevindt. Ook bestaat de neiging om de schuld voor een matige performance te externaliseren. Er wordt dan niet voldoende geleerd van evaluaties.

Dit zou tot de aanbeveling kunnen leiden dat evaluaties vooral door externen moet worden gedaan. Maar ook daaraan is een aantal tekortkomingen verbonden: ten eerste hanteren externen lang niet altijd indicatoren die betekenisvol zijn voor degenen die worden geëvalueerd. Bovendien bestaat vaak ook de neiging om abstracte samengestelde criteria te hanteren. Deze keuze is meestal logisch vanuit de behoefte ‘alles overziend’ tot een eindoordeel te komen. In dit aggregatieproces gaat wel eens veel betekenisvolle informatie verloren. En verder – en daarmee komen we bij de tweede kritiek – verleidt deze geaggregeerde kennis de evaluatoren of de gebruikers van de evaluatie nogal eens tot oordelen in termen van goed of slecht. Als het oordeel dan goed is, leert de geëvalueerde organisatie niets en als het oordeel slecht is, neigt de beoordeelde er vaak toe om zich te immuniseren. Dit kan door aan te geven dat de methode niet goed is of door erop te wijzen dat hun omstandigheden veel slechter zijn dan die van anderen en dat ‘ze er dus niets aan kunnen doen’. Als dat laatste gevoel domineert, is er ook geen leereffect. Combinatie van interne en externe componenten in een evaluatietraject kan hier soelaas bieden. Een interne zelfstudie is dan de kern van het leerproces, maar het proces verloopt beter en sneller onder druk van de externe toets. Indien het ruimtelijke planbureau er inderdaad in slaagt om voor alle overheden te werken, zou deze mogelijk een belangrijke rol kunnen spelen als externe toetsingsinstantie.

Op zichzelf kunnen de eerder genoemde tussenvormen tussen zuiver interne en zuiver externe evaluatie in dit verband ook een bijdrage leveren. In onze gedachtegang is dat echter niet een alternatief voor de hier bepleite combinatie, maar vormt het een interessant element in zo’n combinatie. Anders gezegd: de wetenschap bij betrokken, intern, extern en daartussenin, dat hun evaluaties worden gekoppeld aan en geconfronteerd met evaluaties vanuit een ander perspectief kan het (zelf)kritisch en lerend vermogen vergroten.

Regel 5: Verbind beoordeling en reflectie

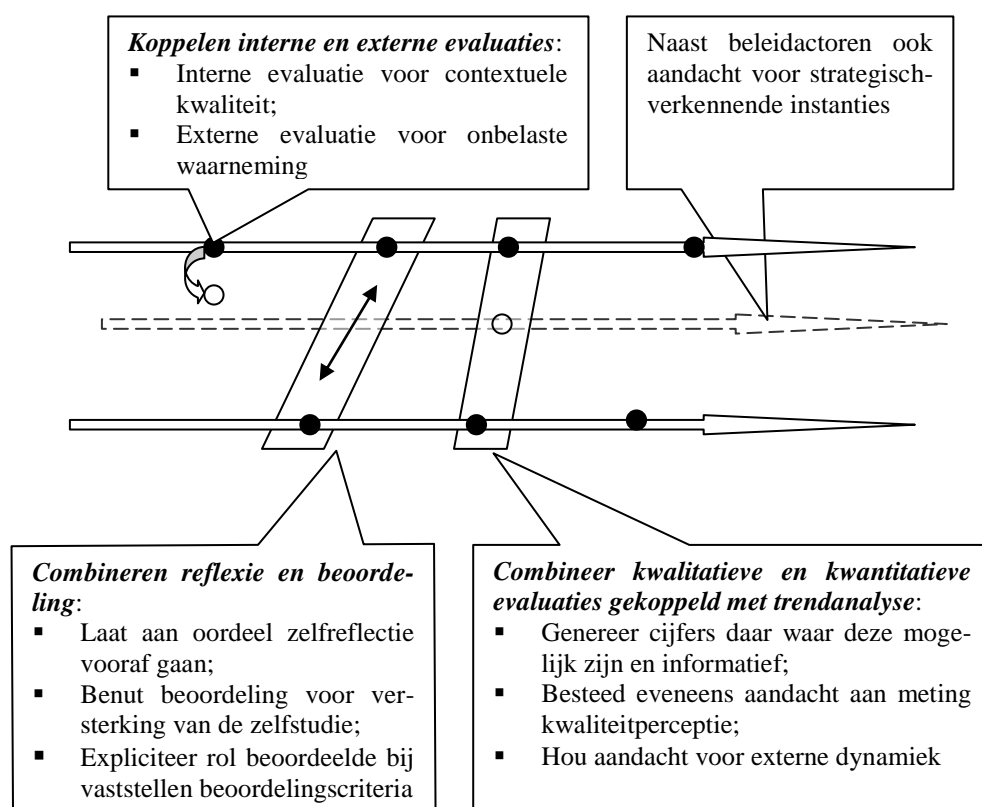
Deze suggestie lijkt wellicht in strijd met de eerdere bevindingen. We constateerden immers dat beoordelende evaluatie tot defensieve reacties aanleiding kan geven die leren belemmeren. Afhankelijk van de verhouding tussen betrokkenen actoren en instanties zal een combinatie van beide invalshoeken dan ook zeker niet altijd mogelijk zijn. Niettemin verdient het aanbeveling ernaar te streven elementen van beide evaluatievormen te combineren. Een beoordeling kan immers rijker en herkenbaarder worden en een grotere kans hebben in beleid door te werken, als die niet alleen is gebaseerd op resultaten, maar ook op inzicht in hoe die resultaten tot stand zijn gekomen. Omgekeerd moge duidelijk zijn dat ‘harde’ gegevens omtrent optredende ontwikkelingen en effecten - in een voldoende ‘veilige’ context - tot reflectie aanleiding zullen zijn. De bij regel 4 genoemde combinatie van interne en externe evaluatie is één van de

vormen voor de hier bepleite koppeling. Daarnaast kan bij verantwoordingsevaluaties expliciet aandacht worden besteed aan de rol van externe condities: reflectie op veranderingen in de omgeving en/of in waardeoriëntaties, kan bijvoorbeeld betekenen dat afrekenen op criteria die een aantal jaren eerder zijn geformuleerd niet meer opportuun zijn, terwijl wel verwacht mag worden dat enigermate op nieuwe ontwikkelingen wordt ingespeeld. Omgekeerd kan het nuttig zijn om van interne reflexieve evaluaties te vragen dat zij zich verhouden tot de geformuleerde eigen doelen en die op hogere niveaus. Langs deze weg is bij verantwoording, bijvoorbeeld in VBTB-kader recht te doen aan voortschrijdend inzicht en contextuele ontwikkelingen. Verantwoording krijgt daarmee een continu en dynamisch karakter.

Regel 6: Combineer kwantitatieve en kwalitatieve evaluaties gekoppeld met trendanalyse

Meer in het algemeen kan de herkenbaarheid en bruikbaarheid van evaluaties op verschillende niveaus en deelterreinen voor betrokken actoren worden bevorderd door kwantitatieve en kwalitatieve bevindingen (waaronder perceptiemetingen) en analyses te combineren. In onze verkenningen kwam naar voren dat sommigen scherp gedefinieerde indicatoren nodig achten om iets van een evaluatie te kunnen leren, terwijl anderen juist van mening zijn dat allerlei kwantificeringspogingen de essentie van wat ruimtelijke kwaliteit (als belevingscategorie) is. Juist het naast elkaar zetten ervan en het op elkaar betrekken kan voor beide 'groepen' tot nieuwe inzichten - en dus tot doorwerking - leiding die anders niet tot stand waren gekomen. Daarbij hoeft het niet alleen te gaan om gegevens die rechtstreeks aan geformuleerde beleidsdoelen gekoppeld kunnen worden. Ook kwantitatieve en kwalitatieve informatie over de veranderende context voor beleid, of anders geformuleerd de feitelijke ruimtelijke ontwikkelingen, de beleving daarvan en (definities van) ruimtelijke kwaliteit kunnen aan leerprocessen en bijstellen van beleid bijdragen.

Schema 2.2 Regels voor het combineren van evaluaties



Schema 2.2 verbeeldt de regels 4, 5 en 6. De horizontale pijlen geven ook hier evaluatie- en onderzoekstrajecten aan die op verschillende niveaus plaatsvinden en waartussen conform de geformuleerde regels koppelingen worden gelegd. De open bolletjes verwijzen naar (elementen van) 'externe' evaluaties.

4. Een evaluatiearrangement voor de ruimtelijke beleidsvorming in Nederland.

Met het totstandbrengen van evaluaties en het versterken van de samenhang tussen enkele soorten evaluaties is al een verbetering aan te brengen in de beleidsprocessen op het terrein van de ruimtelijke ordening. De mogelijkheden tot verbeterde evaluatie zijn daarmee echter nog niet ten einde. We kunnen ons voorstellen dat betrokken ernaar gaan streven om een gezamenlijk evaluatiearrangement te ontwikkelen. Met behulp daarvan kunnen ze naar elkaar toe verantwoording afleggen, kunnen ze het lerend vermogen van de leden van hun organisatie versterken en kunnen ze bovenal de onderlinge samenwerking tot object van evaluatie maken en daarmee ook dat vermogen verbeteren. Het is onze inschatting dat in complexe beleidsnetwerken zoals die bestaan op het terrein van de ruimtelijke ordening, juist hierin grote verbeteringen zijn te halen. We hanteren aangaande een evaluatiearrangement eveneens een drietal regels.

Regel 7: Onderscheid en verbind product- en procesevaluatie

Bij complexe beleidsprocessen kan niet gemakkelijk een eenduidige relatie worden gelegd tussen de dynamica van het proces en de uitkomsten ervan. De vraag hoe het proces verloopt, hoe dat kan worden begrepen en in welke mate dat overeenkomt met ontwikkelde processtrategieën moet daarom nadrukkelijk worden onderscheiden van de vraag tot welke ruimtelijke ontwikkelingen en ruimtelijke kwaliteit het beleid heeft geleid of bijgedragen. Niettemin hebben beide met elkaar te maken. Dat is het geval in de 5N zelf: de combinatie van algemene inhoudelijke eisen en een procesgeoriënteerde sturingsstrategie beoogt ontwikkelingen te bewerkstelligen die de ruimtelijke kwaliteit vergroten en ruimteclaims accommoderen. Daarnaast kan inzicht in het procesverloop mede de uitkomsten ervan verklaren, terwijl omgekeerd (percepties van) feitelijke of mogelijke ruimtelijke ontwikkelingen mede bepalen hoe actoren in die processen opereren. De neiging die wij hier en daar aantreffen om, bij ontstentenis van meetbare productindicatoren, de evaluatie te beperken tot procesindicatoren ('zijn de rode contouren getrokken?' 'hebben gemeenten overlegd?'), achten wij, ook als het (alleen) om verantwoording gaat niet gewenst. De vraag waar al die interactie en bestuurlijke processen toe leiden en wat dat vervolgens voor (bijstelling van) het beleid zou kunnen of moeten betekenen, lijkt - ook bij vage of verschuivende doelen en criteria - een onmisbaar element voor verantwoording en reflectie.

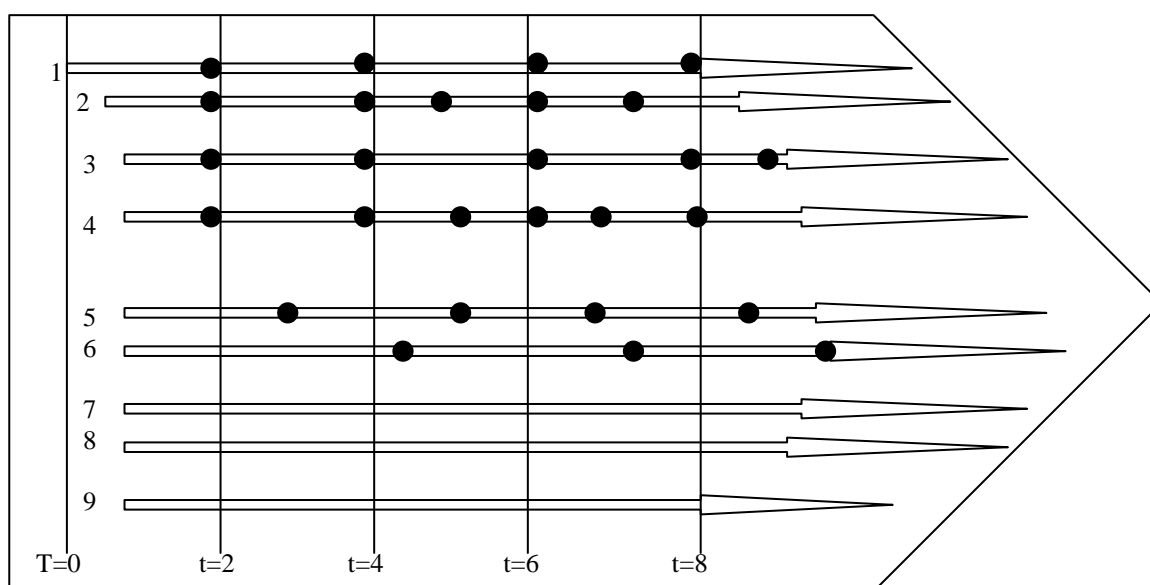
Regel 8: Koppel evaluaties op verschillende niveaus.

In de periode na vaststelling van de 5N zal een behoorlijk aantal evaluaties worden uitgevoerd door een groot aantal partijen. Schema 2.3 geeft hiervan een beeld.

We nemen aan dat veel partijen door zullen gaan met hun eigen evaluaties en dat veel samenwerkingsverbanden eveneens periodiek tot evaluatie zullen overgaan. Al deze evaluaties zullen een eigen evaluatiedoel hebben en vaak zal de vraag centraal staan of de eigen organisatie of het samenwerkingsverband 'het goed doet'. We moeten dus ook niet denken dat bij al de evaluaties de uitvoering van de 5N centraal zal staan. Sterker nog, we mogen verwachten dat deze vraag alleen bij de evaluaties van de rijkspartij die verantwoordelijk is voor de uitvoering van deze nota de centrale vraag zal zijn. Ook bij evaluaties van andere departementen staat hun eigen nota centraal (bijv. het NVVP bij V&W). Bij evaluaties van provincies staat het provinciale beleid centraal en bij de gemeentelijke evaluaties het lokale beleid. Handelingen van andere actoren zijn in dit soort organisatiecentrische evaluaties contextvariabelen. En

vaak ook zijn ze vooral betrekking op variabelen die benut worden om te verklaren waarom het eigen beleid (nog) geen succes is. Met dat al is het echter nog niet gezegd dat deze evaluaties niet van belang zijn in relatie tot de 5N. Integendeel, zij kunnen inzicht geven in doelen van de betrokken actoren, problemen waar zij bij beleidsvorming en -uitvoering tegenaan lopen en/of in feitelijke ruimtelijke ontwikkelingen ter plaatse. Bij centrale evaluaties zou dan ook moeten worden nagegaan welke decentrale evaluaties er zijn of worden gedaan en hoe daarvan gebruik kan worden gemaakt. Omgekeerd zou aan decentrale instanties kunnen worden verzocht om bepaalde centrale aandachtspunten in evaluaties mee te nemen, bijv. de invulling van en de resultaten met betrekking tot de criteria voor ruimtelijke kwaliteit. Op deze wijze wordt, naar wij verwachten, niet alleen het aantal evaluaties (en dus de last van het verzamelen van gegevens) gereduceerd, maar wordt vooral ook bewerkstelligd dat centrale en

Schema 2.3 Overzicht van evaluaties zoals deze de komende tien jaar worden uitgevoerd



Legenda:

1. **Tweede Kamer** laat VROM om de twee jaar verantwoording afleggen over de het gevoerde beleid, inclusief dat uitvoering van de 5N.
2. **VROM** voert naast de tweejaarlijkse monitor door de Tweede Kamer, ook lerende evaluaties uit over de verschillende aspecten van de vijfde nota (en andere beleidsdocumenten)
3. **Financiën** voert VBTB evaluaties uit om de doelmatigheid van overheidsuitgaven vast te stellen.
4. **Andere ministeries** voeren eigen evaluaties uit over ruimtelijk relevante beleid.
5. **Provincies** voeren eigen evaluaties uit van ruimtelijk en gebiedsgericht beleid en van projecten.
6. **Gemeenten** voeren eigen evaluaties uit van ruimtelijk en wijkgericht beleid en van projecten.
- 7/8/9. **Private en maatschappelijke organisaties** die een belangrijke rol spelen bij investeringen in ruimte, maar waarover weinig bekend is in termen van evaluaties.

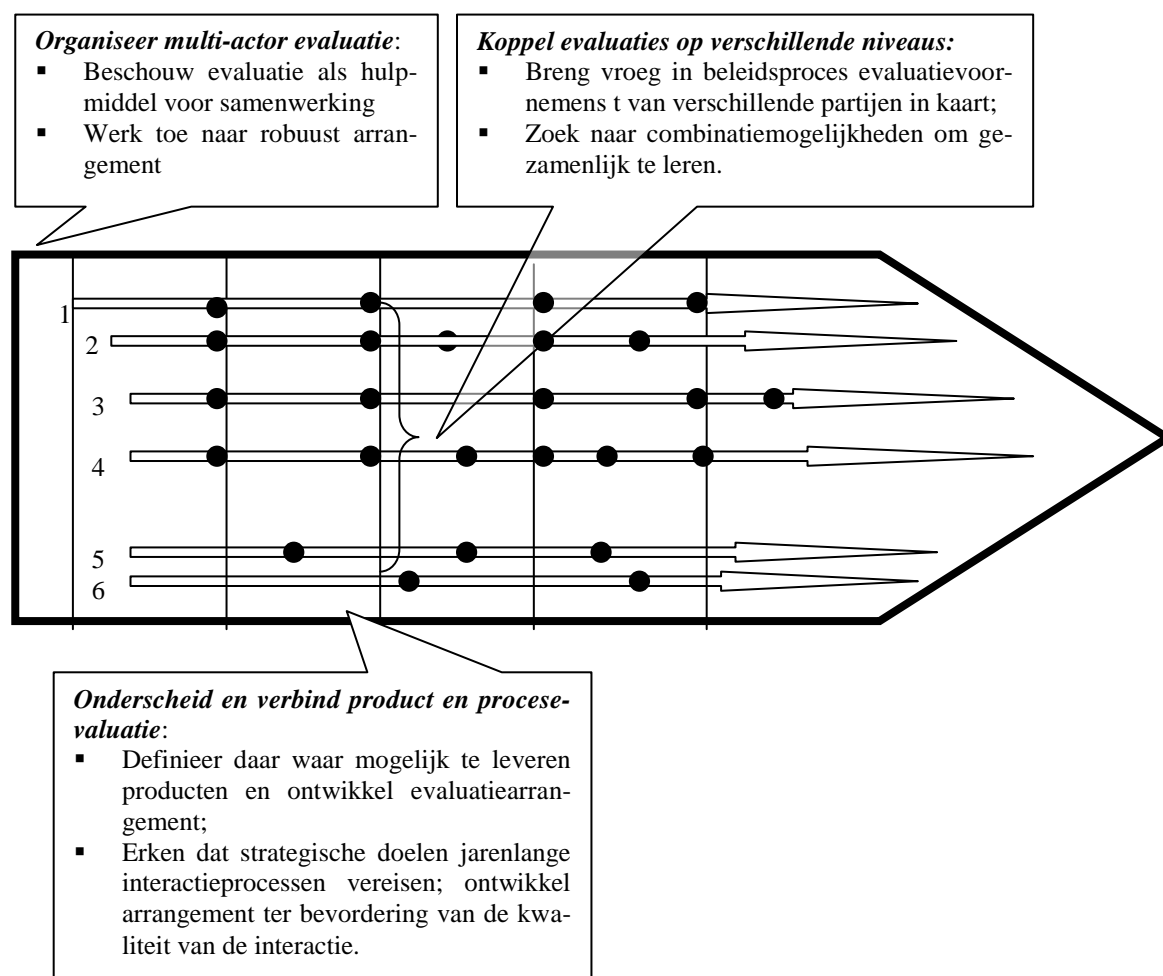
decentrale evaluatie en beleidsontwikkeling meer op elkaar wordt betrokken. Een verdergaande vorm zou kunnen zijn om over en weer representanten in evaluatieteams en/of begeleidingscommissies op te nemen, of zelfs gezamenlijk evaluaties op te zetten en uit te voeren. We komen daar bij regel 9 hieronder op terug. We gaan er overigens van uit dat dergelijke verdergaande vormen van koppeling in de praktijk slechts in een beperkt aantal gevallen realiseerbaar zal zijn.

Regel 9: Organiseer multi-actorevaluatie

Er zijn vele actoren betrokken bij de invulling en uitvoering van ruimtelijk beleid. Er zijn veel beleidsprocessen die ruimtelijke relevantie en/of consequenties hebben. Actoren hebben verschillende perspectieven, zijn geconfronteerd met verschillende (gepercipieerde) randcondities en zijn betrokken in verschillende andere beleidsprocessen. Zoals bij regel 8 hierboven vermeld, kan ervan uit worden gegaan dat dit tot allerlei verschillende evaluaties op verschillende niveaus aanleiding zal geven. De daar bepleite koppeling van evaluaties kan worden bevorderd door op centraal niveau evaluaties op te zetten die nadrukkelijk een multi-actorkarakter hebben. Dat zou om te beginnen betrekking moeten hebben op de vraagstelling en criteria voor de evaluatie: wat achten verschillende betrokkenen (bestuurslagen, ministeries en mogelijk ook maatschappelijke organisaties) de belangrijkste evaluatievragen. Waar willen zij zelf op afgerekend worden en anderen op afrekenen. Welke contextuele informatie is in hun ogen relevant. Oogmerk zou niet moeten zijn om over deze zaken consensus te bereiken, maar om de evaluatieopzet en -inhoud zo rijk te maken dat de verschillende perspectieven er herkenbaar in zijn en op elkaar betrokken kunnen worden. Vervolgens kunnen de bevindingen ook weer door en in discussie tussen actoren betekenisvol worden voor beleidsontwikkeling. Op deze wijze worden leerprocessen en het aanzetten tot beleidsontwikkeling direct in en door het evaluatieproces gestimuleerd. Multi-actorevaluaties kunnen vorm krijgen via permanente of tijdelijke bestuurlijke evaluatieplatforms, door participatie van verschillende actoren in evaluatieteams of begeleidingscommissies, en door gezamenlijk opdrachtgeverschap.

De laatste drie regels zijn in schema 2.4 nog eens weergegeven.

Schema 2.4: Mogelijkheden voor de opbouw van een evaluatiearrangement



Deel 3: Gebruiksaanwijzing

1. Inleiding

In deze gebruiksaanwijzing beschrijven wij vanuit verschillende actorperspectieven welke investeringen in evaluatietrajecten van belang kunnen zijn en hoe met behulp van de betreffende evaluaties drie belangrijke doelen van evaluatie kunnen worden bereikt, te weten het afleggen van verantwoording, het mogelijk maken van zelfreflectie en daaruit voortvloeiende leerervaringen voor wat betreft het eigen beleid en het eigen handelen, en het versterken van het vermogen tot samenwerking. We beginnen elke gebruiksaanwijzing steeds met een positionering van de betreffende actor en (ons beeld van) het daarbij passende evaluatieperspectief. Met deze gebruiksaanwijzingen willen we zo veel mogelijk recht doen aan verschillende actoren en hun zijn rol het ruimtelijk beleid. Daarnaast is het oogmerk van de gebruiksaanwijzingen om aanvullende informatie verstrekken over het door ons gedachte evaluatiearrangement en de bijdrage die het kan leveren aan het realiseren van ambities met betrekking tot ruimtelijke kwaliteit.

De ambities om ruimtelijke kwaliteit te behouden en te ontwikkelen zullen in ieder geval moeten worden omarmd en uitgewerkt door de overheden die op dit terrein actief zijn, zoals de ruimtelijke relevante ministeries van VROM, EZ, Verkeer en Waterstaat, het ministerie van Financiën met zijn specifieke rol in de verantwoording van de rijksuitgaven, de provincies en de gemeenten. We hebben ons hier bij het maken van gebruiksaanwijzingen beperkt tot deze overheidsorganisaties. We denken dat een meer adequate samenwerking tussen deze overheden het sturend vermogen van de overheid in de richting van ruimtelijke kwaliteit aanzienlijk zou vergroten. In dat opzicht is ons voorgesteld evaluatiearrangement al ambitieus te noemen. We zijn ons er evenwel van bewust dat ook andere partijen, waaronder zeker ook maatschappelijke groeperingen, private ondernemingen en burgers belangrijke actoren zijn die eigenlijk in het arrangement betrokken moeten worden. Daarvoor zouden eigen gebruiksaanwijzingen ontwikkeld moeten worden.

We beginnen met de gebruiksaanwijzing voor de Tweede Kamer en het ministerie van VROM. Dit zijn belangrijke actoren voor wat betreft de 5N. Daarmee is niet op voorhand gezegd dat ze ook de meest krachtige actoren zijn daar waar het gaat om het realiseren van ruimtelijke kwaliteit. Wel zullen ze belangrijke initiatiefnemers zijn van ruimtelijke beleid en van beleidsevaluatie. In de evaluatie van nationaal beleid speelt ook het ministerie van Financiën een belangrijke rol. De gebruiksaanwijzing vanuit dit perspectief presenteren we als derde. Daarna presenteren we de gebruiksaanwijzing van andere belangrijke, zoals VROM zou zeggen, sectorale ministeries. We sluiten af met de medeoverheden, juist ook omdat deze zulk een belangrijke rol vervullen in het realiseren van ruimtelijk beleid.

2. Gebruiksaanwijzing vanuit het perspectief van het parlement

2.1 Perspectief

De Tweede Kamer is het orgaan dat voorstellen tot nationaal beleid keurt en dat toezicht houdt op het nationale bestuur. Evaluatieonderzoek zal de Kamer helpen inzicht te krijgen in de uitvoering van nationaal beleid. De Kamer zal overzicht willen hebben op drie gebieden: ruimtelijke kwaliteit, effectiviteit van de eigen middeleninzet van de (rijks)overheid en de kwaliteit van ruimtelijke ordeningsstelsel als geheel.

Weten wat ruimtelijke kwaliteit is

Het is van belang te weten hoe de ruimtelijke kwaliteit van Nederland zich ontwikkelt. Graag zou de Kamer dat in enkele kengetallen zien uitgedrukt. Dat zou het sturen vereenvoudigen. De Kamer weet echter ook dat het aggregeren van kwaliteitsgegevens niet zonder problemen is. Hoe immers is vast te stellen of de ruimtelijke kwaliteit toeneemt van alle gebieden die binnen groene contouren komen te liggen. En hoe is vast te stellen of het lukt tot meervoudig ruimtegebruik te komen? Is het zinvol om in dat kader na te gaan hoeveel vierkante meter grond in Nederland meervoudig wordt gebruikt en weegt daarbij de combinatie van water en toerisme even zwaar als de combinatie van infrastructuur en kantoren. Het lijkt zinvol om alleen op enkele cruciale punten generieke afspraken te maken.

Nu de 5N nadrukkelijk ruimte biedt voor landsdelige variatie op het thema van ruimtelijke kwaliteit, kan de Tweede Kamer de verschillende invullingen van ruimtelijke kwaliteit overlaten aan het ministerie en de vier landsdelen. Indien dat het geval is, zal er ook moeten worden bezien of en welke generiek nationale doelen er dan nog blijven bestaan op nationaal niveau. De overige ambities van ruimtelijke kwaliteit gaan dan formeel ook buiten het VBTB-traject vallen. Ze vallen immers niet uitsluitende onder verantwoordelijkheid van de minister van VROM, maar vloeien voort uit landsdelige afspraken. Wel is het mogelijk dat de Tweede Kamer hier komt tot gemeenschappelijke evaluatiearrangementen met de politieke collega's van de betreffende provincies en gemeenten.

Vaststellen hoe effectief de inzet van rijksmiddelen is

Er worden bij de uitvoering van de 5N rijksmiddelen ingezet om rijksdoelen te bereiken. Juist daar is de behoefte om het Kabinet en het ministerie ter verantwoording te roepen op zijn plaats. De Kamer wil inzicht in de middelen die gemoeid zijn met ruimtelijke investeringen en de effectiviteit van die inzet. Dit wordt belangrijker nu het ministerie van VROM ook enige middelen beschikbaar heeft, specifiek voor het realiseren van ruimtelijke kwaliteit.

Het bestaande VBTB-traject beoogt hiertoe mogelijkheden te scheppen. We stellen evenwel op een punt een verbetering voor. Tot nu toe kijkt VBTB vooral naar de uitgaven van afzonderlijke departementen afgezet tegen de resultaten ervan. Daarmee zal worden doorgedaan. Het verdient echter aanbeveling om ook te bezien in hoeverre investeringen van verschillende departementen elkaar versterken of juist langs elkaar heen schuren. Daarmee krijgt de Tweede Kamer immers een mogelijkheid om te sturen op het samenwerkende vermogen van departementen onderling. In het voorgestelde evaluatiearrangement is dat een belangrijk aangrijpingspunt voor beleidsverbetering.

De kwaliteit van de ruimtelijke beeldvorming

Een derde kennisinteresse betreft het stelsel van beeldvorming zelf. Misschien ligt hier bij een strategische nota als 5N wel het grootste belang van de Kamer. Het gaat immers om de vraag hoe transparant en democratische controlebaar een stelsel is en om de vraag hoe de verschillende investeringsbeslissingen en beheeracties nu op elkaar inspelen. In de taal van de

VBTB spreken we dan van systeemverantwoordelijkheid. Hier ligt een gewichtige uitdaging voor de Kamer. Als landsdelige differentiatie wordt toegestaan en als het leeuwendeel van de investeringen in de ruimte uit private hoek komt, hangt het voortdurend verhogen van de ruimtelijke kwaliteit van Nederland vooral af van het samenspel in het stelsel. De Kamer zou daarmee ook inzicht kunnen krijgen in de vraag waar dat samenspel resulteert in door diverse partijen erkende ruimtelijke kwaliteit en waar deze resulteert in teleurstelling. Dit kan dan weer een impuls vormen om nieuwe rijksinterventies te plegen, met name gericht op het faciliteren van succesvolle samenwerking en het interveniëren in versterde verhoudingen.

Alleen een mix van algemeen en specifieke evaluaties, van product- en procesgerichte evaluatie, gekoppeld aan de monitoring van nieuwe trends, die impact hebben op het ruimtelijke vraagstuk, kan de Kamer tijdig inzicht verschaffen in de keuzes bij dit dilemma. Vandaar ook dat we in deel 2 een arrangement hebben geschetst waarin een pallet aan soorten evaluaties is opgenomen en waarbij we vooral ook pleiten voor versterking van de samenhang daarin. Tabel 3.1 geeft een overzicht van (mogelijke¹⁸) belangen van de Tweede Kamer, gekoppeld aan kennisinteresses, uitdagingen en gepercipieerde belemmeringen.

Tabel 3.1 Perspectief van Tweede Kamer op Evaluatieonderzoek

Belang	<ul style="list-style-type: none"> - Inzicht in realisatie rijksdoelen ruimtelijke ordening (doelbereiking) - Inzicht in samenhang met realisatie andere relevante rijksdoelen (samenhang in rijksbeleid) - Inzicht in totstandkoming en uitwerking van afspraken met landsdelen over realiseren van ruimtelijke kwaliteitstoename. - Controle op doelmatigheid nationale ruimtelijke investeringen. - Stelselinzicht met mogelijkheden tot beloning van succesvol beleid en bijsturen bij inertie (flexibiliteit)
Kennisinteresse	<ul style="list-style-type: none"> - Kwantitatief inzicht in impact van nationale investeringen - Kwalitatief inzicht in stelsel (waar zwak en waar sterk) - Inzicht in leerprocessen bij en samenwerking tussen overheden

Investeren in evaluatie

De Tweede Kamer moet haar aandacht verdelen over vele onderwerpen en kan niet voortdurend toezicht houden op de uitvoering van het nationale ruimtelijke beleid. Een aantal belangrijke toetsmomenten zal moeten volstaan, zoals de (twee-)jaarlijkse monitor conform de VBTB, gekoppeld aan begrotingsdebatten en de in de PKB gehanteerde tijdstippen. Deze maatgevende evaluatiemomenten voor de toets op *ruimtelijke doelbereiking* hebben enerzijds betrekking op een beheerplanologische lijn, waarin vragen naar voren komen als:

- Zijn er contouren getrokken en als dat niet het geval is waarom niet?
- Leidt de optelsom van contouren tot gewenst algemeen ruimtelijk beleid?
- Laten ruimtelijke ontwikkelingen zich vangen in contouren?
- Wat zijn grenzen en mogelijkheden van beheersing van ontwikkelingen?

De doelbereiking heeft anderzijds betrekking op een ontwikkelingplanologische lijn. Daarin is samenwerking met andere partijen een primaire noodzaak. Evaluatie dient daar ook een andere inslag te krijgen. Hier zijn vragen van belang als:

- Genereert het nationaal beleid investeringen van anderen of niet?

¹⁸ Dit voorstel laat ruimte om belangen toe te voegen en tot nadere specificatie te komen.

- Draagt het samenspel van rijksinvesteringen bij aan ruimtelijke kwaliteitsverhoging?
- Versterken de handelingen van rijkspartijen en medeoverheden elkaar?
- Wat zijn grenzen en mogelijkheden van rijksbijdragen aan ontwikkelingen?

Bij beantwoording van beheerplanologische vragen ligt de nadruk op extern beoordelend onderzoek. Bij beantwoording van de ontwikkelingsplanologische vraag is het samen opzetten van evaluaties belangrijk. De Kamer zal misschien betrokken willen worden bij het opstellen van criteria en indicatoren voor het beoordelen van ruimtelijke ontwikkelingen en soms zelfs opdrachtgever willen zijn. Maar ook in het laatste geval blijven onze drie regels voor kwaliteit van evaluaties overeind staan.

Verder vraagt de Kamer om verantwoording over *investeringsstromen* (van alle departementen) en inzicht in de manier waarop investeringsstromen bijdragen aan ruimtelijke kwaliteit. Ten derde is de Kamer geïnteresseerd in mogelijkheden en obstakels van *sturing van ruimtelijke processen*. Ook hier voegen we een voorstel voor gebruik van evaluaties toe die nu nog weinig uit de verf komt en die alles te maken heeft met het gezamenlijk leren van bestuurslagen: we stellen voor dat de Tweede Kamer gezamenlijk met hun politieke counterpart bij de decentrale overheden die verantwoordelijk zijn voor de besteding van die ruimtelijke investeringsgelden, komt tot een onderlinge versterking van evaluatie. Deze interactie zou beide partijen inzicht kunnen verschaffen in de manier waarop verschillende beleidshandelingen op elkaar ingrijpen, wat daaruit resulteert en welke obstakels overheden weerhouden van onderlinge versterking. Gezamenlijk onderzoek met decentrale overheden of andere initiatiefnemers van ruimtelijke investeringsprojecten kan interessante leerervaringen opleveren en de rol van de politieke actoren in beleidsprocessen versterken.

Tabel 3.2: Kennisinteresse bij Tweede Kamer naar soort evaluatie en aard van leerervaring

Kennisobject	Inhoudelijke duiding	Soort evaluatie	Aard leerervaring
Doelbereiking	Inzicht in doelbereiking	Kwantitatief en geaggregeerd	Verbeteren eigen strategie
Sturing	Inzicht in handelingsperspectieven andere actoren	Kwalitatief onderzoek (strategie, percepties, coördinatie)	Gezamenlijk leren
Investeringsstromen	Inzicht in investeringen (publiek en privaat) en wisselwerking	Kwalitatief (motieven, strategie) en kwantitatief (veranderingen in investeringsstromen)	Leren over eigendynamiek investeringen en effect publiek handelen op privaat investeren

2.2 Verantwoording als ambitie van evaluatie

De Tweede Kamer zal in eerste instantie de ‘gangbare’ verantwoordingsmomenten en manieren hanteren (VBTB-procedures, jaarlijkse begrotingsbehandelingen, termijnen van de PKB’s etc.). De Kamer kan eigen evaluaties gebruiken om lessen te trekken vanuit het eigen perspectief (vooral doelbereiking en sturingsmogelijkheden).

Daarnaast kunnen met decentrale actoren gezamenlijke trajecten afgesproken worden voor evaluatie van specifieke ruimtelijke projecten. Deze trajecten kunnen gekoppeld worden aan de tijdsperiode van gangbare beslismomenten (VBTB, PKB’s) maar dit hoeft niet perse. Koppeling legt een zekere druk bij alle actoren om evaluaties daadwerkelijk uit te voeren en op bepaalde momenten gereed te hebben. De evaluaties binnen de trajecten kunnen zowel een beoordelend als een reflexief karakter hebben.

Leermomenten kunnen worden vergroot door andere evaluaties (ook die buiten de eventuele afgesproken trajecten) explicieter aan bestaande procedures van de Kamer te koppelen. Flexibiliteit is daarbij overigens gewenst.

2.3 Evaluatie om ervan te leren

Evaluaties kunnen ook uitgevoerd worden om ervan te leren. Naast de klassieke verantwoordingsevaluatie die de Tweede Kamer met regelmaat bereiken en die vooral dienen voor de controle van het Kabinet, zal de Kamer ook geïnteresseerd zijn in de vraag of en waarom bepaalde elementen van het ruimtelijk beleid niet (of wel) werken. Er is dan aandacht voor het gehele stelsel van actoren die ruimtelijk relevant gedrag vertonen. Bij deze vorm van lerende evaluatie verdient het aanbeveling om interactie tot stand te brengen met relevante actoren. Evaluatieonderzoek vervult hier de functie als aanjager van interactie en verkennen van kracht en zwakte van ruimtelijk beleid, gericht op wederzijds leren of gezamenlijk leren.

Een aanvullend leermoment is te vinden in samenhang tussen publieke en private investeringen. Een groot deel van de ruimtelijke investeringen heeft een privaat karakter. De kennis hierover is beperkt. We kunnen ons voorstellen dat samen met private actoren inzicht ontstaat over de condities waaronder private investeringen private publieke ambities dienen.

2.4 Evaluatie om samenhang in overheidsbeleid te versterken

Evaluatie wordt vooral als beoordelingsinstrument gebruikt. De term *'woensdag gehakt dag'* illustreert dit. Deze handleiding evenwel benut evaluaties ook als mogelijkheid om samenhang te bevorderen. Er wordt vaak gesproken over integraal beleid. Integratie blijkt evenwel niet te verwezenlijken via een integraal besluit. Integratie is met andere woorden niet vooraf af te dwingen. Integratie kan wel resulteren uit interactie tussen betrokken partijen. ***Het is om deze reden dat ook de Kamer er belang bij heeft om deze interactie te bevorderen.***

Evaluatie kan daarbij als hulpmiddel worden ingezet om op gezette tijden met betrokken partijen van gedachten te wisselen over de vraag of de uitkomsten van de interactie in de ogen van alle betrokkenen geslaagd zijn en of hieruit gemeenschappelijke verbeteracties kunnen volgen. Deze ambitie kan op verschillende manieren vorm krijgen. Zo lijkt het interessant vanuit het perspectief van de Kamer om de verschillende ministers die ruimtelijk relevant beleid voeren, uit te dagen een aantal doelen te formuleren waarop ze gezamenlijk willen scoren, bijvoorbeeld daar waar het gaat om het realiseren van meervoudig ruimtegebruik. Als ze deze doelen voorleggen aan de Kamer, is het vervolgens mogelijk om de voortgang in de onderlinge samenwerking in beschouwing te nemen.

Ook lijkt het zinvol dat de Kamer ministers verzoekt om in uitvoeringsovereenkomsten met landsdelen en regio's expliciete afspraken te maken over gemeenschappelijke evaluatie van departementen met medeoverheden. Deze afspraken omvatten dan ten minste twee punten:

- Op welke targets wil het landsdeel de eigen pijlen richten, welke inspanning van rijksoverheden vragen ze daarbij, welke inspanning gaan ze zelf verrichten en waarover willen ze verantwoording afleggen aan het rijk?
- Hoe wordt de evaluatie georganiseerd, wie zit er in de stuurgroep die de evaluatie van begin tot eind begeleidt, hoe wordt het evaluatieonderzoek vormgegeven en hoe komt het evaluatierapport tot stand (bijvoorbeeld consensus in de stuurgroep)?

Ten slotte lijkt het voor grensoverschrijdende vraagstukken gewenst dat politieke vertegenwoordigers interactie starten met vertegenwoordigers uit andere landen om gezamenlijke evaluaties uit te voeren. Daarnaast zou bijzonder zinvol kunnen zijn wanneer bijvoorbeeld de

Tweede Kamer samen met de Provinciale Staten van de vier Deltametropolitane Provincies en de vertegenwoordigers van een aantal gemeenteraden in de Deltametropool op gezette tijden samen evalueren hoe sterk de Deltametropool zich ontwikkelt ten opzichte van concurrerende metropolitane gebieden in Europa, welke gezamenlijke acties zijn opgepakt om de positie te versterken en hoe goed deze acties verlopen. Een gezamenlijke evaluatie vormt dan een hulpmiddel om een integratie van het gefragmenteerde bestuur van de Deltametropool te bevorderen, zonder te verzanden in ingewikkelde structuurdiscussies.

3. Gebruiksaanwijzing vanuit het perspectief van VROM

3.1 Perspectief

VROM draagt specifieke verantwoordelijkheid voor ruimtelijk beleid in den brede en geeft daaraan vorm via de 5N, de PKB en het uitvoeringsplan. Daarbij zijn kerndoelen en criteria geformuleerd (ruimtelijke kwaliteit, combinatie van functies) en wordt de aanzet gegeven tot ontwikkeling van bestuurlijke, juridische en financiële instrumenten. Deze zullen via wetgeving en via de begrotingen van VROM verder tot ontwikkeling komen. De 5N onderscheidt de rol van het eigen ministerie in een drietal soorten beleidstrajecten:

- nationale beheer- en ontwikkelingstrajecten: het - deels onder verantwoordelijkheid van andere ministeries - ontwikkelen van nationale en internationale ruimtelijke projecten;
- regionale en lokale beheer- en ontwikkeltrajecten waarbij het ministerie voor zichzelf een stimulerende rol ziet weggelegd;
- Regionale en lokale beheer- en ontwikkeltrajecten waar de rijksoverheid voor zichzelf een faciliterende rol weggelegd ziet.

De uitvoering van de 5N is gebaseerd op een combinatie van regelgeving, afspraken en prikkels. De 5N heeft oog voor de combinatie van centrale en decentrale beleidsvorming en -uitvoering en voor zowel beheer als ontwikkeling. Vanuit deze positionering zijn bij evaluatie de volgende kennisinteresses van belang:

- feitelijke ruimtelijke ontwikkelingen op geaggregeerd niveau op een breed spectrum van dimensies die gerelateerd zijn of kunnen worden aan ruimtelijke kwaliteit;
- inzicht in doorwerking en effectiviteit van regelgeving, afspraken en (financiële) stimulering;
- globaal inzicht in de richting, de diversiteit en problemen van regionaal en lokaal ruimtelijk beleid;
- inzicht in de effectiviteit van (inter)nationale ruimtelijke investeringen.

Naast specificatie van de kennisinteresses impliceert bovengeschetste positionering ook afhankelijkheid van andere ministeries, bestuurslagen en private actoren.

3.2 Investeren in evaluatie

Vanuit het perspectief van VROM verdient het aanbeveling vooral op de volgende manieren te investeren in evaluatietrajecten. De eerste drie hebben een meer intern karakter, terwijl de laatste drie aandacht geven aan interorganisatorische evaluatie van beleid en beleidprocessen.

- onafhankelijke studies die feitelijke ruimtelijke ontwikkelingen in kaart brengen; daarbij worden onder meer aspecten meegenomen waarop nationaal of regionaal beleid zich nadrukkelijk richt; andere bestuurlijke en maatschappelijke actoren worden uitgenodigd aandachtspunten aan te dragen; de studies monden niet in een beoordeling uit, maar worden wel benut bij andere evaluaties, ontwikkeling van nieuw beleid en bestuurlijk overleg

- Zelfpositionerende evaluaties, waarbij de top van het ministerie met haar eigen ambtenaren scherpe afspraken maakt over de inspanning die worden verwacht in het kader van de 5N en hoe daarover verantwoording wordt afgelegd.
- Zelflerende evaluaties, waarbij betrokkenen bij de implementatie van de 5N aangeven tegen welke punten ze aanlopen, waarover ze graag zouden leren effectiever te zijn, waarna een externe instantie de opdracht krijgt om het betreffende beleid te evalueren en op basis daarvan tot reflectie te komen met de betrokken beleidsmakers.
- Multi-actorevaluaties met betrekking tot convenanten en regionale projecten waarbij het Rijk betrokken is; hierbij worden de in de handleiding bepleite koppelingen van insteken nadrukkelijk nagestreefd.
- Kennis nemen van regionale evaluaties (bijv. gekoppeld aan streekplannen) en bevorderen dat daarin een beperkt aantal aandachtspunten vanuit het landelijk beleid wordt meegenomen; mede op grond daarvan een eigen evaluatie maken van de werking en effectiviteit van het instrumentarium van de Rijksoverheid voor zover gericht op regionale processen; de bevindingen van die zelfevaluatie publiceren en ter discussie stellen.
- Multi-actorevaluaties m.b.t. effecten van ruimtelijk relevant beleid van andere ministeries en de verhouding tot het algemene ruimtelijke beleid.

3.3 Verantwoording

Genoemde evaluatietrajecten leveren gegevens op die voor verantwoording over het departementale en Rijksbeleid kunnen worden benut. De meer specifieke verantwoording zoals die in het kader van de VBTB wordt nagestreefd, kan daarbij plaatsvinden als en voor zover operationele doelen en criteria zijn geformuleerd. In PKB en Uitvoeringsplan is dat ten dele het geval. Op de betreffende punten geven de evaluaties inzicht in resultaten, maar ook in (bedoelde of onbedoelde) processen en condities die daarbij bepalend kunnen zijn geweest. In dit verband kan zowel zelfevaluatie als onafhankelijk externe evaluatie, of een combinatie daarvan zinvol zijn. Aandacht voor contextuele ontwikkelingen en verklaring voor optredende effecten komt ook de verantwoording ten goede.

Aan de in de 5N gekozen sturingsstrategie is echter inherent dat concretisering voor een groot deel later en/of elders totstandkomen. In onderlinge samenhang geven de evaluaties echter ook een basis om over dat sturingsmodel te oordelen en er verantwoording over af te leggen (systeemverantwoordelijkheid). Het kan daarbij enerzijds gaan om op procescriteria gerichte evaluaties, maar ook om een combinatie van inhoudelijke evaluaties die op verschillende niveaus worden gedaan. Onderzoek naar feitelijke ruimtelijke ontwikkeling vormt daarnaast belangrijke input voor verantwoording over het beleidsproces.

Het is van groot belang dat het ministerie op basis van de 5N enkele doelen formuleert waarop zij zelf de komende jaren echt wil scoren en waarop ze dan ook afgerekend wil worden in het VBTB-traject. Deze eigen zelfgeformuleerde verantwoording vooraf geeft anderen belangrijke indicaties over wat het ministerie gaat doen en belangrijk vindt. Het kan hierbij zowel om inhoudelijk resultaat als om procescriteria en systeemverantwoordelijkheid gaan.

Het lijkt daarbij van belang om het onderscheid tussen de drie soorten trajecten in het oog te houden. In het nationale traject (realiseren) ligt de nadruk op het afleggen van verantwoording over behaalde resultaten. Bij faciliteren ligt de nadruk op de ondersteunende rol van het ministerie en zou een belangrijke rol bij het opzetten en uitvoeren van evaluaties gespeeld kunnen en moeten worden door degenen die ondersteund worden. Een van de vragen is hierbij immersin hoeverre zij die ondersteuning ook als zodanig ervaren en waarderen. Deze vorm van omgekeerde verantwoording afleggen, past ons inziens goed bij het strategische karakter van de 5N en in de gekozen verdeling van beleidsinspanningen.

Voor wat betreft de stimulerende rol lijkt het verstandig evaluaties op te zetten in samenwerking met degenen die gestimuleerd moeten worden.

Ten slotte is het mogelijk om voor de gemeenschappelijke verantwoording van verschillende bestuurslagen en instanties, welke een groot deel van de ruimtelijke beleidsaspiraties betreft, gezamenlijke evaluatietrajecten te ontwikkelen. Betrokken overheden leggen daarbij, naast de bestaande interne verantwoording naar de politiek ook verantwoording af naar elkaar aangaande het wel of niet bereiken van ruimtelijke kwaliteit. Gedegen evaluatie kan de neiging tot zwartepieten en afschuiven niet voorkomen, maar kan wel de voorwaarden begunstigen voor wederzijdse leerprocessen. Ons inziens biedt deze vorm van juist in de zoektocht naar ruimtelijke kwaliteit goede diensten. Zij is ook een manier om systeemverantwoordelijkheid vorm te geven.

De verschillende evaluaties kunnen, ook als ze niet expliciet in VBTB-kader zijn opgezet, toch bijdragen aan het VBTB-traject, omdat ze kunnen verhelderen welke (combinaties van) actoren waarop redelijkerwijs afgerekend kunnen (en willen) worden.

3.4 Leren

In feite zijn we met dit laatste al op het terrein van de lerende organisatie terechtgekomen. Lerende organisaties zijn immers bereid om de kritiek van anderen te horen en ter harte te nemen. We bevelen echter aan om onderscheid te maken tussen evaluaties die vooral bedoeld zijn om verantwoording af te leggen naar de eigen politieke toezichhouders en evaluaties die bedoeld zijn om zelf lessen uit te trekken. Vervolgens is het echter wel zaak om beide soorten evaluaties met regelmaat als duo-evaluaties te organiseren. Door de noodzaak om verantwoording af te leggen zal een zelfevaluatie bedoeld om te leren met meer zelfkritisch vermogen en ook veel grondiger worden uitgevoerd. Pas als deze toch niet goed wordt uitgevoerd, zou een externe partij een evaluatie ter hand kunnen nemen.

Bij evaluaties gericht op de eigen lerende organisatie is het van groot belang dat betrokkenen een belangrijke rol spelen in het ontwikkelen van de punten waarop de evaluatie zich gaat richten. Met andere woorden, hun probleempceptie moet verwerkt zijn in het evaluatiedesign. Vervolgens moet de evaluatie evenwel worden uitgevoerd door iemand op wat grotere afstand van de betrokkenen die voldoende affiniteit heeft met de te evalueren processen en beleidsvragen om vertrouwenwekkend te zijn voor degenen die moeten leren (eerste-orde-leren), maar die anderzijds ook zoveel distantie heeft dat de evaluatie ook kritische reflectie op het gedrag en de opvattingen van degenen die moeten leren, kan hebben (tweede-orde-leren)

Het koppelen van soorten evaluaties en evaluaties op verschillende niveaus en het organiseren van debat over evaluaties stimuleren leerprocessen. Als het daarbij ook nog zo te organiseren is dat alle betrokken met regelmaat zowel evaluator als object van evaluatie zijn, ontstaat mogelijk ook meer gevoel voor het moeilijke evenwicht tussen empathie en kritiek, die de basis vormt voor leerrijke evaluaties. Door evaluaties als start en tussenmomenten te gebruiken in beleidsprocessen, zal evaluatie niet alleen een meer integraal onderdeel van beleidsprocessen en bestuurlijk handelen worden, maar ook meer opleveren dan veel van de evaluaties die achteraf worden uitgevoerd.

3.5 Interactie en samenwerking

De 5N gaat in hoge mate uit van de gedachte dat het beleid en de uitvoering daarvan in samenwerking totstandkomen. Het gaat daarbij om samenwerking tussen en binnen bestuursniveaus, en ook samenwerking tussen publieke en private partijen. Met het oog daarop zijn en worden allerlei overleg- en afstemmingsarrangementen gecreëerd, in een deel waarvan

VROM ook is vertegenwoordigd. Deze arrangementen vormen een kader om de hiervoor bepleite multi-actorevaluaties in te positioneren. Evaluatie wordt daardoor minder iets van buitenaf en achteraf, maar wordt ingebed in een proces van beleidsvorming en daarover leren. Daarmee wordt evaluatie niet alleen aangehaakt bij bestaande samenwerking, maar kan het samenwerking en coproductie ook versterken.

In de 5N doen zich een groot aantal mogelijkheden voor om tot evaluatieafspraken te komen die gericht zijn op het bevorderen van samenwerking: Noordzee, Waddenzee (in termen van beheer), en zo verder. Een van de meest uitgesproken uitdagingen kan de Deltametropool worden. Daar dient zich de noodzaak tot samenwerking volop aan en is het vermogen daartoe redelijk beperkt. De 5N wil van de Deltametropool een speerpunt maken. Het lijkt ons van groot belang om direct afspraken te gaan maken tussen rijkspartijen en medeoverheden over wat inhoudelijk de agenda is, wat er eerst wordt aangepakt, wie waarvoor verantwoordelijk is, welke producten er worden geleverd, hoe men met elkaar interacteert (creatieve moment, bestuurlijke moment, wisseling van samenstelling en zo verder) en waar de betrokken gedurende een periode op afgerekend willen worden, wat ze willen leren en welke samenwerkingseisen mogen worden gesteld. Juist door de wetenschap dat er na een bepaalde periode weer een evaluatiemoment volgt, zullen partijen naar onze verwachting net iets meer inspanning plegen ten gunste van samenwerking dan anders. Samenwerking helpt dan namelijk, anders dan in de huidige bestuurlijke praktijk om te scoren.

4. Gebruiksaanwijzing vanuit het perspectief van Financiën

4.1. Perspectief

Het Ministerie van Financiën bevordert en ziet toe op de doelmatige besteding van de Rijksfinanciën. VBTB, de Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid (RPER) en het Handboek Financiële Informatie en Administratie Rijksoverheid (HAFIR) zijn instrumenten daartoe. In de praktijk blijkt het produceren van gegevens die vanuit het perspectief van Financiën bruikbaar zijn voor verantwoording niet eenvoudig te verlopen. Dit komt mede omdat de beoordeling van de doelmatig sterk verbonden wordt met de dreiging dat een negatief oordeel leidt tot een vermindering van het budget van een organisatie.

Voor een deel is deze spanning inherent aan de rol die het ministerie van Financiën speelt ten opzichte van de ‘spending departments’. Toch lijkt het ons van belang dat ook het ministerie van Financiën de voorgestelde samenhang in evaluaties en de mogelijkheid van meer samenhangende evaluatiearrangementen in beschouwing neemt. Mede door evaluaties ook te richten op processen van nationale ruimtelijke investering in plaats van uitsluitend op departementale uitgaven kan zowel de doelmatigheid van deze bestedingen alsook de samenwerking tussen departementen en met medeoverheden ernstig versterkt worden¹⁹.

4.2. Investeren in evaluatie

Vanuit de positie en het perspectief van Financiën lijken de volgende soorten investering in evaluatie zinvol.

- Onafhankelijk onderzoek naar feitelijke ruimtelijke ontwikkelingen op dimensies waarop het nationale beleid (VROM, LNV, V&W, Defensie) is gericht.
- Bevorderen van en participeren in multi-actorevaluaties van algemeen en specifiek nationaal ruimtelijk (relevant) beleid; naast het in beeld brengen van resultaten is inzichtverwerving in de totstandkoming van resultaten (of in de oorzaken van uitblijven daarvan) nadrukkelijk onderwerp van deze evaluaties; naast producten zijn ook processen aan de orde.
- Externe review van zelfevaluaties van andere departementen, met name wat betreft doelrealisatie en financieel beheer.

4.3. Verantwoording

Tot nu toe ligt de nadruk bij het ministerie sterk op verantwoording. Het is opvallend dat deze belangrijke functie van evaluatie vooral wordt ingevuld in de richting van departementale verantwoording. Vanuit de leerstelling van de ministeriële verantwoording is deze insteek begrijpelijk. Als we evenwel het oog richten op de beleidsprocessen die tot doelbereiking moeten leiden, dan valt op dat deze in toenemende mate een interorganisatorisch karakter hebben. Dat wil dus zeggen dat de kwaliteit van de uitkomst afhangt van de inzet van verschillende partijen en ook het vermogen van deze partijen om handelingen onderling af te stemmen.

¹⁹ Men kan stellen dat Financiën een systeemverantwoordelijkheid heeft voor financieel beheer en ook voor verantwoording en evaluatie, zoals VROM een systeemverantwoordelijkheid voor het RO-beleid heeft. Ook BZK heeft een in dit verband relevante systeemverantwoordelijkheid, onder meer voor institutioneel-bestuurlijke aspecten. BZK bleef echter in ons onderzoek buiten beschouwing.

Alleen al vanuit het perspectief van verantwoording afleggen, lijkt het gewenst om naast de departementale organisatiegerichte verantwoording, meer aandacht te besteden aan het verantwoorden van de onderlinge samenhang (of gebrek daaraan) tussen de onderdelen van de rijksoverheid. Met name door óók multi-actor- en zelfevaluaties te benutten, kan de bereidheid tot en de kwaliteit van verantwoording toenemen. Deze kwaliteit kan toenemen doordat zelfevaluatie en multi-actorevaluaties meer aandacht besteden aan de visies van diverse actoren op dezelfde problemen en doordat er meer zicht komt op de door hen ervaren beperkingen. Daarmee wordt inzicht verworven waarover - in redelijkheid en for the time being - wel en waarover geen 'harde' verantwoording kan worden afgelegd. Financiën krijgt daarmee ook meer een basis om zelf over de voortgang van het VBTB-traject verantwoording af te leggen.

4.4. Leren

Als een evaluatie helpt om zwakke plekken in de beleidsvoering van een departement bloot te leggen, betekent dit nog niet direct dat daarmee ook een verbeterproces wordt ingezet. Om tot verbetering te komen is het vooral van belang dat de medewerkers van het betreffende departement lering trekken uit de resultaten. Om te leren dient aan een aantal voorwaarden te worden voldaan. Zo blijkt angst een slechte raadgever. Departementen die een evaluatie vanuit deze optiek ondergaan, zullen zich vooral inspannen om zich te immuniseren tegen kritiek. Het lijkt er daarbij op dat naarmate de evaluatie verder buiten departementen geplaatst wordt, zowel bij het formuleren van criteria als bij het uitvoeren van het onderzoek en het rapporteren erover, neemt de neiging tot immunisering toe en het vermogen om ervan te leren af.

Vanuit deze waarneming valt er veel voor te zeggen om bevordering van het lerend vermogen van departementen als een belangrijk doel van evaluaties te nemen. Dat kan onder meer door bij evaluaties voor een meer interactieve vormgeving te kiezen. Met name participatie in multi-actorevaluaties, zoals eerder beschreven, kan Financiën meer inzicht geven in de betekenis die VBTB-achtige evaluaties hebben voor partijen. Verder biedt deze interactieve aanpak van evaluaties mogelijkheden aan departementen om mee te denken over geschikte manieren om verantwoording af te leggen, gekoppeld aan leerprocessen.

4.5. Samenhang en bevordering samenwerking

In het VBTB-traject is al aandacht voor de noodzaak om ook beleidsstelsels te evalueren om daarmee de kracht van dat stelsel te versterken. Tot nu toe heeft dat niet geleid tot een bijzonder grote aandacht voor de daadwerkelijke toepassing van deze gedachte. Toch biedt onder meer het ruimtelijke beleid wel aanleiding om het denken in termen van samenhang en samenwerking meer centraal te stellen in evaluaties.

Juist vanuit het perspectief van doelmatigheid is een onderlinge versterking van handelingen van diverse partijen van groot belang. Een proef met interdepartementale evaluatie lijkt een mooie eerste stap.

5. Gebruiksaanwijzing vanuit het perspectief van andere departementen

5.1 Perspectief

Naast VROM is een aantal andere departementen verantwoordelijk voor beleid dat duidelijke ruimtelijke dimensies en consequenties heeft. Dat geldt voor V&W (infrastructuur), LNV (landelijk gebied, groene ruimte), EZ (ruimtelijke economie) maar ook voor bijv. Defensie (militaire terreinen). Op deze beleidsterreinen is het perspectief vergelijkbaar met dat van VROM. De gebruiksaanwijzing vanuit het perspectief van VROM geldt dan ook *mutate mutandis* voor de andere departementen. Daarnaast is er echter, uit het oogpunt van kabinetsbeleid en het streven naar samenhang in rijksbeleid, ook een betrokkenheid bij het omvattende RO-beleid als zodanig aan de orde.

5.2 Investeren in evaluatie

Naast investeringen conform paragraaf 3.2 is gezamenlijke evaluatie met VROM over het RO-beleid en de positie van het betreffende beleidsterrein daarbinnen van groot belang. Op deze wijze wordt inzicht verschaft in fricties en feitelijke afwegingen en keuzes in het beleid. Het is voor departementen belangrijk om te onderkennen dat veel van hun ambities ruimtelijke implicaties hebben en dat hun ambities eerder verwezenlijkt kunnen worden, als deze ook bijdragen aan de doelstellingen van ruimtelijke kwaliteit. In dat licht lijkt het van belang om betrokken te zijn in beleidsevaluaties die zich juist richten op het aspect van ruimtelijke kwaliteit. Evaluaties vanuit dit perspectief zullen leiden tot nieuwe inzichten. Vragen die dan van belang zijn, luiden bijvoorbeeld: welke infrastructurele investeringen dragen bij aan de ruimtelijke kwaliteit en waarom is dat het geval? of: welke landbouwontwikkeling versterkt het platteland? Ook vragen als ‘wanneer is de ontwikkeling van onze internationale vervoerassen een bijdrage aan ruimtelijke kwaliteit, wanneer niet en waarom niet?’ Veel van dit soort vragen spelen nooit een rol in beleidsevaluaties, mede ook omdat evaluaties doorgaans departementaal georiënteerd zijn. Maar juist op de raakvlakken van de departementen (mobiliteit als kwaliteit én ruimte als kwaliteit, of ruimtelijk-economische ontwikkeling van werkgelegenheid én ruimtelijke kwaliteit) wordt voor een belangrijk deel bepaald of het rijksbeleid effectief is of niet (met name in de ogen van burgers en medeoverheden!). ***In toenemende mate ligt de effectiviteit immers in het combineren van verschillende kwaliteiten en daarmee het combineren van evaluatiecriteria.***

Het is om deze reden dat we ervoor pleiten om de evaluaties van de verschillende departementen te combineren met een aantal evaluaties die de grenzen van de verschillende organisaties overschrijden. Met name ook in landsdelige en regionale aanpakken liggen goede mogelijkheden om samenhang in beleid te bevorderen.

5.3 Verantwoording

Het blijft belangrijk dat elk departement verantwoording aflegt voor de eigen inzet in beleidsprocessen. Daaraan doen we in het onderhavige voorstel niets af. Wel voegen we daar iets aan toe, in de vorm van ***verantwoording over samenhang***. Als het zo is dat diverse departementen impact hebben op de ruimtelijke ontwikkeling en vorm van Nederland, is het ook van belang dat de verantwoording zich ook uitstrekt tot het vermogen van departementen om onderling samen te werken. In een gefragmenteerde beleidsstelsel is samenhang op rijksniveau een eer-

ste vereiste. Dit geldt uiteraard allereerst voor het ontwikkelen van beleid, maar juist daarom ook voor beleidsevaluatie. Immers daar waar in evaluatie steeds wordt teruggevallen op een departementale insteek, blijft veel van wat er op de raakvlakken van departementen gebeurt, buiten beeld en juist daar gaat veel fout en is ook veel voordeel te behalen. Om dat voordeel te behalen is het allereerst nodig dat verantwoording wordt afgelegd over departementaal grensverkeer.

Met name ook de participatie in evaluaties van dat grensverkeer kan bijdragen aan verantwoording, omdat zo positieve of negatieve wisselwerking tussen de verschillende beleidssectoren van de RO in beeld worden gebracht.

5.4 Leren

Het in beeld brengen van de wisselwerking is een eerste stap. Minstens zo belangrijk is evenwel dat departementen in de komende periode van uitvoering van de 5N (en van hun eigen beleidsnota's niet te vergeten) leren van evaluaties. Met name het opzetten van gezamenlijke evaluaties van het departementaal grensverkeer, de werking en uitwerking daarvan en de opbrengsten en kosten ervan, kunnen nieuwe ideeën genereren die in het beleid beoogde functiecombinaties en ruimtelijke kwaliteitsverhoging dichterbij brengen. Een reguliere gezamenlijke evaluatiepraktijk kan gezien worden als belangwekkend instrument om samenhang en innovatie te bevorderen. Dit kan ons inziens vooral bevorderd worden als betrokkenen gezamenlijk afspraken maken over de manier waarop de grensoverschrijdende activiteiten het best in kaart zijn te brengen en op basis van welke criteria oordelen worden uitgesproken over de kwaliteit van deze grensoverschrijdende activiteiten.

5.5 Interactie en samenwerking

Ten slotte geldt ook voor de ruimtelijke relevante departementen dat een aanzienlijk deel van hun beleidsambities door medeoverheden en andere partijen moet worden gerealiseerd. Ondanks pogingen om tot decentralisatie te komen, blijft er steeds een aanzienlijke interactie bestaan en noodzakelijk tussen rijkspartijen en medeoverheden. De samenhang en mogelijke interferentie tussen ruimtelijke ontwikkelingen op verschillende beleidsterreinen creëert de facto afhankelijkheid tussen departementen en tussen departementen en andere publieke en private partijen. Interactie en samenwerking tussen deze partijen kunnen helpen om met interdependenties om te gaan. Het is daarbij wel van belang om expliciet aandacht te besteden aan deze interactie en deze ook object van reflectie te maken. Regelmatige gezamenlijke evaluatie over nieuwe beleidsintenties en hun onbedoelde effecten voor medeoverheden, kunnen dergelijke zoekprocessen ten goede komen.

6. Gebruiksaanwijzing vanuit het perspectief van regionaal bestuur

6.1 Perspectief

Het regionaal bestuur is op het ruimtelijke terrein verantwoordelijk voor het vervaardigen van streek- en regioplannen. Daarbij kent het regionaal bestuur een normstellende en -toetsende rol ten aanzien van bestemmingsplannen die op lokaal bestuursniveau worden vervaardigd. In de nabije toekomst wordt het stellen en toetsen van contouren (rode, groene) mogelijk ook een onderdeel van deze zogenaamde beheerplanologische beleidslijn. Deze lijn moet expliciete aandacht krijgen in het evaluatiearrangement van de 5N. Daarnaast moet veel van de in de 5N nagestreefde ruimtelijke kwaliteit gerealiseerd worden via de lijn van de ontwikkelingsplanologie. Hierbij gaat het om de ontwikkeling van provinciale projecten en om de vele vormen van gebiedsgericht beleid waarin provincies participeren.

Veel provincies hebben een sterke oriëntatie op de inhoud van (met name beheer-)planologie. Kwaliteit wordt met andere woorden nagestreefd door deze vast te leggen in plannen en van daaruit pogen te voorkomen dat anderen afbreuk doen aan bestaande kwaliteiten. Vastleggen en toezicht zouden we belangrijke oriëntaties kunnen noemen bij provincies. Deze zullen ook in de toekomst blijven bestaan. Daarvoor is een geijkte vorm van evaluatie nodig.

Daarnaast onderkent het provinciale bestuur in toenemende mate de noodzaak om een ontwikkelingspolitieke benadering te entameren²⁰. Juist via de samenwerking met gemeenten, maatschappelijke groeperingen en private partijen moet de hoge ambitie van ruimtelijke kwaliteit worden verwezenlijkt. Dit vraagt om een andere aanpak dan die van de traditioneel ordenende benadering. Provinciebesturen onderkennen dit en willen hun beleid, hun organisatie en daaraan gekoppeld ook hun manier van evalueren daarop aanpassen. In *'Van ordenen naar ontwikkelen'* (IPO 2000) geven provinciale bestuurders ook aan dat in een ontwikkelingspolitieke benadering het vraagstuk van verantwoordelijkheid opnieuw bezien moet worden. Indien partijen gemeenschappelijk op zoek gaan naar ruimtelijke kwaliteit is het ook van belang dat partijen verantwoording afleggen naar elkaar en naar de maatschappij. Het idee van wederzijds verantwoording afleggen krijgt met andere woorden ook hier vorm. Juist dit aspect leent zich bijzonder goed voor vormgeving van de evaluatie van het 5N traject.

Ten derde neemt ook de aandacht voor de procesmatige aspecten van ruimtelijke beheer en ruimtelijke ontwikkeling toe. Het is nodig om deze processen goed in kaart te brengen en te evalueren in termen van het beheren en ontwikkelen van ruimtelijke kwaliteit. De vraag hoe ruimtelijke processen van kwaliteitsreductie en kwaliteitstoename verlopen, moeten centrale aandacht krijgen in evaluaties. Dit is nodig omdat daarmee zowel een bijdrage geleverd kan worden aan provinciale leerprocessen als aan het vermogen tot samenwerking.

6.2 Investeren in evaluatie

Vanuit het perspectief van het regionaal bestuur is het raadzaam om op een aantal manieren in evaluatietrajecten en -programma's te investeren. In de eerste plaats verdient het aanbeveling om te investeren in het uitvoeren van gezamenlijke evaluatieprojecten daar waar in samenwerking met andere overheden (rijk, gemeenten) inhoud wordt gegeven aan de ruimte in Ne-

²⁰ Zie met name het IPO-rapport *'Van ordenen naar ontwikkelen; provincies investeren in de kwaliteit van de ruimte'* (IPO 2000). Hierin bepleiten de opstellers onder meer samenwerking, omdat ook zij onderkennen dat de afweging van claims en het bereiken van ruimtelijke kwaliteit niet de verantwoordelijkheid is en kan zijn van een enkele bestuurslaag.

derland. Dit geldt met name voor convenants die met publieke partners op rijks- en lokaal bestuursniveau zijn gemaakt. In de tweede plaats is het verstandig naast inhoudelijke (dat wil zeggen ruimtelijke) aspecten ook procesmatige aspecten in evaluaties mee te nemen. Met andere woorden: de bestuurlijke processen die leiden tot het formuleren, uitvoeren en toetsen van bijvoorbeeld groene en rode contouren (moeilijkheden en mogelijkheden) moeten worden geëvalueerd. Op die manier kan dan inzichtelijk worden gemaakt hoe men tot een bepaalde ruimtelijke uitkomst is gekomen en - belangrijker nog - hoe hiervan geleerd kan worden. In de laatste plaats moet het regionaal bestuur zich meer bewust worden van de op handen zijnde evaluaties bij andere overheden en doet men op het regionale bestuursniveau er goed aan bij deze evaluaties aan te haken. Op die manier kunnen doelen en perspectieven van het regionaal bestuur worden opgenomen in evaluaties van andere overheden. Tegelijkertijd neemt zij kennis van andere perspectieven die in eigen evaluaties kunnen worden meegenomen.

6.3 Verantwoording

Het afleggen van verantwoording zal een van de centrale functies blijven van evaluaties. Daarbij wordt echter in toenemende mate erkend dat naar de samenleving als geheel toe er vooral ook behoefte is aan een gezamenlijke verantwoording. Of er ruimtelijke kwaliteiten worden geschaad of bereikt, hangt immers af van het samenspel tussen de verschillende overheden en ook met de maatschappelijke groeperingen en private partijen.²¹ Dit behoeft niet in te houden dat er sprake is van een diffuse verdeling van verantwoordelijkheden, waardoor niemand op zijn verantwoordelijkheid kan worden aangesproken. Wel is het van belang om onderscheid te maken tussen de eigen (regionale) verantwoordelijkheid, die van anderen en – last but not least – gedeelde verantwoordelijkheden.

Het is zaak om hierover bij de uitvoeringsovereenkomsten afspraken te maken. Het is immers verstandig om alleen maar verantwoording te willen en moeten afleggen voor die taken en zaken die een partij – zoals de provincie – in alle redelijkheid ook zo zelfstandig kan uitvoeren dat deze op de uitvoering aangesproken kan worden. Bij het afleggen van deze verantwoording zullen bestuurders verschillende lijnen volgen. Eén daarvan heeft betrekking op verantwoording aan het rijk op het punt van specifieke verantwoordelijkheden die men in het kader van het nationale beleid heeft genomen. Daarnaast bestaat uiteraard ook een regionale verantwoordingsbehoefte, waarin overigens de context van nationaal beleid en de landelijke context van belang kan zijn. Belangrijk is dat provincies, net als het rijk, nu ook bereidheid vertonen voor hun beleid ook verantwoording af te leggen naar aanpalende provincies, naar gemeenten en waterschappen en burgers. Dit kan de transparantie en de legitimiteit van het provinciaal ruimtelijk bestuur en beleid aanzienlijk vergroten.

Een derde stap in het domein van verantwoording afleggen, wordt gezet in de richting van het afleggen van verantwoording over gezamenlijk beleid. Er is dan sprake van een gezamenlijke verplichting (bijvoorbeeld regio, rijk, gemeente). Dit is een nieuw fenomeen waar uitdrukkelijk aandacht aan moet worden besteed. Tot nu toe ligt de nadruk immers op de interne verantwoording naar de eigen politieke toezichhouders. Zodra de weg ingeslagen wordt van gemeenschappelijke verantwoording, een weg die in het licht van de zoektocht naar ruimtelijke kwaliteit bijzonder aantrekkelijk is, dient zich wel de vraag aan hoe deze verantwoording is af te leggen. Enerzijds kan er sprake zijn van een gemeenschappelijk evaluatie naar de politieke gremia van de verschillende bestuurslagen, zoals we die al hebben geduid in ons de gebruiks-

²¹ Zoals eerder gesteld, besteden we binnen het kader van onze opdracht om een evaluatiearrangement te maken voor de overheid geen aandacht aan het perspectief van de maatschappelijke groeperingen en private partijen. We zijn evenwel wel van mening dat deze veel meer dan nu het geval is opgenomen moeten worden in de evaluatie, zowel als partij die geëvalueerd wordt, als in de rol van mede-evaluator.

aanwijzing voor de Tweede Kamer, maar er kan ook toe worden overgegaan om de verantwoording direct naar de samenleving vorm te geven. Dit zou een mooie stap voorwaarts zijn in de richting van passende vormen van participatieve democratie.

6.4 Leren

Doordat partijen weten dat ze verantwoording moeten gaan afleggen, wordt in potentie een gunstige voorwaarde geschapen voor effectief gedrag. Of dit gedrag ook optreedt, hangt echter in aanzienlijke mate af van het leervermogen van de (provinciale) organisatie, of beter gezegd van de provinciale politici, bestuurders en ambtenaren. Zij zullen het beheer van de ruimte moeten moderniseren en vormgeven aan een hoogwaardige ontwikkeling van de ruimte. Zoals ook in 'Van ordenen naar ontwikkelen' (IPO 2000) wordt aangegeven, zijn hiervoor vaardigheden nodig die nog niet steeds in ruimte mate voorhanden zijn. Evaluaties van de provinciale praktijk (in brede zin) van ruimtelijk beheer en ruimtelijke ontwikkeling zijn dan, mits goed uitgevoerd uiteraard, een mooie basis voor reflectie en (via training) voor gedragsontwikkeling en versterking van de interne werkwijze van provincies zodanig dat ze beter geëquipeerd zijn voor samenwerking met anderen. Zo blijft de rol van de politici, bestuurders en ambtenaren in formele zin dezelfde, maar kan de rolinvulling aanzienlijk verbeterd worden. Deze verbetering is nodig om slagvaardig te kunnen handelen in een complexe samenleving.

6.5 Samenwerking

Evaluaties zoals in de handleiding geschetst, kunnen naast de verantwoordingsrol en de lerende werking ook de rol hebben dat ze de (alom noodzakelijk geachte) samenwerking met andere partijen agenderen, de kwaliteit van deze samenwerking zichtbaar maken en als consequentie ook bijdragen aan de kwaliteit van diezelfde samenwerking. Evaluaties worden daarmee een belangrijk onderdeel van de beleidssporen. Ze worden niet meer ingezet als een sluitstuk van een beleidsproces (en te vaak als mosterd na de maaltijd), maar als startpunt en tussenmomenten van beleidsprocessen. Evaluaties worden zo geschikt gemaakt als expliciete reflectiemomenten in een langdurig leertraject, waarbij steeds weer de vraag gesteld wordt wie wat hoe heeft gedaan en hoe dat in de onderlinge interactie is uitgekapt, hoe tevreden de eigen en andere partijen daarover zijn en wat daaruit voor verbetervoorstellen zijn de distilleren. Als onderkend wordt dat de ruimtelijke kwaliteit die provinciale politici en bestuurders willen realiseren voor een groot deel afhangt van de handelingen van andere partijen, dan kan het bijna niet anders dan dat er de provinciale partijen veel aan gelegen moet zijn om die andere partijen te helpen een leerproces door te maken op weg naar ruimtelijke kwaliteit. Natuurlijk is het met sommige andere actoren moeilijk om deze kwaliteit te bereiken, maar dat valt pas vast te stellen nadat er diverse pogingen daartoe zijn genomen. Een kritische reflectie op de eigen rol daarbij lijkt gewenst.

7. Gebruiksaanwijzing vanuit het perspectief van het lokale niveau

7.1 Perspectief

Een groot deel van het ruimtelijk beleid krijgt in concreto gestalte op lokaal niveau. Nationaal beleid en eigen ontwikkelingsperspectief zijn daarbij richtinggevende kaders, maar veel wordt ook bepaald door de krachtenvelden waarin lokaal ruimtelijk beleid en de realisatie daarvan gestalte moet krijgen. Om te beginnen is veelal afstemming en onderhandeling nodig met omringende gemeenten en de provincie. Daarnaast moet voor financiering en realisatie overeenstemming worden bereikt met private actoren en vaak ook met de centrale overheid (om in aanmerking te komen voor project- en stimuleringsgelden. Ten slotte hangt, zeker ook op het lokale niveau, ruimtelijk beleid nauw samen met andere beleidsterreinen en hebben omgekeerd ontwikkelingen op andere terreinen (economische ontwikkeling, verkeer en vervoer, huisvesting, toerisme etc.) nadrukkelijk ruimtelijke implicaties. Ruimtelijke ontwikkelingen komen derhalve tot stand in wisselwerking met vele andere partijen en in afweging met veel andere beleidsterreinen.

De samenhangen en krachtenvelden die er vanuit het perspectief van het lokale niveau rondom RO-beleid bestaan, zullen met zich mee brengen dat vaak andere evaluatievragen en criteria interessant zijn dan op andere niveaus het geval is. Zo zal RO-beleid soms meer als een onderdeel van economisch of sociaal beleid worden gezien dan als een overkoepelend beleidsterrein. Wat landelijke kaders betreft, kan dan bijvoorbeeld het Grote-StedenBeleid of een component daarvan meer aanknopingspunten bieden dan het nationale RO-beleid. Het gevaar bestaat dat deze complexiteit tot een zeer groot aantal deels overlappende evaluaties leidt.

7.2 Investeren in evaluatie

Vanuit de eigen probleemstellingen van het lokale niveau en de daar werkzame krachtenvelden, ligt het om te beginnen voor de hand eigen evaluaties te doen plaatsvinden. De aanbevelingen in de Handleiding kunnen daarbij goed dienst doen, omdat ook op lokaal niveau sprake is van multi-actorprocessen. En dan gaat het niet alleen over 'RO-actoren' en zeker ook niet alleen om publieke actoren. Ook private investeerders, maatschappelijke organisaties belangengroepen en (individuele) burgers zijn belangrijke stakeholders.

De bestuurlijke en financiële afhankelijkheid van andere bestuurslagen in het kader van het RO-beleid, vraagt daarnaast om investering in evaluatiearrangementen in deze sfeer. Om te beginnen wordt er op andere niveaus het nodige aan monitoring gedaan. Het is van belang om daar een goed beeld van te hebben en van de mogelijkheden daarvoor aandachtspunten en gegevens aan te dragen. Dat is zeker het geval als het gaat om zaken waar het lokale niveau bestuurlijk en/of financieel op wordt afgerekend. Maar ook kan dergelijke monitoring informatie opleveren die vanuit het lokale perspectief van belang is voor eigen evaluaties en beleidsontwikkeling.

Voorts kan het vanuit het lokale niveau gezien van belang zijn selectief te participeren in op andere bestuurslagen geëntameerde multi-actorevaluaties. Daarmee kan ertoe worden bijgedragen dat lokale vragen en problemen worden meegenomen, wat enerzijds kan helpen om bij anderen meer inzicht in de lokale omstandigheden te geven en anderzijds inzicht op kan leveren waarmee op lokaal niveau aan verbetering van beleid en/of afstemming kan worden gewerkt.

Ten slotte is aandacht nodig voor de vraag hoe economisch met evaluatie om te gaan. Waar kan gebruik worden gemaakt van, of aangehaakt bij evaluaties van anderen? Waar kunnen

anderen worden uitgenodigd om samen evaluaties op te zetten? Wat zijn de prioriteiten vanuit politiek-inhoudelijk perspectief en wat vanuit het ontwikkelen van bestuurlijke en publiek-private samenwerkingsrelaties?

7.3 Verantwoording

De combinatie van de voorgestelde investeringen in evaluatie maakt een verantwoording mogelijk die niet alleen over doelbereiking of effectiviteit gaat, maar waarin ook zichtbaar kan worden gemaakt in hoeverre en op welke punten ontwikkelingen wel en niet op het lokale niveau (door gemeenten) gestuurd konden worden. Daarmee wordt, zowel naar de lokale volksvertegenwoordiging en bevolking als naar bestuurlijke instanties op hoger niveaus een verdieping van het proces van verantwoording mogelijk.

7.4 Leren

Eigen evaluaties en participatie in door anderen geëntameerde evaluaties zijn van belang om nieuwe inzichten te verwerven. Daarbij kan 'externe' monitoring bruikbare gegevens opleveren, maar eigen reflectie is onontbeerlijk. Leerprocessen kunnen betrekking hebben op ruimtelijke oplossingen (bijv. voor meervoudige claims), op het omgaan met problemen (bijv. de verkaveling van verschillende budgetten), of op de ontwikkeling van effectieve en creatieve samenwerkingsrelaties.

7.5 Interactie en samenwerking

Op het lokale niveau komen ruimtelijke ontwikkelingen in wisselwerking tussen vele actoren tot stand. Multi-actorevaluaties, al dan niet gekoppeld aan bestaande overlegarrangementen, kunnen het vinden van verbindingen tussen de belangen, visies en mogelijkheden van de verschillende actoren bevorderen.

Deel 4: Achtergronden

1. Inleiding

In dit deel volgt een aantal op zichzelf staande rapportages over door ons verrichte verkenningen. In het voorgaande is daar op relevante punten reeds naar verwezen. Na een literatuurverkenning in hoofdstuk 2 volgt een beschouwing over gebruik van en ervaringen met kengedallen, productbegrotingen, prestatie­meting en het VBTB-proces als pogingen om effecten en effectiviteit van beleid zichtbaar te maken (hoofdstuk 3). In hoofdstuk 4 wordt kort ingegaan op de ontwikkeling die het doelmatigheids­onderzoek van de Algemene Rekenkamer heeft doorgemaakt en de mechanismen waarvan doorwerking daarvan afhankelijk is. Tenslotte wordt ingegaan op ervaringen met recente evaluaties op het gebied van RO (hoofdstuk 5) en op andere beleidsterreinen (hoofdstuk 6). Bij al deze verkenningen wordt aan het eind expliciet de vraag beantwoord wat lessen en aandachtspunten zijn die eruit volgen voor evaluatie en evaluatiearrangementen met betrekking tot de 5N.

2. Literatuurverkenning

2.1 Inleiding

In dit hoofdstuk schetsen we de ontwikkeling in het denken over evaluatie. Met name die evaluaties en inzichten die relevant zijn voor de evaluatie van de 5N komen aan de orde. Het gaat om nieuwe concepten voor en onderzoek naar beleidsevaluatie waarbij:

- het beleid strategisch en op hoofdlijnen is geformuleerd door de centrale overheid;
- invulling van het beleid voor een belangrijk deel door andere overheden gebeurt;
- zowel de centrale als decentrale overheden de invulling veelal gestalte geven in overleg en samenwerking met andere actoren;
- tussen verschillende overheden en andere actoren aanzienlijke variaties bestaan in percepties, perspectieven, belangen en beleidstheorieën;
- oogmerk is dat zij enerzijds informatie oplevert voor beleidsverantwoording door de centrale overheid en anderzijds bijdraagt aan leerprocessen bij (onderdelen van) centrale en decentrale overheden;
- de context (maatschappelijke opvattingen, internationale ontwikkelingen, fysische en ecologische problemen, technologische ontwikkelingen) aan voortdurende veranderingen onderhevig is.

Toch ligt hiermee niet vast welke wetenschappelijke inzichten voor ons doel relevant zijn. Zowel in de wetenschap als in de praktijk bestaan er verschillende visies op wat evaluatie is of zou kunnen zijn. Deze blijken gerelateerd aan visies op sturing. Enigermate gechargeerd kunnen twee ideaaltypen van evaluatie tegenover elkaar worden gezet.

2.2 Twee visies op evaluatie

Grofweg zijn twee denkmodellen over beleid te onderscheiden: een rationeel-analytische en een sociaal-constructivistische beleidsoriëntatie. Kenmerkend voor de rationeel-analytische oriëntatie is het doelgericht handelen en de (logische) volgtijdelijkheid van een aantal onderscheiden activiteiten. Eerst moeten problemen worden gedefinieerd en de doelen en criteria worden bepaald. Op basis daarvan is een beoordelingskader te construeren, zijn alternatieven te ontwerpen en kan een keuze worden gemaakt. Als laatste stap wordt beleid geëvalueerd. Hierbij staat centraal dat de beoordeling van het uitgevoerde beleid afhangt van de mate waarin het de vooraf gestelde doelen heeft gediend. De rationeel-analytische beleidsoriëntatie kent dus een grotendeels sequentiële en fasegewijze evaluatie, waarbij soms iteratieslagen mogelijk zijn maar die alleen incrementele veranderingen of ontwikkelingen in doelen of beoordelingscriteria toelaten. Ten slotte neemt deze oriëntatie voornamelijk de beleidsinhoud als evaluatieobject; doelbereiking, -doorwerking en -conformiteit staan centraal.

Kortom: de rationeel-analytische beleidsoriëntatie richt zich vooral op de vraag of vooraf gestelde doelen gehaald zijn, wat de oorzaken daarvan zijn en wie daarvoor verantwoordelijk zijn. Meten, beschrijven en inhoudelijk oordelen zijn centrale evaluatieactiviteiten.

Bij de sociaal-constructivistische benadering van evaluaties worden beleidsprocessen opgevat als grillige, dynamische processen, waarin doelen, opvattingen, probleemdefinities en oplossingen voortkomen uit interactie tussen actoren. Ze zijn daarom aan verandering onderhevig. Beleidsontwikkeling wordt opgevat als een interactieve zoektocht: beleidsmakers zoeken ge-

durende het proces naar geschikte probleemdefinities, doelen, oplossingen en afwegingscriteria. Er wordt ruimte gelaten om nieuwe inzichten van personen en externe ontwikkelingen door te laten werken in probleemomschrijvingen, doelformuleringen et cetera. Vanuit deze oriëntatie focust evaluatie meer op het beleidsproces (en minder de beleidsinhoud). Beschrijven, oordelen en leren staan centraal.

Het verschil tussen beide opvattingen is het best samen te vatten in de tweedeling doelrealiseren (rationeel-analytische benadering) versus doelzoeken (sociaal-constructivistische benadering). Evaluaties uitgevoerd in het rationeel-analytische perspectief hebben als ambitie om op gezette tijdstippen het beleid inhoudelijk te meten en beoordelen aan de hand van vooraf opgestelde doelen en normen (doelrealisatie). Aan de andere kant hebben evaluaties uitgevoerd in het sociaal-constructivistische perspectief als ambitie om procesmatige uitspraken over beleidsontwikkeling te doen: is er ruimte voor interactie geweest, is er in samenwerking tot een doelformulering gekomen, etcetera. Met andere woorden: het interactieproces tussen belanghebbende partijen bij beleidsvorming, die mogelijk heeft geleid tot bijstelling van doelstellingen, wordt centraal gesteld. In het rationeel-analytische perspectief domineert de wil om (tussentijds) te beoordelen op vastgestelde doelen en in het sociaal-constructivistische perspectief domineert de wil om recht te doen aan de diversiteit en ontwikkeling in doelen..

aandachtspunt voor evaluatie 5N

- In de 5N zijn elementen van beide beleidsoriëntaties terug te vinden. Weliswaar ligt het accent wellicht wat meer op de sociaal-constructivistische optiek, maar de begrotings- en verantwoordingssystematiek van de Rijksoverheid (VBTB) legt het accent juist op het rationeel-analytische perspectief. Voor een evaluatiemodel voor de 5N betekent dit dat - ondanks bovengenoemde conceptuele tegenstrijdigheden - naar vormen moet worden gezocht die beide perspectieven recht doen.²²

2.3 Leren van evalueren

In de literatuur over en praktijk van evaluatie krijgt de kwestie van het *gebruik* van evaluaties veel aandacht. Het is een belangrijk punt, omdat hiervan de toegevoegde waarde van evaluaties afhangt. Het is ook een lastig punt, omdat de doorwerking van evaluatie niet altijd even gemakkelijk is waar te nemen. Zo maakt Weiss (1980) aannemelijk dat evaluaties niet zozeer directe effecten hebben, maar bijdragen aan de 'state of mind' van besluitvormers, die vervolgens kleine, net even andere, beslissinkjes nemen die nauwelijks als zodanig herkenbaar zijn., maar wel tot koerswijziging leiden. Pogingen om het gebruik vast te stellen via het aantal verwijzingen naar een evaluatierapport in nieuwe beleidsstukken (Mulder et al. 1991), zeggen dan weinig over doorwerking en nog minder over het leervermogen. De bevinding van Mulder c.s. dat gebruik (in de zin van verwijzing) toeneemt naarmate een evaluatie komt op een tijdstip dat er nog politieke discussie over het onderwerp is, lijkt wel een belangrijk aandachtspunt. Belangrijker zijn vervolgens twee vragen:

- I. Waar hangt het gebruik en de doorwerking van af?
- II. Wat is de aard van de doorwerking?

ad I. Verklaring voor gebruik.

²² Overigens kan er ook los van de onderliggende filosofie van beleid en verantwoording gepleit worden voor een combinatie van perspectieven, omdat ze bij politieke en maatschappelijke actoren bestaande opvattingen en kennisinteresses weerspiegelen.

In een overzichtsartikel over gebruik van evaluatie (Shulha and Cousins 1997, p. 196) komt naar voren dat het desbetreffende onderzoek

... produced an extended “shopping list” of likely predictors [of utilization]. These predictors tended to cluster into categories such as: (a) relevance, (b) credibility, (c) user involvement, (d) communication effectiveness, (e) potential for information processing, (f) clients’ need for information, (g) anticipated degree of program change, (h) perceived value of evaluation as a management tool, (i) quality of evaluation implementation, and (j) contextual characteristics of the decision or policy setting.

Bij de contextuele karakteristieken kan worden gedacht aan de timing van rapporten (Mulder et al. 1991; Rist 1994, p. 194-199) of aan het al of niet bestaan van geïnstitutionaliseerde evaluatiepraktijken en -procedures (Leeuw and Rozendal 1994). Onderzoek naar dergelijke samenhangen leveren lang niet altijd (dezelfde) correlaties op. Een algemene indruk die het achterlaat, is dat alle genoemde factoren gebruik kunnen bevorderen, maar noch voldoende, noch noodzakelijke voorwaarde daarvoor zijn. Belangrijker is echter waarschijnlijk dat het niet zo zeer bepaalde ‘objectieve’ karakteristieken van evaluaties of evaluatieprocessen (namelijk de operationalisering die de onderzoekers gebruikten) zijn die het gebruik bepalen, maar de percepties die betrokkenen daaromtrent hebben of krijgen (zie bijvoorbeeld e.g. Rist 1994, p. 200-203). De vraag verschuift daarmee naar: hoe komen die percepties tot stand? Antwoorden daarop zijn vooral gezocht in samenhang met contextuele factoren en met de relatie tussen evaluator en stakeholders. De centrale notie die hierbij naar voren komt, is dat het gebruik van evaluaties sterk afhangt van de mate waarin actoren hun ideeën en ervaringen erin herkennen en ‘zien’ wat ze met de uitkomsten kunnen doen. Dat zou consequenties moeten hebben voor hoe evaluaties worden opgezet en uitgevoerd (Patton 1997).

ad II. Aard van de doorwerking

Hoe worden evaluaties gebruikt? Wat wordt ervan geleerd? En door wie? In het verlengde van bovenstaande inzichten over de mate van gebruik, is het aannemelijk dat ook de aard van het gebruik afhankelijk is van de wijze waarop evaluaties aansluiten bij de belevingswereld van de actoren. Het eerder aangehaalde onderzoek over de doorwerking van Rekenkameronderzoek laat dat ook zien. Hierbij zijn twee kanttekeningen op hun plaats. Ten eerste zij erop gewezen dat ‘aansluiten bij’ niet hetzelfde is als ‘overeenkomen met’. Integendeel, wil een evaluatie bijdragen tot leren of tot anders handelen, dan zal zij nieuwe, onbekende of verrassende elementen moeten bevatten. Essentieel is echter dat zulke elementen op een of andere manier ‘aanhaken’ bij bestaande manieren van denken en doen, en deze daardoor enigermate veranderen. Ten tweede is het van belang op te merken dat dergelijke aansluiting (nieuwe ‘koppelingen’) op heel verschillende manieren tot stand kunnen komen (De Vries 2000, p. 83-86; Van der Meer et al. 2000, p. 260-262). We noemen vier mechanismen.

- *Cognitieve koppeling*: delen van de evaluatie worden vanuit het referentiekader van een of meer actoren als relevant en informatief ervaren. Zo was de bevinding van de Algemene Rekenkamer dat de inzetbaarheid van de Lynx-helikopters van de Koninklijke Marine minder dan 40% was, voor de Marine niet iets wat men als oninteressant terzijde kon of wilde leggen. Het raakte aan de kern van het zelfbeeld van de Marine: een technisch geavanceerd en snel inzetbaar krijgsmachtonderdeel. De kritiek leidde dan ook tot een ingrijpende reorganisatie van het onderhoud (verandering op handelingsniveau). De stelling van de Rekenkamer in hetzelfde rapport dat er bij de aanschaf van de Lynxhelikopter geen concurrentiestelling had plaatsgevonden werd bij de marine als een min of meer wereldvreemde en niet relevante constatering afgedaan: helikopters koop je niet in de supermarkt. Dit laatste voorbeeld laat ook zien hoe belangrijk de context die een actor ervaart, is voor het totstandkomen van aansluiting. De marine zit, althans in de eigen beleving, in een situatie waarin je blij mag zijn als er, op het moment dat je een nieuw wapen-

systeem nodig hebt, ergens in de (westerse) wereld een systeem in ontwikkeling is dat enigermate aan je eisen voldoet. Daar spring je snel op in, zodat je nog een beetje kunt bijsturen in ruil voor financieel bijdragen aan dat ontwikkelingstraject. Het is een complete seller's market.

- *Sociale koppeling*: hier gaat het meer om het aanzien, het vertrouwen, de formele positie van de evaluerende instantie, in relatie tot of gezien door een bepaalde actor. De doorwerking kan, anders gezegd, soms niet zozeer afhangen van wat er gezegd wordt, maar door wie. Hetzelfde oordeel, met dezelfde argumentatie, uitgesproken door een actiegroep kan een ander gewicht hebben dan wanneer het gaat om een gerenommeerd onderzoeksbureau, de interne accountantsdienst, de Tweede Kamer of de rechter. Hierbij spelen institutionele factoren een rol. Maar ook beeldvorming en reputatie.
- *Indirecte koppeling*: ook als actor X niet leert of verandert als gevolg van een van beide voorgaande mechanismen, kan er bij actor X toch nog effect optreden. Dit is het geval als actor X vanuit zijn wereldbeeld zich wel gedwongen of uitgedaagd voelt iets te doen met reacties van bepaalde derde partijen op de evaluatie. Als een departement zijn schouders ophaalt over een evaluatie, maar het parlement maakt er een punt van, kan zo'n situatie zich voordoen.
- *Aansluiting bij ideeën of praktijken in ontwikkeling*: overheidsorganisaties en publieke en publiek-private netwerken zijn gekenmerkt door verschillende perspectieven tegelijkertijd. Binnen en tussen instanties bestaan verschillen in ideeënvorming en de ontwikkeling van praktijken. Evaluaties kunnen die processen beïnvloeden, bijvoorbeeld doordat bepaalde ideeën daardoor een steuntje in de rug of juist tegengas krijgen. Het gewicht daarvan wordt weer door cognitieve en/of sociale koppeling bepaald. Het kan er toe leiden dat lokale ideeën en praktijken algemener of legitiemer worden. Een voorbeeld is een reactie uit het onderwijsveld naar aanleiding van een advies van de Onderwijsraad aan staatssecretaris Adelmund over de basisvorming. De Onderwijsraad suggereerde in oktober dit jaar naar aanleiding van geconstateerde problemen onder meer om minder vakken verplicht te stellen en differentiatie naar niveau tussen leerlingen toe te laten. De reactie: "Dit legitimeert de feitelijke praktijk" (terwijl de Staatssecretaris pas in december reageert).

De conclusie luidt dat doorwerking (gebruik en leereffecten) van evaluaties niet alleen door de opzet en inhoud van die evaluatie wordt bepaald. Opzet en inhoud zijn van belang, maar de impact van evaluaties verloopt via de perspectieven van actoren en hun onderlinge interacties. Inzicht in die perspectieven is behulpzaam bij het opzetten van evaluaties. Ook de formele en informele positie van de evaluerende instantie speelt een rol van belang.

Aandachtspunten voor evaluatie 5N

- Voor doorwerking en leren van evaluatie is aansluiting van de inhoud van de evaluatie op probleemdefinities van betrokken actoren van groot belang (cognitieve aansluiting). Waar probleemdefinities verschillen, moet - uit het oogpunt van effectiviteit van de evaluatie - gepoogd worden die verschillende opties in de evaluatie recht te doen. Dat betekent dat de perspectieven van betrokkenen bij het RO-beleid in het evaluatieproces in kaart en op tafel gebracht moeten worden.
- Ook sociale aansluiting kan doorwerking en leren bevorderen. Dat betekent dat aandacht nodig is voor de relatie tussen evaluatoren en geëvalueerden. Zaken als de gepercipieerde belangen van de evaluatoren, de hun toegeschreven onafhankelijkheid, en het vertrouwen in hun aanpak spelen daarbij een rol, maar ook de wijze waarop geëvalueerden bij de evaluatie betrokken zijn. Wie is opdrachtgever, wie formuleert de vraagstellingen, en hoe komen conclusies over implicaties voor het beleid tot stand?
- Ten slotte is timing van evaluaties van belang: hoe kan in de tijd een goede aansluiting bij beleidsprocessen op verschillende plaatsen in het netwerk worden gerealiseerd?

2.4 Multi-actorevaluatie

De onderkenning van het aansluitingsprobleem heeft bijgedragen aan de ontwikkeling van allerlei nieuwe evaluatiemethodes, waarin betrokken actoren zelf nadrukkelijk een rol spelen. Het gaat daarbij onder meer om ‘utilization focussed evaluation’ (Patton 1997), participatieve evaluatie (Cousins and Earl 1992), responsieve evaluatie (Abma 1996), ‘dialogical evaluation’ (zie ook Edelenbos and Eeten 2001; Ryan and DeStefano 2001) en vierde-generatie-evaluatie (Guba and Lincoln 1989). De basale gedachte achter deze nieuwe vormen van evaluatie is dat er in multi-actorbeleidsprocessen sprake is van pluraliteit van waarden en dat de uiteindelijke effecten van beleid de resultanten zijn van processen van sociale constructie. Er zijn daarom twee redenen waarom de perspectieven van actoren een rol moeten spelen in evaluatieprocessen. De eerste is dat zonder rekenschap van actorperspectieven het verloop van beleidsprocessen niet kan worden begrepen. De tweede is dat evaluaties waarin actoren zich niet op enigerlei wijze herkennen, minder effectief zullen zijn. In deze nieuwe vormen van evaluatie word daarom door de verschillende actoren geparticipeerd in het formuleren van de probleemstelling(en), het verzamelen van gegevens en/of het trekken van conclusies. Zo nemen Guba en Lincoln de door actoren naar voren gebrachte ‘claims’, ‘concerns’ en ‘issues’ tot uitgangspunt waaromheen hun vierde-generatie-evaluatie wordt georganiseerd (Guba and Lincoln 1989, p. 39-40). Soms wordt er tussen actoren onderhandeld over de vraagstelling van de evaluatieopdracht, hetgeen tot gezamenlijk opdrachtgeverschap kan leiden. Soms ook worden onderscheiden perspectieven - voor zover die in onderling overleg niet direct op één lijn kunnen worden gebracht - naast elkaar meegenomen teneinde in een volgende fase van het evaluatieproces nader te worden besproken in overleg tussen betrokken actoren (Guba and Lincoln 1989, p. 42). Deze nieuwe vormen hebben potentieel enkele voordelen.

- De perspectieven van actoren worden rechtstreeks in het evaluatieproces ingebracht en sturen dit mede.
- Actoren komen in het evaluatieproces met andere perspectieven in aanraking en kunnen daarvan leren (koppelingen tussen perspectieven, verandering van relaties).
- Actoren zijn direct betrokken bij de interpretatie van evaluatieresultaten. Daarmee treden er al tijdens het evaluatieproces leereffecten op.

Er kleven ook problemen aan:

- De vraag wie de stakeholders zijn, wordt nogal eens vanuit verschillende perspectieven verschillend beantwoord.
- Het is niet altijd gemakkelijk alle stakeholders in het evaluatieproces te betrekken, bijv. om dat de machtsverhoudingen als te scheef worden ervaren, of omdat onduidelijk is wat er met de resultaten gebeurt.
- Vaak is het moeilijk de stap te zetten van in kaart brengen van verschillende perspectieven naar gedeelde bevindingen die in handelen worden vertaald.

aandachtspunt evaluatie 5N

- Bij de evaluatie van de 5N is zoeken naar een aanpak die de genoemde voordelen te gelde maakt en de nadelen zo veel mogelijk tegengaat, van groot belang, niet alleen vanwege het analytische inzicht dat de beleidseffecten uiteindelijk het resultaat van wisselwerking tussen actoren is, maar ook omdat dat in het gekozen sturingsmodel expliciet tot uitgangspunt is gekozen en het de effectiviteit van de evaluatie zal versterken.

2.5 Joint fact-finding

Inleiding

Een vorm van multi-actorevaluatie is 'joint fact-finding'. De benadering sluit nauw aan bij de door Guba en Lincoln ontwikkelde vierde-generatie-evaluatie. We gaan er in deze paragraaf wat uitvoeriger op in.

Informatie en kennis uit evaluaties worden niet op voorhand door iedere betrokkene gedragen. Evaluaties die als een proces van 'joint fact-finding' worden georganiseerd, zijn in dit verband een interessante optie. Zij zijn er expliciet op gericht om tot gedragen en betekenisvolle evaluatieresultaten te komen.

Joint fact-finding (gezamenlijk kennisvinding) komt voort uit de 'negotiation'- en 'mediation'-literatuur uit de Verenigde Staten (zie bijvoorbeeld Süsskind and Cruikshank 1987; Ehrmann and Stinson 1999). Ehrmann en Stinson (1999, p. 376) omschrijven het als volgt: "Joint fact-finding is a central component of many consensus building processes; it extends the interest-based, cooperative efforts of parties engaged in consensus building into the realm of information gathering and scientific analysis. In joint fact-finding, stakeholders with differing viewpoints and interests work together to develop data and information, analyze facts and forecasts, develop common assumptions and informed opinions, and, finally, use the information they have developed to reach decisions together".

Bij een proces van joint fact-finding ontwikkelen en implementeren stakeholders, besluitvormers en experts in interactie een onderzoeksstrategie en -aanpak voor het beantwoorden van kennisvragen.

Het proces van joint fact-finding moet inzicht bieden in de kennis waarover men het eens is en waarover het men nog oneens is. Men is op zoek naar een gemeenschappelijke basis op grond waarvan gezamenlijk handelen mogelijk wordt. Een mogelijkheid hiertoe is om te streven naar uitonderhandelde kennis ('negotiated knowledge'). Deze vorm van kennis wordt niet langer gezien als de 'echte' feiten, de 'juiste' interpretaties of de 'werkelijke' situatie. Het is in onderhandeling nagaan waarover men het eens kan zijn en waarover (nog) niet ('agree to disagree') om vervolgens gezamenlijk te bepalen op welke wijze eventueel alsnog overeenstemming te bereiken is.

Aandachtspunten in een proces van joint fact-finding

Hoe moet een proces joint fact-finding eigenlijk georganiseerd worden? Hieronder wordt een aantal aandachtspunten besproken, zonder daarbij een bepaalde volgtijdelijkheid na te streven. Deze punten moeten per situatie en beleidsonderwerp specifiek worden ingevuld.

Inventarisatie cruciale kennisvragen

Bij het organiseren van joint fact-finding moet eerst worden nagaan welke cruciale kennisvragen er leven onder de stakeholders. Welke kennis wordt gemist bij het in kaart brengen van de centrale problematiek, welke kennis is nodig om tot een prioritering van de oplossingen van de problematiek te komen? Op deze manier ontstaat een vraaggestuurde kennisontwikkeling. De kennisvragen worden in samenspraak geformuleerd. Bijvoorbeeld: een onafhankelijke procesbegeleider signaleert de noodzaak van de kennisvragen en expliciteert deze vragen, zodat stakeholders hiermee kunnen instemmen.

Aanstellen onpartijdige en gezaggenietende externe deskundige

De kennisvragen worden vervolgens meegegeven aan een inhoudelijke deskundige, die status heeft verworven op het desbetreffende onderwerp en het vertrouwen geniet van de stakeholders en dus als onpartijdig wordt gezien. Op deze manier ontstaat een onafhankelijke productie van kennis, die niet op voorhand argwanend wordt bejegend.

Ontwikkelen en vaststellen onderzoeksdesign

Daarna bepalen de belanghebbenden in onderlinge samenspraak en met de expert de wijze van beantwoording van de kennisvragen. Het onderzoeksdesign waarmee de deskundige op zoek gaat naar antwoorden wordt gezamenlijk ontwikkeld en bepaald. In interactie worden dan de onderzoeksmethode en de uitgangspunten, aannames en vooronderstellingen die daaraan ten grondslag liggen besproken en vastgesteld. Ook de (inhoudelijke, temporele en geografische) systeemgrenzen en de scope van het onderzoek (bijvoorbeeld: wanneer start het onderzoek?, wanneer wordt het afgesloten? welke effecten worden nog in het onderzoek meegenomen? welke onderwerpen nemen we wel en niet mee in het onderzoek?) worden in overleg bekrachtigd.

Het belangrijkste bij het ontwikkelen van het onderzoeksdesign (van bijvoorbeeld het zoeken naar de haalbaarheid en uitvoerbaarheid van oplossingen) is dat beoordelingscriteria in onderhandeling tussen de stakeholders worden ontworpen. Waarden, belangen en normen komen uit een onderling onderhandelingsproces naar voren, en niet van buiten het proces. Met andere woorden: de stakeholders hebben kans gekregen om hun centrale belangen en waarden te laten doorklinken in de opzet van het onderzoek. De criteria genieten meer gezag als stakeholders hun belangen en de borging ervan erin kunnen herkennen. Kortom: de onderzoeksopzet is een uitkomst van een proces van discussie en onderhandeling tussen stakeholders en de externe expert en is niet van tevoren gegeven.

Mocht het zo zijn dat stakeholders er niet uitkomen over welke methodieken te gebruiken, dan kan men besluiten meerdere (concurrerende) methodieken en/of gevoeligheidsanalyses uit te voeren om te analyseren in hoeverre de uitkomsten variëren bij verschillende aannames die ten grondslag liggen bij de verschillende methodieken. Ook kan men beslissen om meerdere onderzoeksmodellen met elkaar te integreren. Ten slotte kan men besluiten om een commissie van wijzen samen te stellen, waarin onafhankelijke experts met verschillende disciplines zitting hebben en als taak hebben hardnekkige kennisconflicten te beslechten.

Tussentijdse bespreking en bekritisering

Een ander aandachtspunt is dat kennisvorming een continu proces van onderhandeling en discussie is. Het is niet zo dat vooraf het onderzoeksdesign wordt gemaakt en dan de experts aan het werk gezet worden, zonder iets van zich te laten horen en pas na enkele maanden met een rapport terugkeren in het proces. Tussentijds worden besprekings- en reflectieronden gehouden, waarin experts uitleg geven over hoe ver het onderzoek is en welke tussenresultaten er gevonden zijn. Op deze wijze krijgen stakeholders de gelegenheid kennisvorming van dichtbij te volgen, te bespreken, te bekritisieren en - belangrijker nog - te internaliseren. Als stakeholders de ontwikkeling van de kennisvorming kunnen volgen, dan wordt de kans vergroot dat de uiteindelijke kennis ook wordt aanvaard. Daarnaast geeft deze vorm van interactieve kennisvinding stakeholders de gelegenheid om voortschrijdende inzichten (over problemen en oplossingen) en nieuwe vragen tijdig in het kennisconstructiesproces te brengen.

De tussentijdse bespreking biedt bovendien experts de gelegenheid om hun kritische reflectie op het in interactie bepaalde onderzoeksdesign te geven. Kennisvorming ondergaat zo continu een wetenschappelijke en/of professionele toets. De expert is niet de hofleverancier van kennis, maar is meer een kennis-mediator, een kennismakelaar en -schakelaar geworden (Edelenbos and Goedhart 2001).

Koppeling kennisvorming aan beleidsvorming

Belangrijk aspect in de organisatie van een proces van joint fact-finding is het koppelen van het proces van kennisvorming aan het proces van beleidsvorming. Bij joint fact-finding worden deze twee processen niet geïsoleerd van elkaar uitgevoerd, maar worden ze juist met elkaar vervlochten. Zo wordt de kans vergroot dat kennis leidt tot gezaghebbende en geconsolideerde besluitvorming (Jasanoff 1990).

aandachtspunt voor evaluatie 5N

- De benadering van joint fact-finding biedt interessante mogelijkheden voor koppeling van onafhankelijk onderzoek aan participatie van betrokken actoren en voor inbedding van evaluatie in beleidsprocessen.

2.6 Verklarende evaluatie

In veel evaluaties wordt nagegaan of vooraf (al dan niet vaag of dubbelzinnig geformuleerde) doelen zijn gerealiseerd, dan wel of bijvoorbeeld de beleidsuitvoering aan een aantal algemene criteria van doelmatigheid, transparantie en/of accountability voldoet. Daar is op zichzelf natuurlijk niets mis mee, want dat levert zinvolle informatie op die tot leren of bijstelling van beleid kan bijdragen. Toch zit hieraan een complicatie die aan de hand van de hiervoor besproken inzichten en evaluatiepraktijken kan worden verduidelijkt. Een concrete bevinding, bijvoorbeeld dat een bepaalde doelstelling slechts in beperkte mate is gerealiseerd, zal vaak voor verschillende actoren een verschillende betekenis hebben. Dat komt, omdat ze er verschillende contexten bij betrekken en verschillende ‘beleidstheorieën’ gebruiken om het resultaat te verklaren. Zo kan het niet alleen gebeuren dat de één het resultaat ‘gegeven de omstandigheden’ nog best aardig vindt, terwijl de ander het ver beneden de maat vindt. Maar ook kan de een het wijten aan de inzet van bepaalde ambtenaren of bestuurders, een ander aan de vaagheid van en onenigheid over het beleid en een derde aan de eerder genoemde omstandigheden, die nu eenmaal niet méér mogelijk maken. In een enigermate gepolitiseerde of gepolariseerde situatie kan dit tot een langdurig en weinig vruchtbaar ‘ideologisch’ debat aanleiding geven. Dat is zeker het geval als in het evaluatieonderzoek zelf geen informatie is verzameld over en geen analyse is gemaakt van mechanismen en processen waardoor het onderhavige resultaat is geproduceerd. In een op verklaring gerichte evaluatie wordt daarover wel evidentie gezocht (Pawson and Tilley 1997). Natuurlijk is dergelijke evidentie, hoe methodisch verantwoord ook totstandgekomen, niet allesbepalend voor leerprocessen. Het blijft zo dat actoren deze nieuwe evidentie in eerste instantie vanuit hun bestaande perspectief zullen interpreteren. Maar verklarend onderzoek kan niettemin resultaten opleveren die in leerprocessen tot nieuwe koppelingen aanleiding geven. Anders gezegd: doorwerking van evaluatie zal nooit om de interpretaties van actoren heen kunnen, maar evaluatie kan wel meer en diverser materiaal aandragen voor die interpretatieprocessen.

aandachtspunt evaluatie 5N

- Onderzoek naar de wijze waarop (verandering van) context, maatschappelijke en fysieke processen en technologische ontwikkelingen mede beleidsuitkomsten bepalen, kan helpen om de stap van al dan niet opgetreden veranderingen naar implicaties voor beleid te maken.
- Vanuit het in par. 2.4 geformuleerde aandachtspunt zouden hierbij onder meer door actoren veronderstelde mechanismen onderwerp van onderzoek moeten zijn.

2.7 Kwaliteit van beleid

In beleidsevaluatie is altijd de vraag naar de kwaliteit van het beleid aan de orde: hoe is het met de kwaliteit gesteld, en hoe kan het beter? De eerder gesignaleerde diversiteit in het denken over beleid en over evaluatie weerspiegelt zich ook in de vraag wanneer een beleidsproces van hoge kwaliteit is. Als we de opvattingen op een lijn zouden zetten dan ligt aan de ene kant daarvan de opvatting dat een beleidsproces kwaliteit heeft naarmate het voorgenomen beleid (van een centraal gestelde actor) en de beoogde doelen van deze actor worden gereali-

seerd. De kracht van deze benadering is die van elegantie en eenvoud: de maker van beleid formuleert – het liefst meetbare – doelen (we zeggen dan wel dat hij het beleid vaststelt). Daarin ligt de potentiële kwaliteit van het beleid al vast.²³ De vraag van de evaluatie is vervolgens of deze kwaliteit in de uitvoering ook bereikt is. Aan de ander kant van de lijn staat de opvatting dat kwaliteit van beleidsprocessen ligt in de uitkomst en dat deze uitkomst in complexe samenlevingen met complexe bestuurlijke structuren, zoals ook het geval is in de Nederlandse ruimtelijke ontwikkeling en inrichting, resulteert uit de inbreng van diverse partijen met een variëteit aan doelen. Ook wordt in deze benadering gesteld dat veel kwaliteit pas ontdekt wordt in het proces in niet terug te vinden is in de doelen vooraf, maar voortvloeit uit de interactie tussen partijen.

Of om het in termen van recente complexiteitstheorieën te stellen: kwaliteit is vooral een fenomeen van emergentie. Kwaliteit wordt ontdekt in de confrontatie tussen verschillende claims, veel meer dan dat deze door een van de partijen vooraf is af te kondigen. De voorstellen in deze studie naar de evaluatie van het uitvoeringsproces van 5N, kunnen op deze zijde van de geschetste lijn worden geplaatst. In deze benadering is er overigens wel volop aandacht voor de kwaliteit van het voorgenomen beleid zoals dat is verwoord in de 5N en in de uitvoeringsprogramma's Nationaal Ruimtelijke Beleid. Deze kwaliteit wordt evenwel in termen van evaluatie niet eenzijdig als uitgangspunt van evaluatie genomen, maar tweezijdig gemeten. Met tweezijdig gemeten bedoelen we dat in de evaluaties zowel aandacht dient te zijn voor de vraag of de doelen van 5N gehaald worden als voor de vraag op 5N en de manier waarop het ministerie daar zelf mee omgaat wel helpt om resultaten van hoge ruimtelijke kwaliteit te behalen.

We maken gebruik van de tweezijdige benadering, om daarmee een antwoord te geven op de bekende bezwaren die al vele jaren worden ingebracht tegen de evaluaties die de beleidsdoelen van een actor centraal stellen. Deze bezwaren zijn de volgende.

- Het voorgenomen beleid hoeft niet uitvoerbaar te zijn; het kan zijn dat het beleid een symbolisch karakter heeft en niet over de kwaliteit uitvoerbaarheid beschikt. Een evaluatie die dan kijkt of doelen zijn gehaald levert weinig inzicht op.
- Het voorgenomen beleid hoeft niet effectief te zijn, omdat de beleidsmakers de impact van de door hen gehanteerde beleidsinstrumenten overschatten. Dit is met name het geval wanneer ze weinig oog hebben voor de complexiteit van situaties waarin ze de instrumenten toepassen. Een evaluatie die dan kijkt of doelen zijn gehaald, levert geen inzicht in de tekortkomingen bij de beleidsmaker zelf.
- Het voorgenomen beleid bestaat niet. Er zijn meerder beleidsmakende instanties; het beleid van andere instanties is vaak ook politiek gelegitimeerd, maar kan tegengesteld zijn aan het geëvalueerde beleid; Evaluatoren concluderen dan vaak dat het geëvalueerde beleid wordt tegengewerkt, terwijl een tweezijdig perspectief tot een in onze ogen meer vruchtbare inzicht kan komen dat de beleidsmakende instanties er niet in geslaagd zijn hun kwaliteitsstrevingen op een vruchtbare wijze op elkaar te betrekken.
- Kwaliteitscriteria verschuiven in de loop der tijd, deels als gevolg van maatschappelijke gebeurtenissen en ontwikkelingen en voor een andere deel door leerprocessen en als gevolg van interactie met andere partijen; Evaluatie leidt nog wel eens tot de conclusie dat beleidsuitvoerders zijn afgedwaald van het goede pad, terwijl het vruchtbaarder lijkt te concluderen dat partijen ruimte hebben gehad om te leren en de vraag voorlicht op het behaalde resultaat in het licht van de dan geldende opvattingen over kwaliteit niet tot een positief oordeel leiden.

²³ In discussies over kwaliteit wordt dan gesproken van inherente kwaliteit; bijvoorbeeld de compacte stad als beleidsconcept is inherent goed, ongeacht de gepercipieerde consequenties.

- Juist in de huidige samenleving met veeleisende burgers die steeds hogere (ruimtelijke) kwaliteitseisen stellen, valt er iets voor te zeggen om beleidsprocessen vooral te beoordelen in hun vermogen tot voortdurende innovatie en niet in de eerste plaats op doelrealisatie.
- Juist op concreet uitvoeringsniveau komen verschillende beleidslijnen samen die soms tegengesteld zijn en in ieder geval tegen elkaar afgewogen moeten worden, zodanig dat hier nog veel beleid (maken van afwegingen) wordt gemaakt en kwaliteitseisen moeten worden gecombineerd.
- Juist de betrokkenheid van meerdere stakeholders het ook zinvol maakt om evaluatie niet langer te zien als een eenzijdige vorm van verantwoording afleggen aan een instantie, maar als een proces van meervoudige verantwoording.

Dergelijke observaties leiden ons tot een manier van denken over kwaliteit en kwaliteitszorg, waarin de zoektocht naar en de meervoudigheid van 'kwaliteit' centraal staan. Mede omdat het bedrijfsleven veel afhankelijker is van de kwaliteitsoordelen van haar klanten is dit soort van kwaliteitsdenken in de private sector sterker ontwikkeld en zijn er ook als meer procesgerichte kwaliteitssystemen voor ontwikkeld.

Het gaat daarbij bijv. om het opzetten van multidisciplinaire ontwerp- en ontwikkelteams, waarbij ook al vaak partijen die aan het eind van de productieketen bij betrokken zijn (concurrent engineering). Ook het organiseren van de reflectie bij alle lagen en eenheden van organisaties op de kwaliteit van de eigen producten en productieproces, alsmede op de mogelijkheden tot voortdurende kwaliteitsverbetering, meestal behandeld in de literatuur onder de titel Total Quality Management is daarvan een belangrijk voorbeeld (Dahlgaard et al. 1998). De nadruk ligt daarbij op het gezamenlijk nemen van verantwoordelijkheid voor kwaliteit in interorganisatorische verbanden, op voortdurende innovatie, op een horizontale én een verticale organisatie van kwaliteitszorg en met gebruikmaking van zich ontwikkelende, meervoudige criteria. Er is daarbij dus niet alleen aandacht voor het verantwoorden van vooraf geformuleerde opdrachten (wat blijft bestaan), maar ook voor het leren ontwikkelen van betere producten en productieprocessen en voor het vermogen om via samenwerking tot verbetering te komen. En daar waar in de publieke sector het accent nog ligt op het eenzijdig verantwoording afleggen aan de politiek en op het meten op vooraf gedefinieerde criteria, stellen wij voor om de aandacht voor leren en samenwerken nadrukkelijk te verhogen. We denken dat deze verbreding van de evaluatie kan helpen de kwaliteit van overheidsproducten en overheidsprocessen beter in beeld te krijgen en te verbeteren.

aandachtspunt voor evaluatie 5N

- De 5N is een strategische nota, die onderkent dat ruimtelijke kwaliteit gerealiseerd wordt in interactie tussen partijen. Ook onderkent de 5N dat ruimtelijke kwaliteit een meervoudige betekenis heeft en niet centraal is af te kondigen. Dit moet consequenties hebben voor het evaluatiearrangement dat de uitvoering van de 5N begeleidt. Ten eerste zal er sprake moeten zijn van een duurzaam arrangement in de zin dat de gehele planperiode met regelmaat evaluatie plaats vindt. Ten tweede zal er sprake moeten zijn van een meervoudig arrangement, waar de evaluatieve vermogens van de verschillende partijen die invulling geven aan de kwaliteit van ruimte worden gecombineerd. Ten derde is er naast het afleggen van verantwoording voor het bereiken van enkele duidelijke doelen, ook aandacht nodig voor de op innovatie gerichte leerprocessen en voor de samenwerkingsproces dat onontbeerlijk zijn voor het realiseren van kwaliteit.

3. Ervaringen met kengetallen en VBTB

3.1 Inleiding

In de nota Van Beleidsbegroting Tot Beleidsverantwoording (VBTB) (Tweede Kamer 1998-1999, p. 4) staat dat structuur, inhoud en presentatie van de begroting en verantwoording van een ministerie zullen worden aangepast. Voor de begroting gaat gelden: *"Wat willen we bereiken, wat gaan we daarvoor doen en wat mag dat kosten?"*, voor de verantwoording: *"Hebben we bereikt wat we hebben beoogd, hebben we gedaan wat we daarvoor zouden doen en heeft het gekost wat we dachten dat het zou kosten?"* Aansprekende zinnen die een welbewuste en daadkrachtige overheid uitstralen. Dit voornemen past goed in het huidige tijdsgewricht waar meten is weten en sturing op output wenselijk en noodzakelijk wordt geacht.

De in het voorgaande neergelegde wensen ten aanzien van doelstellingen, instrumenten, kosten en verantwoording zijn niet nieuw. In de jaren zestig en zeventig is hard gewerkt aan het rationaliseren van beleidsprocessen. Beleidsanalytische methoden en technieken werden te hulp geroepen om beleidsalternatieven helder ten opzichte van elkaar te positioneren en onderling te vergelijken. Doelstellingenanalyses werden noodzakelijk geacht voor de verdere rationalisering van beleid. Het was de tijd van de prestatiebegroting, waarbij het ging om de in fysieke eenheden uitgedrukte prestaties - het volume-aspect van de begroting - te koppelen aan de kosten die met die prestaties samenhangen. Eind jaren tachtig, begin jaren negentig trad een nieuw tijdperk in, geïnspireerd op verwachtingen van NPM (New Public Management). Kengetallensystemen, programmabudgettering, productbegroting zijn pogingen die de afgelopen tien jaar expliciet onder de aandacht zijn gekomen. Recent is daar de VBTB dus bijgekomen.

In dit hoofdstuk gaat het om het in beeld brengen van de ervaringen op het terrein van (ex ante en ex post) evaluatie met behulp van het zichtbaar (proberen te) maken van de prestaties van overheden. Allereerst wordt stilgestaan bij de ervaringen van de bestuurslaag die het dichtst bij de burger staat, namelijk de gemeenten. Gemeenten zijn in Nederland al meer dan een decennium bezig om kengetallen bij bestuur en beheer te introduceren. Vervolgens wordt stilgestaan bij de ambities die op rijksniveau hebben gespeeld en nog steeds spelen om prestaties in het proces van begroten en verantwoorden mee te nemen. Vandaar dat in een afzonderlijk deel wordt teruggeblikt op de ervaringen met de departementale doelstellingenanalyse en prestatiebegroting. Daarna wordt ingegaan op het VBTB-proces en de eerste resultaten daarvan. Uit zowel de gemeentelijke ervaringen als de ervaringen op rijksniveau zal blijken dat bij het inweven van een grotere prestatie-oriëntatie tot op heden geen wezenlijke voortgang is geboekt.

3.2 Ervaringen van gemeenten: kengetallen en productbegrotingen

In 1990 is door de ErasmusUniversiteit Rotterdam een onderzoek verricht naar de theoretische en praktische mogelijkheden en beperkingen van kengetallen bij bestuur en beheer bij decentrale overheden (Haselbekke et al. 1990). Uit een literatuurstudie die in het kader van dit onderzoek is verricht, kwamen tal van potentiële voordelen van kengetallensystemen naar voren (Haselbekke et al. 1990, p. 138-139):

- Beter kwantitatief inzicht
- Impuls tot bedrijfsmatig werken
- Bijdrage aan kwaliteit dienstverlening

- Bijdrage aan de verhoging productiviteit
- Toeneming beheersbaarheid
- Meer duidelijkheid binnen de organisatie
- Grotere vrijheid bij taakuitvoering
- Beter gemotiveerde werknemers
- Betere mogelijkheden tot sturing

Behalve voordelen werden er ook potentiële nadelen onderkend (ibid. p.139):

- Bureaucratisering
- Schijn duidelijkheid
- Conflicten bij disfunctioneel gebruik
- Mogelijkheden tot symbolisch beleid

Wellicht belangrijker dan deze waarnemingen gebaseerd op de literatuur, waren de inzichten die het praktijkdeel van het onderzoek opleverde. In de praktijk werd een bonte variëteit aan kengetallen en kengetallensystemen waargenomen. Initiatieven voor kengetallensystemen komen vrijwel altijd van de directie Financiën, meestal in samenspraak met de politieke leiding. Het blijkt dan ook dat veel van dergelijke systemen erop zijn gericht om de effectiviteit en de efficiency van gemeentelijk handelen in beeld te brengen. Efficiency geeft een beoordeling van het voortbrengingsproces; er wordt mee uitgedrukt of er doelmatig (genoeg) wordt omgegaan met schaarse alternatief aanwendbare middelen. De effectiviteit of doeltreffendheid geeft aan in hoeverre de prestaties tegemoetkomen aan de doelstellingen. Doelstellingen vinden in de regel hun weerslag in normen. Dit kunnen technische normen zijn, normen ontleend aan de resultaten van de eigen organisatie of normen afgeleid uit processen bij vergelijkbare organisaties. Met name dit laatste blijkt in de praktijk een sterke aantrekkingskracht uit te oefenen op degenen die kengetallensturing ambiëren. Benchmarking is hier een typisch voorbeeld van. In de praktijk blijkt deze vorm van het gebruik van kengetallen ook de grootste problemen met zich mee te brengen. Verschillen in waarden van kengetallen kunnen namelijk ook door andere factoren worden verklaard dan verschillen in doelmatig en/of doeltreffend werken. Er zijn verschillende categorieën waardoor verschillen verklaard kunnen worden (Haselbekke et al. 1990, p. 67068).

De comptabiliteitsvoorschriften laten ruimte om de begroting op een niet uniforme wijze in te richten (of kunnen niet voorkomen dat van de voorgeschreven uniformiteit wordt afgeweken); kengetallen kunnen daarom wel dezelfde naam hebben, maar verschillende verschijnselen in beeld brengen en omgekeerd.

- Gemeenten verkeren in (zeer) verschillende omstandigheden wat betreft omvang, ligging, bodemgesteldheid, bevolkingsopbouw, sociaal-economische structuur; al deze factoren zorgen voor verschillen in de waarde van kengetallen; voor een correcte interpretatie van deze verschillen moeten de gemeten waarden steeds worden geschoond voor dit soort 'omgevingsfactoren'.
- Gemeenten kunnen doelbewust verschillend beleid voeren. Verschillend beleid zal tot verschillende kengetallen aanleiding (kunnen) geven.
- Tenslotte treden er verschillen op in de gemeten waarden van kengetallen, omdat gemeenten verschillen kunnen vertonen in de mate waarin zij doelmatig en/of doeltreffend werken.

Deze factoren leiden tot het inzicht dat sturingsambities, beheerambities en bijbehorende evaluatiestudies, die zich richten op doelmatigheid en doeltreffendheid op meer belemmeringen stuiten dan wordt verondersteld. (Bordewijk and Klaassen 2000) hebben het feitelijk gebruik van prestatiekengetallen bij gemeenten in kaart gebracht evenals de stand van zaken met be-

trekking tot de productbegroting. Bij een productbegroting is de gedachte dat het politieke bestuur bepaalt *wat* de dienst moet doen, welke producten de dienst moet leveren, terwijl de managers bepalen *hoe* dat gebeurt, met inzet van welke bedrijfsmiddelen, analoog aan externe uitbesteding. Het betekent dus een andere manier van begroten. Niet langer heeft de autorisatie betrekking op de *input*, op de bedragen die mogen worden besteed voor bepaalde overheidstaken en de middelen waarmee die gerealiseerd gaan worden, maar het gaat om de *output*, om de producten die door de organisatie geleverd worden en de prijs die daarvoor vanuit de collectieve middelen betaald wordt.

In het onderzoek zijn negen middelgrote gemeenten betrokken. Van deze gemeenten zijn de begrotings- en rekeningstukken onderzocht en zijn interviews gehouden met degenen die in de gemeente verantwoordelijk zijn voor de introductie van prestatiekengetallen en/of de productbegroting. Uit de begrotings- en rekeningstukken van de onderzochte gemeenten bleek dat geen van de gemeenten beschikt over een productbegroting in de hiervoor beschreven zin. In al deze gemeenten spreekt men wel van een productbegroting, maar nergens wordt consequent per product aangegeven welk bedrag beschikbaar is, zodat een prijs per product kan worden berekend. Nergens wordt ook consequent afgerekend op basis van geleverde prestaties.

De prestatiegegevens in de onderzochte begrotings- en rekeningstukken hebben alleen een illustratief karakter, en zijn niet functioneel in de zin dat ze fungeren als ijkpunten voor sturing of verantwoording. Per productgroep worden kengetallen in de begroting en in het jaarverslag opgenomen. Het betreft verhoudingsgetallen (bedrag per inwoner) en kostendekkingpercentages. Daarnaast worden outputgegevens per productgroep gepresenteerd. Het betreft hier outputgegevens als aantal ingeschreven cursisten, het aantal polderinstallaties, het aantal hectaren watergangen, aantal wijkcentra, aantal wijkposten, etcetera.

Geen van de onderzochte gemeenten beschikt dus over een productbegroting die werkelijk is opgebouwd uit de prijs per producteenheid en het aantal eenheden, terwijl ook andere elementen van New Public Management slechts in beperkte mate worden aangetroffen. In feite is er bij de onderzochte gemeenten dan ook geen sprake van outputsturing. Dat is opmerkelijk, aangezien al deze gemeenten wel een dergelijke productbegroting nastreven, en ook overigens onder medewerkers planning en control welhaast consensus lijkt te bestaan over de wenselijkheid van een productbegroting. De vraag is hoe deze discrepantie is ontstaan. Waarom lukt het niet, om tot echte productbegrotingen te komen? Maar ook: waarom blijft men zich dat toch als ideaal stellen? In de studie worden drie verklaringen gegeven, die hieronder kort de revue passeren.

De homogeniteit van het product

In de eerste plaats kan worden geconstateerd dat het in veel gevallen lastig is om het product goed te definiëren, of de kwaliteit te meten. Het heeft alleen zin de begroting te baseren op aantallen producten en de prijs daarvan, wanneer die producten homogeen zijn, zodat de dienst geen mogelijkheid heeft om tot lagere productiekosten te komen door de aard van het product te wijzigen. Bij een volledig homogene productie kunnen de activiteiten in het productieproces eenduidig aan de input worden gerelateerd (het berekenen van standaardkosten voor activiteiten). Bij de samenhang van activiteiten met de geproduceerde goederen en diensten gaat het om de relatering van activiteiten aan de prestaties. Bij een volledige samenhang is normering van de activiteiten voor een bepaalde productie mogelijk. Worden beide stappen samengenomen, dan ontstaat de volgende classificatie van productietypen:

		Is de output te identificeren?	
		Ja	Nee
Is de productie homogeen?	Ja	I	II
	Nee	III	IV

Type I, met homogene productie en identificeerbare output, staat voor outputbudgettering. Er kan volledige kostprijscalculatie plaatsvinden (bijvoorbeeld voor uitgifte paspoorten). Type II, waarbij de output niet of niet volledig te identificeren is, maar de productie wel homogeen is, staat voor procesbudgettering (normering van de activiteiten, bijvoorbeeld bekostiging op basis van verzorgde lesuren). Type III, met identificeerbare output maar geen homogene productie, staat voor projectbudgettering (een kengetallenproject binnen een gemeente). Type IV, niet identificeerbare output en geen homogene productie, staat voor inputbudgettering (vergoeding van ingezette productiemiddelen).

Niet alle gemeentelijke taken behoren tot type I, maar tot de types II, III, of IV. Deze lenen zich niet voor outputbudgettering. Weinig gemeenten maken dit onderscheid, waardoor gemeentelijke organisatieonderdelen hun inspanningen moeten uitdrukken in producten zonder dat die term van toepassing is, of zonder dat sturing op basis van die producten mogelijk is.

Afrekenen op basis van prestaties

Een systeem van verrekening op basis van geleverde prestaties bij een van tevoren vastgestelde prijs per eenheid kan tot problemen leiden, wanneer de te leveren productie niet van tevoren kan worden vastgesteld door de gemeenteraad, maar afhankelijk is van maatschappelijke ontwikkelingen en de vraag van het publiek, zoals bij de inzameling van afval. Het gaat er niet om een van tevoren vastgesteld tonnage op te halen. Het is vooral belangrijk dat er geen vuilniszakken blijven staan. Een grotere afvalproductie leidt dan tot een overschrijding in de concernrekening.

Prikkelssystemen

Het optreden van efficiencywinst en de mogelijkheid voor een organisatieonderdeel om die geheel of gedeeltelijk te behouden, is in het systeem van afrekenen op basis van prestaties de prikkel om coöperatief te zijn. Wanneer het organisatieonderdeel de efficiencywinst geheel mag behouden, zal de gemeente in totaliteit, (en daarmee de lokale samenleving) nooit kunnen profiteren van de lagere voortbrengingskosten, bijvoorbeeld in de vorm van lagere tarieven of prijzen van producten. Men komt op die wijze in een volledig starre gebudgetteerde situatie terecht.

Wanneer efficiencywinst geheel of gedeeltelijk worden afgeroomd, perverteert het systeem echter ook. Immers, in het jaar $t+1$ dat volgt op het jaar waarin de efficiencywinst is behaald, wordt het budget op een lager niveau vastgesteld (het budget van jaar t verminderd met het deel van de efficiencywinst dat is afgeroomd). De mogelijkheden om in jaar $t+1$ efficiencywinst te behalen (ceteris paribus) nemen af. In het jaar $t+2$ verslechtert de situatie opnieuw, omdat het budget voor het organisatieonderdeel weer vermindert met de afgedragen efficiencywinst. De loyaliteit van het organisatieonderdeel zal steeds verder afnemen naarmate het moeilijker wordt een efficiencywinst te behalen (de prikkel valt steeds meer weg). Als ten slotte de prikkel is “opgesoupeerd”, kan het systeem van bezuinigen op basis van ingehouden efficiencywinsten niet meer in stand worden gehouden door interne motivatie.

3.3 Ervaringen bij het Rijk: lessen uit het verleden

Vanaf de jaren zeventig is er bij de rijksoverheid een concrete belangstelling te constateren voor het aanleggen van criteria bij de beoordeling van afzonderlijke uitgaven. In de Miljoenennota 1970 wordt uitgebreid ingegaan op het gebruik van het Planning Programming Budgeting System (PPBS) ofwel programmabudgettering voor beleidsombuigingen en -aanpassing. Programmabudgettering bestaat uit een combinatie van een tweetal methoden, namelijk de indeling van de begroting naar programma's en de analyse van baten en offers. Ook krijgt Performancebudgeting ofwel prestatiebegroting voor het eerst uitgebreid aandacht. Prestatiebegroting is minder omvattend dan programmabudgettering. Programmabudgettering leent zich voor uitspraken over doelmatigheid én kan tevens bijdragen aan prioriteitenstelling. Bij prestatiebegroting concentreert men zich op het relateren van tot stand gebrachte fysieke resultaten aan de daarmee gemoeide kosten; door middel van een confrontatie met normgetallen resulteert een indicatie van de bereikte doelmatigheid. In de Miljoenennota 1971 worden de voornemens op de verschillende terreinen concreter. Als verzamelmechanisme voor fysieke prestaties van overheidsprogramma's wordt de prestatiebegroting gezien als uitgangspunt voor het opstellen van meerjarenramingen. Beleidsanalyse wordt in de nota geplaatst naast de prestatiebegroting en de meerjarenramingen. Bij zowel prestatiebegroting als meerjarenramingen worden de doelstellingen als gegeven beschouwd. Men ziet het in die tijd dan ook als noodzakelijk om rijksbreed tot een volledige inventarisatie van doelstellingen te komen (de departementale doelstellingsanalyse). De Werkgroep Beleidsdoelstellingen (opgericht tijdens de eerste vergadering van de COBA op 3 november 1971) heeft zich beziggehouden met *"het inventariseren en het zoveel mogelijk rubriceren van de doelstellingen van het overheidsbeleid en - indien mogelijk - het aangeven van indicatoren voor een nadere concretisering met het oog op het operationeel maken van de doelstelling"* (COBA 1978). Langs deze weg hoopte men tot een checklist te komen waaraan toekomstige beleidsanalyses konden worden getoetst. Daarnaast diende het resultaat van de doelenanalyse, een departementale programmastructuur, zijnde een hiërarchisch geordende structuur van hoofd-, sub-, enkelvoudige doelstellingen, instrumenten en activiteiten, als hulpmiddel bij departementale beleidsvoorbereiding en evaluatie. Een volledige, departementale programmastructuur zou meer dan één doel dienen: de relevantie en actualiteit van doelstellingen kunnen eraan afgemeten worden, ze biedt een referentiepunt om instrumenten die beogen bij te dragen aan eenzelfde doelstelling tegen elkaar af te wegen, ze is een middel om interferentie tussen programma's op het spoor te komen, ze bevordert de communicatie tussen beleidsvoorbereidende en beleidsuitvoerende afdelingen en ze vergroot de slagvaardigheid van het beleid door een betere informatie over doelstellingen en de beschikbare instrumenten (Werkgroep beleidsdoelstellingen 1973).

De hooggespannen verwachtingen rechtvaardigen de breedte waarmee de doelenanalyses werden opgezet: in januari 1974 neemt de COBA het principebesluit dat alle departementen een start maken met de doelstellingsanalyse. Voor de ondersteuning van deze activiteit werd een handleiding opgesteld, toen bleek dat de wijze waarop de exercitie moest plaatsvinden lang niet altijd duidelijk was (COBA 1976). De in de handleiding aanbevolen benadering loopt langs twee lijnen: uitvoering van een deductieve en een inductieve analyse om tot een uiteindelijke structuur in hoofd- en subdoelstellingen, enkelvoudige doelstellingen, instrumenten en activiteiten te komen.

De resultaten van de doelstellingsexercitie waren mager. Slechts vier ministeries (BiZa, Defensie, VROM en CRM) hebben een doelenstructuur opgesteld. Andere departementen kwamen niet verder dan één of meer hoofddoelen. Eind 1979 blaast de Werkgroep Beleidsdoelstellingen het streven naar een volledige doelstellingsstructuur van alle departementen af. Wilmer constateert dat de uit een departementale analyse voortvloeiende doelstellingen niet op directe wijze een basis kunnen vormen voor een departementale programmastructuur (Wilmer 1980). Een directe relatie tussen doelen en middelen ontbreekt. Doelenanalyses leiden hoogstens tot een fijnzinnig inzicht

in hetgeen men (uiteindelijk) wenst te bereiken, maar niet tot indicaties over het instrumentarium dat moet worden ingezet en hoe dit moet geschieden. Daarnaast bezitten de analyses een hoog abstractieniveau. Gaandeweg ontstaat twijfel over de praktische toepassingsmogelijkheden en wordt men zich bewust van mogelijke negatieve consequenties van het expliciteren van doelstellingen. In 't Veld wijst op de onderschatting van de complexiteit van het openbaar bestuur: "*In ieder geval gaat de departementale handleiding van de COBA mank aan het euvel dat van variëteit van de procedure geen sprake is, dat dus betrekkingen tussen procedure, organisatie en beleid zijn verwaarloosd of miskend*" (In 't Veld 1978, p. 41).

Na het mislukken van de doelstellingsanalyse verschuift de aandacht op rijksniveau naar indicatoren en kengetallen. De indicatoren en kengetallen zijn bedoeld om inzicht te krijgen in het verloop en werking van beleid- en productieprocessen. Een belangrijk forum vormt het Platform Beleidsanalyse. Het platform is een initiatief van het Ministerie van Financiën en dient voor het uitwisselen van kennis en ervaring op het gebied van beleidsonderzoek tussen rijksoverheid, de andere overheden, het bedrijfsleven, de non-profitsector en de wetenschappelijke wereld. In 1990 komt het platform met het rapport '*Maatwerk door meetwerk; over het meten met indicatoren en kengetallen bij beleids- en productieprocessen in de kwartaire sector*'. Het gaat in op methodologische kwesties (signalering en terugkoppeling, afbakening productieprocessen, te meten elementen van productieprocessen, normstelling voor de beoordeling en inhoud van begrippen als effectiviteit efficiency en productiviteit). De Miljoenennota 1991 maakt van enkele proefprojecten, gericht op het werken met (begroting)kengetallen. Deze interdepartementale proefprojecten zijn van verschillende aard (bijvoorbeeld kwaliteit van beleid van brandweeropleidingen, de budgetteringsproblematiek van justitiële inrichtingen, activiteiten van de Inspectie Gezondheidsbescherming, doelmatigheid van de behandeling van aanvragen voor subsidies bij EZ). De proefprojecten leggen bestuurlijke, organisatorische en inhoudelijke problemen bij het werken met kengetallen bloot. De problemen zijn weliswaar oplosbaar, maar niet volgens een standaardformaat voor alle organisaties in alle omstandigheden. Daarnaast werd gesignaleerd dat er kengetallen en indicatoren te onderscheiden zijn die een nauwe relatie met de begroting hebben, bijvoorbeeld voor zover ze inzicht geven in de doelmatigheid en doeltreffendheid van beleid- en productieprocessen. Dergelijke kengetallen werden relevant geacht voor de begroting, maar zijn in de praktijk daarin niet opgenomen. Met name dit punt, gevoegd bij het streven om te komen tot resultaatgericht management, vormt de aanleiding voor het VBTB-proces²⁴.

3.4 Ervaringen bij het Rijk: het VBTB-proces

In 1998 is besloten tot het VBTB-proces (Tweede Kamer 1998-1999). Dit is een nieuwe poging om budgetten, output en maatschappelijke effecten met elkaar in verband te brengen. Niet zozeer het budget per begrotingsartikel, maar de te bereiken beleidsdoelen komen centraal te staan. De begrotingen moeten dus meer inzicht geven in de doelen van het beleid, de effecten die dit beleid oplevert, alsmede de kosten die daarmee zijn gemoeid.

VBTB is een invulling / explicitering van het model van resultaatgericht management (RM) dat in 1991 in de nota 'Verder bouwen aan beheer' is geïntroduceerd. De volgende tabel, ontleend aan Berkhout en Sanders (2000:7), geeft een overzicht van de relatie tussen RM en VBTB, alsmede de invulling die VBTB heeft gekregen:

²⁴ Ook speelde de vrij fundamentele kritiek van de in de Tweede Kamer vertegenwoordigde fracties mee. Deze kritiek komt erop neer dat binnen een begroting, tussen de afzonderlijke begrotingen en tussen de begroting en verantwoording te weinig logische samenhang wordt ervaren (Van der Burg et al. 1999, p. 226).

Resultaatgericht management	VBTB
Centrale sturing op hoofdlijnen	<ul style="list-style-type: none"> - Beleidsagenda met prioriteiten in dat begrotingsjaar in relatie met regeerakkoord en het beleid op de artikelen. - Beleidsartikelen met doelen en autorisatie op dat niveau. - Overzichten die per beleidsthema inzicht geven in doelen, activiteiten, prestaties en beleidsuitgaven van ministeries. - Beheersinformatie in de verdiepingsbijlage.
Meer integraal management	<ul style="list-style-type: none"> - Bijeenbrengen programma- en apparaatuitgaven. - Bijeenbrengen meerdere instrumenten op één beleidsartikel - Opnemen bedrijfsvoeringparagraaf gebaseerd op risicobeleid en uitzonderingen
Afspraken over de productie en daarvoor te maken kosten	<ul style="list-style-type: none"> - Autorisatie van in begroting opgenomen gelden, via directe koppeling aan een toelichting met prestaties, effecten, doelgroepgegevens en veronderstellingen in ramingen en beleid i.p.v. organisatie- en instrumentgeoriënteerde artikelen.
Systematische verantwoording	<ul style="list-style-type: none"> - Derde woensdag in mei als tegenhanger van 'Prinsjesdag'. - Beleidsmatige verantwoording (departementaal jaarverslag). - Financieel Jaarverslag van het Rijk versus Miljoenennota. - Samenhang verbeteren tussen begroting en verantwoording door een vaste, herkenbare indeling van alle stukken. - Toelichting aan hand van prestatie- en effectgegevens over de ondernomen activiteiten en de te bereiken doelen.
Periodieke doorlichting, evaluatieonderzoeken	<ul style="list-style-type: none"> - Eens in de 5 jaar iedere algemene beleidsdoelstelling (centraal op de beleidsartikelen in de begroting) evalueren, gericht op de effecten van het beleid, de beleidsuitvoering en de juistheid en actualiteit van de beleidsinhoud(theorie)

In eerder onderzoek hebben Berkhout en Sanders geïnventariseerd wat de kritische succesfactoren zijn voor de implementatie van VBTB (2000, p. 6).

1. Politiek moet inzien dat het een langdurig en omvattend proces is. Bij politiek, ministeries en toezichthouders is een cultuuromslag vereist.
2. Er is structureel politiek draagvlak nodig voor implementatie van het rapport en voor het (durven) sturen op resultaten.
3. Politiek moet zich beperken tot hoofdlijnen, zichzelf disciplineren in wat ze willen weten.
4. De departementen moeten vertrouwen hebben in het proces en in de afspraken.
5. Er is structureel draagvlak bij de ambtelijke leiding nodig.
6. De departementen moeten tijd beschikbaar hebben voor VBTB naast het reguliere proces.
7. Reeds tijdens het implementatietraject is integratie van financiën en beleid noodzakelijk.
8. De resultaten van beleid moeten kunnen worden gemeten.
9. De departementen moeten over een maatschappelijke antenne beschikken.
10. Het proces moet ook en wat educatief en speels karakter hebben.

Van der Knaap (1999, p. 237)voegt daaraan met betrekking tot het benutten van prestatiegegevens de volgende elementen toe.

1. Heldere verwachtingen aangaande de verschillende typen prestatiegegevens en hun mogelijkheden.
2. De (technische) kwaliteit van prestatiegegevens.

3. De bestuurlijke inbedding van prestatiegegevens (in beleids- en bedrijfsvoeringsprocessen en de daaraan gekoppelde cycli van planning en control).

De Rekenkamer heeft de afgelopen twee jaar de inspanningen die worden gepleegd op het terrein van de VBTB intensief gevolgd en daarover regelmatig gerapporteerd. In de Voortgangsrapportage VBTB nader beschouwd (Algemene Rekenkamer 1999-2000) doet de Rekenkamer verslag van een aantal bevindingen. De indruk was toen dat, ondanks departementale verschillen, er in zijn algemeenheid duidelijk vorderingen zijn gemaakt ten aanzien van de nieuwe begrotingsindeling in beleidsartikelen die aan de VBTB-eisen voldoet, en in het uitwerken van de beleidsartikelen inclusief het formuleren van de bijbehorende beleidsdoelstellingen.

Desondanks concludeert men dat in termen van VBTB (wat willen we bereiken, wat moeten we daarvoor doen en wat mag dat kosten) er in de voorbeeldbegrotingen die ministeries hebben gemaakt, nog een aantal onduidelijkheden zit. Zo is op onderdelen het ambitieniveau niet voldoende expliciet gemaakt. En voor het onderdeel bedrijfsvoering constateert de Rekenkamer dat er sprake is van een significante verlaging van het ambitieniveau.

3.5 Lessen voor de toekomst

Uit het voorgaande valt af te leiden dat bij het rijk een hoog ambitieniveau is geformuleerd wat betreft het concreter maken van beleidsdoelen en dat men dit sterker wil koppelen aan het proces van verantwoording. Het lijkt erop dat er op het punt van het formuleren van beleidsdoelstellingen vorderingen worden gemaakt, maar dat was ten tijde van de departementale doelstellingenanalyse ook het geval. Veel complexer wordt het koppelen van instrumenten aan beleidsdoelstellingen. Eenvoudige effectiviteitsratio's of efficiencymaatstaven kunnen in verantwoordingsstukken worden opgenomen. Het is echter geen uitgemaakte zaak waar dit toe zal leiden. Immers doelmatigheid en doeltreffendheid van beleid zijn aspecten die over het algemeen meeromvattende analyses vereisen. De ervaringen bij de gemeenten, als het gaat om afrekenen op basis van producten, zouden aanleiding moeten zijn tot terughoudendheid. Minder nog dan de gemeenten kent het Rijk organisatieonderdelen die homogene producten leveren. Vaak is de productie zelf uitbesteed of gedecentraliseerd naar agentschappen of lagere overheden, hoewel het ook wel voorkomt dat het Rijk zelf als producent optreedt (aantal paspoortcontroles door de Marechaussee). Maar het heeft geen zin om van de directie Constitutionele Zaken van het Ministerie van BZK een kwantitatief meetbare output te verwachten, en het is bepaald gevaarlijk wanneer men op dat departement het aantal opgeheven gemeenten als output gaat beschouwen. Terecht wordt dan ook gewaarschuwd tegen 'een preoccupatie met het vraagstuk van outputmeting' (Mol 2000).

Het rijk dient zich niet alleen te spiegelen aan de schijnbare brede verspreiding van het instituut productbegroting bij gemeenten, maar zich ook te realiseren dat het meestal om niet meer dan een vorm van window dressing gaat, waarbij achter de façade van de productbegroting de werkelijke budgettaire beslissingen incrementeel van aard zijn en zich richten op de beschikbare middelen. Op de input dus. Ook bij de begrotingsdiscussies op rijksniveau staat de verdeling van de beschikbare middelen centraal, en spelen te bereiken doelstellingen in kwantitatieve zin maar een ondergeschikte rol.

3.6 VBTB en de 5N

In het te ontwerpen evaluatiestramien zal vanwege de rijksbreed omarmde VBTB-procedure rekening gehouden moeten worden met het daarin verwoorde streven te evalueren aan de hand van meetbare doelstellingen. Het formuleren van doelstellingen op een hoog aggregatie-

niveau heeft de charme van de eenvoud. Immers, met een relatief gering aantal doelstellingen kan men transparantie (en dus overzicht) en volledigheid bereiken. De charme van de eenvoud heeft een keerzijde in de zin dat het een trade off heeft met functionaliteit. Dat laatste betreft tenminste twee aspecten. Ten eerste is het op een hoger abstractieniveau steeds onwaarschijnlijker dat zicht wordt verkregen op de effectiviteit van de individueel ingezette instrumenten. Zou het al lukken een gedeelde opvatting te genereren over de mate waarin een doelstelling is bereikt, dan nog blijft het de vraag waaraan dit in evaluatieve zin aan gekoppeld zou moeten worden. Koppeling met het totaal van het instrumentarium lijkt het meest voor de hand liggende. Immers de veronderstelling van de beleidsverantwoordelijken zal zijn dat het totaal van het instrumentarium de effecten bewerkstelligt. Een dergelijke koppeling is in alle gevallen onbevredigend als het gaat om sturing en verantwoording. Het bereiken of nagenoeg bereiken van de doelstellingen geeft geen indicatie welke van de onderliggende instrumenten effectief was, in hoeverre er sprake was van elkaar tegenwerkende instrumenten en zelfs de vraag of de doelstelling niet bereikt zou zijn zonder het instrumentarium. Duidelijk zal zijn dat iedere confrontatie met de kosten van de ingezette middelen ook obsoleet is.

Een VBTB-gedachtelijn toegepast op ruimtelijk beleid kan, als het dan echt moet, slechts gelden in situaties waarin een hoge mate van decompositie heeft plaatsgevonden. Enkelvoudige doelstellingen, een veronderstelde causaliteit tussen één of meer instrumenten en het te beïnvloeden fenomeen, alsmede een afwezigheid van rivaliserende verklaringen zijn noodzakelijke ingrediënten voor evaluaties met de ambitie om verantwoording af te leggen over effectiviteit en efficiëntie en om sturingsveranderingen inhoudelijk te motiveren. Een VBTB gedachtelijn impliceert voor een evaluatiekader voor de 5N:

- Het uitvoeren van evaluatiestudies op het niveau van projecten, het uitvoeren van effectiviteitsstudies van beleidsinstrumenten, onderzoek naar de kosteneffectiviteit van (handhavings)instrumenten, vormen de eerste laag voor oordeelsvorming ten behoeve van bijsturing en verantwoording. Dit is het niveau van de multicriteria-analyses, de scorecards, de kosteneffectiviteitsanalyses e.d.;
- Onderzoeken naar negatieve interferenties van beleidsinstrumenten / projecten vormen een tweede laag. Twee projecten, of twee handhavingstrajecten kunnen individueel (kosten)effectief zijn, maar behoeven dat in combinatie niet meer te zijn;
- Evaluatie van coördinatie tussen de spelers op het terrein van de ruimtelijke ordening (interbestuurlijke afstemming).

Tegenover deze invalshoek staat oordeelsvorming op een hoger aggregatieniveau. Dit betreft het zo systematisch mogelijk in beeld brengen van de in de 5N neergelegde doelstellingen en ambitieniveaus en de ontwikkeling daarvan in de tijd. Het gaat daarbij nadrukkelijk niet meer om een instrumenteel perspectief (in de zin van het bijsturen door de intensiteit van het ruimtelijk instrumentarium te variëren of andere instrumenten in te zetten), maar om monitoring (signalering). Afwijkende ontwikkelingen dienen niet op dit hoge aggregatieniveau tot verandering van inzet van het ruimtelijk instrumentarium te leiden in verband met de bovengesignaleerde causaliteitsproblematiek. Wijzigingen in de inzet van het ruimtelijk instrumentarium moeten aanhaken bij de inzichten uit de specifieke evaluatiestudies op projectniveau en het niveau van beleidsmaatregelen.

4. Doorwerking onderzoek Algemene Rekenkamer

In zekere zin kan de Algemene Rekenkamer worden beschouwd als een geïnstitutionaliseerde leerfunctie van de overheid. Als onafhankelijk Hoger College van Staat controleert de Rekenkamer de rechtmatigheid en doelmatigheid van het beleid en de organisatie van de overheid. Naast het tegengaan van malversaties mag verondersteld worden dat de wetgever met dit instituut beoogt een mechanisme te creëren waardoor het functioneren van de overheid en het vormgeven en uitvoeren van beleid kan verbeteren.

De taakstelling en werkwijze van de Rekenkamer is in de loop van de tijd veranderd (Stevens 1979; Dolmans 1987; Dolmans 1989; De Vries 2000). Die veranderingen hebben van doen met gewijzigde inzichten over doorwerking en effectiviteit van controle en evaluatie. Aanvankelijk beperkte de Rekenkamer zich tot de controle op rechtmatigheid van overheidsuitgaven. Het belangrijkste instrument om 'doorwerking' van haar bevindingen te bevorderen was gelegen in de formele bevoegdheid de Rekening van het Rijk goed te keuren, dan wel er goedkeuring aan te onthouden. Dat laatste gebeurt zelden of nooit, maar ten aanzien van afzonderlijke uitgavenposten zegt de Rekenkamer soms dat ze niet rechtmatig zijn, of, vaker, dat de rechtmatigheid ervan niet kan worden vastgesteld. Mede op instigatie van de Rekenkamer heeft de Operatie Comptabel Bestel plaatsgevonden die met name de laatste categorie in omvang heeft gereduceerd.

Vanaf het begin van de jaren 80 ging de Rekenkamer naast het rechtmatigheidsonderzoek steeds meer doelmatigheidsonderzoek doen in aansluiting op de behoefte om meer inzicht in de effecten en effectiviteit van het beleid te verwerven (Leeuw 1992). Bij sommige doelmatigheidsonderzoeken heeft de Rekenkamer zelf optredende effecten in kaart gebracht en deze naast beleidsdoelstellingen gelegd. Nogal eens deed zich daarbij het probleem voor dat die doelstellingen, in de visie van de Rekenkamer, niet scherp genoeg geformuleerd waren om een oordeel over doelbereiking te kunnen formuleren. Dat leidde dan tot een pleidooi om beleidsdoelstellingen helder en toetsbaar te formuleren. Daarnaast ontwikkelde de Rekenkamer een 'normatiek' die als kader voor de beoordeling van organisatie, informatievoorziening en beheer rondom beleidsvorming en -uitvoering werd gehanteerd, in de veronderstelling dat een positief resultaat op deze punten effectief beleid waarschijnlijk maakt (Bemelmans-Videc 1994). Onderzoek (Van der Meer 1999; De Vries 2000; Van der Meer et al. 2000) laat zien dat de doorwerking bij beide werkwijzen afhangt van de aansluiting tussen analyse van de Rekenkamer en probleem- en werkelijkheidsdefinities van beleidsactoren.

In dezelfde periode ontwikkelde de Rekenkamer expliciet een beleid om het gebruik en de doorwerking van haar onderzoek te bevorderen. Steeds nadrukkelijker ging de verbetering van het functioneren van de overheid deel uitmaken van de missie van de Rekenkamer. In het Verslag 2000 schrijft de Rekenkamer: "Het is de taak van de Rekenkamer om het functioneren van de overheid en de organisaties die met het Rijk verbonden zijn te toetsen. Zo probeert de Rekenkamer bij te dragen aan het verbeteren van de overheid." (Algemene Rekenkamer 2001). Gekoppeld aan dat streven ging de Rekenkamer meer aandacht besteden aan de timing van onderzoek (om de kans te vergroten dat het in erop volgende politieke besluitvorming zou worden benut), kwam het parlement nadrukkelijker als doelgroep in beeld en werden meer rapporten gepubliceerd, zodat ook media-aandacht en publieke discussie zouden worden gestimuleerd. Beide laatstgenoemde policies zijn er onder meer op gericht te voorkomen dat bevindingen van de Rekenkamer te snel in bureauladen verdwijnen. In het hierboven reeds aangehaalde onderzoek (met name De Vries 2000) blijkt dat aandacht van media en parlement inderdaad aan de doorwerking van Rekenkameronderzoek kan bijdragen. Echter, ook in dit geval is de mate en aard van die doorwerking mede afhankelijk van probleemdefinities die, bijv. in parlementaire kringen, bestaan of ontstaan. Daarbij is de doorwerking veelal indirect, geleidelijk en sluipend (vgl. Weiss 1980; zie ook Van der Knaap 1997, die voor onder meer de Europese Rekenkamer tot soortgelijke bevindingen komt).

Verantwoording die in het kader van het VBTB-proces wordt nagestreefd, sluit aan bij de visie van de Algemene Rekenkamer. Zij ondersteunt Financiën en denkt mee over de vormgeving van het proces. Daarnaast is de Rekenkamer doende om criteria voor ‘bruikbaarheid’ van de eigen onderzoeksresultaten te ontwikkelen. Ook wordt meer contact gezocht met departement om een beter beeld te vormen van wat er (voor hen) allemaal speelt. Beide punten kunnen gevolgen hebben voor opzet van onderzoekstrajecten en –rapporten.

5. Ervaringen met evaluaties RO

5.1 Evaluaties van de Vierde Nota RO

In deze paragraaf behandelen we ervaringen met evaluatie van het RO-beleid in het kader van de Vierde Nota RO. Daarvoor besteden we eerst kort aandacht aan de inhoud van de Vierde Nota RO en de sturingsfilosofie die erin terug te vinden is. Vervolgens behandelen we ervaringen met het evalueren van de Vierde Nota. In par. 5.2 wordt aandacht besteed aan het werk van de Werkgroep 5N RO van de Tweede Kamer die ook het beleid in het kader van de Vierde Nota evalueerde.

De Vierde Nota: inhoud en sturingsfilosofie

De Vierde Nota Ruimtelijke Ordening (VINO, 1988) en de Vierde Nota Extra (VINEX, 1992) legden een sterke nadruk op de versterking van de concurrentiepositie van Nederland. De nadruk op de economische ontwikkeling (Nederland Distributieland) werd vertaald met aandacht voor Mainports (Rotterdamse haven en Schiphol) en het stimuleren van bedrijvigheid. Aan de andere kant was er aandacht voor de kwaliteit van het milieu. Zo ontstond een soort dubbele oriëntatie die in veel beleidsonderdelen van de nota's terug te vinden is: nadruk op economische groeipotenties en behoud van leefomgeving.

Deze spanning heeft veel ruimtelijke projecten de afgelopen tien jaar gedomineerd. Vooral de discussie over de mainports- de uitbreiding van Schiphol en de Rotterdamse haven, een beleidsonderdeel dat gaandeweg steeds meer nadruk kreeg, is deze spanning zeer sterk terug te vinden (zie bijv. Klijn and Koppenjan 1999; Van Eeten 1999). Maar de spanning is ook in het beleid van de VINEX-locaties terug te vinden, waar enerzijds wordt gestreefd naar verstevigen van stedelijke omgeving en stedelijk draagvlak en anderzijds naar beperking van de automobiliteit. Een andere hoofdlijn in de RO-nota's is de nadruk op de (her)ontwikkeling van de steden (Nederlands Stedenland) (VROM 1988; VROM 1992; zie ook Teisman 2001; VROM 2001c).

- Stedelijke knooppunten; aanwijzen van knooppunten moet bepaalde steden voorrang geven bij publieke voorzieningen en investeringen voor infrastructuur.
- Het sleutelprojectenbeleid; sleutelprojecten betreffen grootschalige investeringen in stedelijke gebieden. Men zou het ook als een soort concretisering van het knooppunten- beleid kunnen zien.
- VINEX-locaties; concentreren van woningbouw tegen de grote steden. Dit zorgt voor meer draagvlak voor het voorzieningenniveau van deze steden en voor minder autogebruik doordat bewoners dicht bij het werk wonen, zo is de gedachte.
- Het ABC-locatiebeleid; poging om het mobiliteitsprofiel van bedrijven af te stemmen op de ideale vestigingslocatie. Bedrijven met veel werknemers zouden zich moeten vestigen op zogenaamde A-locaties (centrumlocatie goed bereikbaar met openbaar vervoer). Alleen bedrijven met weinig werknemers en met een sterk goederenvervoer karakter zouden zich moeten vestigen op C-locaties (snelweglocaties). B-locaties zitten hiertussenin. De uitvoering van het ABC-locatiebeleid ligt bij gemeenten en zou vooral door strenge parkeernormen gestimuleerd moeten worden.

De sturingsfilosofie van de Vierde Nota's is (vooral wat betreft de VINEX) sterk gericht op uitvoering. Naast een restrictief beleid waarin werd aangegeven waar ontwikkelingen zich moesten voltrekken, poogt de rijksoverheid de uitvoering van het beleid te verzekeren door overleg met lagere overheden. De keuze van locaties voor woningbouw geschiedt door de gemeenten waarna het rijk (lees VROM) uit het aanbod de VINEX-locaties selecteert en ver-

volgens met lagere overheden, vooral op stadsgewestniveau, convenanten afsluit. In deze convenanten is vastgelegd waar gebouwd wordt, maar ook wat (woningdifferentiatie) hoeveel en in welk tempo (fasering). Bovendien zijn de rijksbijdragen vastgelegd. Vervolgens kan de rijksoverheid zich terugtrekken in de rol van toezichhouder en de uitvoering aan gemeenten overlaten. In feite plaatst VROM de eigen beleidsontwikkeling als een soort beleidsbepalende fase voor een uitvoeringsfase van lagere overheden. Deze uitvoeringsgerichtheid geeft het beleid wel duidelijke doelstellingen in termen van kwantiteit (aantal te realiseren woningen, type van die woningen etc.) en kwaliteit (locatiekeuzen, gewenste relatie wonen openbaar vervoer etc.).

De Vierde Nota RO geëvalueerd

De hoeveelheid evaluaties en onderzoeken die te maken hebben met beleidsonderdelen van de Vierde Nota RO is enorm. Veel onderzoek richt zich op deelaspecten van het beleid zoals het functioneren van het ABC-beleid, de convenanten in het kader van de VINEX-gebieden, de Mainports (die gaandeweg een steeds prominentere plaats kregen op de beleidsagenda), of het knooppuntenbeleid. Een omvangrijk deel van de onderzoeken vond plaats binnen het evaluatieprogramma VINO/VINEX wat door de RPD is opgezet. Binnen dat kader maakte de RPD onderscheid tussen doorwerkingsonderzoek, conformiteitsonderzoek en doelbereikingsonderzoek (Herweijer et al. 1990):

- *doorwerkingsonderzoek*: in hoeverre wordt het gedachtegoed uit de rijksnota's opgepakt door verschillende doelgroepen en hoe werkt het door in hun besluiten? Hier gaat het dus vooral om de verspreiding van het gedachtegoed en of het gedachtegoed daadwerkelijk bij andere actoren bekend is en in afwegingen wordt meegenomen. Merk op dat dit niet noodzakelijkerwijs betekent dat actoren ook conform rijksdoelstellingen handelen. Beleid heeft immers wel doorgewerkt indien andere actoren bereideneerd er van afwijken (Mastop and Faludi 1993);
- *conformiteitsonderzoek*: in hoeverre stemmen de inhoud van handelen en besluiten overeen met beleidstoezeggingen en afspraken in rijksnota's? Hierbij gaat het dus om concrete handelingen van andere actoren (maar feitelijk van de rijksactoren zelf zoals verder op zal blijken);
- *doelbereikingsonderzoek*; in hoeverre zijn de geformuleerde gewenste ruimtelijke ontwikkelingen gehaald?

Recentelijk krijgt ook het *doelmatigheidsonderzoek* (in hoeverre is de energie die in bepaalde beleidshandelingen wordt gestoken goed besteed?) steeds meer aandacht.

Bovengenoemde indeling legt vooral nadruk op de focus van het onderzoek. Dus de vraag staat centraal of het onderzoek zicht richt op de doelen van beleid (doelbereikingsonderzoek maar ook het doelmatigheidsonderzoek) of op het proces waarin uitkomsten totstandkomen (doorwerking en conformiteitsonderzoek). In het laatste onderzoek staan vooral gedrag en opvattingen van actoren centraal die eventuele gewenste beleidsuitkomsten realiseren.

Daarnaast kan onderscheid worden gemaakt in het soort van onderzoek. Vrij veel onderzoek wat zicht richt op doelbereiking heeft het karakter van monitoring: het systematische verzamelen van (grote hoeveelheden) gegevens om zo een indruk te krijgen van uitkomsten en effecten van onderzoek. Recente voorbeelden zijn de "Balans Ruimtelijke Kwaliteit" uitgevoerd door de RPD en het onderzoek naar de relatie woningmarkt - grondmarkt. Het laatste onderzoek bevat een inventarisatie van grondprijzen en een redenering welk effect van de grondprijsstijging aan een 'normale' prijsstijging valt toe te rekenen. Op basis daarvan wordt geconstateerd dat er toch nog een aanzienlijke extra waarde overblijft, die dus ten goede komt aan de grondeigenaren. De Balans Ruimtelijke Kwaliteit uitgevoerd door de RPD wil bijdragen aan "een volwaardige plaats voor monitoring in het proces van verkennen, formuleren, uitvoeren en actualiseren/vernieuwen van beleid (de beleidscyclus)" (VROM 1999, p. 7). Het

rapport bevat een grote hoeveelheid gegevens over ruimtegebruik en veranderingen daarin en koppelt deze waar mogelijk aan geformuleerde doelen in het ruimtelijke beleid. Het rapport maakt daarbij onderscheid tussen centrale doelen, strategische doelen en uitvoeringsdoelen. Centrale doelen zijn doelen die ruimtelijk beleid wil meehelpen te bereiken, maar die daar niet alleen door worden bepaald, zoals geleiding van mobiliteit. Strategische doelen zijn ruimtelijke doelen en zij geven aan op welke manier de ruimtelijke ordening een bijdrage wil leveren aan het bereiken van de centrale doelen. Zo kan een bundeling van verstedelijking (strategisch doel) bijdragen aan het versterken van stedelijk draagvlak (centraal doel). Uitvoeringsdoelen zijn concrete doelen die gekoppeld zijn aan beleidsprogramma's waarmee de RPD andere actoren benadert. Een voorbeeld van zo'n strategisch doel is het restrictief beleid dat poogt bepaalde gebieden te vrijwaren van verstedelijking. In de visie van "Balans Ruimtelijke Kwaliteit" draagt dit operationele doel weer bij aan het strategische doel bundeling verstedelijking. Het rapport poogt een eerste inventarisatie te geven van een aantal ruimtelijke ontwikkelingen met betrekking tot de te realiseren doelen.

Aan de andere kant is er meer op verklaring gericht onderzoek. Dit kan zich vooral richten op de doelen van het beleid maar ook op het proces. De twee onderzoeken die in de paragrafen hierna worden besproken - tOETs uitgevoerd door CPB, SCP, RIVM/DLO en RPD ((CPB et al. 2001); zie par. 5.3) en Notie van Ruimte uitgevoerd door de Tweede Kamer ((Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000); zie par. 5.2) - zijn voorbeelden van verklarend onderzoek dat zich vooral richt op de doelen van beleid. Zij pogen zoals het tOETs-onderzoek de kracht te beproeven van bepaalde sturingsconcepties en instrumenten en gebruiken daarvoor beschikbare gegevens over ruimtelijke ontwikkelingen en eerder gedrag van ruimtelijke actoren.

Verklarend onderzoek kan zich ook meer richten op het proces. Een voorbeeld is het onderzoek van de KUN/PDI(UVA) in opdracht van de RPD uit 1995 naar de doorwerking van de beleidstheorie over kwaliteit en handelen van andere ruimtelijke actoren. In tegenstelling tot het tOETs-onderzoek en het rapport Notie van Ruimte poogt het onderzoek inzicht te krijgen in het handelen van actoren en de interacties tussen deze actoren zoals bij het afsluiten van VINEX-convenanten. Opvallend is overigens dat het onderzoek behoorlijk positief is over de interdepartementale coördinatie bij de totstandkoming van de VINEX-convenanten, terwijl later onderzoek (zie bijvoorbeeld par. 5.2) diezelfde interdepartementale coördinatie nog al kritisch oordeelt. Kennelijk is interdepartementale coördinatie bij beleidsvoorbereiding iets anders dan bij de uitvoering van ruimtelijk beleid. Tabel 5.1. vat bovenstaande indeling nog eens samen.

Tabel 5.1. Voorbeelden van Evaluatieonderzoek in de ruimtelijke ordening

	Doelen van beleid	Proces
Monitoring	Systematische gegevensverzameling over doelbereiking (kwantificerend, indicatoren). Voorbeelden: -Balans Ruimtelijke verkenningen (1999) -Relatie Woningmarkt - Grondmarkt (1999-2000)	Systematisch monitoring van activiteiten en inspanningen
Verklarend onderzoek	In kaart brengen obstakels doelbereiking . Voorbeelden: - Notie van Ruimte (2000) - tOETs (2001)	In kaart brengen handelingsperspectieven actoren. Voorbeelden: - Evaluatie verstedelijkingsbeleid (1995)

5.2 Werkgroep 5N

Een ambitieuze en uitgebreide evaluaties van de Vierde Nota is het rapport ‘*Notie van Ruimte*’, uitgevoerd door de Werkgroep Vijfde Nota Ruimtelijke Ordening van de Tweede Kamer. In deze paragraaf wordt aandacht besteed aan:

- de aanpak en centrale bevindingen van het onderzoek;
- een reflectie op de door het evaluatierapport gevonden dynamiek en knelpunten in het ruimtelijk beleid;
- een korte behandeling van de reacties op het rapport en de lessen voor de evaluatie van de 5N.

Aanpak en bevindingen

In het evaluatie rapport “Notie van Ruimte” wordt een overzicht van het VINEX-beleid gegeven en wordt per onderdeel (internationale concurrentiekracht, stedelijk gebied, landelijk gebied, geleiding stad en land) bekeken wat er van het beleid terecht is gekomen. In die zin draagt het onderzoek sterk het karakter van een doelbereikingsonderzoek. Tevens besteedt het aandacht aan instrumenten van de VINEX en hun effectiviteit.

De evaluatie gaat volgens de inleiding (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 9) over:

- De feitelijke wijzigingen in de ruimtelijke inrichting;
- Het procesverloop, de interactie met andere partijen in de ruimtelijke ordening;
- De werking en reikwijdte van instrumenten van nationale overheid.

Voor de evaluatie wordt gebruik gemaakt van bestaande evaluaties (waaronder die van de RPD), beschikbare literatuur en (beperkt) eigen uitgezet onderzoek. Bij het RPD-onderzoek constateert de werkgroep dat de nadruk van het onderzoek vooral lag op doorwerkings- en conformiteitsonderzoek en vrijwel niet op doelbereiking. Gezien de latere opmerkingen van de werkgroep over helderheid en toetsbaarheid van beleidsdoelen lijkt zij dit te betreuren, maar dat is niet expliciet uit het rapport op te maken.

Al in de inleiding spreekt de werkgroep uit wat haar is opgevallen bij de evaluatie. Zij stelt: “De belangrijkste les voor volgende nota’s is dat het geformuleerde ambitieniveau plus de verantwoordelijkheid voor de realisatie van de ambities, duidelijk moeten doorklinken in de uitvoeringsstrategie”. Daaraan voegt zij nog eens toe dat het aan meetbare handvaten voor het beleid ontbreekt. Zij constateert dat naast ex post evaluatie aandacht is voor doelmatigheids-onderzoek en schrijft: “Doelmatigheidsonderzoek is een andere nieuwe vorm: krijgen we waar voor onze energie en ons geld. Deze vraag zal naar verwachting steeds belangrijker worden, zeker nu het monitoringprogramma met de Balans Ruimtelijke Kwaliteit steeds verder vorm krijgt. Dit stelt echter ook eisen van meetbaarheid en evalueerbaarheid aan de formulering van beleid, en helderheid over doelen en gewenste effecten gekoppeld aan deadlines. Ook hieraan heeft het in de VINO en VINEX vaak ontbroken” (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 12).

Het onderzoek is lovend over de enthousiasmerende werking van het VINEX-beleid en de uitvoeringsgerichtheid. De werkgroep stelt zelfs in haar conclusies “de grote uitdaging voor de Vijfde Nota is om hetzelfde enthousiasme los te maken als destijds” om er vervolgens onmiddellijk achteraan te laten volgen “maar dat niet meer door de vingers laten glippen” (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 293). Met die toevoeging wordt ook duidelijk dat de werkgroep kritisch is over wat er daadwerkelijk is bereikt. De stedelijke knooppunten en het ABC-beleid kwamen niet uit de verf, de VINEX-wijken worden wel gerealiseerd, wat de werkgroep gezien de locaties een indrukwekkende prestatie vindt, maar de

koppeling met vervoersbeleid is vrijwel niet totstandgekomen, terwijl er ook twijfels bestaan over de kwaliteit van de wijken.

Knelpunten in het ruimtelijk beleid volgens de werkgroep 5N

In het evaluatierapport Notie van Ruimte zijn enkele centrale gedachten te vinden over mogelijke oorzaken van het op onderdelen mislukken van de Vierde Nota. Deze zijn samen te vatten onder vier belangrijke thema's: onduidelijke concepten, sturingsfilosofie, verkokering en de noodzaak van (financiële) dragers van ruimtelijke ontwikkelingen. Hieronder worden ze systematisch nagelopen.

De eerste rode draad in Notie van Ruimte is de kritiek op de *onhelderheid van de gehanteerde concepten* in het VINEX-beleid. Dit punt komt in vrijwel alle evaluatiehoofdstukken terug en wordt keer op keer aangezet. Vooral als het gaat om de gehanteerde concepten om de stimulering van de economie te bevorderen en te sturen zoals het idee van stedelijke knooppunten, ABC-beleid, maar ook als het gaat om een begrip als ruimtelijke kwaliteit, constateert de werkgroep dat de concepten onhelder zijn, niet gedefinieerd in meetbare termen of te vrijblijvend. Over het begrip ruimtelijke kwaliteit schrijft zij: *'De rijksoverheid ... heeft verzuimd het kwaliteitsbegrip op locatieniveau helder en inzichtelijk te omschrijven. ... Het kwaliteitsbegrip is fragmentarisch en objectieve beleidsgerelateerde criteria ontbreken.'* (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 178). De concepten van stedelijke knooppunten en Mainports zijn volgens de werkgroep neergezet zonder een duidelijk beeld waar dit ruimtelijk toe moet leiden. De basisgedachte van het ABC-beleid heeft communicatieve kracht volgens de werkgroep, maar er wordt geen duidelijkheid geschapen over de juiste match tussen locatie en bedrijf. Kortom, het beleid heeft gefaald omdat de gehanteerde concepten te vaag waren, beleidsmatig niet duidelijk genoeg en niet genoeg in toetsbare criteria gegoten. Deze kritiek is in feite een variant op de oude stelling dat beleid faalt omdat de doelen niet duidelijk genoeg zijn geformuleerd.

Een tweede draad in de evaluatie vormt de beschouwing over de *sturingsfilosofie*. De werkgroep constateert dat deze filosofie vooral communicatief is: het stimuleren van anderen om tot handelen over te gaan. Aan de ene kant vindt de werkgroep dit een vrijblijvende strategie. Het duurt lang voor bestemmingsplannen zijn aangepast, er zit veel in de pijplijn en de RPD en inspecties slagen er onvoldoende in hun stempel op de daadwerkelijke uitvoering te zetten. De werkgroep erkent wel dat er beperkingen zitten aan het sturende vermogen van het rijksniveau. Opvallend is ook dat de werkgroep de convenanten die VROM heeft gesloten met medeoverheden vaak te inflexibel vindt. Bovendien - en hier manifesteren zich belangen van de evaluator - stelt de werkgroep dat de onderhandelingen voor de convenanten zich gedeeltelijk buiten de lijnen van de democratische controle af spelen (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 263). Deze spanning poogt zij op te lossen door in haar aanbevelingen te pleiten ook voor een duidelijk onderscheid tussen ruimtelijk beleid en uitvoeringsstrategie (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 296). Het ruimtelijk beleid moet standvastig zijn, zo betoogt de werkgroep en visievormende ontwikkelingsbeelden kunnen daarin voorzien. Tegelijk moet de uitvoering flexibel zijn. Een werkprogramma met investeringsprojecten dat tweejaarlijks wordt herzien, zou de flexibiliteit moeten verzorgen. Daarnaast zou het rijk een rol moeten spelen in bijzondere opgaven die bestuurlijke grenzen overstijgen.

Een derde draad betreft de opmerkingen over *verkokering*. De werkgroep laat kritische geluiden horen over de afstemming tussen de departementen. Vooral het knooppuntbeleid en de afstemming wonen en verkeer moeten het ontgelden. Over de stedelijke knooppunten wordt opgemerkt: *'De in de PKB aangekondigde prioriteiten voor stedelijke knooppunten betreffen hoofdzakelijk investeringen van andere rijksdiensten, met name bij V&W, het toenmalige*

WVC en binnen het eigen departement DGM. Dit is niet waargemaakt' (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 131). En over de VINEX-wijken stelt de werkgroep dat de geledingen binnen de rijksoverheid er niet in geslaagd zijn normen en werkwijzen af te stemmen zodat openbaar vervoer en stedenbouw gelijk opliepen. Hierdoor is de mobiliteitsreductie in de VINEX-locaties niet gehaald (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 177). De conclusie over verkoking is des te meer interessant omdat een hiervoor gememoreerd onderzoek uit 1995 {Evaluatie van stedelijk beleid, 1995} nog positief was over de departementale samenwerking bij de afsluiting van de VINEX-convenanten, terwijl de werkgroep haar kritiek spuit over de realisatie van VINEX-locaties. Kennelijk lukt het de departementen wel om gezamenlijke uitgangspunten te formuleren in de onderhandelingen met lagere overheden maar niet om die uitgangspunten vervolgens in de eigen lijnorganisaties gerealiseerd te krijgen. Hier botsen dus een procesoriëntatie in de VINEX-onderhandelingen met een projectmatige oriëntatie in de realisatiefase van de locaties zelf en een lijnoriëntatie binnen de betrokken departementen.

Een vierde draad is ten slotte het thema *proces of projectgericht* en de beschouwingen over *draggers van ruimtelijke ontwikkelingen*. De werkgroep constateert dat de sleutelprojecten als voorbeelden van gerichte investeringen, meer hebben bijgedragen dan de knooppunten (Werkgroep Vijfde Nota Ruimtelijke Ordening 1999-2000, p. 132). Zij vindt het positief dat de rijksoverheid zich in verschillende projecten heeft gemanifesteerd als investeringspartner. Dit investeringsperspectief blijkt bij gemeenten beduidend beter aan te slaan dan het perspectief van de verbetering van de dagelijkse leefomgeving. Wel vraagt zij zich af wat er van ruimtelijke projecten terechtkomt als een drager ontbreekt. Zij pleit dan ook voor voldoende financiële middelen om projecten van nationaal belang waar te maken. Procedurele gerichtheid en integraal karakter acht de werkgroep onvoldoende om gewenste ruimtelijke ontwikkelingen te realiseren. De werkgroep tendeeft ernaar om ruimtelijke projecten, inclusief financiële investeringen en vastomlijnd projectmanagement (en projectprocedures) centraal te stellen.

Het rapport van de werkgroep 5N RO leidt tot de volgende aanbevelingen.

- Formuleer heldere concepten en operationaliseer deze in meetbare beleidsdoelen.
- Zorg voor duidelijke ruimtelijke investeringen die beleidsdoelen ondersteunen en die als drijvers voor gewenste ruimtelijke ontwikkelingen kunnen fungeren.
- Zorg voor interdepartementale afstemming van ruimtelijke investeringen.
- Leg algemene visie vast en monitor uitvoering elke twee jaar en pas haar aan aan veranderende omstandigheden en knelpunten en gebreken in de gekozen strategieën.
- Geef het Rijk een speciale positie in grote belangrijke investeringsprojecten die bestaande schaalniveaus overstijgen.

Lessen voor evaluatie van de 5N

De aanbevelingen van de werkgroep 5N roepen vragen op over het evalueren van ruimtelijke ontwikkelingen, die ook als leerervaringen kunnen fungeren voor toekomstige evaluaties:

- De aandacht van de werkgroep voor duidelijke concepten en de op doelbereiking gerichte evaluatie is in lijn met de VBTB.
- Een vraag is of dit streven past bij de opzet en sturingsfilosofie van de 5N.
- De werkgroep pleit voor drijvers van ruimtelijke investeringen. De vraag is evenwel waar projecten vandaan komen. Evaluatieonderzoek in het kader van de 5N moet niet alleen aandacht besteden aan bestaande projecten maar kan zich ook richten op de vraag waarom bepaalde investeringen niet worden voorgesteld.

- De werkgroep constateert tevens dat departementale investeringen beter van de grond komen dan interdepartementale. Dit vraagt om aandacht voor de samenhang en herkomst van ruimtelijke investeringen. Deze aandacht is waarschijnlijk relatief makkelijk in gangbare bestaande evaluatieprocedures van de rijksbegroting te verwerken. Het vereist een overzicht van projecten, financiële benodigdheden en herkomst van de gelden. Het wordt dan misschien duidelijker welke projecten financieel kans van slagen hebben en waar financiële tekorten aanwezig zijn. Tevens kan hiermee zichtbaarder worden welke departementen investeren in bepaalde projecten.
- De werkgroep hecht aan heldere meetbare criteria. Deze maken evaluatie op doelbereiking mogelijk. Aan de andere kant constateert de werkgroep dat de werkelijkheid snel verandert, doelen achterhaald raken en als gevolg van dynamiek in de uitvoering beleid verandert, waarop gereageerd zou moeten kunnen worden. Hoe de spanning tussen beide opgelost kan worden en wat de rol van onderzoek is in het oplossen van die spanning, wordt niet echt duidelijk. Waarschijnlijk is het nodig om in evaluaties van de 5N meer aandacht aan deze spanning te besteden. Te denken valt aan een soort voortschrijdende evaluatie waarin bevindingen worden vertaald in discussiepunten die aanleiding kunnen geven tot het bijstellen van doelen tijdens de rit. Of, anders geformuleerd, als de formuleringen in de 5N worden beschouwd als randvoorwaarden waarbinnen de ontwikkelingen van het ruimtelijk beleid zich zouden moeten afspelen, dan zouden evaluaties tevens aandacht moeten besteden aan de beperkingen voor nieuwe ontwikkelingen die de randvoorwaarden veroorzaken, de onhaalbaarheden als gevolg van dynamische economische en ruimtelijke ontwikkelingen en gewenste nieuwe doelen. Evaluaties zouden zo dus meer gebruikt kunnen worden als expliciete terugkoppelingen naar het oorspronkelijk geformuleerde beleid en een discussie moeten uitlokken over de wenselijkheid van aanpassingen.
- Welke mechanismen kunnen *interdepartementale coördinatie* versterken en tegelijkertijd versterken dat er dragers voor ruimtelijke projecten ontstaan. Vanuit de interessante observatie dat het kennelijk nog wel lukt om de voorbereiding van ruimtelijke visie en projecten te coördineren, maar niet de projectmatige aansturing en de aansturing in de lijn, is een interessante vraag hoe deze afstemming tussen procesmanagement, projectmanagement en lijnmanagement kan worden verbeterd. En welke rol kan (evaluatie)onderzoek en periodieke monitoring in die afstemming spelen. Misschien kan periodieke evaluatie zorgen voor een sterkere terugkoppeling tussen de verschillende vormen van management en een expliciet moment van discussie vormen over spanningen Dit betekent echter ook dat bij toekomstige evaluaties over de 5N de afzonderlijke inspanningen van de verschillende departementen explicieter in beeld gebracht zouden moeten worden.

5.3 Ex ante evaluatie 5N

Inleiding

In juni 1999 heeft het ministerie van VROM aan het CPB, het SCP, het Milieu- en Natuurplanbureau (RIVM/DLO) en het RPD-planbureau opdracht gegeven gezamenlijk een ex ante evaluatie uit te voeren naar de Vijfde Nota Ruimtelijke Ordening (CPB et al. 2001).

Hoewel de 5N zeven criteria voor ruimtelijke kwaliteit introduceert, heeft deze ex ante evaluatie deze criteria niet als toetsingscriteria gebruikt. De reden hiervoor is dat in de 5N nog niet helder is hoe met deze criteria wordt omgesprongen (o.a. slag naar concrete operationalisering is nog uitgebleven). De studiegroep voert een brede en globale analyse uit van een aantal belangrijke concepten in de 5N, te weten (CPB et al. 2001, p. 4):

- kwaliteit in stad en land: een ontwikkelingsbeeld van Nederland in contrasten;

- stedelijke netwerken: een ontwikkelingsbeeld van meerkernige samenhangende gebieden met integrale afstemming van functies in een netwerksamenleving;
- water: een ontwikkelingsbeeld van een duurzaam ingerichte Delta met veerkrachtige en goed beheersbare watersystemen.

Niet alleen wordt de inhoud van de 5N langs bovenstaande meetlat ex ante geëvalueerd, maar ook worden de sturingsfilosofie en instrumenten in beschouwing genomen. De volgende bevindingen worden in het evaluatierapport besproken.

Kwaliteit in stad en land: contouren

De contourenbenadering wordt in de 5N als een nieuwe beleidsinstrument geïntroduceerd. De ex ante evaluatie voorspelt dat het trekken van rode contouren een lastig proces zal worden. Immers, voor gemeenten is een ruim getrokken rode contour van strategisch belang. Het maakt het mogelijk om door middel van uitbreiding van het bebouwd gebied groei te realiseren. Bij het stellen van rode contouren zal er volgens de studiegroep een dilemma'situatie ontstaan. Aan de ene kant zullen ruime rode contouren leiden tot aantasting van 'open ruimte' buiten de rode contour. Ook zal het instrument weinig sturend werken op het intensiveren en transformeren van bestaand stedelijk gebied, een doelstelling die centraal staat in de 5N. Aan de andere kant zullen krappe rode contouren weinig ruimte bieden voor de dynamiek van de stad en zullen leiden tot hogere grondprijzen binnen en vlak buiten de contouren (2001, p.5). Overheden zullen hiertoe snel een verzoek indienen bij de Minister, zo is de verwachting van de studiegroep. De problemen van grondspeculatie die zijn opgetreden bij de aanwijzing van locaties voor verstedelijking in de Vierde Nota Extra zullen naar verwachting met het trekken van rode contouren niet worden opgelost. Immers, de zoeklocaties zijn voor verstedelijking zijn in meer of mindere mate bekend. De evaluatie stelt voor om in dat licht een 'open-ruimteheffing' als aanvullend instrument te hanteren.

De studiegroep stelt verder dat het vaststellen van rode contouren nog geen oplossing biedt voor ruimtelijke coördinatieproblemen. Prikkel bij de verschillende besturen op gemeentelijk, provinciaal en rijksniveau ter beïnvloeding van het hun gedrag zouden een optimale allocatie van de verschillende functies niet vanzelf tot stand brengen. De studiegroep stelt voor om de ruimtelijke afwegingen van individuele en collectieve waarden te maken op een schaalniveau (stadsgewest of kaderwetgebied) waarop (externe) effecten afspelen. De evaluatie spreekt van financiële prikkels om samenwerking op de meest geëigende schaalniveau te bevorderen, bijvoorbeeld door taken te koppelen aan budgetten en samenwerking te belonen.

De ex ante evaluatie meldt (CPB et al. 2001, p. 6) dat de groene contour een effectief en helder instrument is om het beoogde beleidsdoel van behoud van natuur, culturele of landschappelijke waarden van een gebied te bereiken. De studiegroep meent dat dit instrument te verkiezen is boven prijsmaatregelen. Een heffing kan niet voorkomen dat in een gebied op ongewenste locaties wordt gebouwd. Echter, het instrument wordt in de 5N niet helder uitgewerkt als het gaat om het vaststellen van de omvang van de groene contouren: de afweging van achterliggende belangen bij het vaststellen van het zoekgebied voor een groene contour is niet duidelijk geformuleerd in de 5N.

Tenslotte meldt het evaluatierapport dat er nog te weinig richting is gegeven aan het concept balansgebied²⁵. Uitwerking vinden de evaluatoren belangrijk omdat de meerwaarde van rode contouren hier direct mee zou samenhangen.

²⁵ In de PKB deel 3 (VROM 2001a, p.5) wordt het concept balansgebied als overbodig terzijde geschoven, nadat nadere invulling aan het beleid van rode en groene contouren is gegeven.

Beleidsconcept stedelijke netwerken

Ten aanzien van het concept stedelijke netwerken heeft de evaluatiegroep een analyse uitgevoerd. Het rapport komt tot de constatering dat stedelijke netwerken als planologisch concept een beperkte waarde hebben ten opzichte van de huidige stadsgewestelijke aanpak. Op dit moment vindt al afstemming plaats op het niveau van provincies en kaderwetgebieden. De studiegroep acht een stadsgewestoverschrijdende samenwerking aanbevelingswaardig, maar vindt een bredere samenwerking nodeloos complex en minder direct.

Ook achten de evaluatoren de voorgestelde investeringen in het openbaar vervoer niet nodig, omdat het op zichzelf gezien een beperkt effect voor de substitutie van de auto naar het openbaar vervoer tot gevolg zou hebben. Met de investeringen (subsidies voor openbaar vervoer, aangevuld met andere maatregelen als stringente parkeernormen, prijsbeleid, en dergelijke) kan de ruimtelijke interactie binnen de stedelijke netwerken weliswaar toenemen, maar de absolute omvang ervan zou te klein zijn om stedelijke netwerken die het stadsgewestelijk niveau zouden overschrijden als één geheel te laten functioneren. Men verwacht dat het effect ervan op de (internationale) concurrentiepositie van Nederland ook beperkt zal zijn.

Beleidsstrategie voor water

Ten aanzien van het onderwerp water doet de evaluatiegroep een aantal inhoudelijke uitspraken o.a. over de maatschappelijke kosten en baten van ruimte voor water, de bescherming van zoetwatervoorraden en de watertoets. Ten aanzien van functiecombinaties melden de evaluatoren dat de 5N te makkelijk ervan uitgaat dat water altijd te combineren is met andere functies als recreatie, wonen en werken. Zo zou een combinatie met natuur, met name hoogwaardige natuur, alleen mogelijk zijn als er in voldoende mate gunstige randvoorwaarden voor natuur(ontwikkeling) worden gecreëerd.

Sturing in de ruimtelijke ordening en het openbaar bestuur

De regie op het ruimtegebruik komt volgens de 5N te liggen bij het middenbestuur. De verhouding tussen provincie en kaderwetgebied is volgens de evaluatiegroep niet op alle punten even duidelijk (nevenschikking of bovenschikking in relatie tot de toetsende en handhavende taken van de provincie). Daarnaast zou integraal regionaal ruimtelijk beleid worden bemoeilijkt door de grote diversiteit van regionale bestuurlijke organisaties en gedecentraliseerde rijksdiensten. Bovendien zou de rol van de provincies in de Ruimtelijke Ordening in het verleden beperkt zijn door centralisatie en het ontstaan van directe relaties tussen gemeenten en het Rijk. De vraag is over de provincies in staat zijn de beoogde taakverzwaren uit te voeren.

Verder zien de evaluatoren een probleem voor de democratische legitimiteit van het openbaar bestuur. De bestuurlijke ondoorzichtigheid die zou ontstaan door netwerkvorming binnen het openbaar bestuur, met wisselende betrokkenheid van alle denkbare lagen van EU tot deelraad, en die tussen bestuurslichamen en private partijen, zou daartoe bijdragen. Duidelijker zou moeten worden gemaakt wie op welk moment verantwoordelijk is voor welk beleid, zo vinden de evaluatoren.

In tegenstelling tot de verticale sturing is de horizontale sturing minder uitgewerkt in de 5N. De evaluatoren zijn van mening dat niet mag worden gerekend op automatische horizontale doorwerking van het ruimtelijk beleid. Dit is een complex sturingsvraagstuk, waar meer aandacht aan moet worden besteed (o.a. in de PKB).

Lessen voor het opstellen van een evaluatiearrangement voor de 5N

De volgende lessen kunnen uit de ex ante evaluatie tOETs worden gehaald:

- evaluatie van globale concepten: de ex ante evaluatie neemt afstand van evaluatie aan de hand van concrete en gedetailleerde criteria (de 5N noemt er vijf). Dit maakt het mogelijk om enkele concepten in brede zin te evalueren en verschillende andere onderwerpen (die ermee samenhangen) mee te nemen in de evaluatie, zoals bijvoorbeeld de relatie tussen het trekken van rode contouren, grondspeculatie en strategisch gedrag van gemeenten;
- inhoud en sturing: de ex ante evaluatie richt zich niet alleen op inhoudelijke evaluatie van bepaalde concepten, maar ook op de sturingsfilosofie die hiermee samenhangt; daarnaast geeft de evaluatie ook handreikingen voor veranderingen en daarmee dus niet alleen beschrijvend, beoordelend, maar ook prescriptief;
- terugkoppeling tussen- en eindresultaten van de evaluatie: de tussenrapportages zijn in een stuurgroep besproken, maar hierin hebben alleen vertegenwoordigers van de uitvoerders van de evaluatie zitting. Een stuurgroep aangevuld met vertegenwoordigers van doelgroepen maakt het mogelijk vroegtijdig te leren van de evaluatie (fourth generation) en doorwerking te faciliteren.
- er is waardering voor een vroegtijdige aandacht voor evaluatie in de uitvoering van de 5N. Dit maakt evaluatie tot deel van beleidsontwikkeling, niet enkel tot sluitstuk.

5.4 Koepelplan Noord-Brabant

In februari 2001 heeft het College van Gedeputeerde Staten van Noord-Brabant het ontwerp gepubliceerd van het Koepelplan Reconstructie Zandgronden, *Reconstructie aan zet* (Beckers and Haarmann 2001). In de provincie Noord-Brabant vindt een debat plaats over de gewenste inrichting van een vitaal platteland en de wijze waarop dit kan worden gerealiseerd. In Noord-Brabant wordt deze discussie gevoerd ten aanzien van het reconstructiegebied dat Oost- en Midden-Brabant tot Breda omvat.

In het koepelplan geeft de provincie aan wat zij de komende twaalf jaar van de zeven regionale reconstructiecommissies verwacht en wat de provincie zelf zal doen om het platteland van Oost- en Midden-Brabant klaar te maken voor de 21e eeuw. Het koepelplan bevat de opdrachten en kaders waarbinnen de commissies moeten werken.

Onder aansturing van het provinciebestuur zal er in Noord-Brabant één koepelplan voor alle regio's tezamen komen, op basis waarvan vervolgens in elke regio door tussenkomst van een regionale uitwerkingscommissie nadere accenten zullen worden gelegd. Binnen de kaders zoals vastgelegd in de uitwerkingsplannen, zal ten slotte op het lokale niveau daar waar wenselijk een uitvoeringscommissie de verdere detaillering vormgeven en met een plan van aanpak voor de feitelijke uitvoering komen.

De provincie en haar partners kiezen voor een procesmatige aanpak van de Reconstructie. Gaandeweg de planvorming en -uitvoering kunnen nieuwe inzichten, beleidsontwikkelingen en ontwikkelingen in de externe omstandigheden leiden tot aanpassingen in de taakstelling voor de reconstructiecommissie, die op hun beurt kunnen leiden tot aanpassing van het reconstructieplan en/of daarop gebaseerde uitwerkingsplannen.

De Provinciale Commissie Landelijk Gebied (PCLG) heeft aan Telos, het Brabants Centrum voor Duurzaamheidvraagstukken, gevraagd om een onafhankelijk oordeel en advies te geven over het proces van het Ontwerp-Koepelplan, als bijdrage aan duurzame ontwikkeling. We beschrijven hieronder de gevolgde werkwijze, de bevindingen en de reacties van actoren. Ten

slotte identificeren we punten in deze casus die relevant zijn voor vormgeving aan evaluatie in het kader van de 5N.

Aanpak en werkwijze

Telos heeft als evaluator gebruik gemaakt van twee methoden. Eerst heeft zij interviews gehouden met diverse stakeholders die betrokken zijn bij de reconstructie van de zandgrondbieden, op lokaal, regionaal en rijksniveau. In de tweede plaats heeft zij externe deskundigen (wetenschappers, adviseurs) aangeschreven om hun visie en advies te geven met betrekking tot het ontwerp-Koepelplan. De externen zijn deskundig op terreinen als ruimtelijke ordening, planologie, bestuurskunde, beleid, en strategie.

Op een werkconferentie 'Van denken naar doen' (29 maart 2001) geven de externe deskundigen hun visies op het ontwerp-Koepelplan. Het doel van de werkconferentie is om na te gaan of de aanpak en sturingsfilosofie, beschreven in het Ontwerp-Koepelplan, voldoende garanties bieden voor een gedragen en evenwichtige ruimtelijke implementatie op termijn. Voor de werkconferentie zijn uitgenodigd de leden van de PCLG, voorzitters van de regionale reconstructiecommissies, leden van de provinciale projectgroep en gebiedscoördinatoren. De externe deskundigen spraken hun visies uit over drie dilemma's.

Dilemma in de probleemopvatting: *transformatieproces of RO-project?* Wordt het op te lossen probleem beschouwd als een ingrijpend integrale transformatieproces, of als een groot ruimtelijk ordeningsproject? Als het reconstructievraagstuk als een transformatieproces zal worden opgevat, betekent dit een integrale aanpak waarin verschillende problemen (economisch, fysiek en sociaal) moeten worden opgelost. Uiteenlopende beleidsdomeinen zullen daarbij betrokken zijn, met als gevolg dat veel actoren betrokken zullen zijn met verschillende belangen, en dientengevolge veel verschillende visies op wat de problemen en de oplossingen zijn. Dit kan leiden tot onoverzichtelijkheid en legt een grote nadruk op de procesaanpak. Wordt de reconstructie van de zandgronden daarentegen opgevat als een groot RO-project, dan ligt het wat eenvoudiger: het probleem is helder, immers het is een ruimtelijk probleem, en er zijn vastomlijnde procedures voor ontwikkeld. Niet zozeer een procesaanpak, maar een procedurele aanpak ligt voor de hand.

Dilemma in het proces: *open of gesloten procesgang?* Wordt het proces opgevat als een open proces met veel vrijheid voor de betrokkenen voor eigen invulling, of als een gesloten proces waarin rollen van betrokkenen, doelen en de verwezenlijking daarvan al van te voren bekend worden verondersteld? Hier staat de keuze centraal hoe het probleem moet worden aangepakt. In een open proces worden alle betrokken actoren uitgenodigd hun visie op het probleem en de oplossingen te articuleren. Daarna wordt gestart met het gemeenschappelijk zoeken naar oplossingen. De overheid is hierbij niet meer de enige partij die oplossingen implementeert. Ook andere organisaties (bedrijven, maatschappelijke organisaties) nemen hun verantwoordelijkheid en investeren in de oplossingen. Een gesloten proces daarentegen vertrekt veel meer vanuit heldere taakstellingen, beoogde resultaten, de wegen waarlangs resultaten moeten worden behaald en betreft actoren selectief. De overheid staat boven de partijen.

Dilemma in de uitwerking: *zekerheid versus flexibiliteit.* Wordt er een grote nadruk gelegd op (planning van) wettelijk vereiste procedures of wordt zo veel mogelijk flexibiliteit in het proces ingebouwd? Hier staat de keuze centraal hoeveel flexibiliteit wordt ingebouwd bij de aanpak van het beleidsvraagstuk. In een aanpak die stuurt op "zekerheid" spelen de procedures een hoofdrol. Er wordt een scherpe planning voor de procedures gehanteerd en de procedures zitten al vroeg in het proces. Deze aanpak is echter heel wat minder "zeker" dan die op het eerste gezicht lijkt. De grote druk op de procedures organiseert veel tegendruk en weerstand en kan gemakkelijk leiden tot organiseren van veel "hindermacht". Een flexibele aanpak laat

meer ruimte voor maatwerk, nieuwe ideeën en inbreng van andere actoren. De wettelijk vereiste procedures komen bij deze aanpak achter in het proces en zijn dan een formele bevestiging geworden van de afspraken die alle betrokkenen met elkaar hebben gemaakt. Nadeel hiervan lijkt een langere voorbereidingstijd en op het eerste oog geringere beheersbaarheid.

Bevindingen

Dilemma 1

Uit de visies van de externe deskundigen en de daarop volgende discussie tussen stakeholders komt naar voren dat het proces het karakter moet hebben van een transformatieproces voor de lange termijn, met integrale duurzame ontwikkeling als inzet. Bij een transformatieproces richting duurzame ontwikkeling gaat het om fundamentele vraagstukken, waarbij traditionele oplossingen vaak niet meer werken. De ontwikkeling en uitvoering van plannen in het kader van de reconstructie is dan ook een creatief en innovatief proces, gebaseerd op een gemeenschappelijke doelstelling en onderling vertrouwen, met behulp van een decentrale aanpak.

Dilemma 2

De evaluatie over dilemma 2 brengt naar voren dat de nadruk in het Koepelplan op de beleidsinhoud ligt en dit ten koste gaat van de ontvankelijkheid voor alternatieve probleemdefinities en alternatieve oplossingen. Het proces dient dan ook meer voorzien te worden van ‘receptoren’, die deelnemers gevoelig maken voor onverwachte en onvoorspelbare nieuwe mogelijkheden, feiten en ontwikkelingen.

De wijze waarop de plannen thans zijn gepresenteerd versterkt de indruk van een top-down-benadering, met een beslissende rol voor de (provinciale) overheid. Door de gedetailleerde uitwerking lijkt het alsof het geheel dichtgetimmerd is. Of er feitelijk voldoende beleidsruimte voor het regionale maatwerk is, zal moeten blijken. De hoge dynamiek en de grote complexiteit vragen echter om een ander sturingsconcept: van hiërarchische sturing naar netwerksturing, op basis van coproductie. Een dergelijk op vertrouwen gebaseerd veranderingsproces veronderstelt zowel verantwoordelijkheid als aansprakelijkheid. Daarbij dient duidelijk te zijn wat vast (niet-onderhandelbaar) en wat variabel (veranderbaar) is.

Dilemma 3

Uit de evaluatie over dilemma 3 komt naar voren dat, hoewel het Koepelplan voorziet in een voor zeven reconstructiegebieden vergelijkbare inhoud en aanpak, de gebieden onderling verschillen naar (voor)geschiedenis, aard van de problematiek en mogelijkheden voor oplossingen. Daardoor zullen de resultaten en termijnen verschillend zijn. Dit vraagt om differentiatie, maatwerk en aangepaste procedures. Er bestaat echter een spanning tussen het voldoen aan (inter)nationale regels, beleidsnota's en afspraken enerzijds en het scheppen van ruimte voor maatwerk, nieuwe regels en onorthodoxe oplossingen anderzijds.

Verder gaat de evaluatie in op de verhouding tussen de betrokkenheid, de machtsmiddelen en het kennisniveau op het provinciale niveau en op het regionale en lokale niveau en bij andere belanghebbende partijen. Deze is asymmetrisch en onevenwichtig. Extra aandacht verdient ‘het primaat van de politiek’. Dat geldt zowel het provinciale als het lokale niveau. Formeel beslissen provinciale staten over de uitkomsten van de planvorming en de gemeenteraden over tal van uitvoeringsmaatregelen. Hun betrokkenheid in de loop van het proces behoeft nadere uitwerking in het Koepelplan. Daarnaast moet er in de aanpak veel meer aandacht komen voor de stille kracht in het transformatieproces: de burger-consument, in de rol van kiezer, bewoner, vrijwilliger, recreant, werknemer, verkeersdeelnemer en koper van voedsel. De aanpak is nu nog te sterk aanbodgestuurd.

Naast de kritische beoordelingen zijn er ook aanbevelingen in de evaluatie opgenomen. Ze worden hieronder kort besproken.

Aanbeveling 1: integraliteit

Geadviseerd wordt om de voorgestelde RO-aanpak sterker te koppelen aan de sociaal-economische ontwikkeling van de regio's (transformatieproces). Op regionaal niveau zouden de reconstructiecommissie aangevuld moeten worden met vertegenwoordigers uit andere sectoren. Dat betekent overigens niet dat men continu met iedereen rond de tafel moet gaan zitten. Een heldere fasering, taakverdeling en afstemming voorkomt verzanding en Babylonische spraakverwarring.

Aanbeveling 2: maatwerk

Regionale verschillen in uitgangspositie en aard van de problematiek vereisen maatwerk in de aanpak en uitkomsten. Het Europees, nationaal en provinciaalbeleid stellen echter in sterke mate eisen aan de regionale ontwikkelingen. De mate waarin de regio's ruimte krijgen voor maatwerk wordt verschillend ingeschat. Uit de gedetailleerde inventarisatie van het bestaande beleid in het ontwerp koepelplan, ontstaat het beeld dat er sprake is van een keurslijf. In het vervolg van het proces dient de ruimte voor maatwerk een continu aandachtspunt te zijn. De provincie kan als intermediaire bestuurslaag de rol van 'voorvechter en bewaker van de beleidsruimte' op zich nemen door er bij de hogere overheden (EU en Rijk), maar ook bij zichzelf, op toe te zien dat voldoende beleidsruimte geboden te worden.

Aanbeveling 3: open- en geslotenheid

Het proces rondom de ontwikkeling van de reconstructiegebieden is redelijk dichtgetimmerd. Geadviseerd wordt om vooraf duidelijk aan te geven welke fasen open en welke gesloten zijn. Het gaat er niet om dat zo veel mogelijk een open proces wordt georganiseerd, maar dat openheid en geslotenheid elkaar afwisselen. De mate van openheid kan per onderwerp, situatie of ronde in het proces verschillend ingevuld worden. De momenten van openheid/geslotenheid moeten op interactieve wijze worden ingevuld. In gezamenlijk overleg met stakeholders moet adequaat met het dilemma open of gesloten procesgang worden omgegaan.

Aanbeveling 4: protocol

In de koepelplan wordt gekozen voor coproductie van beleid. Dat vereist andere spelregels dan voor traditionele beleidsontwikkeling gebruikt worden. Geadviseerd wordt om hierover vooraf helderheid te scheppen door de spelregels vast te leggen in een protocol. Houdt daarbij rekening met:

- de primaat van de politiek (meebetalen geeft gemeenteraden invloed);
- de rollen;
- de bevoegdheden;
- de verantwoordelijkheden;
- de aansprakelijkheid.

Aanbeveling 5: organisatie

Het gevaar bestaat dat de centrale projectorganisatie te ver van de gebieden af komt te staan. Dit kan het leveren van maatwerk in de weg staan. Geadviseerd wordt om het zwaartepunt naar regionaal niveau te verplaatsen door op regionaal niveau ondersteunende projectorganisatie in te richten. Deze organisatie zou het plan- en uitvoeringsproces moeten ondersteunen. Voorwaarde voor een goed functioneren van deze regionale eenheden is dat ze hun antennes (receptoren) open hebben staan voor externe ontwikkelingen (bijv. nationale en Europese regelgeving en sociaal-economische ontwikkeling).

Aanbeveling 6: flexibiliteit

De aanpak om eerst gebiedsgericht streefbeelden of visies te formuleren, vervolgens uitvoeringsprogramma's en daarna te zoeken naar de noodzakelijke middelen, verdient aanbeveling. Het gebruik van het instrument van gebiedscontracten voor de uitvoering is een interessante optie om invulling te geven aan de gemeenschappelijke verantwoordelijkheid van de participanten in het reconstructieproces.

De reconstructiecommissies zouden continu moeten volgen in hoeverre het streefbeeld en het daarvan afgeleide uitvoeringsprogramma, bereikt wordt. Er dient continu aandacht te zijn voor de bijstelling van het streefbeeld en uitvoeringsprogramma op basis van nieuwe inzichten.

Reacties van actoren, doorwerking

De evaluatie van het ontwerp-Koepelplan is binnen de provinciale organisatie overwegend goed ontvangen. Vroegtijdige betrokkenheid van verschillende ambtelijke afdelingen, provinciale bestuurders en provinciale politici hebben ertoe geleid dat hun visies, percepties en belangen in het evaluatierapport konden worden opgenomen. Zo is intern breed draagvlak ontstaan.

Hoe de reacties bij de lokale stakeholders is geweest, is onduidelijk. Het rapport is met name voor interne doeleinden geschreven. Het feit dat externe stakeholders niet betrokken waren bij de uitvoering van de evaluatie toont dit aan.

Lessen voor evaluatie 5N

Puntsgewijs kunnen de volgende lessen worden getrokken voor de evaluatie van de 5N:

1. bij de evaluatie worden zowel procesmatige als inhoudelijke punten over de reconstructie van de zandgronden in de provincie Noord-Brabant tegen het licht gehouden: een mooie combinatie van proces- en productevaluatie;
2. bij de uitvoering van de evaluatie maakt Telos (opdrachtnemer van de evaluatie) gebruik van kennis van zowel externe en interne deskundigen als van de PCLG die een oordeel over het ontwerp-koepelplan moeten uitspreken. Door het organiseren van de werkconferentie wordt aan stakeholders de gelegenheid gegeven meningen, beelden en percepties over het koepelplan te uiten, die vervolgens meegenomen worden in het evaluatierapport. Wel is opvallend dat de werkconferentie vooral een interne aangelegenheid is; lokale stakeholders, zoals gemeenten, bedrijven, maatschappelijke groepen, e.d. zijn niet uitgenodigd;
3. interessant is dat in de adviezen gepleit wordt voor een koppeling van een traditionele RO-aanpak (gericht op beheersing van ruimtelijke activiteiten) aan een aanpak van ruimtelijke ontwikkeling (gericht op transformatie van gebieden);
4. ten slotte is interessant dat flexibiliteit een belangrijk aspect in het advies is. Er moet zo veel mogelijk flexibiliteit in het ruimtelijke proces worden ingebouwd om in te kunnen spelen op met lokale ontwikkelingen. Taak- en doelstellingen moeten niet te rigide worden benaderd: in overleg moeten taakstellingen op decentraal niveau kunnen worden aangepast naar nieuwe inzichten.

5.5 Waardevolle cultuurlandschappen

Inleiding

In 1993 heeft LNV een elftal 'Waardevolle CultuurLandschappen' (WCL's) aangewezen (tezamen bijna 500.000 ha, of 1/7 van het Nederlandse grondoppervlak) en daarvoor verschillende doelen geformuleerd:

- sectoraal (bijv. voor 'landbouw': duurzame en concurrerende landbouw en voor 'natuur': stabiele en beheersbare ecologische hoofdstructuur en ontwikkeling van algemene natuurwaarden);
- integraal (bijv. vergroting van de toegankelijkheid en belevings- en gebruiksmogelijkheden van natuur en landschap voor recreanten);
- facet (bijv. terugdringen milieubelasting);
- proces (bijv. provincie als regisseur).

Het beleid moest bottom-up, onder regie van de provincie gestalte krijgen. Het WCL-beleid in de periode 1994-'99 is geëvalueerd door Alterra. We bespreken de werkwijze, bevindingen en de reactie van de Staatssecretaris in haar brief aan de Kamer op de evaluatie.

Aanpak evaluatie

De evaluatie gaat ervan uit dat de effectiviteit van het beleid niet rechtstreeks kan worden vastgesteld, ten eerste omdat meetbare doelen ontbreken en ten tweede omdat optredende ontwikkelingen veelal niet eenduidig aan het gevoerde beleid kunnen worden toegeschreven. Gekozen wordt derhalve voor een 'monitoring'-benadering waarin triangulatie wordt toegepast: effecten-, uitvoerings-, middelen-, doel- en procesevaluatie. In concreto zoeken de evaluatoren een 'totaalbeeld' dat is opgebouwd uit oordelen over organisatie, planvorming, mobilisatieproces, realisatie en (perspectieven voor) continuïteit. Voor elk van deze variabelen is een aantal 'meetbare indicatoren' vastgesteld. De waarden voor die indicatoren (voor 'organisatie' is dat bijv. de vraag of er (tijdig) een (actief) projectbureau is opgezet; voor 'realisatie' onder meer het aantal slaapplekken voor toeristen) zijn vastgesteld aan de hand van beschikbare gegevensbestanden en - voor zover dat niet mogelijk was - door de beelden van een breed spectrum van 'experts' en 'betrokkenen' te inventariseren. Ook hebben er groeps gesprekken van experts en betrokkenen plaatsgevonden onder meer over de vraag in hoeverre effecten aan het WCL-beleid kunnen worden toegeschreven. Een en ander leidt tot tabellen waarin per indicator, per variabelen en ten slotte per WCL scores worden toegekend: 'goed', 'redelijk' of 'onvoldoende'.

Bevindingen en doorwerking

Vervolgens wordt het WCL-beleid als instrument beoordeeld. Dit levert suggesties op voor aanpak en politiek-inhoudelijk beleid in de toekomst. Het is niet helemaal duidelijk hoe de evaluatoren aan die aanbevelingen komen. Ongetwijfeld zijn zij bij hun verkenningen in het veld op allerlei knelpunten gestuit. Maar daarnaast lijken er ook bepaalde politieke keuzes (bijv. dat het Europese beleid eerder dan het Nederlandse het referentiekader zou moeten zijn), organisatorische inzichten (bijv. dat WCL's niet te klein en niet te groot moeten zijn - hoewel deze aanbeveling geen verband heeft met de beoordelingen waar de evaluatoren zelf toe gekomen zijn) en opvattingen over sturing en evaluatie (de aanbeveling om beter meetbare doelen te formuleren) een rol te spelen. De Staatssecretaris neemt onder meer de aanbevelingen van meetbare doelen ('in de geest van deze tijd') en die m.b.t. de grootte van WCL's over.

Betekenis voor evaluatie 5N

- Het WCL-beleid en de evaluatie daarvan is interessant omdat het qua sturingsmodel verwantschap vertoont met dat van de 5N.
- De spanning tussen meetbare doelen en een strategisch beleid waarin bottom-up en in co-productie invulling tot stand moet komen, is ook in dit geval zeer duidelijk. Evaluatoren en Staatssecretaris kiezen voor scherper geformuleerde doelen, maar doen tegelijkertijd niets af aan het gekozen sturingsmodel. De vraag is welke 'beleidstheorie' daaronder zit en door wie die gedragen wordt. Interessante vraag is ook wie, gezien het sturingsmodel, die meer meetbare doelen zou moeten formuleren.
- Er is natuurlijk niets mis mee als evaluatoren in een rapport allerlei prikkelende aanbevelingen - waar ze ook vandaan komen - opnemen. Toch is de herkomst wellicht van belang om gefundeerd te kunnen bepalen wat de beleidsimplicaties ervan zouden kunnen of moeten zijn.
- De beschikbare stukken geven geen inzicht in wat de evaluatie voor betrokkenen op provinciaal en WCL-niveau betekent en wat op die niveaus derhalve de (leer)effecten zullen en kunnen zijn. Interactie met actoren over de bevindingen zou hieromtrent meer inzicht kunnen geven.

6. Ervaringen met evaluatie en onderzoek in andere domeinen

6.1 Gasboringen in de Waddenzee

Inleiding

Het boren naar gas in de Waddenzee is een taai beleidsvraagstuk waar beleidsmakers, politici, natuur- en milieuorganisaties en de Nederlandse AardolieMaatschappij (NAM) al jaren hun hoofd over buigen. De NAM is van mening dat gaswinning geen schade voor natuur- en milieuvrienden met zich mee hoeft te brengen. De NAM wil dan ook graag zo snel mogelijk het zogenaamde Waddengas winnen. Natuur- en milieuorganisaties vinden gaswinning onacceptabel. Zij vrezen dat onherstelbare schade aan het ecosysteem zal optreden.

De problematiek rond gasboringen in de Waddenzee wordt bepaald door de spanningsvolle relatie tussen natuur- en milieuaspecten en economische aspecten. Dit komt tot uitdrukking in enerzijds de hoofddoelstelling (in de PKB bepaald) van een 'duurzame bescherming en ontwikkeling van de Waddenzee als natuurgebied', en anderzijds de eeuwigdurende concessieverlening aan de oliemaatschappijen NAM, Mobil en Elf Petroland.

Er heerst nogal wat onduidelijkheid over de potentiële milieueffecten van gasboringen in de Waddenzee. In 1995 is in het Plan van Aanpak besloten een grootschalig onderzoek, de Integrale Bodemdalingstudie Waddenzee (IBW), uit te voeren naar de relaties tussen winning van gas en effect op natuur en milieu. De NAM zou dit onderzoek organiseren en financieren.

Aanpak en werkwijze van het onderzoek

Doel van de studie

De Integrale Bodemdalingstudie Waddenzee (IBW) (NAM 1998) heeft als doel inzicht te verkrijgen in de cumulatieve effecten van gaswinning op de Waddenzee. Het gaat daarbij om bodemdalingseffecten ten gevolge van zowel winning uit nieuwe als bestaande gasvelden. Bovendien wordt gekeken naar de relatieve bodemdaling als gevolg van zeespiegelstijging.

Uitvoering van de studie

De NAM heeft een groot aantal wetenschappelijke instanties ingehuurd om de integrale bodemdalingstudie te laten uitvoeren. Het betreft het Instituut voor Bos- en Natuuronderzoek (IBN-DLO), het Nederlands Instituut voor Onderzoek der Zee (NIOZ), de Universiteit Utrecht, het Waterloopkundig Laboratorium (WL), Alkion, de Rijksuniversiteit Groningen en Grondmechanica Delft. De onderzoekers die hebben meegewerkt aan de studie, staan bekend als dé deskundigen op de relevante onderzoeksgebieden. De deskundigen hebben in de studie in sterke mate gebruik gemaakt van kwantitatieve voorspellingsmodellen voor de periode van 2000-2050.

Onafhankelijke begeleidingscommissie

Een onafhankelijke begeleidingscommissie is in het leven geroepen om de inhoud en de kwaliteit van de studie te bewaken. De commissie bestond uit vertegenwoordigers van de NAM, de meest betrokken departementen (ministerie van Economische Zaken, ministerie van Landbouw, Natuurbeheer en Visserij, ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer en ministerie van Verkeer en Waterstaat), de provincies Friesland en Groningen, de eilandgemeenten en de kustgemeenten. De voorzitter en secretaris waren afkomstig

van de Directie Noord-Nederland van Rijkswaterstaat. De NAM was met twee afgevaardigden vertegenwoordigd in de begeleidingscommissie.

Natuur- en milieuorganisaties hebben geen deel uitgemaakt van de begeleidingscommissie. Voor de rijksoverheid was dit een bewuste keuze. Deze is van mening dat alleen partijen die direct betrokken zijn, moeten worden opgenomen in een dergelijke commissie. Andere partijen krijgen later nog voldoende inspraakmogelijkheden. De NAM heeft overwogen om de Waddenvereniging te benaderen, maar durfde het uiteindelijk toch niet aan haar te betrekken. Enerzijds was men met de Waddenvereniging in een juridische strijd verwickeld. Anderzijds had de Waddenvereniging in het verleden geweigerd mee te werken. De NAM hield rekening met de mogelijkheid dat de Waddenvereniging ook ditmaal niet zou willen meewerken. De NAM of de ministeries hebben echter nooit aan de Waddenvereniging of andere 'groene' partijen gevraagd of zij wilden deelnemen aan de begeleidingscommissie.

Scenario's

Zoals we al eerder hebben vermeld, wordt het vraagstuk van gasboringen in de Waddenzee in sterke mate gekenmerkt door onzekerheid over de te verwachten effecten van bodemdaling op natuur en milieu van de Waddenzee. Ook ten aanzien van de daadwerkelijk aan te treffen gasvoorraad in de Waddenzee heerst grote onzekerheid. Uiteindelijk bepaalt de aan te treffen 'gasbel' de bodemdaling: hoe meer gas er in de bodem zit en wordt gewonnen, des te meer zal de bodem dalen en effect hebben op natuur en milieu. Door proefboringen kan de onzekerheid rond de gasvoorraad worden gereduceerd. De NAM heeft echter geen toestemming gekregen om proefboringen uit te voeren, omdat dit het Wad te veel schade zou bezorgen. Om de onzekerheid van de aan te treffen gasvoorraad (en de bodemdaling) beheersbaar te maken, hebben de wetenschappers daarom gebruik gemaakt van twee scenario's. Het eerste scenario is de 'maximum case', waarbij is uitgegaan dat de verwachte gasvoorraad (prospect) volledig wordt aangetroffen. Het tweede scenario, de 'base case', wordt het meest waarschijnlijke scenario geacht. In dat laatste geval wordt veertig tot vijftig procent van de 'prospects' in de bodem aangetroffen.

Werkwijze

Hoewel de NAM opdrachtgever is, hebben de onderzoekers zelf de onderzoeksvragen geformuleerd. Wel is een deel van de basisgegevens aangeleverd door de NAM. Het betreft hier gegevens over de hoeveelheid gas die men in de bodem verwacht aan te treffen. In 1997 heeft het ministerie van Economische Zaken TNO-NITG (Nederlands Instituut voor Toegepaste Geowetenschappen) een schatting laten doen over de te verwachten gasvoorraden. TNO-NITG kwam toen met conservatiever cijfers: er zou minder gas zitten. Dit zou betekenen dat de effecten van boringen op de bodemdaling beperkter van omvang zouden zijn.

In de IBW zijn de effecten van bodemdaling op de geomorfologie, kwelders, bodemfauna en vogels in kaart gebracht. Voor de effecten op de geomorfologie zijn vier modellen ingezet. Er zijn juist vier modellen gehanteerd om de resultaten te kunnen controleren en de robuustheid ervan te kunnen achterhalen. De gedachte is dat naarmate er ongeveer dezelfde resultaten uit de modellen tevoorschijn komen, de robuustheid van die resultaten toeneemt.

Voor een kwantificering van de effecten van bodemdaling op ongewervelde bodemdieren zijn modelstudies uitgevoerd (NAM 1998). Ook voor het in beeld brengen van de effecten op vogels zijn modelberekeningen toegepast. Deze gehanteerde modellen zijn zeer specialistisch en doorgaans alleen begrijpelijk voor deskundigen op het betreffende terrein.

Bevindingen

Ten aanzien van de geomorfologie wordt geconcludeerd dat voor de maximum case van bestaande en nieuwe velden in de periode 2000-2050 de omvang van de bodemdaling zodanig

is, dat onder de huidige zeespiegelstijging in een aantal decennia vrijwel volledige compensatie door sedimentatie van zand uit de Noordzeekustzone en slib mag worden verwacht.

De algemene conclusie van de IBW is dat gaswinning geen noemenswaardige negatieve effecten op de bodemdaling en natuur en milieu heeft. Men spreekt van tijdelijke en verwaarloosbaar kleine (geen meetbare) effecten. Bovendien zijn de effecten in verhouding tot de zeespiegelstijging kleinschalig. De effecten van de bodemdaling op natuur en milieu vallen naar verwachting binnen de natuurlijke variatie. Het concept-eindrapport is voorgelegd aan vier hoogleraren. Zij concludeerden dat het onderzoek zorgvuldig is uitgevoerd en ondersteunen de conclusies.

Reacties van actoren, doorwerking

Natuur- en milieuorganisaties en vele leden van de Tweede Kamer hebben de uitkomsten van de studie met grote scepsis ontvangen. De Waddenvereniging heeft bij het kabinet aangedrongen op een onafhankelijke contra-expertise alvorens Paars II een beslissing zou nemen over de voortzetting van mijnbouwactiviteiten in het Waddengebied. In Nederland zou, volgens de Waddenvereniging, zo langzamerhand geen onafhankelijke instantie of wetenschapper meer voorhanden zijn die nog onbevooroordeeld en in alle rust onderzoek kan doen. Dat het onderzoek naar bodemdaling, dat in 1999 is uitgevoerd, onder zeer strikte geheimhouding is uitgevoerd, maakt vele natuur- en milieugroepen in Nederland wantrouwig. Ze hebben het idee dat er iets gebeurd is, dat het daglicht niet kan verdragen.

Ook het feit dat de NAM een aantal basisgegevens heeft aangeleverd, waarmee de onafhankelijke onderzoekers vervolgens aan de slag zijn gegaan, heeft niet bijgedragen tot een intersubjectivering van het onderzoek. Deze basisgegevens zouden volgens milieuorganisaties te subjectief van aard zijn. Volgens de Waddenvereniging maakt het ontbreken van een onafhankelijke controle op gegevens die door de NAM zijn aangeleverd de kwaliteit van het NAM-onderzoek twijfelachtig. Ook betwijfelt zij of er een goede basis is voor een zorgvuldige besluitvorming door het kabinet. De onafhankelijke begeleidingscommissie vindt dat dit wel het geval is.

De gehanteerde (berekennings)modellen vinden natuur- en milieuorganisaties onvoldoende transparant en begrijpelijk voor een niet-deskundige. Ook de vertegenwoordigers van de NAM en de leden van de begeleidingscommissie hebben weinig inzicht in de werking van de modellen. De aannames en vertrekpunten van de modellen zijn onvoldoende inzichtelijk. Hierdoor wordt het onduidelijk hoe tot bepaalde onderzoeksresultaten is gekomen. Het gevolg is dat de motivatie en argumentatie van de conclusies in twijfel worden getrokken.

Natuur- en milieuorganisaties zijn van mening dat door eerst de ‘maximum case’ te presenteren en vervolgens de ‘base case’, de indruk wordt gewekt dat het allemaal heel erg meevalt met de effecten van bodemdaling op natuur en milieu. Immers, zo is de teneur in het rapport, de prospects worden toch niet volledig waargemaakt. De onderzoekers en de leden van de begeleidingscommissie zijn het hier niet mee eens. Ze menen echter dat de ‘maximum case’ noodzakelijk was met het oog op onzekerheden bij het aantreffen van gas in de bodem van de Waddenzee. Op deze manier wordt rekening gehouden met het ‘worst case’-scenario.

Reflectie

Er kan een aantal verklaringen worden gegeven voor de aangetroffen dynamiek in het onderzoek naar de natuur- en milieueffecten van gasboringen in de Waddenzee.

Gebrek aan gezag

De voorspellingsmodellen naar de bodemdaling van de Waddenzee die in de IBW worden gebruikt, zijn erg ingewikkeld en daardoor alleen begrijpelijk voor een klein aantal deskundigen (binnen een bepaald wetenschappelijk vakgebied). Doordat de berekeningen niet door een ieder kunnen worden gecontroleerd, trekken bepaalde partijen, waaronder natuur- en milieuorganisaties, de uitkomsten van de studie in twijfel. Belanghebbenden stellen bepaalde aannames, uitgangspunten en betooglijnen ter discussie. Het gezag van de wetenschappelijke onderzoekers wordt door natuur- en milieuorganisaties openlijk aangevochten als de resultaten van de studie bekend worden. Ze zouden, volgens natuur- en milieuorganisaties, naar de mond van de NAM praten.

Belangenverstrengeling

De schijn van partijdigheid wordt onherroepelijk gewekt als de IBW niet alleen wordt uitgevoerd door (of beter gezegd onder leiding van) de NAM, maar de NAM bovendien de externe wetenschappelijke experts een aantal basisgegevens aanlevert. De NAM levert als 'informatie-monopolist' de (voornamelijk seismologische) gegevens over bodemstructuur en verwachte gasvoorraden aan.

Buitensluiten andersdenkenden

Inhoudelijke sturing door de NAM heeft er bovendien toe geleid dat waarden, perspectieven en belangen van andere partijen (natuur- en milieuorganisaties) niet of onvoldoende zijn meegenomen in de opzet en uitvoering van het onderzoek. Er wordt bewust voor gekozen om andersdenkenden over gasboringen in de Waddenzee zoals de Waddenvereniging of diverse natuur- en milieuorganisaties buiten de begeleidingscommissie te houden. Hiermee wordt de groep betrokkenen klein gehouden en kunnen bepaalde basiswaarden, belangen en perspectieven overgeaccentueerd worden.

Er is onvoldoende gezamenlijke, interactieve bepaling van aannames en uitgangspunten voor het onderzoek: door eenzijdige bepaling van aannames en uitgangspunten (door overheid en NAM en niet gezamenlijk met natuur- en milieuorganisaties en/of leden van de Tweede Kamer) worden de resultaten van het onderzoek niet geaccepteerd. Aannames bepalen immers in sterke mate de uitkomsten van onderzoek. Het kader van de onderzoeksvragen van de IBW is door de NAM geformuleerd, de begeleidingscommissie heeft uitgebreid gediscussieerd over de uitgangspunten van de studie. Daarmee hebben zij aangegeven wat wel is meegenomen in de studie en wat niet. Wanneer een andere partij dit kader zou hebben opgesteld of eraan zouden hebben meegewerkt, zou het er waarschijnlijk anders uit hebben gezien.

Juiste antwoord, verkeerde informatievraag

Voor een aantal partijen die betrokken zijn bij de discussie over gasboringen in de Waddenzee is de centrale vraag of gaswinning kan geschieden zonder blijvende schade toe te brengen aan het ecosysteem. Voor andere partijen (zoals natuur- en milieuorganisaties) is deze vraag niet relevant; voor hen gaat het om de vraag of een activiteit als gaswinning – die altijd enige verstoring met zich meebrengt - überhaupt wel thuis hoort in een natuurgebied als de Waddenzee. Aangezien de IBW slechts antwoord geeft op de eerste vraag, worden de uitkomsten van de studie niet geaccepteerd als zijnde toereikende informatie om tot een afgewogen besluit te komen. De IBW komt niet tegemoet aan alle informatiebehoeften die leven onder belanghebbenden. Hierdoor wordt een op zich 'juist antwoord gegeven op een verkeerde (of andere) informatievraag'.

Lessen voor evaluatie 5N

Puntsgewijs kunnen uit de casus 'gasboringen in de Waddenzee' de volgende lessen worden afgeleid voor de organisatie van de evaluatie van de 5N.

- Vroegtijdig betrekken van stakeholders (ook met andere percepties en belangen) in de opzet van de evaluatie: zorg voor gemeenschappelijk opdrachtgeverschap waarbij gezamenlijk bepaald wordt welke vragen moeten worden beantwoord, welke methodologie moet worden gehanteerd, welke aannames moeten worden genomen, et cetera.
- Zorg ervoor dat evaluatoren voldoende gezag genieten onder stakeholders, zodat het niet kan voorkomen dat uitkomsten van evaluaties op voorhand niet worden gedragen.
- Zorg voor een begeleidingscommissie van de evaluatie waarin alle belanghebbende zitting hebben: zorg voor een gezamenlijke feitenvindingsproces door bijvoorbeeld (tussentijdse) onderzoeksresultaten met belanghebbenden te bespreken zodat kritische reflectie wordt georganiseerd en kansen worden vergroot dat resultaten van de evaluaties internaliseren en doorwerken.

6.2 Kennisimpasse: onderzoek als middel in de strijd

Inleiding

In de klassieke visie op (evaluatie)onderzoek leidt onderzoek tot meer informatie over de aard van een probleem en dus tot een betere basis voor gezamenlijk handelen. Uit het literatuuroverzicht van multi-actorevaluatie is al gebleken dat deze voorstelling van zaken te naïef is. Vanuit de aanvaarding dat kennis sociaal geconstrueerd is, kunnen tegenstrijdige opvattingen bestaan over de aard van de kennis, gehanteerde methoden en uitkomsten etc. Kortom iedere actor interpreteert het beschikbare onderzoek vanuit zijn eigen perceptie.

Dat iedere actor onderzoek gebruikt vanuit zijn eigen percepties is geen nieuw gegeven (Lindblom and Cohen 1979). Maar het betekent ook dat onderzoek een voorzetting kan worden van de belangenstrijd tussen actoren met andere middelen. Onderzoek heeft dan geen informatie- of enlightenment-functie, maar is munitie in de strijd tussen actoren. De casus van de emissie van zink uit bouwmaterialen vormt een illustratie van een dergelijk proces {deze sterk gereduceerde beschrijving is gebaseerd op Klijn/Koppenjan/Van Bueren 2000}. Het laat bovendien zien hoezeer (evaluatie)onderzoek verbonden is met het beleidsproces waarin de actoren zich bevinden.

bevindingen en reacties van actoren

Halverwege de jaren tachtig werden diffuse zinkemissies een probleem. Zink verscheen op de lijst van prioritair stoffen, die bijzondere aandacht verdienen omdat de normen voor deze stoffen sterk worden overschreden. Voor al deze stoffen dient een basisdocument te worden opgesteld, een soort state of the art-rapport over de omvang van het probleem, de oorzaken en mogelijke oplossingen. Het Basisdocument Zink werd opgesteld door het RIVM en verscheen in 1992. Het rapport bracht weinig nieuws en de zinkindustrie zag aanvankelijk dan ook geen probleem. Mede omdat de handhaving van de normen tot dan toe nog geen grote prioriteit had. Dit veranderde echter toen zinken bouwmaterialen in het kader van het Duurzaam-Bouwenbeleid (DUBO) werd ontraden in een handleiding van de Stuurgroep Experimenten Volkshuisvesting (SEV).²⁶ De waterschappen juichen deze ontwikkeling toe, aangezien zij te maken hebben met overschrijding van de normeringen voor zink en met de zuiveringskosten

²⁶ De basis voor een dergelijke status wordt gevormd door een LevensCyclusAnalyse (LCA). Deze bepaalt de milieubelasting van een product in vergelijking tot andere producten. Uitkomst van zo'n LCA is nogal afhankelijk van de gehanteerde aannames en de meegenomen effecten. Zo ligt de nadruk meestal sterk op kwantificeerbare milieueffecten en worden sommige milieueffecten (zoals de winning) niet meegenomen. Dit maakt dat er nog al wat kritiek is op deze LCA's.

als gevolg van de overschrijding van de normen. De zinkproducenten zijn echter zeer ongelukkig met deze ontwikkelingen. Door de DUBO-lijst ontstaat druk op gemeenten om af te zien van het gebruik van zinken bouwmaterialen. Bovendien is het slecht voor de afzet en imago van het product.

Protesten bij het ministerie van VROM tegen dit beleid leverden niets op, omdat VROM dit tot het beleidsterrein van provinciale en lokale overheden rekende. Ook geeft de minister van VROM aan dat preventie, het voorkomen van emissies, een belangrijk uitgangspunt is van het waterkwaliteitsbeleid en dat dit dus gevolgen kan hebben voor productieprocessen en materiaalkeuzen.

Na een mislukte juridische procedure tegen de SEV-lijsten ging de industrie met het ministerie van VROM in discussie over het onderwerp dat wel tot de competentie van dit ministerie behoorde: de normstelling. Het strijdtoneel wordt nu verlegd naar het kennisfront: de vaststellingen van de normen, de onderbouwing ervan en het gedane onderzoek ernaar. In 1995 bracht de zinkindustrie een Addendum uit op dit Basisdocument, waarin zij ageerde tegen de onderzoeksmethode van het RIVM waarop de normen zijn gebaseerd. Deze methode hield geen rekening met verschillen in zinkconcentraties die van nature in het milieu voorkomen: deze verschillen van gebied tot gebied. Ook hield de gehanteerde methode geen rekening met de mogelijkheid van soorten om zich aan te passen aan veranderende zinkconcentraties. In het Addendum stellen zij een andere methode voor, die wel met beide aspecten rekening houdt. De zinkindustrie vecht niet alleen de normen voor de schadelijkheid van de concentratie aan (en de onderzoeksmethode waarop die zijn gebaseerd) maar ook de gegevens over de hoeveelheid zink die daadwerkelijk van bouwmaterialen afspoelt naar het oppervlaktewater. Deze emissies, zo stelt de industrie, worden veel te hoog gesteld.

Het RIVM, verantwoordelijk voor de normstelling, is niet onder de indruk. De minister van VROM vraagt hierop de Gezondheidsraad om advies. Het is dan inmiddels december 1995. Het advies van de Gezondheidsraad verschijnt in februari 1998. In de tussentijd breekt echter een ware rapportenoorlog los over aard, inhoud en ernst van het zinkprobleem. Vrijwel alle partijen laten zelf onderzoek doen, of huren specialisten in om andere onderzoeksmethoden aan te bevelen of bestaande methoden gebruikt door andere partijen te bekritisieren

De rapporten worden geproduceerd door de zinkindustrie, die zich inmiddels georganiseerd heeft in de Stichting Zink en Milieu, en de onderzoeksinstituten van VROM, het RIVM, en van V&W, het RIZA. Ook diverse onderzoeksbureaus, zoals TNO, Tauw en DHV, produceren rapporten in opdracht van één van deze partijen. In de rapporten herhalen partijen de eigen standpunten en wordt weinig nieuwe informatie wordt toegevoegd. Wel worden de schattingen van de omvang van de diffuse zinkemissies naar beneden bijgesteld: deze blijken niet zo hoog als in het Basisdocument is verondersteld. De sfeer is grimmig en wordt gekenmerkt door wederzijds wantrouwen, wat alleen maar toeneemt. Onderzoek van het andere kamp wordt bij voorbaat als partijdig verworpen.

Begin 1998 verschijnt dan eindelijk het langverwachte advies van de Gezondheidsraad. Deze spreekt een Salomonsoordeel uit: er blijken geen juiste methoden te bestaan om milieukwaliteitsnormen op te baseren. Wel vindt de Gezondheidsraad dat de overheid uit voorzorg normen mag stellen en beleid mag voeren om emissies te reduceren. Het kabinet zegt hierop toe om in overleg met het bedrijfsleven maatregelen te nemen om diffuse emissies te reduceren. In overleg met de Stichting Duurzaam Bouwmetaal, de opvolger van de Stichting Zink en Milieu, wordt een twee-sporentraject afgesproken: een normstellingstraject en een productinnovatietraject. Dit levert enige dooi op, alhoewel deze slechts langzaam intreedt.

Maar de meeste vooruitgang wordt geboekt in de schattingen van het aandeel van zink uit bouwmaterialen in de totale emissie van zink in het milieu. RIZA en industrie hebben in gezamenlijk opdrachtgeverschap een onderzoek laten uitvoeren naar de diffuse zinkemissies. Eind 1999 komt dit rapport uit. De schatting van 105 ton per jaar blijkt niet alleen veel lager dan in het Basisdocument was verondersteld (4125 ton per jaar) maar ook beduidend lager dan die van verscheidene andere latere onderzoeken. Het gezamenlijk opdrachtgeverschap

leidt ook tot meer interactie tussen de betrokken partijen en een voorzichtig verlaten van de eigen stellingen.

Reflectie: onderzoek als strijdmiddel

De zink-casus laat zien hoe onderzoek onderdeel kan worden van de strijd tussen partijen. Op dat moment verdwijnt de waarde van het onderzoek vrijwel geheel. Actoren zijn niet langer bereid om ander onderzoek dan het eigen onderzoek te accepteren en beschouwen ander onderzoek als partijdig. Onderzoek vervult in dat geval noch een functie in het evalueren van doelbereiking, noch in het bevorderen van leren. De casus geeft aanleiding tot enkele interessante observaties: de niet-aansluitende argumentatiestructuren, de (gepercipieerde) partijdigheid van het onderzoek en een mogelijke oplossing in een dergelijke situatie.

Onderzoek en niet-aansluitende argumentatiestructuren

De keuze van de zinkindustrie de gehanteerde normeringsmethodiek aan te vechten leidt tot nauwelijks op elkaar aansluitende argumentatiestructuren. Voor V&W en het RIZA zijn normen maatstreepjes op weg naar het uiteindelijke doel: zo laag mogelijke emissies om zo schoon mogelijke water te bereiken. Ook voor VROM is de inzet emissiereductie. De suggestie dat de methodiek om deze normen vast te stellen de toets der wetenschappelijke kritiek niet kan doorstaan, raakt voor hen niet de kern van het vraagstuk. Zij zijn er immers van overtuigd dat het cruciale issue is dat de emissies te hoog zijn. Deze discongruentie tussen argumenten verklaart ten dele de inhoudelijke impasse: partijen praten langs elkaar heen. De zinkindustrie hanteert een wetenschappelijk-inhoudelijke rationaliteit om haar gelijk te halen; de overheidspartijen redeneren vanuit een politiek-bestuurlijke noodzaak maatregelen te nemen, ook als de wetenschap nog niet zo ver is. Er is sprake van een dialoog tussen doven (Van Eeten 1999). Dit wordt versterkt door het feit dat er in het overleg tussen zinkindustrie en departementen geen mensen aan tafel zitten die wel over de 'wetenschappelijke inhoud' van het vraagstuk gaan, het RIVM en RIZA. Als de Gezondheidsraad met zijn rapport komt, blijkt het bedrijfsleven met zijn inhoudelijke strategie zijn doelen niet kunnen bereiken. Gelijk hebben zonder het te krijgen wordt het. Het streven van de overheid naar emissiereducties wordt legitiem geacht, ook als de wetenschappelijke normering wetenschappelijk niet waterdicht is.

'Partijdige' wetenschap: de sociale impasse

De inhoudelijk-wetenschappelijke strategie van het bedrijfsleven leidt tot een sociale impasse. Deze strategie is gebaseerd op de gedachte dat objectief wetenschappelijk onderzoek deze strijdige claims van betrokken partijen zou kunnen beslechten. Maar deze gedachte negeert het inzicht dat de werkelijkheid zich moeilijk laat kennen en onderzoeksresultaten sterk worden bepaald door gehanteerde aannames en methodieken. Het gevolg is dat onderzoek kennisconflicten in dergelijke situaties eerder versterkt dan oplost. Onderzoek komt beschikbaar in een arena waarin betrokken partijen vanuit belangen naar onderzoeksresultaten kijken en geen wetenschappelijke maar strategische en beleidsmatige criteria hanteren.

Bovendien gebeurt er nog iets anders. Door de keuze de wetenschappelijke methode van de kennisinstellingen RIVM en RIZA aan te vechten, pretendeert de zinkindustrie beter dan deze instellingen te weten wat de stand van de wetenschap is en bedreigt zij de core business van deze instellingen. Een dergelijke claim is strijdig met de dominante perceptie bij de gesprekpartners over betrouwbaar en onafhankelijk wetenschappelijk onderzoek. Dergelijk onderzoek wordt door 'onafhankelijke' overheidsinstellingen verricht en niet door het bedrijfsleven ingehuurd consultants. Dat deze onafhankelijke overheidsinstellingen ook een broodheer hebben, en in Nederland wel een zeer dominante 'marktpositie' hebben, doet aan deze overtuiging niet af.

Lessen voor evaluatie: gedeeld opdrachtgeverschap

Het probleem van partijdig onderzoek wordt in deze casus pas opgelost op het moment dat procesmanagement en procesdesign wordt toegepast. Pas door afspraken te maken over gezamenlijk opdrachtgeverschap en hoe met resultaten om te gaan (procesontwerp) wordt de impasse doorbroken. Naast gezamenlijk opdrachtgeverschap kan ook worden gedacht aan toetsing van resultaten door onafhankelijk forum en afspraken om (nieuwe) meningsverschillen tot onderzoekspunten te maken zoals geschiedde door de Gezondheidsraad in te schakelen. Al deze stappen leiden eerder tot een situatie waarin onderzoek fungeert als katalysator voor verdere interactie en zoeken naar overeenstemming.

Aandachtspunten voor evaluatie 5N

- de casus wijst op het gevaar dat onderzoek uitsluitend nog een rol vervult in de ‘strijd om beleid’. In dat geval verliest het onderzoek zowel zijn waarde voor het evalueren van doelbereiking als voor het vergroten van leren
- de casus onderstreept het belang van cognitieve en sociale aansluiting van (evaluatie)onderzoek bij de perspectieven van actoren
- de rol en positie van onderzoekers/evaluatoren in het beleidsnetwerk vereist daarbij specifieke aandacht
- voor een goed gebruik in de zin dat onderzoek leidt tot verdere handelingen en discussies tussen actoren is een zorgvuldige regie nodig, zeker bij duidelijke gearticuleerde belangen(tegenstellingen)
- gemeenschappelijk opdrachtgeverschap kan daarbij een interessant arrangement zijn

6.3 ISV

Introductie ISV

Achtergronden

Het Stedelijke Vernieuwingsbeleid is een reactie op maatschappelijke en beleidsmatige ontwikkelingen die zich de afgelopen twee decennia hebben voltrokken (VROM 1997). Gedacht kan worden aan problematiek als selectieve migratie, economische verschuivingen en veranderingen op de woningmarkt. Het ruimtelijk beleid gericht op deze problematiek heeft zich ontwikkeld van stadsvernieuwing naar stedelijke vernieuwing, en dit in toenemende mate binnen de context van het Grote-Stedenbeleid.

Het startpunt van het ISV ligt eigenlijk bij de ontwikkeling van het Grote-Stedenbeleid. De problematiek in de vier grootste steden was op een gegeven moment zo urgent dat deze steden hiervoor expliciet aandacht hebben gevraagd bij de rijksoverheid. Dit resulteerde in 1994 in het Deltaplan voor de grote steden. In dit ‘plan’ pleitten de vier grote steden voor een bestuurlijk contract tussen het kabinet en de grote steden. Het eerste paarse kabinet nam dit ‘Grote-Stedenbeleid’ (GSB) op in het regeerakkoord en in 1995 werd onder de naam ‘Een keer ten goede’, een convenant gesloten tussen het ministerie van Binnenlandse Zaken en de vier grote steden.

Dit convenant concentreerde zich op vijf thema’s: werk, onderwijs, veiligheid, zorg en leefbaarheid. Vooral op wijk- en buurtniveau zou op een vernieuwende manier moeten worden samengewerkt met bewoners, het onderwijs, de woningcorporaties, de politie, de ondernemers, het openbaar ministerie en het welzijnswerk. De samenwerking met marktpartijen, zoals winkeliers, bedrijven, vastgoedbeheerders, makelaars, ontwikkelingsmaatschappijen en beleggers werd in die tijd minder sterk belicht.

Op bestuurlijk vlak kende het convenant een drietal uitgangspunten (Priemus 1996, p. 171):

- het terugdringen van verkokering bij het rijk en het bevorderen van verdere decentralisatie;
- het versterken van de regiefunctie van het stedelijk bestuur ter bevordering van de samenhang tussen (gedecentraliseerde) beleidssectoren;
- het door herschikking vergroten van de financiële armslag van de steden.

Inmiddels is het Grote-Stedenbeleid volwassen geworden. Aan de vier grote steden zijn 21 middelgrote steden toegevoegd en er is een aparte minister van Grote-Steden- en Integratiebeleid (GSI) aangesteld. De vijf thema's 'werk', 'onderwijs', 'veiligheid', 'zorg' en 'leefbaarheid' zijn omgevormd tot een sociale component, een economische component en een fysiek-ruimtelijke component. Het terugdringen van verkokering en het bevorderen van samenhang tussen beleidssectoren is omgevormd tot 'integraliteit'. Deze integraliteit binnen het GSB komt tot uitdrukking in een stadsvisie (meerjaren ontwikkelingsprogramma (MOP)) die de (25) grootste steden in Nederland moeten opstellen. Vanuit deze stadsvisie wordt de totale opgave binnen de sociale, de economische en de fysieke kolom in beeld gebracht. Vervolgens wordt er voor elke kolom beleid geformuleerd in de vorm van een meerjarig gebiedsgericht programma. Zo'n programma kan beschouwd worden als een plan van aanpak. Uiteraard moeten de programma's van de verschillende kolommen op elkaar zijn afgestemd.

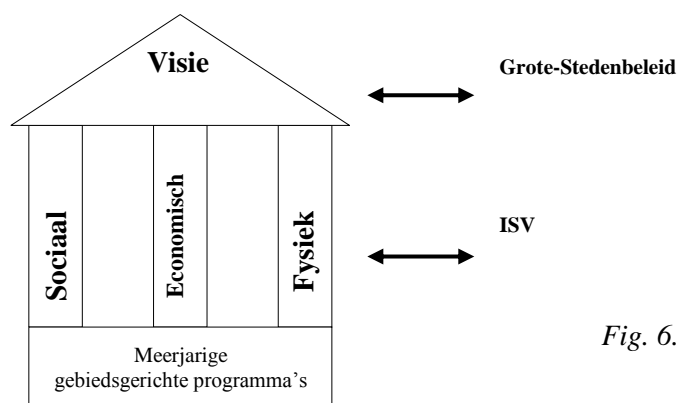


Fig. 6.1: Relatie tussen ISV en GSB.

Bron: ministerie
van VROM

Voor de fysieke opgave moet dus ook een programma worden geschreven en dit programma heeft betrekking op stedelijke vernieuwing. Op grond van dit programma komen de steden in aanmerking voor het Investeringsbudget Stedelijke Vernieuwing (ISV). Daar moet bij worden opgemerkt dat het ISV ook geldt voor gemeenten die niet onder het GSB vallen. Deze gemeenten hoeven geen sociaal en economisch programma te schrijven maar moeten, als zij daartoe worden aangewezen, wel een programma voor de fysieke kolom schrijven (programmagemeenten). Als een gemeente niet is aangewezen om een programma te schrijven, kan zij wel projecten indienen waarvoor geld uit het ISV verkregen kan worden (projectgemeenten).

Het ISV-budget: samenstelling, verdeling en beheer

Het ISV-budget wordt toegekend op grond van het ontwikkelingsprogramma. In het budget zijn subsidies samengebracht die voorheen via afzonderlijke regelingen werden uitgekeerd. Het gaat om middelen voor stadsvernieuwing, herstructurering van vooral naoorlogse wijken, bodemsanering, geluidswerende voorzieningen, historische woonomgeving en grondkosten van VINEX-locaties. In het regeerakkoord zijn extra middelen beschikbaar gesteld voor de vitaliteit van de steden. Deze middelen zijn onder meer bestemd voor lokale milieuhinder, stadseconomie en grootschalig groen. Ook zij zijn opgenomen in het ISV-budget. Al deze

subsidies spelen echter alleen een rol bij de ‘voeding’ van het budget. Vanaf de ‘verlening’ is het ISV-budget één integraal budget zonder schotten.

Het ISV-budget wordt vastgelegd voor een periode²⁷ van vijf jaar (het ISV-investerings-tijdvak). Dit ‘indicatieve’ budget wordt vooraf bekend gemaakt om gemeenten in staat te stellen een programma te ontwikkelen dat kan rekenen op commitment van corporaties en private partijen.

De verdeling van de budgetten over de gemeenten is gebaseerd op de gemeentelijke behoefte aan rijksbijdragen. Deze behoefte is bepaald aan de hand van een aantal verdeelindicatoren. Deze verdeelindicatoren zijn weer afgeleid van de oorspronkelijke behoefteramingen van de verschillende ‘voedingsbronnen’ van het ISV: stadsvernieuwingsgeld, herstructureringsgeld, gelden voor bodemsanering, gelden voor grootschalig groen etc. Deze verdeling wordt de rijksverdeelsleutel genoemd.

Van het totale ISV-budget wordt 10% gereserveerd voor andere budgetten. Ten eerste is dat het zogenaamde innovatiebudget. Dat budget is bestemd voor gemeenten die gedurende het ISV-tijdvak (5 jaar) met uitzonderlijke of innovatieve plannen komen (IPSV). Het tweede budget is bestemd voor incidentele knelpunten (knelpuntenpot) die in het eerste ISV-tijdvak optreden en voor bijzondere problemen die te maken hebben met de overgang van de ‘oude’ regelingen naar het ISV.

Het ISV-budget wordt voor de rechtstreekse gemeenten beheerd door het rijk en voor de niet-rechtstreekse gemeenten door de provincie. De rechtstreekse gemeenten zijn alle GSB-gemeenten aangevuld met de ‘aanleungemeenten’ Alkmaar, Amersfoort, Emmen, Lelystad en Zaanstad. Deze gemeenten zijn allemaal programmameenten. Bij de niet-rechtstreekse gemeenten kan een onderscheid worden gemaakt tussen programmameenten en projectgemeenten. Welke niet-rechtstreekse gemeenten programmameente zijn, wordt door de provincie bepaald. De overige niet-rechtstreekse gemeenten in de provincie zijn dan automatisch projectgemeente.

Rechtstreekse gemeenten	Niet-rechtstreekse gemeenten
Rijk beheert budget	Provincie beheert budget
Allemaal programmameenten	Programmameenten Projectgemeenten
Alle 25 GSB-gemeenten + de vijf aanleungemeenten / G30 nieuwe stijl	Worden aangewezen Overige gemeenten zijn vrij projecten in te dienen

Zoals aangegeven is het ‘indicatieve’ budget bekend, voordat gemeenten hun ontwikkelingsprogramma opstellen. Het budget wordt verleend nadat het rijk respectievelijk de provincie een positief oordeel hebben uitgesproken over het gemeentelijk ontwikkelingsprogramma. Aan het eind van het vijfjarig tijdvak moeten de gemeenten vervolgens een verantwoordingsverslag indienen. Als de prestaties van de gemeente voldoende zijn geweest, kan zij daarna weer in aanmerking komen voor een subsidie voor het volgende tijdvak.

Wet Stedelijke Vernieuwing en Beleidskader Stedelijke Vernieuwing

Om voor toekenning van het investeringsbudget in aanmerking te komen, stelt het rijk als voorwaarde dat een gemeente een meerjarig ontwikkelingsprogramma opstelt waarin zij haar beleid voor de fysieke omgeving uiteenzet. Bij de GSB-gemeenten is dit programma dus een onderdeel van het MOP. Belangrijk is dat de gemeente haar beleid voor de fysieke omgeving weloverwogen opstelt in samenhang met economisch en sociaal beleid. Ook is het belangrijk dat de private sector intensief bij de ontwikkeling en uitvoering van programma's en plannen

²⁷ De eerste periode is van start gegaan in 2000 en duurt tot 2005. Alle rechtstreekse gemeenten hebben hun programma voor 1 januari 2000 moeten inleveren. Voor alle niet-rechtsreekse gemeenten was 1 juli de uiterste inleverdatum.

is betrokken en dat die sector in financiële zin zijn steentje bijdraagt. Meer formeel worden daarom aan het ontwikkelingsprogramma twee soorten voorwaarden gesteld. De eerste soort is in principe duurzaam en is daarom in de Wet Stedelijke Vernieuwing²⁸ opgenomen. De tweede soort is meer tijdgebonden en wordt daarom vastgelegd in de algemene maatregel van bestuur: het Beleidskader Stedelijke Vernieuwing (VROM et al. 1999).

Het ontwikkelingsprogramma is niet alleen een subsidievoorwaarde. Het is er vooral voor de gemeenten zelf: als basis voor de communicatie naar burgers en investeerders en als leidraad voor het dagelijks beleid. Daarom zijn in het wetsvoorstel voorwaarden opgenomen:

- vaststelling van het ontwikkelingsprogramma door de gemeenteraad;
- onderbouwen van het ontwikkelingsprogramma door een analyse van de problemen en potenties in de fysieke leefomgeving;
- gedegen betrokkenheid bij de voorbereiding van het ontwikkelingsprogramma van de relevante lokale partijen en van de gemeenten waarvan de belangen rechtstreeks in het geding zijn;
- weergave in het ontwikkelingsprogramma van de toetsbare beoogde resultaten (zo mogelijk in termen van beoogde effecten van beleid);
- vermelding in het ontwikkelingsprogramma van de samenhang met aanpalende beleids-terreinen op het gebied van de fysieke leefomgeving en een vermelding in het ontwikkelingsprogramma van de geschatte omvang van de benodigde investeringen in de eerste vijfjaarsperiode.

De voorwaarden die in het Beleidskader stedelijke vernieuwing worden opgenomen, hebben betrekking op een twaalfstal prestatievelden die ten minste in het ontwikkelingsprogramma aan de orde moeten komen. Van deze prestatievelden hebben de eerste zes een procesmatig karakter en de laatste zes een inhoudelijk karakter. De twaalf prestatievelden zijn:

1. ontwikkelen van een aanpak gericht op de langere termijn;
2. integreren van beleidsterreinen;
3. integreren van schaalniveaus;
4. benutten van kansen en ontwikkelingspotenties;
5. genereren van private investeringen;
6. aangeven van prioriteiten, fasering en aandachtsgebieden;
7. verbeteren van fysieke voorwaarden voor economische activiteiten;
8. creëren van een gevarieerd aanbod aan woonmilieus;
9. creëren van een gewenste omgevingskwaliteit;
10. bevorderen van duurzaamheid;
11. nastreven van zorgvuldig ruimtegebruik;
12. versterken van de betrokkenheid bij de dagelijkse leefomgeving.

Het formuleren van de prestatievelden betekent niet dat een gemeente op elk van de prestatievelden beleid moet ontwikkelen. De gemeente moet echter wel uitdrukkelijk nagaan of op de genoemde terreinen in de gemeente een opgave ligt. In het ontwikkelingsprogramma moet vervolgens goed gemotiveerd worden waarop de beleidskeuzen en prioriteiten zijn gebaseerd en dus ook waarom eventueel ten aanzien van bepaalde prestatievelden wordt afgezien van een beleidsinspanning.

²⁸ Met terugwerkende kracht per 1 januari 2000 in werking getreden.

Evaluatie van het ISV-beleid

Evaluatie en verantwoording

Binnen de ISV-systematiek wordt een onderscheid gemaakt tussen de evaluatie van het ISV-beleid en de verantwoording van het ISV-beleid. De evaluatie (op grond van monitoring en visitatie) heeft als doel te leren van opgedane ervaringen om daarmee het ISV-beleid te verbeteren. Het ISV-beleid is ook nadrukkelijk als een groeimodel gepresenteerd. De verantwoording heeft als doel de gemeenten af te rekenen op de door hen geleverde prestatie in het ISV-tijdvak. Daarbij is vooral van belang in hoeverre de van tevoren gestelde doelen daadwerkelijk zijn bereikt.

Evaluatie

Bij de evaluatie van het ISV-beleid kan een onderscheid worden gemaakt tussen twee niveaus: het rijk en de gemeenten²⁹. Het rijk voert de evaluatie uit door middel van de Monitor Stedelijke Vernieuwing. Deze monitor heeft alleen betrekking op de rechtstreekse gemeenten (G30 nieuwe stijl) en kan gezien worden als de ‘thermometer’ van de landelijke Stedelijke Vernieuwing. Hij geeft een beeld van de aspecten die, op grond van het Beleidskader Stedelijke Vernieuwing, relevant worden geacht voor de Stedelijke Vernieuwing, en signaleert daarin trends. In een volgende paragraaf wordt op deze monitor dieper ingegaan. Naast deze monitor wordt op landelijk niveau een systeem van visitatie opgezet waarin zelfanalyse en de onderlinge uitwisseling van ervaringen een belangrijke plaats hebben. Deze visitatie vormt een kwalitatieve aanvulling op de kwantitatieve gegevens van de Monitor Stedelijke Vernieuwing.

Op lokaal niveau zijn de gemeenten verantwoordelijk voor het inrichten van een stelsel voor kwaliteitszorg voor de interne bewaking en verbetering van de kwaliteit van stedelijke vernieuwing. De gemeenten zijn vrij in de manier waarop ze dit kwaliteitszorgsysteem vormgeven, maar veelal betekent dit dat gemeenten het ontwikkelingsprogramma verder uitwerken tot een uitvoeringsprogramma. Vervolgens wordt er binnen de gemeentelijke organisatie een programmacoördinator of programmamanager aangesteld. Deze coördinator is de spil als het gaat om het bewaken en meten van de voortgang³⁰ van het uitvoeringsprogramma en het inhoudelijk en financieel terugkoppelen naar het bestuur.

Een slimme coördinator maakt daarbij zo veel mogelijk gebruik van de kennis die in de gemeentelijke organisatie aanwezig is, waardoor het ISV beter raakt ingebed in de gemeentelijke organisatie, en de organisatie beter leert van de ervaringen met het ISV. Dit biedt onder meer de volgende voordelen.

- De organisatie heeft houvast bij het opstellen van personele planning en interne afstemming en prioritering van werkzaamheden.
- De organisatie is breed op de hoogte van projecten en weet wie verantwoordelijk is.
- Het bevordert het projectmatig werken.
- Het bestuur van de gemeente is beter op de hoogte van de voortgang en knelpunten bij projecten.

Verantwoording

De verantwoording (aan het Rijk respectievelijk de provincie) voor de uitgaven uit het ISV-budget vindt plaats aan het eind van het vijfjarige investeringstijdvak (2005). Dit gebeurt aan

²⁹ In hoeverre de provincies bij de evaluatie van niet-rechtstreekse programmameenten betrokken zijn, is niet geheel bekend. Bij monde van het Interprovinciaal Overleg (IPO) hebben de provincies aangegeven met hun monitoring te willen aansluiten bij de Rijksmonitor.

³⁰ Veelal gebruiken gemeenten voor het meten van de voortgang het (proces)format van de Monitor Stedelijke Vernieuwing als bouwsteen voor het eigen kwaliteitszorgsysteem.

de hand van een door de gemeente opgesteld verantwoordingsverslag. De gemeente heeft in haar ontwikkelingsprogramma meetbare doelstellingen moeten aangeven. In het verantwoordingsverslag geeft de gemeente aan in hoeverre de in het ontwikkelingsprogramma geformuleerde doelstellingen zijn gerealiseerd, en wat er de oorzaken van zijn dat de resultaten - naar boven of beneden - afwijken van de oorspronkelijke doelstellingen. Het is met name hierom dat gemeenten wettelijk verplicht zijn het in de vorige paragraaf beschreven kwaliteitszorgsysteem in te voeren.

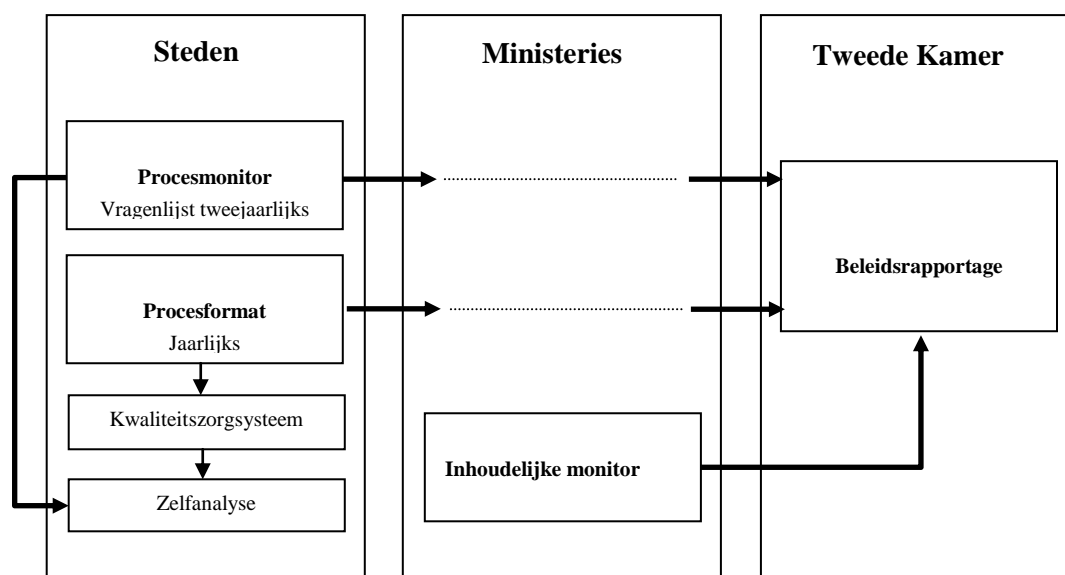
Het niet volledig realiseren van de doelstellingen is op zich niet voldoende grond voor een sanctie. Als een gemeente overtuigend kan aangeven dat er gegronde redenen zijn voor het niet halen van de doelstellingen, is er niets aan de hand. Blijkt echter dat een gemeente 'toerekenbaar' slecht heeft gepresteerd dan kan, nadat bestuurlijk overleg heeft plaatsgevonden, een sanctie worden getroffen. Omdat een financiële sanctie (het opschorten van uitbetalingen) de gemeente kan belemmeren in het aangaan van overeenkomsten met investeerders zal de sanctie in eerste instantie de vorm hebben van aanvullende voorwaarden voor het nieuwe ISV-tijdvak. Op grond van de Algemene Wet Bestuursrecht is het overigens altijd mogelijk tussentijds, wanneer een gemeente ernstig in gebreke blijft, een financiële sanctie te treffen.

Over het algemeen zal de eindverantwoording voor beide partijen geen verrassingen opleveren. Immers door de monitoring en visitatie hebben alle betrokkenen steeds een goed beeld van de voortgang. De formele eindverantwoording kan daarom worden beschouwd als de laatste fase van de communicatie tussen de bestuurslagen over de voortgang en resultaten van het stedelijke vernieuwingsbeleid.

De Monitor Stedelijke Vernieuwing

Deze paragraaf gaat nader in op het (kwantitatieve) evaluatie-instrument van het rijk. De Monitor Stedelijke Vernieuwing is zoals gezegd de 'thermometer' van de landelijke stedelijke vernieuwing. Op de langere termijn is het de bedoeling dat de monitor de maatschappelijke effecten van de stedelijke vernieuwing in kaart brengt. Vooralsnog, in deze eerste periode, is de monitor echter gerelateerd aan de prestatievelden uit het Beleidskader.

fig.6.2



Monitor Stedelijke Vernieuwing

De Monitor is ontwikkeld door een werkgroep met vertegenwoordigers van de deelnemende ministeries (VROM³¹, LNV en EZ) en gemeenten en wil inzicht bieden in de inhoudelijke en procesmatige voortgang van de stedelijke vernieuwing. Hiertoe bestaat de monitor uit een inhoudelijke monitor en een procesmonitor. De inhoudelijke monitor brengt elke twee jaar voor de zes inhoudelijke prestatievelden uit het Beleidskader Stedelijke Vernieuwing de fysieke ontwikkelingen in de steden in kaart aan de hand van statistische gegevens uit een groot aantal bronnen. De inhoudelijke monitor bevat 26 indicatoren. Om de bestuurslast te beperken, gaat de verzameling van statistische gegevens grotendeels buiten de gemeenten om: 24 van de 26 indicatoren worden uit landelijke bronnen geput (VROM, 2001).

De procesmonitor meet de voortgang van het beleidsproces. Dat gebeurt enerzijds via een tweejaarlijkse vragenlijst van negen vragen die de gemeenten invullen. Daarmee wordt voor de zes procesmatige prestatievelden uit het beleidskader de ontwikkeling in kaart gebracht. Anderzijds vindt een jaarlijkse inventarisatie plaats via het zogeheten procesformat. Met behulp van dit format wordt voor de zes inhoudelijke prestatievelden een 'gemiddelde' invoeringsfase van de lopende en voorgenomen projecten bepaald. De invoeringsfase is gebaseerd op een schatting van de gemeenten zelf. De monitor is geen instrument om gemeenten te controleren, en legt daarom geen meetlat langs de voortgangsinformatie.

De inhoudelijke monitor en de procesvragenlijst kennen een tweejaarlijkse afnamecyclus. De nulmeting heeft begin 2001 plaatsgevonden (VROM 2001b). In 2003 vindt een tussenmeting plaats en in 2005 de eindmeting. Het procesformat wordt wel elk jaar geactualiseerd. Dit procesformat vormt vooral in de eerste jaren een belangrijke schakel in de beleidsinformatie: het is immers weinig realistisch om te verwachten dat, nu er een stedelijke vernieuwingsbeleid gevoerd wordt, de fysieke werkelijkheid in de steden er ineens heel anders uit zal zien. Gezien het vraagstuk waar het beleid zich op richt, mag worden verwacht dat de resultaten van het beleid op z'n vroegst over enkele jaren in de steden en dus via de inhoudelijke monitor zichtbaar zullen worden. De monitoring van de procesvoortgang, zoals die met behulp van het procesformat in kaart wordt gebracht, zal daarom het instrument zijn dat zeker gedurende de eerste jaren de belangrijkste beleidsinformatie zal leveren.

Afstemming met andere monitorsystemen

Zoals aangegeven is het ISV nauw verwant aan het GSB: het vormt naast sociale en economische programma's de derde pijler van het GSB. Naast de ISV-monitor bestaat er daarom ook nog een overkoepelende GSB-monitor, en monitoren voor de economische pijler en de sociale pijler. Deze laatste is overigens nog in ontwikkeling.

De verantwoordelijkheden voor de verschillende monitorsystemen zijn als volgt verdeeld:

GSB:	ministerie van BZK
ISV (fysieke component):	ministerie van VROM
Economische component:	ministerie van EZ
Sociale component:	ministerie van SoZa

De coördinerende verantwoordelijkheid voor de gehele monitorstructuur ligt bij de Minister van Grote-steden- en Integratiebeleid. De verschillende monitoren zijn zodanig op elkaar afgestemd dat bijna geen overlap bestaat tussen de indicatoren welke door de verschillende monitoren worden gemeten. Het is de verantwoordelijkheid van elk ministerie erop toe te zien dat de eigen monitor aanvullend, en niet overlappend is binnen de totale monitorstructuur.

³¹ Het ministerie van VROM coördineert de Monitor Stedelijke Vernieuwing.

Bevindingen

De nulmeting

In januari 2001 is Rigo Research en Advies in opdracht van het ministerie van VROM gestart met de nulmeting van de Monitor Stedelijke Vernieuwing. Dit onderzoek is op 26 mei 2001 aan de Tweede Kamer aangeboden. De inhoudelijke nulmeting biedt een statistisch beeld van de uitgangspositie van de G30 bij de aanvang van het stedelijke vernieuwingsbeleid. Pas bij de vervolgmetingen krijgen deze gegevens betekenis: dán onstaat zicht op ontwikkelingen en zal de monitor zijn primaire functie (signalering in de tijd) kunnen gaan vervullen.

De procesmonitor biedt daarentegen al wel interpreteerbare informatie. Zo komt bijvoorbeeld uit het procesformat – overigens geheel volgens verwachting – het beeld naar voren dat het beleid zich gemiddeld nog in de startfase bevond. Tegelijkertijd blijkt dat er wel degelijk verschillen bestaan tussen aandachtsvelden; met name op de traditionele volkshuisvestelijke terreinen is het beleid wat verder gevorderd dan op nieuwe terreinen als bijvoorbeeld ‘verdunding van woongebieden’. Hieronder wordt wat dieper ingegaan op de uitkomsten van de nulmeting.

Uitkomsten inhoudelijke monitor

De uitkomsten van de inhoudelijke nulmeting zijn vooral een bevestiging van reeds lang bestaande en bekende verschillen tussen de grote steden (G30) en de rest van het land, en tussen landsdelen. Vergelijking van de rechtstreekse gemeenten met overig Nederland leert dat de G30 een ongunstige score hebben waar het de omgevingskwaliteit – zowel fysiek als sociaal – betreft. In het bijzonder gaat het dan om zaken als groenvoorziening, overlast en sociale cohesie. Een algemeen beeld is dat in het Westen, en dan vooral in de G4, de druk op zowel de woningmarkt als de markt voor bedrijventerreinen groot is en de kwaliteit van het wonigaanbod en van de woonomgeving laag, zowel ten opzichte van andere delen van Nederland als ten opzichte van niet-rechtstreekse gemeenten. De discrepanties tussen vraag en aanbod op de woningmarkt zijn in de G30 gemiddeld groter dan daarbuiten. Er wordt in onvoldoende mate geboden waar de consument om vraagt: groenstedelijke woonmilieus met woningen in het luxe koopsegment. Daarnaast bieden de G30 ook in kwantitatieve zin onvoldoende aanbod, en dat resulteert in een relatief hoog prijsniveau van de koopwoningen. Over het algemeen zijn de problemen die met de G30 worden geassocieerd het grootst in de G4. Naar landsdeel is de gebruikelijke volgorde (in afnemende problematiek): West, Zuid, Oost en Noord. In Noord en Oost is de omgevingskwaliteit het sterkst en is er minder discrepantie tussen vraag en aanbod dan in het Westen en Zuiden van het land. Vooral in Noord lijkt de balans op de woningmarkt echter door te slaan in een aanbodoverschot en een bijbehorend laag prijsniveau van de koopwoningen. Ook qua werkgelegenheid is het beeld in het Noorden en Oosten minder gunstig dan in het Westen en Zuiden (VROM, 2001).

Uitkomsten procesmonitor

De nulmeting laat zien dat gemeenten aandacht hebben voor het vraagstuk van de doorwerking van papier naar praktisch. Samenhangende en ambitieuze visies zijn weliswaar noodzakelijk, maar nog geen voldoende voorwaarde voor een voortvarende vertaling in concrete projecten. Dit vergt veel van de gemeentelijke organisatie. Er wordt gewerkt aan methoden en technieken om het gedachtegoed uit het MOP te verankeren in handelen.

De nulmeting laat zien dat gemeenten op verschillende manieren bezig zijn deze slag te maken. Zo is men in veel gemeenten bezig om organisatieaanpassingen³² door te voeren die een meer integrale en gebiedgerichte werkwijze beogen (kantelen). Overigens geven veel

³² Deze wijzigingen van de gemeentelijke organisatie zijn meestal geen direct gevolg van het ISV. Vaak zijn deze wijzigingen al ingezet om beter tegemoet te komen aan het op dezelfde leest als het ISV geschoeide, en al langer bestaande GSB.

gemeenten te kennen daarbij tegen belemmeringen in wet en regelgeving aan te lopen. Binnen de gemeentelijke organisatie zorgt ook de arbeidsmarktproblematiek regelmatig voor belemmeringen.

In het ISV is de afstemming tussen verschillende beleidsterreinen (horizontaal) en verschillende schaalniveaus (verticaal) van groot belang. Instrumenten die de zogenaamde horizontale afstemming willen bevorderen en borgen – effectmonitoring, kwaliteitszorg, programmamanagement, worden overal gehanteerd, maar bevinden zich in het algemeen nog in de ontwikkelingsfase. De verticale afstemming kan nog verder worden ingevuld. Bij het opstellen van de MOP's werd veelvuldig overlegd, maar minder helder is hoe gemeenten bij de uitwerking en uitvoering van de programma's een verband leggen tussen de verschillende schaalniveaus.

In de wijken, waar het uiteindelijk allemaal gebeurt, laten de G30 een veelkleurig palet aan werkwijzen en benaderingen zien, die onverlet laten dat de wijk bij uitstek gezien wordt als het niveau waarop de stedelijke vernieuwing wordt ingevuld. Met de logische en ook terechte nadruk die de wijk als integratiekader krijgt, mag de samenhang met ontwikkelingen elders niet verloren gaan (VROM, 2001).

Uitkomsten procesformat

De meeste projecten, en daarmee de voortgang op de inhoudelijke prestatievelden, bevinden zich in de fase van besluitvorming en contractvorming; de verschillen tussen de steden zijn gering. In het licht van de recente start van het ISV is dat ook niet verbazingwekkend.

Wat meer in detail bezien, bevestigt de voortgangsmeting de indrukken uit een eerdere kwalitatieve analyse van de gemeentelijke MOP's (RIGO Research en Advies BV 2001). Vrij traditionele thema's als de nieuwbouw van woningen en de huisvesting van senioren komen in veel projecten aan bod. Aandachtsgebieden als intensief ruimtegebruik, verdunning van woongebieden en beperking van hinder, worden minder vaak expliciet meegenomen in de projecten. Vervolgmetingen kunnen laten zien of, en in welke mate, dat in de komende jaren wel gebeurt (VROM, 2001).

Reflectie en reactie

Vooralsnog heeft op het functioneren van de Monitor Stedelijke Vernieuwing tijdens de nulmeting weinig reflectie plaatsgevonden, of is hierover in ieder geval weinig bekend. Wel zijn er enkele algemene reacties en opmerkingen.

Enkele gemeenten tekenen aan dat bij het schetsen van een landelijk beeld en het signaleren van trends, het specifieke van de lokale situatie verloren gaat. Een ambitieuze en vernieuwende aanpak wordt niet zichtbaar en kan dus ook niet gewaardeerd worden. Veel gemeenten hebben daarom van de mogelijkheid gebruik gemaakt om de vragenlijsten en procesformats van een toelichting te voorzien. Daaruit komt een beeld naar voren dat de veelkleurigheid van de gemeentelijke ISV-praktijk, en de inzet waarmee gemeenten aan de slag zijn, onderstreept.

De verschillen tussen wijken binnen een gemeente zijn veelal aanzienlijk groter dan de verschillen tussen gemeenten. Opmerkelijk is bijvoorbeeld dat, blijkens de indicatoren, de waardering voor de woonomgeving zo weinig uiteenloopt tussen de diverse delen van Nederland. Die geringe verschillen zijn een gevolg van het hoge schaalniveau waarnaar geaggregeerd is. Uit andere bronnen blijkt dat er tussen wijken wel degelijk grote verschillen in belevingsaspecten voorkomen. Deze worden op landsdeelniveau uitgemiddeld.

De statistische gegevens die verzameld zijn ten behoeve van de inhoudelijke monitor zijn opgeslagen in een ISV-database. Deze gegevens zijn uiteraard ook voor de steden interessant. Het ministerie van VROM zal een 'afslag' van het gegevensbestand dan ook aan de steden beschikbaar stellen. Als gevolg van de afspraak met de gemeenten dat in de ISV-Monitor

geen individuele steden herkenbaar zullen zijn, krijgen de steden uitsluitend hun eigen cijfers teruggestuurd. Dat stelt hen in staat zich te vergelijken met de gepubliceerde gemiddelden van de G30, de G4 en andere categorieën.

lessen voor evaluatie 5N

Welke lessen kunnen worden getrokken uit deze beschouwing van de ISV-monitor? De ISV-monitor is nog erg vers. Er is slechts een nulmeting geweest en deze heeft vooral inhoudelijk de stand van zaken in de steden bevestigd. Dit leverde weinig stof voor interessante bevindingen. Ook met het functioneren van de ISV-monitor als 'instrument' is nog maar weinig ervaring opgedaan. Lessen op grond van deze beschouwing hebben dus niet zo'n stevig fundament. Toch zijn er binnen de systematiek van de ISV-monitor enkele aspecten aan te wijzen die aandachtspunten kunnen zijn bij een evaluatie van de 5N. Ik noem deze aandachtspunten in een vrij willekeurige volgorde.

Inhoud & proces

Het is opvallend dat bij de ISV-monitor expliciet aandacht is voor zowel de inhoud als het proces. Dit is ook inherent aan de sturingsfilosofie achter het ISV (integratie, interactie en innovatie), waarin een goed proces als voorwaarde wordt gezien voor bevredigende inhoudelijke resultaten. Dit geldt ook voor het beleid in de 5N.

Flexibiliteit instrument

Het ISV is gepresenteerd als een groeimodel dat zich in de toekomst verder zal ontwikkelen. Het is dan ook van belang dat het monitorsysteem flexibel is en kan meebewegen met toekomstige beleidsveranderingen. Het is nadrukkelijk de bedoeling dat de ISV-monitor een dergelijke flexibiliteit kent. Tegelijkertijd bestaat al snel het gevaar dat een monitorsysteem geïnstitutionaliseerd raakt. Het gaat immers om een vrij ingewikkelde organisatie en procedures van gegevensverzameling. Ook bij een evaluatie van de 5N zal met dit spanningsveld tussen flexibiliteit en institutionalisering van het instrument moeten worden omgegaan.

Prikkel voor zowel afnemer als leverancier

Evalueren is nodig, maar kost ook bepaalde inspanningen. Een evaluatie zal soepeler verlopen als deze evaluatie voor beide partijen (evaluator/afnemer en geëvalueerde/leverancier) de benodigde inspanning waard is. In het geval van de ISV-monitor is de verzamelde informatie vooral van belang voor het Rijk. Toch is het voor gemeenten ook aantrekkelijk om aan de evaluatie mee te werken. Zij krijgen hun eigen gegevens teruggeleverd en afgezet tegen het gemiddelde. Zij krijgen dus informatie over hun relatieve positie. Dergelijke informatie is in een netwerksamenleving met toenemende onzekerheden belangrijke informatie voor een actor. Daarnaast biedt de ISV-monitor aan de gemeenten ook een format om een intern kwaliteitszorgsysteem op te zetten.

Kortom, het is belangrijk dat een evaluatie prikkels biedt aan beide partijen. Een interessante manier om aan deze prikkel vorm te geven is het terugleveren van bewerkte informatie aan de geëvalueerde partij. Het is sowieso belangrijk om stil te staan bij de enorme berg aan informatie die bij evaluaties verkregen wordt, en zich af te vragen hoe deze informatie in onze kennisintensieve maatschappij nog op andere manieren van nut kan zijn.

Afstemming met andere monitorsystemen

Evalueren en monitoren is hot. De samenleving is continu in beweging en we moeten de vinger aan de pols houden. Het aantal monitorsystemen en evaluaties is de afgelopen jaren dan ook explosief gegroeid. Het grote gevaar hiervan is dat bij partijen al een zekere evaluatiemoedigheid begint op te treden. Ook bij het ISV / GSB was dit het geval. Er is daarom veel aandacht besteed aan de coördinatie tussen, en integratie van monitors die gerelateerd zijn aan het

GSB. Zo is er bijna geen overlap tussen de indicatoren van de verschillende monitors en worden de gegevens zo veel mogelijk op één moment bij de gemeenten opgevraagd. Ook wordt gestreefd naar het instellen van contactpersonen (verantwoordelijk voor het gehele GSB-monitorcluster) die een vast en herkenbaar aanspreekpunt vormen voor de gemeenten. Daarnaast wordt geprobeerd de gemeenten zoveel mogelijk te ontlasten door veel indicatoren buiten de gemeenten om te betrekken. Zo zijn bij de ISV-monitor slechts 3 indicatoren rechtstreeks van gemeenten afkomstig.

Een dergelijke coördinatie tussen verschillende aan elkaar gerelateerde monitorsystemen is echter ingewikkeld en staat de flexibiliteit van de verschillende afzonderlijke monitors in de weg.

Ook voor een evaluatie van de 5N geldt dat het niet de enige evaluatie is. Zij zal om de aandacht moeten strijden in concurrentie met vele andere monitors en evaluaties. Het is goed om na te denken op welke manier een evaluatie van de 5N gecombineerd of geïntegreerd kan worden met andere (bestaande of toekomstige) evaluaties en monitors, waarbij overigens gewaakt moet worden over de flexibiliteit van het instrument.

Houdt oog voor locatiespecifieke omstandigheden

Eén van de belangrijke stokpaardjes van het ISV is het hameren op maatwerk. Helaas klinken er geluiden dat de ISV-monitor onvoldoende recht zou doen aan dit maatwerk. De gegevens zouden naar een te hoog abstractieniveau worden geaggregeerd en aan betekenis verliezen. Wellicht dat de voorgenomen visitaties deze kritiek in de toekomst kunnen ondervangen. Echter, niet alleen stedelijke vernieuwing is een kwestie van maatwerk, maar alle ruimtelijke ingrepen zijn een vorm van maatwerk. Ook de 5N hanteert een sterk gebiedsgerichte benadering. Bij een evaluatie van de 5N zal dus een evenwicht moeten worden gevonden tussen enerzijds recht doen aan de specifieke situatie en anderzijds de hanteerbaarheid en vergelijkbaarheid van de verkregen informatie.

Bijlage 1: Leden Begeleidingscommissie

Dhr. B. van Bleek, Ministerie van VROM (RPD)

Mw. E. Driessen, Ministerie van VROM (RPD)

Mw. MM.A. Galle, Ruimtelijk Planbureau

dr. P. van der Knaap , Ministerie van Financien

Mw. J.H.P. Vrolijk, Ministerie van VROM (RPD)

Mw. R. Wersch, Provincie Gelderland

Secretaris: Dhr. E. Taale (RPD)

Bijlage 2: Geïnterviewde personen

drs. G. Beukema, directeur Interprovinciaal Overleg

drs. J.A. de Boer, adjunctdirecteur Strategie en Ontwikkeling, Interprovinciaal Overleg

drs. G.N.J.A. Bukkems, Algemene Rekenkamer, Hoofd Bureau Europees en Rijksbreed Doelmatigheidsonderzoek

mr. B.G. Bults, Ministerie van Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieu (RPD, afdeling Bestuursaangelegenheden)

drs. J.A.A. Dikmans, Ministerie van Verkeer en Waterstaat

dhr. S. Depla, lid Tweede Kamer

drs. P.L.B.A. van Geel, gedeputeerde RO, Provincie Noord-Brabant

prof.dr. M. Hajer, Universiteit van Amsterdam

ir. J.R. Lambrechtsen, Inspecteur Ruimtelijke Ordening in de provincies Zeeland, Noord-Brabant en Limburg

mr. P.G.A. Noordanus, voorzitter VROM-Raad

ing. K.J. de Ruiter, plaatsvervangend directeur Directie Groen Ruimte en Recreatie, Ministerie van Landbouw, Natuurbeheer en Visserij

ir. F. Uit den Boogaard, Ministerie van Financiën (Inspectie)

dhr. P. van Veen, Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieu (FEZ)

drs. K. Vijlbrief, Ministerie van Economische Zaken (ICES)

Bijlage 3: Geraadpleegde literatuur

- Abma, T. A. (1996). Responsief evalueren: discourses, controversen en allianties in het post-moderne. Rotterdam.
- Algemene Rekenkamer (1999-2000). Voortgangsrapportage VBTB nader beschouwd, Tweede Kamer.
- Algemene Rekenkamer (2001). Verslag 2000. Den Haag, Algemene Rekenkamer.
- Beckers, T. and W. Haarmann (2001). Reconstructie beschouwd: een advies over het ontwerp-koepelplan reconstructie zandgronden. Tilburg, Telos, Brabants Centrum voor Duurzaamheidsvraagstukken.
- Bemelmans-Videc, M. L. (1994). Nota Codificatie doelmatigheidsonderzoek Algemene Rekenkamer. Den Haag, Algemene Rekenkamer.
- Berkhout, A. G. B. and M. R. T. W. Sanders (2000). "Van Beleidsdocument Tot Beleidspraktijk: zeven dilemma's bij de implementatie van beleid." Beleidsanalyse(3/4): 5-22.
- Bordewijk, P. and H. L. Klaassen (2000). Wij laten ons niet kennen: een onderzoek naar het gebruik van kengetallen bij negen grotere gemeenten. Den Haag, VNG-uitgeverij.
- COBA (1976). "Het instrument subsidie: een leidraad voor subsidieonderzoek." Beleidsanalyse 5(2): 2-37.
- COBA (1978). "Het zesde jaarverslag van de Commissie voor de ontwikkeling van de beleidsanalyse." Beleidsanalyse 7(2): 2-60.
- Cousins, J. B. and L. M. Earl (1992). "The case for participatory evaluation." Educational Evaluation and Policy Analysis 14(4): 397-418.
- CPB, RIVM, RPD-Planbureau and SCP (2001). Ex ante evaluatie van de Vijfde Nota Ruimtelijke Ordening.
- Dahlgaard, J. J., K. Kristensen and G. J. Kanji (1998). Fundamentals of Total Quality Management. London, Chapman & Hall.
- De Vries, G. J. (2000). Beleidsdynamica als sociale constructie: een onderzoek naar doorwerking van beleidsevaluatie en beleidsadvisering. Delft, Eburon.
- Dolmans, L. (1989). Naar supervisie en meer aandacht voor doelmatigheid: de Algemene Rekenkamer tussen 1945 en 1988. Van Camere van der Rekeninghen tot Algemene Rekenkamer. A. M. et.al. Den Haag, Sdu: 377-431.
- Dolmans, L. J. F. (1987). "De Algemene Rekenkamer en de doelmatigheid van de rijksdienst." Beleid en Maatschappij(1): 18-29.
- Edelenbos, J. and M. v. Eeten (2001). "The missing link: processing variation in dialogical evaluation." Evaluation 7(2): 204-210.
- Edelenbos, J. and R. Goedhart (2001). "Het probleemoplossend (on)vermogen van interactieve beleidsvorming." Openbaar Bestuur 11(12): 13-17.
- Ehrmann, J. R. and B. L. Stinson (1999). Joint fact-finding and the use of technical experts. The consensus building handbook: a comprehensive guide to reaching agreement. L. Süsskind, S. McKearnan and J. Thomas-Larner. London, Sage: 375-400.
- Financiën, M. v. (2001). Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid. Den Haag, Ministerie van Financiën: 53.
- Guba, E. G. and Y. S. Lincoln (1989). Fourth generation evaluation. Newbury Park, London, New Delhi, Sage.
- Haselbekke, A. G. J., H. L. Klaassen, A. P. Ros and R. J. i. t. Veld (1990). Prestaties tellen: kengetallen als instrument voor een bedrijfsmatig(er) bestuur en beheer van decentrale overheden. Den Haag, VNG-Uitgeverij.
- Herweijer, M., G. J. A. Hummels and C. W. W. Van Lohuizen (1990). Evaluatie van indicatieve plannen: handleiding en begrippen. Den Haag, Ministerie VROM: 106.
- In 't Veld, R. J. (1978). Over grenzen van bestuur. Den Haag, Vuga.

- IPO (2000). Van ordenen naar ontwikkelen; provincies investeren in de kwaliteit van de ruimte. Den Haag.
- Jasanoff, S. (1990). The fifth branch: advisers as policy makers, Harvard University Press.
- Klijn, E. H. and J. M. F. Koppenjan (1999). "De politiek en interactieve besluitvorming: van institutionele spelbreker naar spelbepaler." Beleidswetenschap **13**(1): 47-68.
- Leeuw, F. L. (1992). "Doelmatigheidsonderzoek bij de Algemene Rekenkamer." Bestuurswetenschappen **46**(1): 8-24.
- Leeuw, F. L. and P. J. Rozendal (1994). Policy evaluation and the Netherland's government: scope, utilization and organizational learning. Can governments learn? Comparative perspectives on evaluation and organizational learning. R. C. R. R. C. S. F.L. Leeuw. New Brunswick/London, Transaction Publishers. 67-86.
- Lindblom, C. E. and D. K. Cohen (1979). Usable Knowledge: social science and social problem solving. New Haven/London, Yale University Press.
- Mastop, J. M. and A. Faludi (1993). "Doorwerking van strategisch beleid in dagelijkse beleidsvoering." Beleidswetenschap **7**(1): 71-90.
- Mol, N. P. (2000). "Vijf stellingen over financieel overheidsbeheer, in: Openbare Uitgaven." Openbare Uitgaven **32**(4): 152-161.
- Mulder, H.-P., G. Walraven, A. d. Groot, F. S. Terpstra, P. Rozendal, O. Delsman, A. v. Loon, R. Venderbosch and F. L. Leeuw (1991). "Gebruik van beleidsevaluatieonderzoek bij de rijksoverheid." Beleidswetenschap **5**: 203-228.
- NAM (1998). Integrale bodemdalingsstudie Waddenzee. Assen, NAM.
- Patton, M. Q. (1997). Utilization-focused evaluation. Thousand Oaks, London, New Delhi, Sage.
- Pawson, R. and N. Tilley (1997). Realistic evaluation. London, Thousand Oaks, New Delhi, Sage.
- Priemus, H. (1996). Stedelijke vernieuwing: op weg naar de duurzame stad. In staat van vernieuwing: maatschappelijke vernieuwingsprocessen in veelvoud. N. J. M. Nelissen, T. Itink and A. W. Van de Ven. Bussum, Coutinho: 155-175.
- RIGO Research en Advies BV (2001). Beloften van vernieuwing, RIGO Research en Advies BV.
- Rist, R. C. (1994). The preconditions for learning: lessons from the public sector. Can governments learn? Comparative perspectives on evaluation and organizational learning. F. L. Leeuw, R. C. Rist and R. C. Sonnichsen. New Brunswick London, Transaction Publishers: 189-205.
- Ryan, K. and L. DeStefano (2001). "Dialogue as a democratizing evaluation method." Evaluation **7**(2): 188-203.
- Shulha, L. M. and J. B. Cousins (1997). "Evaluation use: theory, research and practice since 1986." Evaluation Practice **18**(3): 195-208.
- Stevens, T. (1979). De Rekenkamer. Leiden, Stenfert Kroeze.
- Süsskind, L. and J. Cruikshank (1987). Consensual approaches to resolving public disputes, Basic Books.
- Teisman, G. R. (2001). "Sturen met strategisch beleid: sturingsopvattingen in ruimtelijk beleid, resultaten en lessen." Bestuurskunde **10**(3): 129-138.
- Tweede Kamer (1998-1999). nota Van Beleidsbegroting Tot Beleidsverantwoording., 26 573, nr. 2.
- Van der Burg, C., H. W. O. L. M. Korte and R. P. Oosterom (1999). "Van beleidsbegroting tot beleidsverantwoording." Openbare Uitgaven **31**(5): 226-233.
- Van der Knaap, P. (1997). Lerende overheid, intelligent beleid. Den Haag, Phaedrus.
- Van der Knaap, P. (1999). "Prestatiegegevens en beleidsevaluatie bij de rijksoverheid: van ontwikkeling naar gebruik, in: Openbare Uitgaven." Openbare Uitgaven **31**(5): 234-245.
- Van der Meer, F. B. (1999). "Evaluation and the social construction of impacts." Evaluation **5**(4): 387-406.

- Van der Meer, F. B., G. J. D. De Vries and G. A. N. Vissers (2000). "Evaluatie en leerprocessen bij de overheid: de rol van institutionele condities." Beleidswetenschappen **14**(3): 253-277.
- Van Eeten, M. (1999). Dialogues of the deaf: defining new agenda's for environmental deadlocks. Delft, Eburon.
- VROM (1988). Vierde Nota over de Ruimtelijke Ordening. Den Haag, Ministerie van VROM.
- VROM (1992). Vierde Nota over de Ruimtelijke Ordening Extra. Den Haag, Ministerie van VROM.
- VROM (2001a). Vijfde Nota over de Ruimtelijke Ordening, PKB deel 3: Kabinetsstandpunt. Den Haag, Ministerie VROM.
- VROM, M. v. (1997). Nota stedelijke vernieuwing. Den Haag, Ministerie van VROM.
- VROM, M. v. (1999). Balans ruimtelijke kwaliteit 1999. Den Haag, Ministerie van VROM (RPD).
- VROM, M. v. (2001b). Monitor stedelijke vernieuwing: nulmeting. Den Haag, Ministerie van VROM.
- VROM, M. v. (2001c). Ruimte maken, ruimte delen: Vijfde Nota Ruimtelijke Ordening 2000/2020. Den Haag, Ministerie van VROM.
- VROM, M. v. (2001d). Vijfde Nota over de Ruimtelijke Ordening: uitvoeringsprogramma nationaal ruimtelijk beleid. Den Haag, Ministerie van VROM: 106 + bijl.
- VROM, M. v., M. v. EZ and M. v. LNV (1999). Beleidskader stedelijke vernieuwing. Den Haag.
- Weiss, C. H. (1980). "Knowledge creep and decision accretion." Knowledge: creation, diffusion, utilization **1**(3): 381-404.
- Werkgroep beleidsdoelstellingen (1973). "Interim rapport." Beleidsanalyse **2**(3): 2-12.
- Werkgroep Vijfde Nota Ruimtelijke Ordening (1999-2000). Notie van ruimte: op weg naar de Vijfde Nota Ruimtelijke Ordening. Den Haag, Tweede Kamer 27210.
- Wilmer, C. (1980). "Evaluatie van de departementale doelstellingsanalyse." Beleidsanalyse **9**(2): 22-24.