

Statenvoorstel 58/12 A

Voorgestelde behandeling :
PS-vergadering : 12 oktober 2012
Statencommissie : Commissie voor Economische Zaken en
Bestuur, 14 september 2012

Onderwerp

Beleidsvaluaties provincie Noord-Brabant (vervolgonderzoek);
onderzoeksrapport Zuidelijke Rekenkamer

Datum

15 augustus 2012

Documentnummer

3079855

Directie

Griffie

Bijlage(n)

2

Aan Provinciale Staten van Noord-Brabant,

Samenvatting

Op 14 mei 2012 heeft de Zuidelijke Rekenkamer haar onderzoeksrapport “Beleidsvaluaties provincie Noord-Brabant (vervolgonderzoek)” gepubliceerd en verzonden aan alle leden van Uw Staten.

In dit rapport is, in vervolg op eerdere rapporten uit 2006 en 2008, nagegaan wat de huidige stand van zaken is ten aanzien van het instrument beleidsvaluatie en de eerder door Gedeputeerde Staten gedane toezeggingen naar aanleiding van de eerdere genoemde rapporten.

De Rekenkamer constateert aan de hand van een vijftal hoofdvragen, dat er in de randvoorwaardelijke sfeer belangrijke stappen zijn gezet door de provincie. Wel dienen er nog verbeteringen te worden aangebracht. Dit is verwoord in een tweetal aanbevelingen aan Provinciale Staten en een drietal aanbevelingen aan Gedeputeerde Staten.

Uit de reactie van Gedeputeerde Staten blijkt, dat het college behoefte heeft aan een hernieuwd (vervolg)onderzoek en advisering ten aanzien van beleidsvaluatie vanuit een multi-actorenperspectief. En daarnaast vindt zij de beperking tot een onderzoek ‘technische kwaliteit’ van beleidsvaluaties een gemiste kans.

1. Inleiding

Algemeen

Op 14 mei 2012 heeft de Zuidelijke Rekenkamer haar onderzoeksrapport “Beleidsvaluaties provincie Noord-Brabant

(vervolgonderzoek)” gepubliceerd en verzonden aan alle leden van Uw Staten.

In dit rapport is, in follow-up van eerdere rapporten uit 2006 en 2008, nagegaan wat de huidige stand van zaken is ten aanzien van het instrument beleidsevaluatie en de door Gedeputeerde Staten gedane toezeggingen naar aanleiding van de eerdere genoemde rapporten.

Datum

15 augustus 2012

Documentnummer

3079855

Onderzoeksvragen

Het onderzoek richt zich op vijf, vooraf vastgestelde, hoofdvragen:

Stand van zaken gebruik instrument beleidsevaluaties

1. In hoeverre is binnen de provincie Noord-Brabant sprake van een onderzoeksbeleid?
2. Op welke wijze is binnen de provincie Noord-Brabant invulling gegeven aan de organisatorische inbedding van het instrument beleidsevaluatie?
3. Op welke wijze vindt binnen de provincie Noord-Brabant de aansturing en beheersing plaats van beleidsevaluaties die door, of in opdracht van, GS zijn uitgevoerd?
4. Wat is de stand van zaken met betrekking tot de kwaliteit van het proces en het product, de bruikbaarheid en de benutting van beleidsevaluaties die door, of in opdracht van, GS zijn uitgevoerd binnen de provincie Noord-Brabant?

Nakomen toezeggingen

5. In hoeverre en op welke wijze zijn GS hun gedane toezeggingen uit de rapporten *Focus op beleid* (2006; 10 stuks) en *Quickscan Evaluatie Informatieve weg* (2008; 2 stuks) nagekomen?

Conclusies

Het onderzoek leidt volgens de ZRK tot een aantal conclusies bij de geformuleerde hoofdvragen (hoofdstuk 3, paragraaf 3.1. van deel I); samengevat:

1. Hoewel de provincie Noord-Brabant sinds 2007 beschikt over een onderzoeksbeleid, heeft dit volgens de rekenkamer op onderdelen verbetering.
2. Binnen de provincie wordt *in opzet* op een juiste wijze invulling gegeven aan de organisatorische inbedding van het instrument beleidsevaluatie. Wel vindt de rekenkamer dat onvoldoende sprake is van centrale sturing. *De uitvoering* van het onderzoeksbeleid heeft beslist verdere verbetering. De rekenkamer plaatst onder andere kanttekeningen bij:
 - a. het feit dat men er niet in is geslaagd om een onderzoeksprogrammering te realiseren zoals aangekondigd;
 - b. het feit dat men er niet in is geslaagd om de focus (van aandacht, sturing en toezicht) te verbreden van 217a naar andere onderzoeken;

- c. de kwaliteit van de 217a-onderzoeksplannen;
 - d. de kwaliteit van de onderzoeken;
 - e. het ontbreken van inzicht in de totale geraamde en feitelijke kosten van onderzoek, en in de geraamde en feitelijke inzet van eigen personeel bij onderzoeken en de kosten daarvan.
3. De aansturing en beheersing van de onderzochte beleidsevaluaties hebben op een correcte wijze, dat wil zeggen conform de nota Evaluatie van beleid, plaatsgevonden.
 4. Voor wat betreft het proces (uitvoering) en het product (rapportage) zijn de onderzochte evaluaties echter vaak van onvoldoende kwaliteit. Hierdoor boet ook de bruikbaarheid van de uitkomsten in.⁴
 5. GS zijn hun tien toezeggingen uit het rekenkamerrapport *Focus op beleid*, op één na, *in opzet* op een juiste wijze nagekomen. Vier daarvan zijn *in de praktijk* echter onvoldoende nagekomen. Ook aan hun toezeggingen uit het rapport *Quickscan Evaluatie Informatieve weg* hebben GS invulling gegeven.

Datum

15 augustus 2012

Documentnummer

3079855

Aanbevelingen

In het licht van vorenstaande conclusies formuleert de ZRK de volgende aanbevelingen (hoofdstuk 3, paragraaf 3.2. van deel I):

Aanbevelingen aan PS:

1. Vraag GS om het onderzoeksbeleid op korte termijn te actualiseren en voor besluitvorming aan PS voor te leggen. Vraag GS om hierbij ook de aandachtspunten uit voorliggend onderzoek te betrekken en zorg te dragen voor afstemming met andere kaders die zich richten op dezelfde onderwerpen.
2. Zie scherper toe op nakoming van de door GS gedane toezeggingen.

Aanbevelingen aan GS:

1. Verleen de eenheid K&O bij haar definitieve instelling voldoende bevoegdheden en sanctiemogelijkheden voor een effectieve sturing op de (kwaliteit van de) programmering, opzet en uitvoering van onderzoek.
2. Zie toe op een kwalitatief goede uitvoering van het onderzoeksbeleid. Investeer in professionaliteit, besteed daarbij aandacht aan het kritisch begeleiden van onderzoeken die in- en extern worden uitgevoerd en wees realistisch of onderzoeken binnen de gegeven randvoorwaarden (naar inhoud, beschikbare middelen en benodigde tijd) uitgevoerd kunnen worden. Op deze wijze zal de bruikbaarheid (verder) toenemen.
3. Presenteer in de onderzoeksplannen voor elk gepland onderzoek een raming/inschatting van de verwachte totale kosten en geef bij de verantwoording over de realisatie voor elk uitgevoerd onderzoek inzicht in de feitelijke kosten. Geef daarbij inzicht in zowel de externe als interne kosten.

Reactie Gedeputeerde Staten

GS hebben een bestuurlijke reactie gegeven op basis van het concept bestuurlijke rapport (hoofdstuk 4 van deel I); samengevat:

1. het college heeft behoefte aan een hernieuwd (vervolg)onderzoek en advisering ten aanzien van beleidsevaluatie vanuit een multi-actorenperspectief.
2. het college vindt het een gemiste kans, dat het onderzoek beperkt is tot een onderzoek naar de ‘technische kwaliteit’ van beleidsevaluaties.

2. Het voorstel.

Wij stellen u voor te besluiten conform bijgaand ontwerpbesluit.

Presidium van Provinciale Staten van Noord-Brabant,
de voorzitter de griffier

prof. dr. W.B.H.J. van de Donk mw. drs. C.J.M. Dortmans

Auteur: J.W.H.M. Dankers toestel (06) 55686522

Bijgevoegd treft u de volgende bijlagen aan:

1. Beleidsevaluaties provincie Noord-Brabant (vervolgonderzoek); deel I, Bestuurlijk rapport; d.d. 14 mei 2012 (reeds in uw bezit)
2. Evalueren = Leren; Deskundigen aan het woord; verslag van gesprekken met externe deskundigen; d.d. juni 2012 (reeds in uw bezit)

Datum

15 augustus 2012

Documentnummer

3079855

Ontwerp-besluit 58/12 B

Voorgestelde behandeling :
PS-vergadering : 12 oktober 2012
Statencommissie : Commissie voor Economische Zaken en
Bestuur, 14 september 2012

Onderwerp

Beleidsvaluaties provincie Noord-Brabant (vervolgonderzoek);
onderzoeksrapport Zuidelijke Rekenkamer

.

Datum

15 augustus 2012

Documentnummer

3079855

Directie

Griffie

Bijlage(n)

2

Provinciale Staten van Noord-Brabant,

- gelezen het voorstel van het presidium d.d. 15 augustus 2012;
- gelet op het advies van de Statencommissie voor Economische Zaken en Bestuur d.d. 31 augustus 2012;

besluiten:

1. De conclusies en aanbevelingen van de Zuidelijke Rekenkamer met betrekking tot het onderzoek "Beleidsvaluaties provincie Noord-Brabant (vervolgonderzoek)" te onderschrijven.
2. Gedeputeerde Staten de opdracht te geven alle aanbevelingen nader uit te werken en toe te passen.
3. Gedeputeerde Staten opdracht te geven om de uitwerking van het tweede besluit uiterlijk in het eerste kwartaal 2013 ter bespreking voor te leggen aan de Statencommissie voor Economische Zaken en Bestuur.

's-Hertogenbosch, 21 september 2012
Provinciale Staten van Noord-Brabant,
de voorzitter

de griffier

**Beleidsbeoordelingen provincie Noord-Brabant
(vervolgonderzoek)**

Deel I Bestuurlijk rapport

14 mei 2012

ISBN 978-90-8768-034-3

Inhoudsopgave

Leeswijzer	3
1 Inleiding	4
1.1 Aanleiding	4
1.2 Doelstelling en onderzoeksvragen	5
1.3 Onderzoeksaanpak	5
2 Instrument beleidsevaluaties: opzet en werking	8
2.1 Onderzoeksbeleid	8
2.2 Organisatorische inbedding inclusief aansturing en beheersing (opzet)	11
2.3 Aansturing en beheersing beleidsevaluaties (praktijk)	17
2.4 Kwaliteit uitvoering beleidsevaluaties (praktijk)	17
2.5 Nakomen toezeggingen door GS	20
3 Conclusies en aanbevelingen	22
3.1 Conclusies	22
3.2 Aanbevelingen	23
4 Reactie Gedeputeerde Staten	25
5 Nawoord rekenkamer	27

Leeswijzer

De Zuidelijke Rekenkamer heeft in de periode december 2011 - maart 2012 onderzoek verricht naar de stand van zaken van het gebruik van het instrument beleidsevaluatie door de provincie Noord-Brabant. Dit in vervolg op haar onderzoeken uit 2006 en 2008 naar dit onderwerp.

De resultaten van het onderzoek worden in twee deelrapporten weergegeven. Het voorliggende rapport (Deel I) is een bestuurlijke samenvatting van het onderzoek. In hoofdstuk 1 worden de aanleiding van het onderzoek en de onderzoeksopzet beschreven. Hoofdstuk 2 bevat voor de vijf hoofdvragen van het onderzoek de bevindingen op hoofdlijnen en de conclusies waartoe het onderzoek heeft geleid. In hoofdstuk 3 worden de conclusies nog eens op een rij gezet en worden de naar aanleiding daarvan geformuleerde aanbevelingen beschreven. De hoofdstukken 4 en 5 bevatten achtereenvolgens de bestuurlijke reactie van Gedeputeerde Staten (GS) op het onderzoek en het nawoord van de rekenkamer.

Het onderzoek omvat daarnaast een tweede rapport (Deel II). In dit rapport van bevindingen worden de onderzoeksresultaten uitgebreid weergegeven. Dit Deel II kan geraadpleegd worden op de website van de rekenkamer: www.zuidelijkerekenkamer.nl.

1 Inleiding

1.1 Aanleiding

Een adequate inzet van het instrument beleidsevaluatie is belangrijk voor zowel GS als PS (Provinciale Staten). Met de invoering van het dualistische bestuursmodel in 2003 is een scheiding aangebracht in de functies en de bevoegdheden van GS en PS. De uitvoering van beleid is bij GS geconcentreerd en de kaderstelling en controle van beleid bij PS. Voor deze taken van GS en PS is een goed functionerend en bruikbaar systeem van (beleids)evaluaties van belang. (Beleids)evaluaties zijn namelijk een belangrijke bron van informatie voor deze taken. Zo kan de informatie uit evaluaties worden benut voor bijvoorbeeld de ontwikkeling en (bij)sturing van beleid. Uiteraard moet voor dit laatste het beleid wel dusdanig geformuleerd zijn dat ook daadwerkelijk geëvalueerd kan worden. De informatie uit beleidsevaluaties is ook van belang in het kader van de verantwoording van GS aan PS (en de samenleving) over wat zij met de ingezette publieke middelen hebben gedaan en welke resultaten dat heeft opgeleverd.

Om kwalitatief goede informatie te genereren, dienen zowel de opzet als de feitelijke uitvoering en benutting van beleidsevaluaties aan bepaalde eisen te voldoen.

Gelet op het belang van het instrument beleidsevaluatie voor de provincie (inclusief PS) heeft de Zuidelijke Rekenkamer, als één van haar eerste onderzoeken, in 2006 onderzoek verricht naar het gebruik van het instrument beleidsevaluatie door de provincie. Op 15 september 2006 publiceerde de rekenkamer haar onderzoeksresultaten in het rapport *Focus op beleid; het gebruik van het instrument beleidsevaluatie door Gedeputeerde Staten van Noord-Brabant*. De hoofdconclusie luidde dat de wijze waarop binnen de provincie Noord-Brabant gebruik wordt gemaakt van het instrument beleidsevaluatie nadrukkelijk verbetering behoeft. Naar aanleiding van het onderzoek deden GS tien toezeggingen.

In het verlengde van het rapport uit 2006 hebben GS in 2008 de rekenkamer verzocht om de kwaliteit van de evaluatie Informatieve weg te beoordelen met het oog op het opdoen van een leerervaring op het gebied van evaluatieonderzoek. De rekenkamer heeft dit verzoek ingewilligd. Op 5 november 2008 publiceerde ze haar onderzoeksresultaten in het rapport *Quickscan Evaluatie Informatieve weg*. In het rapport wordt onder andere geconcludeerd dat aan de voorwaarden om te komen tot een kwalitatief goede uitvoering van de evaluatie Informatieve weg naar de mening van de rekenkamer onvoldoende is voldaan. Naar aanleiding van deze quickscan deden GS twee toezeggingen.

Mede gelet op het belang van het instrument beleidsevaluatie voor de provincie is de rekenkamer, in vervolg op haar onderzoeken uit 2006 en 2008, nagegaan wat de

huidige stand van zaken is en in hoeverre GS hun toezeggingen uit 2006 en 2008 zijn nagekomen.

1.2 Doelstelling en onderzoeksvragen

Met dit vervolgonderzoek stelt de rekenkamer zich ten doel inzichtelijk te maken in hoeverre de provincie Noord-Brabant beschikt over een in opzet en werking bruikbaar systeem van beleidsevaluaties dat ingezet kan worden om kwalitatief goede informatie te genereren over (de doelmatigheid en doeltreffendheid van) de realisatie van beleidsdoelstellingen. Daarnaast wil ze inzicht geven in hoeverre GS hun toezeggingen naar aanleiding van de rekenkameronderzoeken uit 2006 en 2008 zijn nagekomen.

Het onderzoek richt zich op vijf hoofdvragen. Vier daarvan richten zich op de stand van zaken van het gebruik van het instrument beleidsevaluaties. Deze vragen hebben betrekking op het onderzoeksbeleid, de organisatorische inbedding van het instrument beleidsevaluatie, de aansturing en beheersing, en de kwaliteit, bruikbaarheid en benutting van beleidsevaluaties. De vijfde hoofdvraag richt zich op het nakomen van de toezeggingen uit de rapporten *Focus op beleid* en *Quickscan Evaluatie Informatieve weg*. Hoofdstuk 2 bevat de beantwoording van de hoofdvragen.

1.3 Onderzoeksaanpak

Dit vervolgonderzoek heeft betrekking op de periode 15 september 2006 tot 1 januari 2012. De vijf hoofdvragen zijn beantwoord door middel van documentanalyse, een dossieronderzoek en verdiepende interviews met betrokkenen binnen de provinciale organisatie. Een overzicht van de geraadpleegde documenten en gesprekspartners is opgenomen in bijlage 1, respectievelijk bijlage 2 van Deel II.

Door middel van dossieronderzoek is een aantal beleidsevaluaties beoordeeld (beantwoording hoofdvraag 3 en 4). In de periode 1 januari 2008 tot 1 januari 2012 zijn achttien 217a-evaluaties afgerond en vierentwintig niet-217a evaluaties afgerond en geagendeerd voor een statencommissie (zie bijlage 3 van Deel II). Het gaat om ex post evaluaties die door of in opdracht van GS zijn uitgevoerd en waarin, gedurende de uitvoering van beleid of achteraf, de uitvoering van het beleid wordt getoetst en beoordeeld. De 217a-evaluaties zijn onderzoeken die in het kader van artikel 217a van de Provinciewet zijn uitgevoerd. Dit artikel verplicht GS om periodiek onderzoek uit te (laten) voeren naar de doelmatigheid en doeltreffendheid van het door hen gevoerde bestuur. De rekenkamer heeft de volgende zeven evaluaties geselecteerd:

Titel evaluatierapport	Uitvoerder
<i>217a</i>	
• Aandacht voor armoede. Evaluatie armoedebeleid Noord-Brabant (2008)	Extern
• Evaluatie ex art. 217a Provinciewet van beleid en subsidie van voormalige stortplaatsen in de provincie Noord-Brabant (2009)	Extern
• Evaluatie van de provinciale regelingen Integrale Dorpsontwikkelingsplannen (IDOP) (2010)	Extern
<i>niet-217a</i>	
• Evaluatie GSB Provincie Noord-Brabant (2008)	Extern
• Evaluatie Beleidsnota Uitvoering Flora- en Faunawet 2006 “Gewogen Belangen” (2009)	Intern
• Evaluatie Uitvoeringsprogramma Verkeersveiligheid 2010 (2010)	Extern
• Vertrouwen in Brabant. Evaluatierapport van het bestuursakkoord 2007-2011 (2011)	Deels intern deels extern

Bij de selectie heeft de rekenkamer de volgende criteria gehanteerd:

- beleidscomponent aanwezig;
- verhoudingsgewijs gelijk aantal 217a en niet-217a evaluaties;
- spreiding over beleidsvelden/statencommissies;
- spreiding door de tijd;
- zowel intern als extern uitgevoerde evaluaties.

Eind 2011 is binnen de provinciale organisatie de aanzet gegeven voor een eenheid Kennis & Onderzoek (K&O). Deze eenheid in oprichting zal zich onder andere gaan bezighouden met het instrument beleidsevaluatie. Tijdens interviews die voor dit onderzoek zijn gevoerd, is gebleken dat deze eenheid aan een strategische visie werkt, waarvoor op dat moment werd voorzien dat deze zou worden ingebracht voor de vergadering van de commissie voor Economische Zaken en Bestuur (EZB) van 1 juni 2012.¹ De rekenkamer heeft geconcludeerd dat haar onderzoek meer toegevoegde waarde zal hebben voor PS, maar ook voor GS en de eenheid K&O als het onderzoeksrapport gelijktijdig met de strategische visie beschikbaar is. De rekenkamer heeft daarom besloten haar onderzoek zo in te richten dat het eerder kon worden afgerond dan oorspronkelijk gepland. Daartoe heeft de rekenkamer, in afwijking van de startnotitie, de zeven geselecteerde evaluaties via een quickscan op hoofdlijnen beoordeeld. Daardoor worden voor hoofdvraag 4 alleen opvallende zaken over deze evaluaties aangegeven, die de provincie als aandachtspunten kan betrekken bij toekomstige evaluaties.

Voor hoofdvraag 1 tot en met 4 zijn dezelfde normen toegepast als in de onderzoeken uit 2006 en 2008. Het betreffen algemeen aanvaarde eisen voor onderzoeksbeleid en beleidsevaluaties. Deze waren destijds vervat in een checklist. GS hebben de punten uit deze checklist van de rekenkamer één op één

¹ In de week dat de rekenkamer het conceptrapport van bevindingen afrondde, werd bekend dat deze planning is aangepast. De visie zal naar verwachting worden ingebracht voor de commissievergadering van 31 augustus 2012.

overgenomen in de, door hen in 2007 vastgestelde, nota *Evaluatie van beleid; Toepassing van beleidsevaluatie binnen de provincie Noord-Brabant*. Voor de beoordeling van de hoofdvragen, en de kwaliteit van de geselecteerde evaluaties (hoofdvraag 4) in het bijzonder, heeft de rekenkamer dan ook de richtlijnen van de provincie als uitgangspunt genomen. Voor de quickscan van de evaluaties heeft de rekenkamer uit deze (50) richtlijnen 16 aandachtspunten gedestilleerd. In bijlage 4 van Deel II zijn deze aandachtspunten weergegeven.

Bij hoofdvraag 5, het door GS nakomen van hun toezeggingen, vormen de toezeggingen van GS uit 2006 en 2008 het uitgangspunt voor de beoordeling.

2 Instrument beleidsevaluaties: opzet en werking

In dit hoofdstuk wordt antwoord gegeven op de vijf hoofdvragen van het onderzoek. De antwoorden zijn afgeleid uit de bevindingen van het onderzoek zoals die in hoofdstuk 2 en 3 van Deel II zijn beschreven. In de paragrafen 2.1 tot en met 2.5 worden voor respectievelijk hoofdvraag 1, 2, 3, 4 en 5 de bevindingen uit het onderzoek op hoofdlijnen en de conclusies die daaruit voortkomen, beschreven.

2.1 Onderzoeksbeleid

De eerste hoofdvraag van het onderzoek luidt: *in hoeverre is binnen de provincie Noord-Brabant sprake van een onderzoeksbeleid?*

De rekenkamer constateert dat de provincie Noord-Brabant haar onderzoeksbeleid heeft geformuleerd in de nota *Evaluatie van beleid; Toepassing van beleidsevaluatie binnen de provincie Noord-Brabant*. Deze nota is op 26 juni 2007 door GS vastgesteld en op 19 oktober 2007 en 22 februari 2008 besproken in de commissie voor Bestuur en Middelen (BM). In de nota geven GS aan dat het om een kaderstellende nota gaat. De nota is niet voor een PS-vergadering geagendeerd en niet door PS vastgesteld.

Verder beschikt de provincie over een aantal kaders die ook relevant zijn voor voorliggend onderzoek:

- Voor 217a-onderzoek hebben PS in 2004 formele kaders vastgesteld in de *Doelmatigheids- en doeltreffendheidsverordening provincie Noord-Brabant* (DDV). De rekenkamer constateert dat de meeste elementen die in de nota *Evaluatie van beleid* voor het onderzoeksplan 217a worden genoemd, overeenkomen met de formele kaders.
- Eind 2006 hebben PS *Sturen met Kaders* (SmK) vastgesteld. Met SmK versterken PS in de voorbereiding van beleidskaders hun kaderstellende rol en in de evaluatie van het beleid hun controlerende rol. Eind 2009 is de invoering en toepassing van SmK geëvalueerd, waarna de beleidslijn op enkele punten is aangescherpt in *Sturen met Kaders; evaluatie*. In overeenstemming met SmK is in de nota *Evaluatie van beleid* vastgelegd dat PS als uitwerking van het bestuursakkoord een meerjarig programmatisch overzicht van kaderstellende nota's dienen op te stellen. Zowel in het programmatisch overzicht als in de kaderstellende nota's moet het moment worden vermeld waarop deze nota's worden geëvalueerd conform de checklist uit het rekenkamerrapport *Focus op beleid*. GS dienen jaarlijks, op basis van het meerjarig overzicht en in afstemming met de onderzoeksplannen van PS en de rekenkamer, een onderzoeksplan 217a vast te stellen.
- In het *Handboek Documentaire Informatievoorziening* (DIV) zijn de organisatiebrede richtlijnen voor dossiervorming opgenomen. Daarnaast hebben

GS, mede naar aanleiding van opmerkingen van de rekenkamer, begin 2007 een nieuw besluit Informatiebeheer vastgesteld dat op 11 mei 2007 door de commissie BM voor kennisgeving is aangenomen. Het besluit regelt de verantwoordelijkheid voor het beheer, de toegankelijkheid, archiefvorming en -ordening van archiefbescheiden en informatiebestanden.

In de nota Evaluatie van beleid wordt onder andere ingegaan op:

- Doel en noodzaak van evaluatie.
- De toepassing van evaluatie, waaronder:
 - o Planning en overzicht van onderzoek.
 - o Reservering van menskracht en/of financiële middelen.
- De onderzoeksorganisatie.
- Het waarborgen van de kwaliteit van evaluatie.
- De benutting van evaluatie-uitkomsten.

Bijlage 1 van de nota bevat het Doelmatigheids- en doeltreffendheidsonderzoekstatuut. In bijlage 2 van de nota is een leidraad opgenomen die is opgesteld in de vorm van een checklist met kwaliteitscriteria waaraan evaluaties en de verslaglegging daarover dienen te voldoen. De provincie heeft, zoals eerder vermeld, in deze leidraad de checklist van de rekenkamer één op één overgenomen. De leidraad dient bij de voorbereiding en uitvoering van alle (beleids)onderzoek te worden toegepast.

Kwaliteit nota Evaluatie van beleid

Richtlijn 1.1 van de leidraad geeft vier elementen waar in het onderzoeksbeleid op moet worden ingegaan. De rekenkamer constateert dat de provincie in haar onderzoeksbeleid op twee van de vier elementen slechts ten dele ingaat. In de nota Evaluatie van beleid ontbreken namelijk:

- richtlijnen over wanneer het beleid wel en wanneer niet geëvalueerd dient te worden. Hierdoor bestaat het risico van een enigszins willekeurige inzet van het instrument beleidsevaluatie. Dit risico wordt nog versterkt doordat elke directie afzonderlijk en zelf bepaalt wanneer beleid wel en wanneer niet wordt geëvalueerd.
- uitgewerkte kwaliteitscriteria voor de verslaglegging over evaluaties (op een enkele uitzondering na).

Daarnaast stelt de rekenkamer vast dat de nota Evaluatie van beleid, met het daarin in de bijlagen opgenomen statuut en leidraad, een concept lijkt te zijn, onvolledig en niet (meer) actueel is en de duidelijkheid ervan op punten te wensen overlaat:

- Concept en onvolledig.

Enkele voorbeelden:

- o In de nota wordt bijvoorbeeld gesteld dat aan de leidraad aspecten worden toegevoegd die nodig zijn en gevolgd moeten worden om een goede dossiervorming van evaluaties te waarborgen. Deze toevoeging heeft (nog) niet plaatsgevonden. De rekenkamer constateert verder dat in de nota, op

een enkel punt na, geen expliciete eisen zijn vastgelegd over welke informatie/documenten in elk geval in het onderzoeksdossier moet worden opgenomen. Hiermee komen GS één van hun toezeggingen uit het rekenkamerrapport inzake dossiervorming ook slechts ten dele na; in de nota wordt niet expliciet ingegaan op vastlegging van informatie in het onderzoeksdossier over de benutting. Overigens stelt de rekenkamer ook vast dat in het Handboek DIV en het besluit Informatiebeheer ook geen regels zijn vastgelegd voor de inhoud van onderzoeksdossiers. Doordat er op dit terrein geen regels zijn vastgelegd en de verantwoordelijkheid voor de volledigheid van de (onderzoeks)dossiers bij de afzonderlijke beleidsafdelingen ligt, bestaat het risico op een willekeurige toepassing. Tegenover de rekenkamer is aangegeven dat de provincie voornemens is te kijken in hoeverre standaardisatie van dossiers mogelijk is, waarbij gekeken zal worden welke documenten standaard in een dossier opgenomen zouden moeten zijn.

- o De commissie BM heeft op 22 februari 2008 ingestemd met een aantal wijzigingen/aanvullingen die in het statuut moeten worden doorgevoerd. De portefeuillehouder heeft daarbij toegezegd om deze te verwerken in het statuut. De rekenkamer constateert dat dit niet is gebeurd.
- o De nota bevat geen eisen voor de wijze waarop achteraf over de uitvoering van het onderzoeksplan 217a verantwoording dient te worden afgelegd.
- Niet (meer) actueel.
De nota is op verschillende punten niet (meer) actueel. Dit is het gevolg van het uit elkaar lopen van opzet en praktijk en veranderde omstandigheden/nieuwe ontwikkelingen. Zie paragraaf 2.2.
- Duidelijkheid laat op punten te wensen over.
 - o De nota is op alle onderzoeken van toepassing, maar lijkt toch vooral geënt op 217a en ex post onderzoeken.
 - o De leidraad bevat de opsomming van eisen die aan evaluaties worden gesteld. De nota omvat echter enkele eisen die niet in de leidraad naar voren komen.
 - o De richtlijnen die in de nota zijn opgenomen voor het onderzoeksplan 217a, komen niet geheel overeen met de eisen die daarvoor zijn opgenomen in de DDV. De DDV omvat namelijk twee elementen die niet expliciet terugkomen in de nota en de nota omvat weer een aantal eisen die niet zijn opgenomen in de DDV.

Conclusie

Op basis van bovenstaande concludeert de rekenkamer voor de eerste hoofdvraag dat hoewel de provincie Noord-Brabant sinds 2007 beschikt over een onderzoeksbeleid, dit volgens de rekenkamer op onderdelen verbetering behoeft.

2.2 Organisatorische inbedding inclusief aansturing en beheersing (opzet)

De tweede hoofdvraag van het onderzoek luidt: *op welke wijze is binnen de provincie Noord-Brabant invulling gegeven aan de organisatorische inbedding van het instrument beleidsevaluatie?*

De rekenkamer constateert dat de provincie Noord-Brabant hieraan op verschillende wijze invulling heeft gegeven. Mede met het oog op een systematische inzet van evaluatie van beleid en het genereren van kwalitatief goede informatie over de doelmatigheid en doeltreffendheid van de realisatie van haar beleidsdoelstellingen:

- hebben PS op 20 februari 2004 in het kader van artikel 217a de DDV vastgesteld;
- hebben PS op 3 november 2006 SmK vastgesteld en ingestemd met de implementatie daarvan;
- hebben GS begin 2007 een nieuw besluit Informatiebeheer vastgesteld;
- hebben GS op 26 juni 2007 de nota Evaluatie van beleid vastgesteld;
- is in de geest van deze nota een onderzoeksorganisatie ingericht, zijn jaarlijks onderzoeksplannen opgesteld en uitgevoerd en is bij de uitvoering van de 217a-onderzoeken toegezien op een juiste toepassing van SmK en de nota (inclusief leidraad met kwaliteitscriteria voor onderzoek);
- is eind 2009 SmK geëvalueerd en hebben GS en PS eind 2010 de aangescherpte beleidslijn *Sturen met Kaders; evaluatie* besproken en omarmd;
- is de aangescherpte beleidslijn geïmplementeerd;
- is in het kader van de benoeming van ‘kennis en onderzoek’ als provinciaal speerpunt, in de zomer van 2011 de aanzet gegeven voor een centrale eenheid K&O;
- worden, via de opleidingsmodule Stukkenloop, beleidsevaluaties onder de aandacht gebracht van alle nieuwe medewerkers.

Onderzoeksorganisatie

De onderzoeksorganisatie is in de praktijk iets anders ingericht dan omschreven in de nota Evaluatie van beleid. Er is een Onderzoekscoördinatieteam (Oct) gevormd dat eind 2011 is opgeheven. Het team bestond, onder verantwoordelijkheid van de concerncontroller, uit een medewerker van de concernstaf welke voorzitter was en één onderzoekscoördinator uit elk van de vijf directies. Bij het vormgeven van het Oct is aangesloten bij het decentrale organisatiemodel van de provincie, waarbij de verantwoordelijkheid voor beleid en onderzoek (inclusief het feitelijk vormgeven van het onderzoeksdoel) bij de directies ligt. In de nota is het idee van meer centrale regie en het bundelen van kennis en ervaring terug te vinden, maar dat bleek, zo werd in gesprekken tegenover de rekenkamer aangegeven, in de praktijk moeilijk te realiseren.

Het Oct stelde jaarlijks in samenwerking met de (beleids)directies het onderzoeksplan 217a op. Voor de onderzoeken uit deze plannen werd een

begeleidingsteam samengesteld onder leiding van een beleidsinhoudelijke directiemedewerker. In deze teams zat ook altijd een lid van het Oct die in onderzoeksmethodisch-technische zin toezicht hield op de uitvoering en kwaliteit van de onderzoeken. Voor deze taken beschikte het Oct niet over doorzettingsmacht/kon niets 'afdwingen', maar had een coördinerende, faciliterende/begeleidende en adviserende rol. Verder was er expliciet voor gekozen om in eerste instantie de focus van het Oct te leggen op 217a-onderzoeken, omdat het 'anders teveel werd'. De rekenkamer stelt vast dat de provinciale organisatie er in de periode 2007-2011 niet in is geslaagd om deze focus te verbreden tot ander onderzoek.

Als richtlijn is in de nota Evaluatie van beleid opgenomen dat onderzoeken in principe onder interne regie samen met één of meer externe partners worden uitgevoerd. In de praktijk worden de meeste onderzoeken volledig uitbesteed. Daarvoor worden verschillende redenen gegeven: het op deze manier beter kunnen waarborgen van onafhankelijkheid en objectiviteit, het ontbreken van een onderzoeksafdeling binnen de provinciale organisatie en in het algemeen onvoldoende onderzoekskennis bij beleidsinhoudelijke medewerkers.

In 2011 was de organisatie naar eigen zeggen klaar om een stap te zetten in de richting van meer centrale regie en het bundelen van kennis en ervaring. Het belang van betere evaluaties en een goede informatiestructuur wordt inmiddels onderkend door de directieraad/GS. 'Kennis en onderzoek' is benoemd tot speerpunt om de wisselwerking tussen kennis en beleid verder te versterken en te professionaliseren. Eind 2009 werd in het 217a-onderzoeksplan 2010 de trendbreuk van een decentrale naar een meer centrale organisatie aangekondigd. In lijn daarmee is in de zomer 2011 de aanzet gegeven voor een eenheid K&O. De taken van het Oct worden ondergebracht bij deze eenheid, welke overigens een bredere opdracht heeft dan het Oct had. Het Oct is opgeheven. Tijdens de uitvoering van voorliggend onderzoek moest het definitieve takenpakket van K&O nog worden opgesteld. De taak zal vooral interveniërend en faciliterend van aard zijn.

De Commissaris van de Koningin is bestuurlijk opdrachtgever van de eenheid en de eenheid valt onder verantwoordelijkheid van de directeur Ruimtelijke Ontwikkeling & Handhaving die ambtelijk opdrachtgever is. De eenheid wordt concernbreed gepositioneerd. In de eenheid K&O is, net als in het Oct, uit elke directie een vertegenwoordiger aanwezig, zodat de link tussen kennis en onderzoek en de verschillende beleidsterreinen blijft bestaan. Deze vertegenwoordigers zijn gemiddeld 1,7 dagen per week voor K&O beschikbaar.

Planning en overzicht

Met SmK besluiten PS dat ingezet en uitgevoerd beleid ook daadwerkelijk geëvalueerd moet worden, om de beleidscirkel te sluiten. In de nota Evaluatie van beleid komt dit terug. Zo is bijvoorbeeld de richtlijn opgenomen dat bij vaststelling van nieuw beleid aangegeven dient te worden of, en zo ja (globaal) hoe het beleid te zijner tijd wordt geëvalueerd. Uit de evaluatie van SmK blijkt echter dat niet

elke beleidscirkel wordt (af)gesloten met een evaluatierapportage en het tijdstip van evaluatie niet altijd in de zogenoemde startnotities is vastgelegd. Naar aanleiding van deze evaluatie geven PS eind 2010 GS opdracht om dit in de toekomst wel te realiseren.

Ondanks het feit dat de wens er was, de toegevoegde waarde werd onderkend en het voornemen in de loop van de tijd verschillende keren is geuit om het onderzoeksplan 217a te relateren aan een meerjarenonderzoeksplanning, is dat er tot op heden nog niet van gekomen. Wel is er (door het Oct) een aantal stappen gezet om de wijze van programmeren te verbeteren. Deze werden voor het eerst gezet in 2007 voor de aanvulling op het 217a-onderzoeksplan 2006. Op basis van een organisatiebrede inventarisatie van geplande onderzoeken uit de beleidscycli in de begroting en onderzoeksplannen van de directies, en rekening houdend met geplande onderzoeken van PS en de Zuidelijke Rekenkamer, werd een aantal onderzoeken geselecteerd die zich leenden voor 217a. Ook de 217a-onderzoeksplannen 2007 tot en met 2011 kwamen op deze wijze tot stand, waarbij het plan een verzameling van losse beleidsevaluaties was. Het 217a-onderzoeksplan voor 2012 is iets anders tot stand gekomen. Het bevat onderwerpen die conform deze 'oude', en onderwerpen die conform een nieuwe insteek zijn geselecteerd. Ook is er meer variatie in type evaluaties; naast ex post ook ex ante en ex durante. Via de nieuwe insteek zijn onderzoeken geselecteerd die voortvloeien uit de kernopgaven van Brabant en het nieuwe Bestuursakkoord en daarmee uit de beleidscyclus, en zich richten op majeure beleidsthema's. Vanuit het team K&O is desgevraagd aangegeven dat ze in de toekomst een concernbreed onderzoeksprogramma op strategisch niveau op willen stellen, zowel voor 217a als niet-217a.

De rekenkamer constateert dat GS de 217a-onderzoeksplannen 2007 tot en met 2012, op één punt na, conform de vereiste uit de DDV, de nota Evaluatie van beleid en de evaluatie van SmK opstellen. Dit ene punt betreft het volgende: in de onderzoeksplannen wordt bij de onderzoeksomschrijving niet altijd voor alle onderzoeken op alle verplichte onderwerpen (object, reikwijdte, onderzoeksmethode, doorlooptijd, wijze van uitvoering en tijdsperiode) ingegaan. Op 19 januari 2007 hebben GS aan de commissie BM toegezegd de tijdsperiode bij de onderzoeksomschrijving op te nemen en bij gelegenheid deze aanvulling op te nemen in de DDV. De rekenkamer stelt vast dat dit vijf jaar later nog niet is gebeurd. Daarnaast constateert de rekenkamer dat in het onderzoeksplan 217a pas drie jaar na de toezegging van GS voor het eerst voor één van de vier geplande onderzoeken wordt ingegaan op de tijdsperiode.

De rekenkamer is van mening dat de duidelijkheid van de onderzoeksplannen 217a op een aantal punten te wensen overlaat. Zo bevatten de onderzoeksplannen 2010 en 2011 bijvoorbeeld wel een meerjarige onderzoeksplanning voor de periode 2010-2015, maar ontbreekt deze, zonder opgaaft van redenen, in het onderzoeksplan 2012. Ook wordt bijvoorbeeld bij deze meerjarige onderzoeksplanningen niet consequent

aangegeven welke 217a-onderzoeken zijn en zijn niet altijd alle geplande 217a-onderzoeken opgenomen.

De rekenkamer acht het een goede ontwikkeling dat met ingang van het 217a-onderzoeksplan 2009 inzicht wordt gegeven in de stand van zaken van de onderzoeken uit het voorgaande onderzoeksplan.

Met het 217a-onderzoeksplan 2012 voldoen GS aan de eis van PS naar aanleiding van de evaluatie van SmK dat het onderzoeksplan 217a bij de begrotingsbehandeling aan PS wordt aangeboden. Deze eis is een aanscherping van de regels uit de DDV/nota Evaluatie van beleid. GS wachten met de aanpassing van de DDV omdat de verwachting is dat door de ontwikkelingen rondom K&O er ook andere aanpassingen nodig zijn.

In tegenstelling tot wat in de nota Evaluatie van beleid is opgenomen, is er geen Databank Onderzoek gekomen. In deze databank zouden alle onderzoeken opgenomen moeten worden om kennis over onderzoeken binnen de provincie te delen, om onderzoekservaringen uit te wisselen en om onderzoeksvragen te bundelen.

Kosten van onderzoek

Conform het gestelde in de nota Evaluatie van beleid worden menskracht en/of financiële middelen gereserveerd en komen ten laste van de bij het beleid behorende begrotingspost. De directies nemen, conform de nota en de DDV, in hun jaarplannen de begrote budgetten op en plannen menskracht in. Omdat in de begroting en jaarstukken de begrote en gemaakte onderzoekskosten niet afzonderlijk zijn opgenomen, kan via deze documenten geen inzicht worden verkregen in de onderzoekskosten. Ook in het onderzoeksplan 217a wordt in het algemeen niet ingegaan op de verwachte kosten, omdat deze, zo is aangegeven, op het moment van het opstellen van het plan vaak nog niet bekend zijn. Bij uitbesteding van onderzoek worden met de uitvoerende externe partijen contracten gesloten, waarmee de kosten in principe vastliggen. De meeste onderzoeken worden extern uitgevoerd. Bij zowel extern als intern (uitgevoerd) onderzoek geldt echter dat voor de inzet van menskracht vanuit de provinciale organisatie geen financiële inschatting wordt gemaakt, zodat er voor (uitgevoerd) onderzoek geen volledig inzicht kan worden gegeven in de beoogde en gerealiseerde financiële middelen en de met de personele inzet gepaard gaande kosten.

Tot eind 2011 had de provincie geen inzicht in de totale kosten van onderzoek. Deze kosten zijn eind 2011 door het team K&O in kaart gebracht. Het gaat daarbij om de inkoop van onderzoek (inclusief beleidsevaluaties) en advies. De tabel op pagina 15 geeft een overzicht voor 2010. Daarnaast is er sprake van (subsidie)relaties met verschillende instellingen, zoals het PON en Telos, die onderzoeks- en adviesdiensten leveren aan de provinciale organisatie.

Inkoop onderzoek/advies ten dienste van:	2010 in € miljoen
- Beleidsvoorbereiding/opstellen van beleid	0,8
- Beleidsuitvoering	6,6
- Monitoren en evalueren van beleid	1,3
Totaal	8,7

Kwaliteit en gebruik

Voor de 217a-onderzoeken geldt dat, conform de nota Evaluatie van beleid, de leidraad wordt toegepast bij de voorbereiding en uitvoering van de onderzoeken. Daarnaast wordt ondersteuning en advies gegeven. Dit alles gebeurde door het Oct. De rekenkamer stelt vast dat in tegenstelling tot voorzien in de nota er geen jaarlijkse beoordeling op kwaliteit, bruikbaarheid en gebruik van een aantal door de provincie uitgevoerde evaluaties wordt uitgevoerd. In gesprekken is tegenover de rekenkamer aangegeven dat niet duidelijk is in hoeverre de leidraad bij niet-217a onderzoeken wordt toegepast. Verder wordt in één gesprek gesteld dat de nota niet echt ‘leeft’ in de organisatie.

In gesprekken is tegenover de rekenkamer verder aangegeven dat door de nota en (de begeleiding van) het Oct de kwaliteit van de evaluaties is verbeterd, als je de periode 2007 vergelijkt met de periode direct daarna. Ook is 217a-onderzoek op de kaart gezet en wordt er serieuzer en professioneler met 217a omgegaan. Hoewel de kwaliteit is toegenomen, is deze, zo wordt ook gesteld, niet altijd even hoog. Bij bespreking van het 217a-onderzoeksplan 2010 plaatste ook de commissie BM een kritische kanttekening bij de resultaten van de onderzoeken. Uit voorliggend onderzoek blijkt eveneens dat nog stappen gezet moeten worden op het gebied van kwaliteit (zie paragraaf 2.4). In gesprekken zijn de volgende knelpunten genoemd: de formulering van de beleidsdoelstellingen, de cultuur die binnen de provinciale organisatie heerst(e): men heeft moeite om kritisch te zijn op zichzelf en op elkaar, en de ‘eigenwijsheid’ in de organisatie.

Ook het gebruik van (het instrument) beleidsevaluatie is volgens de gesprekspartners toegenomen, maar kan beter. De provinciale organisatie is zich bewuster geworden van nut en noodzaak van evaluaties. Maar, zoals eerder opgemerkt, blijkt uit de evaluatie van SmK dat lang niet elke beleidskring wordt (af)gesloten met een evaluatierapportage en het tijdstip van evaluatie niet altijd in de zogenoemde startnotities is vastgelegd. De leidraad schrijft aan GS een aantal acties voor ten behoeve van de benutting van de uitkomsten van de evaluaties. In gesprekken wordt gesteld dat evaluaties met name worden gebruikt om daarna nieuw beleid te formuleren. Als knelpunt voor het gebruik van evaluaties worden genoemd:

- dat evaluaties vaak beschouwd worden als een instrument om te laten zien hoe goed het bestaande beleid was en/of om het opstellen van nieuw beleid te legitimeren;
- dat te weinig wordt nagedacht over de bredere toepasbaarheid van de onderzoeksuitkomsten en dat er te weinig coördinatie op de onderzoeken zit;

- dat evaluaties te weinig flexibel worden ingezet; zich veelal (conform opzet) richten op afspraken die een aantal jaren daarvoor zijn gemaakt en daarbij voorbij wordt gegaan aan bijvoorbeeld een veranderde beleidsopvatting of omgeving.

Binnen de provinciale organisatie is in het kader van de organisatieontwikkeling een aantal trajecten ingezet om te komen tot 'een meer ondernemende en lerende organisatie'. De verwachting van de ambtelijke organisatie is dat deze trajecten een cultuurverandering tot stand zullen brengen en een bijdrage zullen leveren aan het (verder) vergroten van de kwaliteit en het gebruik van beleidsevaluaties. Genoemd wordt bijvoorbeeld de Agenda van Brabant die PS op 11 juni 2010 hebben vastgesteld. Op hoofdlijnen is het doel van deze agenda dat de provincie Noord-Brabant bij de Europese top van kennis- en innovatieregio's blijft behoren. Om dit te bereiken is ook een nieuw profiel van de provincie geformuleerd en is 'kennis en onderzoek' als speerpunt benoemd om de wisselwerking tussen kennis en beleid verder te versterken en te professionaliseren. Om eenheid en coördinatie, waaronder concernbrede aanbevelingen, te borgen is gekozen om K&O op te zetten. De verwachtingen bij de provincie over de impuls die K&O aan het instrument beleidsevaluatie zal geven, zijn hoog.

Stukkenloop

Alle nieuwe medewerkers zijn verplicht om aan de module Stukkenloop deel te nemen. In deze module wordt onder andere ingegaan op beslispunten om de routing van de documenten te bepalen: wanneer gaat een stuk wel en wanneer niet naar een commissie en/of PS en waarvoor (kennisname, besluitvorming). Als medewerkers niet deelnemen, worden ze daarop in eerste instantie zelf aangesproken en indien nodig hun leidinggevende.

In het najaar van 2011 hebben GS een aantal nieuwe afspraken rondom de stukkenloop vastgesteld die vanaf januari 2012 gelden. GS en de directieraad willen met de nieuwe afspraken de kwaliteit en tijdigheid van collegebesluiten en uitgaande stukken verder verbeteren.

Afsluitend

De rekenkamer constateert dat GS niet geheel voldoen aan richtlijn 1.2 uit hun leidraad, volgens welke zij beleidsevaluaties in de praktijk conform het onderzoeksbeleid dienen uit te voeren (zie ook paragraaf 2.4).

Conclusies

Op basis van bovenstaande concludeert de rekenkamer voor de tweede hoofdvraag:

- Binnen de provincie wordt *in opzet* op een juiste wijze invulling gegeven aan de organisatorische inbedding van het instrument beleidsevaluatie. Wel vindt de rekenkamer dat onvoldoende sprake is van centrale sturing.
- *De uitvoering* van het onderzoeksbeleid behoeft beslist verdere verbetering. De rekenkamer plaatst onder andere kanttekeningen bij:

- het feit dat men er niet in is geslaagd om een onderzoeksprogrammering te realiseren zoals aangekondigd;
- het feit dat men er niet in is geslaagd om de focus (van aandacht, sturing en toezicht) te verbreden van 217a naar andere onderzoeken;
- de kwaliteit van de 217a-onderzoeksplannen;
- de kwaliteit van de onderzoeken;
- het ontbreken van inzicht in de totale geraamde en feitelijke kosten van onderzoek en in de geraamde en feitelijke inzet van eigen personeel bij onderzoeken en de kosten daarvan.

2.3 Aansturing en beheersing beleidsevaluaties (praktijk)

De derde hoofdvraag van het onderzoek luidt: *op welke wijze vindt binnen de provincie Noord-Brabant de aansturing en beheersing plaats van beleidsevaluaties die door, of in opdracht van GS zijn uitgevoerd?*

De rekenkamer constateert dat de aansturing en beheersing van de door haar onderzochte evaluaties op hoofdlijnen spoort met hetgeen in de nota Evaluatie van beleid hierover is opgemerkt. Zo is er onder andere gewerkt met begeleidingsteams en hebben GS in de meeste gevallen een besluit genomen over de wijze waarop conclusies en aanbevelingen van de evaluaties dienen te worden benut.

Conclusie

Op basis van bovenstaande concludeert de rekenkamer voor de derde hoofdvraag dat de aansturing en beheersing van de onderzochte beleidsevaluaties op een correcte wijze, dat wil zeggen conform de nota Evaluatie van beleid, hebben plaatsgevonden.

2.4 Kwaliteit uitvoering beleidsevaluaties (praktijk)

De vierde hoofdvraag van het onderzoek luidt: *wat is de stand van zaken met betrekking tot de kwaliteit van het proces en het product, de bruikbaarheid en de benutting van beleidsevaluaties die door, of in opdracht van, GS zijn uitgevoerd binnen de provincie Noord-Brabant?*

De rekenkamer beantwoordt deze vraag, zoals eerder opgemerkt, op hoofdlijnen op basis van een quickscan van zeven evaluaties en de bijbehorende dossiers.

Onderzoeksopzet

Wat betreft de onderzoeksopzet van de evaluaties signaleert de rekenkamer het volgende. Positief is dat de aanleiding en het gebruiksdoel van bijna alle beleidsevaluaties duidelijk geformuleerd zijn en de onderzoeken meestal duidelijk zijn afgebakend. Bij het merendeel van de evaluaties kan daarentegen een

kanttekening geplaatst worden bij de formulering van de doelstelling en de onderzoeksvragen. Bij twee evaluaties is de doelstelling niet helemaal helder en bij twee andere evaluaties is de doelstelling moeilijk toetsbaar. Ook ontbreekt bij twee evaluaties een deel van de doelstelling of onderzoeksvragen. Bij de evaluaties waarin het doel onder andere is om de doelmatigheid van beleid te beoordelen, is dit moeilijk terug te vinden in de onderzoeksvragen. Doelstelling en onderzoeksvragen sluiten op dit punt niet volledig op elkaar aan. Dit heeft ook consequenties voor de onderzoeksuitvoering.

Onderzoeksuitvoering en rapportage

Doordat er geen duidelijke onderzoeksvragen over doelmatigheid worden geformuleerd in die evaluaties waarin dit wel de doelstelling is, wordt hier in de bevindingen en conclusies ook niet op overtuigende wijze op ingegaan. Zo is bijvoorbeeld in twee evaluaties de conclusie over doelmatigheid voornamelijk speculatief. In de ene evaluatie wordt gesproken in termen van ‘waarschijnlijk zou’ en ‘neigt naar’, in de andere van ‘had waarschijnlijk’. In weer een andere evaluatie, waar sowieso geen (duidelijke) conclusie is geformuleerd, wordt in de bevindingen niet specifiek ingegaan op de doelmatigheid.

Bij veel van de onderzochte evaluaties bestaat er (enige) twijfel of een deel van de bevindingen logisch en/of betrouwbaar is. In deze gevallen is meestal niet goed te volgen hoe tot een oordeel wordt gekomen. Zo is bijvoorbeeld in één evaluatie niet te volgen hoe de onderzochte projecten beoordeeld zijn en door wie. Daarnaast is het opmerkelijk dat men in deze evaluatie zelf aangeeft dat getwijfeld wordt aan de betrouwbaarheid van gegevens van een midterm review, maar dat men besluit deze gegevens toch te gebruiken omdat er geen beter alternatief is. Bij een andere evaluatie kan de lezer, op basis van het onderzoeksrapport, niet reconstrueren hoe men voor de onderzochte projecten tot de oordelen is gekomen voor de onderscheiden onderdelen inhoud, procedures en interactie. Ook is niet altijd duidelijk hoe deze oordelen leiden tot het samenvattende oordeel. Bij één project is de beoordeling bijvoorbeeld Inhoud: aandacht, Procedures: negatief en Interactie: negatief. Het samenvattende oordeel luidt vervolgens: ‘Ondanks een moeizaam proces heeft het een goed product opgeleverd’.

De vraag of bevindingen betrouwbaar zijn, kan ook gekoppeld worden aan het gebruik van het onderzoeksinstrument. Er is een aantal beleidsevaluaties waarbij interviews/gesprekken (ook via bijeenkomsten/workshops/enquêtes) de belangrijkste of enige bron van informatie zijn. Daardoor ontstaat er bij de rekenkamer soms twijfel of op deze manier een objectief beeld geschetst wordt. Zo geven bij één evaluatie betrokken provinciale projectleiders zelf aan of de doelen van verschillende activiteiten zijn behaald, zonder dat inzicht wordt gegeven in de wijze waarop de projectleiders tot dit oordeel komen. Bij een andere evaluatie worden alleen personen geïnterviewd, waardoor het de vraag is in hoeverre men op basis daarvan een (objectief) beeld kan geven van effect en doelmatigheid.

Op basis van het dossieronderzoek zet de rekenkamer bij de meeste evaluaties vraagtekens bij de volledigheid van de beantwoording van de onderzoeksvragen. Dit wordt bij twee evaluaties in de hand gewerkt doordat de doelstelling en/of onderzoeksvragen niet goed geformuleerd zijn. In drie andere evaluaties wordt op sommige onderzoeksvragen zeer kort/summier antwoord gegeven of wordt aangegeven dat de vraag niet beantwoord kan worden.

Dossiervorming

Twee van de door de rekenkamer onderzochte dossiers zijn vrij summier, ondanks het uitdrukkelijke verzoek van de contactpersoon van het voorliggende onderzoek om de dossiers door te lopen op compleetheid.

Benutting

Voor de meeste evaluaties is uit de dossiers op te maken dat GS een besluit nemen over de wijze waarop conclusies en aanbevelingen van een onderzoek dienen te worden benut. In een aantal gevallen wordt dit door GS slechts globaal gedaan of worden aanbevelingen niet opgevolgd zonder duidelijke uitleg waarom niet. Bij twee evaluaties blijkt uit het dossier dat de uitkomsten van het onderzoek gecommuniceerd zijn aan betrokkenen. In de door PS vastgestelde DDV is de richtlijn opgenomen dat GS de rapportage en het plan van verbetering ter behandeling aanbieden aan PS. In de nota Evaluatie van beleid staat dat de rapporten ter bespreking aan de betrokken functionele commissie worden gestuurd. De rekenkamer stelt vast dat GS alle onderzochte evaluaties aan de betrokken commissie hebben aangeboden. Vijf evaluaties zijn vervolgens ter bespreking, en de andere twee ter kennisname, geagendeerd in de betrokken commissie. Geen van de evaluaties is geagendeerd voor een PS-vergadering.

Conclusie

Op basis van bovenstaande concludeert de rekenkamer voor de vierde hoofdvraag dat de door de rekenkamer onderzochte beleidsevaluaties voor wat betreft het proces (uitvoering) en het product (rapportage) vaak van onvoldoende kwaliteit zijn. Hierdoor boet ook de bruikbaarheid van de uitkomsten in.²

² De daadwerkelijke benutting maakte geen onderdeel uit van het onderzoek van de rekenkamer.

2.5 Nakomen toezeggingen door GS

De vijfde hoofdvraag van het onderzoek luidt: *in hoeverre en op welke wijze zijn GS hun gedane toezeggingen uit de rapporten 'Focus op beleid' en 'Quickscan Evaluatie Informatieve weg' nagekomen?*

Toezeggingen uit het rekenkamerrapport 'Focus op beleid'

De rekenkamer constateert dat GS:

- (1) met de nota Evaluatie van beleid hun toezegging zijn nagekomen om een onderzoeksbeleid te formuleren waarin expliciet invulling wordt gegeven aan (de opzet van) beleidsevaluaties. Met deze nota komen ze in opzet ook hun toezegging na om dit beleid te verankeren in de organisatie. In de praktijk wordt dit deel van de toezegging deels nagekomen: tot eind 2011 was nog niet echt sprake van centrale sturing en het bundelen van kennis en ervaring, de focus is blijven liggen op alleen de 217a-onderzoeken en het Oct kon niets 'afdwingen';
- (2) met de nota Evaluatie van beleid in opzet hun toezegging zijn nagekomen om te komen tot een meer planmatige en systematische inzet van het instrument beleidsevaluatie.³ In de praktijk wordt deze deels nagekomen daar GS het voornemen om te komen tot een meer planmatige inzet tot op heden nog niet hebben gerealiseerd (relateren onderzoeksplan 217a aan meerjarenonderzoeksplanning);
- (3) met de nota Evaluatie van beleid in opzet hun toezegging zijn nagekomen om bij de besluitvorming in de beleidvaststellingsfase extra aandacht te besteden aan de meer systematische en planmatige inzet van het instrument beleidsevaluatie (gesloten beleidscyclus op de diverse beleidsterreinen). GS hebben ervoor gekozen om dit punt niet alleen mee te nemen in de nadere uitwerking van het onderzoeksplan 217a (zoals toegezegd), maar van toepassing te verklaren voor het totale onderzoeksbeleid. In de praktijk zijn GS ook deze toezegging slechts deels nagekomen daar, zoals uit de evaluatie van SmK blijkt, in tegenstelling tot wat geacht wordt, niet elke beleidscirkel wordt (af)gesloten met een evaluatierapportage en het tijdstip van evaluatie niet altijd in de startnotitie is vastgelegd.
- (4) met de nota Evaluatie van beleid en in de praktijk hun toezegging zijn nagekomen om op basis van een breder inzicht van evaluaties die jaarlijks in de organisatie worden uitgevoerd, ieder jaar een aantal beleidsevaluatie-onderzoeken expliciet te benoemen in het onderzoeksplan 217a. Ze hebben echter niet elk jaar, zoals toegezegd een vijftal 217a-onderzoeken benoemd;
- (5) met de nota Evaluatie van beleid hun toezegging zijn nagekomen om bij de ontwikkeling van richtlijnen voor de aanpak en uitvoering van

³ De oorspronkelijke toezegging van GS bevat in aanvulling hierop de zinsnede: "zonder dat dat leidt tot een zodanig tijdsbeslag op de organisatie dat dit ten koste zou gaan van de uitvoering van het primair proces". De rekenkamer heeft in het kader van haar onderzoek niet kunnen vaststellen of ook aan dit element van de toezegging door GS is voldaan.

evaluatieonderzoeken het door de rekenkamer gepresenteerde normenkader en de daarvan afgeleide checklist mee te nemen;

- (6) met de nota Evaluatie van beleid in opzet hun toezegging zijn nagekomen om de uitkomsten van beleidsevaluaties te benutten. In hoeverre GS in de praktijk daadwerkelijk invulling (hebben ge)geven aan deze toezegging, maakte geen onderdeel uit van het onderzoek van de rekenkamer;
- (7) met de nota Evaluatie van beleid deels hun toezegging zijn nagekomen om bij de ontwikkeling van de richtlijnen voor de aanpak en uitvoering van evaluatieonderzoeken ook aandacht te schenken aan de vastlegging van informatie in het dossier over de benutting van de uitkomsten van beleidsevaluaties en de mogelijke consequenties voor andere beleidsterreinen. En aan het laatste aspect ook aandacht te besteden bij de behandeling van de uitkomsten in het college van GS. In de nota wordt alleen expliciet aandacht gevraagd voor mogelijke consequenties voor andere beleidsterreinen;
- (8) hun toezegging zijn nagekomen om voor het jaar 2006, op basis van een organisatiebrede inventarisatie, een aantal onderzoeken te selecteren die in de tweede helft 2006/begin 2007 worden uitgevoerd. Eind 2006 hebben GS namelijk op basis van een inventarisatie vier onderzoeken toegevoegd aan het 217a-onderzoeksplan 2006;
- (9) hun toezegging zijn nagekomen om de organisatiebrede richtlijnen ten aanzien van dossiervorming en de toepassing daarvan kritisch tegen het licht te houden. GS hebben in 2007 namelijk door een externe partij een onderzoek laten uitvoeren naar (de wijze waarop in de praktijk met) de richtlijnen uit het Handboek DIV wordt omgegaan;
- (10) hun toezegging zijn nagekomen om toe te (laten) zien op verplichte en/of periodieke deelname van medewerkers aan de module 'stukkenloop'.

Toezeggingen uit het rekenkamerrapport 'Quickscan Evaluatie Informatieve weg'

De rekenkamer constateert dat GS:

- (1) hun toezegging om de aanbevelingen van de rekenkamer te gebruiken bij de lopende oriëntatie van de provincie op een verbetering van (de kwaliteit van) het evaluatieonderzoek in de toekomst, hebben ingevuld door de in 2007 geïntroduceerde werkwijze voort te zetten;
- (2) hun toezegging om de rekenkamer te informeren over de voortgang van het proces om te komen tot verhoging van de kwaliteit van onderzoek, hebben ingevuld door de rekenkamer te informeren over de voorgenomen en afgeronde 217a-onderzoeken door toezending van respectievelijk de onderzoeksplannen 217a en de onderzoeksrapporten.

Conclusie

Op basis van bovenstaande concludeert de rekenkamer voor de vijfde hoofdvraag dat GS hun tien toezeggingen uit het rekenkamerrapport *Focus op beleid*, op één na, *in opzet* op een juiste wijze zijn nagekomen, maar vier daarvan *in de praktijk* onvoldoende zijn nagekomen. Ook aan hun toezeggingen uit het rapport *Quickscan Evaluatie Informatieve weg* hebben GS invulling gegeven.

3 Conclusies en aanbevelingen

In dit hoofdstuk worden de conclusies nog eens op een rij gezet (paragraaf 3.1) en worden naar aanleiding daarvan aanbevelingen geformuleerd (paragraaf 3.2).

De rekenkamer constateert alles overziend dat in vergelijking met haar onderzoek uit 2006 de provincie Noord-Brabant in randvoorwaardelijke sfeer enkele belangrijke stappen heeft gezet. Zo hebben GS in 2007 een onderzoeksbeleid vastgesteld en een onderzoeksorganisatie ingericht voor het programmeren, uitvoeren van en toezicht houden op 217a-onderzoeken. De provincie beschikt op dit moment over een systeem van beleidsevaluaties dat ingezet kan worden om kwalitatief goede informatie te genereren over (de doelmatigheid en doeltreffendheid van) de realisatie van beleidsdoelstellingen. Dit systeem behoeft echter zowel in opzet als in de praktische toepassing nog verbetering. Zo geldt onder andere dat de provinciale organisatie er niet in is geslaagd om onderzoeken en de programmering daarvan meer centraal aan te sturen, om het voornemen ten aanzien van een betere wijze van programmeren te realiseren en om de focus die lag op 217a-onderzoek te verbreden tot ander onderzoek. Verder is de kwaliteit van de 217a-onderzoeksplannen en onderzoeken weliswaar iets verbeterd, maar is op onderdelen beslist verdere verbetering nodig. Sinds eind 2011 beschikt de provincie over een team K&O dat een en ander sterk moet verbeteren.

3.1 Conclusies

De conclusies bij de vijf hoofdvragen van dit onderzoek luiden als volgt:

1. Hoewel de provincie Noord-Brabant sinds 2007 beschikt over een onderzoeksbeleid, behoeft dit volgens de rekenkamer op onderdelen verbetering.
2. Binnen de provincie wordt *in opzet* op een juiste wijze invulling gegeven aan de organisatorische inbedding van het instrument beleidsevaluatie. Wel vindt de rekenkamer dat onvoldoende sprake is van centrale sturing.

De uitvoering van het onderzoeksbeleid behoeft beslist verdere verbetering. De rekenkamer plaatst onder andere kanttekeningen bij:

- het feit dat men er niet in is geslaagd om een onderzoeksprogrammering te realiseren zoals aangekondigd;
- het feit dat men er niet in is geslaagd om de focus (van aandacht, sturing en toezicht) te verbreden van 217a naar andere onderzoeken;
- de kwaliteit van de 217a-onderzoeksplannen;
- de kwaliteit van de onderzoeken;
- het ontbreken van inzicht in de totale geraamde en feitelijke kosten van onderzoek, en in de geraamde en feitelijke inzet van eigen personeel bij

onderzoeken en de kosten daarvan.

3. De aansturing en beheersing van de onderzochte beleidsevaluaties hebben op een correcte wijze, dat wil zeggen conform de nota Evaluatie van beleid, plaatsgevonden.
4. Voor wat betreft het proces (uitvoering) en het product (rapportage) zijn de onderzochte evaluaties echter vaak van onvoldoende kwaliteit. Hierdoor boet ook de bruikbaarheid van de uitkomsten in.⁴
5. GS zijn hun tien toezeggingen uit het rekenkamerrapport *Focus op beleid*, op één na, *in opzet* op een juiste wijze nagekomen. Vier daarvan zijn *in de praktijk* echter onvoldoende nagekomen. Ook aan hun toezeggingen uit het rapport *Quickscan Evaluatie Informatieve weg* hebben GS invulling gegeven.

3.2 Aanbevelingen

Op basis van haar bevindingen en daarop gebaseerde conclusies formuleert de rekenkamer de volgende aanbevelingen.

Aanbevelingen aan PS:

- 1 Vraag GS om het onderzoeksbeleid op korte termijn te actualiseren en voor besluitvorming aan PS voor te leggen. Vraag GS om hierbij ook de aandachtspunten uit voorliggend onderzoek te betrekken en zorg te dragen voor afstemming met andere kaders die zich richten op dezelfde onderwerpen.
- 2 Zie scherper toe op nakoming van de door GS gedane toezeggingen.

Aanbevelingen aan GS:

- 1 Verleen de eenheid K&O bij haar definitieve instelling voldoende bevoegdheden en sanctiemogelijkheden voor een effectieve sturing op de (kwaliteit van de) programmering, opzet en uitvoering van onderzoek.
- 2 Zie toe op een kwalitatief goede uitvoering van het onderzoeksbeleid. Investeer in professionaliteit, besteed daarbij aandacht aan het kritisch begeleiden van onderzoeken die in- en extern worden uitgevoerd en wees realistisch of onderzoeken binnen de gegeven randvoorwaarden (naar inhoud, beschikbare middelen en benodigde tijd) uitgevoerd kunnen worden.⁵ Op deze wijze zal de bruikbaarheid (verder) toenemen.

⁴ De daadwerkelijke benutting maakte geen onderdeel uit van het onderzoek van de rekenkamer.

⁵ Zie in dit verband ook een eerdere publicatie van de Zuidelijke Rekenkamer: Criteria voor het ontvankelijk verklaren van onderzoeksvragen door Provinciale Staten Limburg, maart 2011.

- 3 Presenteer in de onderzoeksplannen voor elk gepland onderzoek een raming/inschatting van de verwachte totale kosten en geef bij de verantwoording over de realisatie voor elk uitgevoerd onderzoek inzicht in de feitelijke kosten. Geef daarbij inzicht in zowel de externe als interne kosten.

4 Reactie Gedeputeerde Staten

Op 9 mei 2012 ontving de rekenkamer de volgende reactie van GS van Noord-Brabant:

Allereerst willen wij opmerken dat uw rapport op een gunstig moment verschijnt. De thematiek van het rapport onderstreept de bredere beweging die wij als bestuur in gang hebben gezet rondom kennis en onderzoek. Kennis is ons inziens cruciaal om te komen tot toekomstbestendig provinciaal beleid in een omgeving die soms al even onkenbaar als onberekenbaar is. Een omgeving ook, waarin wij ons geconfronteerd zien met complexe maatschappelijke opgaven. Om die opgaven een stap verder te brengen is kennis cruciaal. Met het bestuurlijk speerpunt 'Kennis en Onderzoek' onderstrepen wij dat belang. We zien 'Kennis en Onderzoek' niet alleen als een instrument om het leervermogen van de provinciale organisatie te versterken, maar ook als middel om te komen tot publieke innovatie. Het enorme belang dat wij allen hechten aan de Brabantse kenniseconomie vraagt niet enkel om investeringen van het bedrijfsleven en onderwijs, maar betekent ook het nodige voor onze eigen werkwijzen.

Ten aanzien van de wijze waarop u in uw rapport beleidsevaluaties beoordeelt, hebben wij een tweetal opmerkingen die te maken hebben met respectievelijk de inhoudelijke en de technische afbakening van het onderzoek. Deze opmerkingen resulteren in een aantal vragen, die we bij u onder de aandacht willen brengen.

Om met de eerste opmerking te beginnen: in een omgeving waarin economische en maatschappelijke ontwikkelingen elkaar steeds sneller opvolgen is het effectief kunnen anticiperen op de relevante ontwikkeling van groot belang. Dat geldt ook voor beleidsevaluaties. Het gaat hierbij wat ons betreft niet alleen om het kunnen verantwoorden van gevoerd beleid aan Provinciale Staten en - in bredere zin - de Brabantse samenleving, maar ook om het leren gericht op toekomstig beleid. Wij constateren daarbij dat beleid steeds minder vanuit een één-actor benadering wordt ingezet. Het is niet meer de provincie die alleen bepaalt (zo dat al ooit het geval was); steeds meer wordt beleid samen met andere partners in en buiten Brabant vormgegeven en uitgevoerd. We realiseren ons immers terdege dat we in een tijd leven waarin het steeds minder mogelijk is de toekomst met louter provinciale plannen te beheersen. De wereld om ons heen verandert dermate snel, dat het tegenwoordig meer op veerkracht, allianties en improvisatievermogen neerkomt, dan op klassieke plannen en systemen.

De trend naar *governance* (meer partijen die betrokken zijn bij besluitvorming) is al geruime tijd ingezet, maar heeft met de Agenda van Brabant en het bestuursakkoord Tien voor Brabant een extra accent gekregen. Dit zou ons inziens echter ook gevolgen moeten hebben voor de evaluatie van beleid. De klassieke evaluatiemethoden (doelevaluaties) zijn gebaseerd op één-actorbenaderingen,

waarbij het formuleren van beleidsdoelstellingen en de verantwoordelijkheid voor het doelbereik ligt bij één (provincie)actor. In een dergelijke context is de afrekenbaarheid van beleid redelijk eenduidig vast te stellen. Dat wordt al lastiger in een situatie waarin meerdere partners betrokken zijn bij het formuleren en uitvoeren van beleid. Afrekenbaarheid blijft ook dan van belang, maar in een governance-omgeving zouden ook andere dimensies van beleid aan de orde moeten komen, zoals ‘proces’ en ‘leren’. Wij vinden het om deze reden opmerkelijk dat u bij de beoordeling van evaluaties hier geen rekening mee lijkt te houden.

Wij zouden het dan ook zeer op prijs stellen als u aan kunt geven op welke wijze wij evaluaties kunnen uitvoeren, rekening houdend met een multi-actorenperspectief. Hoe is het repertoire aan traditionele doe-evaluaties te verbreden en wellicht te combineren met meer pluricentrische evaluaties of zelfs responsieve evaluaties?

In het verlengde van het voorgaande, betreft de tweede opmerking de beperking van uw onderzoek tot de ‘technische kwaliteit’ van beleidsevaluaties. Dat vinden wij een gemiste kans. Evaluaties voeren wij uit om beleid te kunnen verantwoorden en de uitkomsten te benutten voor nieuw beleid. De opvattingen van PS, en daarmee de politieke ‘weging’ van de door ons uitgevoerde beleidsevaluaties lijkt ons daarbij van wezenlijk belang.

Afsluitend. Beleidsonderzoek, waaronder ook beleidsevaluaties, staat voor een omvangrijke veranderingsopgave. Uw rapport onderstreept de uitdaging die wij reeds hebben gesignaleerd en de noodzaak om daar stevig op in te zetten. Wij zullen de aanbevelingen die in uw rapport doet dan ook nadrukkelijk meenemen in het traject rondom ‘Kennis en Onderzoek’ en de daarop volgende organisatorische verankering van de kennis- en onderzoeksfunctie. Zoals aangegeven is het speerpunt ‘Kennis en Onderzoek’ breder dan alleen beleidsevaluatie. Beleidsevaluatie is het sluitstuk van de beleidscyclus terwijl wij ons met kennis en onderzoek richten op het optimaliseren van de gehele beleidscyclus. Tegen deze achtergrond zouden wij het waarderen als u, wanneer u in de nabije toekomst nogmaals komt met een onderzoek over beleidsevaluatie, dit inbedt in het bredere spectrum.

5 Nawoord rekenkamer

In haar reactie schetsen GS de context van het onderzoekbeleid dat zij de komende periode willen vormgeven en de rol van het speerpunt 'kennis en onderzoek' daarbij. Met het oog hierop heeft de rekenkamer haar onderzoek zo ingericht dat de uitkomsten ervan betrokken kunnen worden bij het visiedocument dat de provincie thans ontwikkelt en dat GS in augustus 2012 aan PS willen voorleggen.

GS maken daarnaast enkele opmerkingen over de afbakening van het onderzoek. De rekenkamer merkt daarover het volgende op. Zoals vastgelegd in de startnotitie die voorafgaand aan het onderzoek aan de provincie voor commentaar is aangeboden, was het doel om na te gaan wat de stand van zaken is met betrekking tot (de inzet van) het instrument beleidsevaluatie en wat GS hebben gedaan met eerder gedane toezeggingen. Desalniettemin heeft de rekenkamer besloten om een verzoek van de provincie te honoreren om ook aandacht te besteden aan het bredere onderzoeksspectrum. De rekenkamer heeft in dat verband op zich genomen om nog vóór de bespreking van het visiedocument met PS een extra publicatie te verzorgen, waarin recente inzichten van externe deskundigen op dit terrein worden bijeengebracht. De rekenkamer zal deze publicatie onder de aandacht brengen van PS, GS en de ambtelijke organisatie van de provincie Noord-Brabant en van Limburg, waar de rekenkamer eveneens een vervolgonderzoek Beleidsevaluaties heeft ingesteld.

De rekenkamer constateert met genoegen dat GS het belang van het onderwerp, de uitdaging waarvoor ze daarbij staan en de noodzaak om daar stevig op in te zetten onderkennen en aan het onderwerp een nieuwe impuls willen geven. De rekenkamer is in dat verband verheugd dat GS toezeggen om de aanbevelingen van de rekenkamer mee te nemen in het traject rondom 'kennis en onderzoek'. Zij nodigt PS uit om GS te verzoeken hieraan concrete invulling te geven en in een plan van aanpak vast te leggen, zodat de noodzakelijke verbeteringen thans ook echt gerealiseerd worden.

Vastgesteld door de Zuidelijke Rekenkamer op 14 mei 2012.



drs. L. Markensteyn
voorzitter



drs. P.W.M. de Kroon
directeur-secretaris

Evaluëren = Leren

Deskundigen aan het woord



Evaluëren = Leren

Deskundigen aan het woord



Verslag van drie gesprekken die de Zuidelijke Rekenkamer voerde met externe deskundigen en marge van haar vervolgonderzoeken naar beleidsevaluatie in de provincie Limburg en Noord-Brabant

Eindhoven, juni 2012

Inleiding

En marge van haar vervolgonderzoeken naar de stand van zaken van het gebruik van het instrument beleidsevaluatie door de provincie Limburg en Noord-Brabant heeft de Zuidelijke Rekenkamer in het voorjaar van 2012 gesprekken gevoerd met drie externe deskundigen. Doel van deze gesprekken was om van hen te vernemen wat de stand van het denken over beleidsevaluatie is en voor welke uitdagingen provincies staan die de effectiviteit en efficiëntie van hun functioneren en presteren willen onderzoeken.

De neerslag van deze gesprekken bieden wij hierbij als een aparte publicatie aan onder de titel 'Evalueren = Leren'. Wij vertrouwen erop dat de vele suggesties, kanttekeningen en commentaren van de geïnterviewden betrokken zullen worden bij de verdere vormgeving van beleidsevaluatie door de beide provincies.

De gesprekken vonden plaats rondom de volgende vragen:

- Wat zijn recente ontwikkelingen en trends op het terrein van beleidsevaluatie?
- Hoe veranker je (de aandacht voor) beleidsevaluatie in je organisatie?
- Hoe zorg je ervoor dat de evaluaties kwalitatief aan de maat zijn, hoe geef je een meerjarig onderzoeksprogramma vorm, hoe ga je om met vraagstukken van uitbesteding versus zelf uitvoeren?
- Wat moet je doen om de uitkomsten van evaluaties echt te benutten?
- Wat is de rol voor de volksvertegenwoordiging op het terrein van beleids-evaluatie, hoe organiseer je de aandacht van de volksvertegenwoordiging/ Provinciale Staten voor het onderwerp, hoe kunnen zij sturing geven, hoe zie je het terug in de formele documenten van planning & control?
- Welke vorm van beleidsevaluatie past bij een beleid dat voortdurend in beweging is en waar naast de provincie vele andere actoren in de ontwikkeling en uitvoering van gezamenlijke programma's en projecten actief zijn (Europa, rijksoverheid, gemeenten, bedrijfsleven, maatschappelijke organisaties)?

De drie deskundigen hebben een langdurige expertise op het terrein van beleids-evaluatie. Hun deskundigheid bestrijkt de wereld van onderzoek, wetenschap en publieke sector.

Het zijn Peter van der Knaap, Froukje Boonstra en Frans Leeuw.

Peter van der Knaap: directeur doelmatigheidsonderzoek bij de Algemene Rekenkamer. In een eerdere functie was hij onder meer hoofd Beleidsevaluatie en Instrumentatie van het ministerie van Financiën en verantwoordelijk voor de rijksbrede kaderstelling en uitvoering op het vlak van evaluatieonderzoek en resultaatgericht begroten (VBTB). In 1997 promoveerde hij aan de Erasmus Universiteit Rotterdam op het proefschrift 'Lerende overheid, intelligent beleid', een onderzoek naar de bijdragen van beleidsonderzoek aan het leervermogen van de overheid.

Froukje Boonstra: gepromoveerd op een onderzoek naar besluitvorming in gebiedsgericht beleid aan de toenmalige Katholieke Universiteit Nijmegen. Was coördinator beleidsondersteunend onderzoek van het toenmalige ministerie van Landbouw, Natuur en Voedselkwaliteit voor het thema ILG (Investeringsbudget Landelijk Gebied) en gebiedsprocessen. Momenteel senior bestuurskundige bij Alterra, kennisinstituut voor de groene leefomgeving van Wageningen University & Research centre, waar zij zich richt op het ondersteunen van bestuurlijke vernieuwing in het landelijk gebied, via onderzoek, advies en procesfacilitatie.

Frans Leeuw: directeur van het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het ministerie van Veiligheid en Justitie. Daarnaast is hij hoogleraar 'Recht, Openbaar Bestuur en Sociaal-wetenschappelijk Onderzoek' aan de Maastricht University. In eerdere functies was hij onder meer hoogleraar aan de Universiteit Utrecht met als leeropdracht 'Onderzoek van maatschappelijke effecten van overheidsbeleid', directeur doelmatigheidsonderzoek bij de Algemene Rekenkamer en Hoofdinspecteur, Inspectie van het Onderwijs.

Peter van der Knaap

10 mei 2012

We leven in een netwerksamenleving waarin beleidsinterventies van allerlei partijen over elkaar heen buitelen. Dat heeft consequenties voor onderzoek naar maatschappelijke effecten van beleid en vraagt om een ‘slimme’ evaluatiebenadering. Een benadering waarbij je steeds opnieuw voor de vraag komt te staan wat voor soort evaluatieonderzoek, in welk stadium, gegeven de specifieke omstandigheden, het meest geëigend is.

Als je voor je beleid afhankelijk bent van andere actoren, dan kun je in ieder geval beginnen met de vraag hoe de processen zijn ingericht, hoe de overdracht van bijvoorbeeld dossiers tussen de betrokken instanties is geregeld. Dat soort onderzoek moet eigenlijk vast onderdeel van het beleidsproces zijn. Juist bij veel schakels in de keten moet je alert zijn op de verhouding tussen de schakels onderling. De zwakste schakel bepaalt immers de kracht van de ketting. Dat betekent dat je aan de voorkant goed nadenkt hoe je het beleid inricht en er op bedacht bent welke flessenhalzen je kunt tegenkomen. En als je het beleid vervolgens aan het uitrollen bent, moet je in de gaten houden of, en waar de knelpunten zich in de praktijk voordoen. Dat soort inzichten heb je nodig om te kunnen bijsturen. Heb je dat niet, dan vraag je om problemen.

Een aantal jaren geleden zijn we bij de Algemene Rekenkamer begonnen met een ander soort onderzoek waarmee we achteraf heel gericht kijken naar verklaringen voor behaalde resultaten. Daaruit hebben we geleerd dat dit soort informatieproblemen in netwerken en samenwerkingsverbanden een belangrijke verklaring vormen voor achteraf tegenvallende resultaten. Dat onderstreept dus het belang van een *ongoing evaluation* of evaluatie *ex durante*. Je hebt dat nodig om goed te kunnen sturen. Begrijp me goed, ik vind niet dat het in de plaats mag komen van effectevaluaties. Het gaat er alleen aan vooraf, als een noodzakelijke voorwaarde voor goede sturing.

Heilige graal

Er is natuurlijk veel op te merken over voor- en nadelen van verschillende vormen van evaluatie. De heilige graal, dat zijn *randomized controlled trials*. Je hebt een voormeting en een nameting en je hebt een experimentele groep en een controlegroep. In tegenstelling tot wat veel mensen denken kun je de beleidspraktijk vaak wel zo inrichten, het ministerie van Onderwijs, Cultuur en Wetenschap is wat dat betreft goed bezig. Maar dat kan niet altijd en het kan ook kostbaar zijn. Je moet in veel gevallen wachten voordat je weet wat eruit komt. En je kunt erover discussiëren of het altijd maatschappelijk aanvaardbaar is. Er zijn tal van argumenten om het niet te doen, maar toch kom je het soms tegen en dat getuigt van moed en visie. Als Algemene Rekenkamer doen wij het niet, omdat wij geen beleid maken.

Wat wij wel doen, is achteraf kijken, bijvoorbeeld naar de verschillen in kosten-effectiviteit van beleidsinstrumenten. Zo hebben we laatst onderzoek gedaan naar de kosteneffectiviteit van verschillende maatregelen op het gebied van energiebesparing.

Geïnformeerd

Het is verstandig om kleinschalig te beginnen. De commissie Dijsselbloem, die het parlementair onderzoek deed naar de onderwijsvernieuwingen, kwam ook met dat advies. Maar al te vaak wordt nieuw beleid heel breed uitgerold, zonder dat ook maar iets bekend is over de vraag of het werkt of niet. De conclusie die je daaraan moet verbinden, is: ga in ieder geval goed om je heen kijken wat er over een onderwerp al bekend is. Dat heet met een sjiek woord *evidence informed* beleid. Zodat wat je doet minimaal is gestoeld op aannames die geworteld zijn in kennis over wat wel en niet werkt. En richt het dan als het even kan zo in dat je zinvolle opmerkingen kunt maken over de effectiviteit.

Op dat punt kan er veel meer dan vaak wordt gedacht. Een voorbeeld: je hebt een regeling die van toepassing is op mensen van 55+ en je weet na verloop van tijd hoeveel mensen daar gebruik van hebben gemaakt. Dan kun je vervolgens, ten minste als je een grote populatie hebt, met wat correcties nagaan hoe het zit met 53-jarigen en 54-jarigen. En kun je uitspraken doen over de kans dat ook zij van die regeling gebruik zouden willen maken. Het is een vorm van *impact evaluation*. Hessel Oosterbeek van de Universiteit van Amsterdam is iemand die daarin is gespecialiseerd. Eigenlijk zijn het natuurlijke experimenten, weliswaar niet als zodanig opgezet, maar met een onderzoeksbril kun je er wel zo naar kijken. Het is ook een vorm van evaluatie die veel meer toepasbaar is dan de klassieke vooropgezette controlegroep of experimentele groep; minder wetenschappelijk gevalideerd, maar je kunt er een heel eind mee komen en ook nog eens tegen minder kosten. Hetzelfde geldt voor *syntheseonderzoek*, dat verschillende (deel)onderzoeken combineert. Dan kijk je wat er elders al gedaan is, wat de ervaringen zijn met programma's die wel of niet tot succes hebben geleid.

Misschien kun je zeggen dat dat voor provincies allemaal wat hoog gegrepen is. Maar vergis je niet, op het gebied van ruimtelijke ordening en openbaar vervoer bijvoorbeeld, is er op provinciaal niveau zeer veel informatie beschikbaar. De inzichten die je daar uit kunt halen, zou je al dan niet kunnen bundelen binnen het Interprovinciaal Overleg of in andere samenwerkingsverbanden. Vandaag de dag hoeft je ook niet meer de fysieke bibliotheek in, je hoeft het niet allemaal zelf te onderzoeken of in een experimentele vorm te gieten. Het is van het grootste belang om vooral in de voorkant van het beleid te investeren. Dat is de beste manier

om enige mate van zekerheid te krijgen dat de interventie die je wilt plegen in ieder geval aangrijpt op de belangrijkste factoren.

Verkeersveiligheid is in dat verband een aardig voorbeeld. We weten uit onderzoek dat decentrale wegbeheerders niet altijd de meest moderne inzichten toepassen bij de inrichting van rotondes. Terwijl die informatie er wel is. Wat mij betreft is er geen excuus om die beschikbare kennis niet toe te passen. Je hoeft echt geen tonnen uit te geven om aan dat inzicht te komen, dat is gewoon beschikbaar.

Zinnvolle evaluaties beginnen met serieuze aandacht voor wat je wilt bereiken. *Factfull politics*, geïnformeerde besluitvorming en het zodanig vormgeven en uitwerken van je beleid dat je je niet achteraf vertwijfeld hoeft af te vragen wat nou eigenlijk ook alweer het doel van het beleid was en hoe je dat kunt operationaliseren. Dat moet je voor een belangrijk gedeelte al uit je bestaande stuurinformatie kunnen halen. Dat betekent dat als het over verkeersveiligheid gaat, dat je vanaf het begin ongevallen bijhoudt en dat je op basis daarvan kunt zeggen of een bepaald beleid heeft geleid tot een significante daling van ongelukken.

Een andere belangrijke vraag is hoe je beter kunt nagaan wat het specifieke effect is van de inzet van jouw middelen als provincie, temidden van middelen die door andere partijen zijn ingebracht, zoals Europa, rijksoverheid, gemeenten, bedrijfsleven en maatschappelijke organisaties. Succes heeft vele vaders. Als het fout gaat, ligt dat anders. Tegelijkertijd, als je pas achteraf ontdekt dat andere ketenpartners niet hebben geleverd en dat je daarom bepaalde doelen niet hebt gehaald, dan heb je zelf je werk al die tijd niet goed gedaan. Dan is het alsof je aan de lopende band staat en zegt: ik breng de spatborden aan en of de auto aan het eind van de productielijn ook echt kan rijden, daar ga ik niet over. Modern bestuur kan zo niet werken. Ook al ben je niet de regisseur van de keten - wat de provincie overigens vaak wel is - dan nog heb je een verantwoordelijkheid om in de gaten te houden hoe het met het geheel zit. Op het moment dat de andere ketenpartners verzaken of het beoogde effect niet wordt bereikt, dan zul je toch echt je hand moeten opsteken. En al helemaal als je de regie wél hebt; dan kun je niet achteraf zeggen: we constateren nu vier of vijf jaar later dat het niet is gelukt, maar by the way ... het ligt niet aan ons maar aan 'hullie'. Dat kan niet. Je zult zeker moeten stellen dat de keten in zijn geheel goed werkt, anders ben je in je reguliere sturing al niet goed bezig.

Focus op beleidsresultaat

Langzamerhand zien we veranderingen in het soort onderzoek dat we doen. Dat heeft te maken met veranderingen in de beleidsomgeving. Dat er één partij is die

op basis van zo goed mogelijke informatie kan besluiten wat er moet gebeuren, dat vervolgens in zijn eentje kan uitvoeren en achteraf ook nog zelf vaststelt of het heeft gewerkt, dat kom je in de beleidspraktijk bijna niet meer tegen.

Maar deze ontwikkeling lijkt ook wel een 'excustruus' te worden voor overheden en bestuurders waar ze zich achter kunnen verschuilen. Hoe meer partijen een gedeelde verantwoordelijkheid hebben, hoe minder vaak er eentje is die zegt: ik ga er over. Op allerlei terreinen waar sprake is van decentralisatie, lijkt de rijksoverheid zich terug te trekken op een soort systeemverantwoordelijkheid: wij zijn er voor de voorlichting, voor het overmaken van het geld en misschien nog voor het maken van afspraken. Maar wat er uiteindelijk bereikt wordt, dat is de verantwoordelijkheid van andere overheden, van Brussel of van uitvoerende instanties in het veld.

Dat is natuurlijk een vreemde figuur. In de Tweede Kamer gaat het niet over een convenant dat je hebt afgesloten, maar over het beleidsresultaat zelf, over zwerfafval of over voedselveiligheid. Dan kun je nog zo hard roepen dat zwerfafval iets is voor de gemeenten, maar als je er geld aan uitgeeft en je hebt ambities en je kiest voor dat systeem, dan mag je ook verwachten dat je daar verantwoording over aflegt. Vaak hoor je dan: als we de samenwerking evalueren of we evalueren de mate waarin partners tevreden zijn, dan is dat ook voldoende. Nee dus, dat is nooit voldoende. Ewald Engelen, de financieel geograaf, zegt in dat verband: hoe ingewikkelder het systeem, hoe meer 'beleidslulkoek', waarbij het alleen nog maar gaat om het afsluiten van convenanten en het beleggen van meetings, terwijl het echte onderwerp, zoals bijvoorbeeld het revitaliseren van de lokale economie, naar de achtergrond verdwijnt. En als je dan ook nog eens de procesevaluatie zo inricht dat het op de achtergrond blijft, dan is iedereen heel tevreden, maar je weet één ding zeker: de burger is al lang afgehaakt. En je weet ook niet, en dat is misschien nog wel kwalijker, of wat je doet enig effect heeft. Je zult dus altijd de resultaten moeten kunnen beoordelen.

Dat zien wij als een ontwikkeling voor ons als Algemene Rekenkamer. Dat wij blijven hameren op effecten en daar onderzoek naar doen. En voor onszelf, als instelling die evaluatieonderzoek doet, is de belangrijkste ontwikkeling dat we steeds meer de ontvanger van beleid centraal stellen. Of dat nou bedrijven zijn of burgers, of de natuur, Waarbij de centrale vraag is: wat is het probleem en sluit het beleid aan op dat probleem?

Een voorbeeld: vorig jaar hebben we onderzocht of maatregelen, die waren bedoeld om de veiligheid van winkeliers in het midden- en kleinbedrijf te verbeteren,

of die aansluiten bij wat die winkeliers zelf belangrijk vinden. Dat is een nieuw type onderzoek. We hebben niet zozeer gekeken naar de formele doelstellingen van het Rijk of van veiligheidspartners. Nee, we hebben gekeken hoe het zit met de ontvanger. Kan hij ermee uit de voeten? En vervolgens: wat kunnen we zien wat er in de praktijk gebeurt? Daalt het aantal overvallen? We kijken dus niet alleen of het aansluit, maar ook naar de feiten die daarbij horen. We pellen niet meer de beleidstoren van boven naar beneden af, zoals we altijd hebben gedaan, maar we draaien het om. We proberen naar het beleid te kijken door de ogen van de mensen waar het beleid zich op richt. Dat is een vorm van *responsieve beleidsevaluatie*. De doelgroep van beleid heeft natuurlijk maling aan convenanten en kwaliteitsafspraken. Hen gaat het om het effect. Het is een minder formalistische evaluatie, waarbij je ook open moet staan voor dingen die misschien niet zo bedoeld zijn maar in de praktijk wel goed blijken te werken. Voor rekenkamers en traditionele evaluatoren blijkt dat overigens een lastig waar te maken ambitie.

Organiseren van aandacht

Hoe organiseer je de aandacht voor evaluatie in de bredere zin? Ik denk dat het helemaal onderaan begint bij de kwaliteit van beleidsmedewerkers en een morele verantwoordelijkheid van bestuurders. Soms gaat het daar al mis. Dan zie je bestuurders die vooral geïnteresseerd zijn om de wereld te laten zien waartoe ze zo al in staat zijn en die geen behoefte hebben om na te gaan of datgene wat ze doen ook echt iets uithaalt.

Wat betreft de beleidsmedewerkers kom je bij de vraag of je op de wereld bent om problemen te helpen beheersen of op te lossen, of dat je er bent om je bestuurder te dienen. Idealiter ligt dat in het verlengde van elkaar. Dan is er ook een goede voedingsbodem om je aan de voorkant al bezig te houden met een inhoudelijke vraag als: wat is er eigenlijk bekend – als dit het probleem is dat we willen beheersen of kleiner willen maken – over instrumenten die succesvol zouden kunnen zijn? Om dat te kunnen doen heb je voldoende capaciteit op niveau nodig. Daar maak ik me wel eens zorgen over, zeker bij decentralisaties. Op het terrein van natuurbescherming constateerden we bijvoorbeeld een aantal jaren geleden dat provincies en gemeenten heel weinig mensen hadden met voldoende verstand van natuurbescherming, terwijl zij wel projectaanvragen moeten beoordelen. Dan staat de kwaliteit van de beoordeling onder druk en loopt de natuurbescherming gevaar. Je *standard operating procedures* moeten natuurlijk wel op orde zijn.

Durven leren

Wat ontzettend helpt, is het gebruik van nieuwe media en het organiseren van transparantie. Zodat iedereen met een paar simpele muisklikken kan opvragen wat

er over een onderwerp al bekend is. Het begint bij *evidence informed* besluitvorming. Als je dat niet op orde hebt, dan kun je organiseren wat je wil, je kunt zaken verplicht stellen, maar dan blijf je altijd aan de achterkant bezig. Dat is misschien wel dankbaar werk voor een rekenkamer, maar het is natuurlijk niet de goede manier om duurzaam aandacht voor evaluatie te krijgen.

Het organiseren van interne tegenspraak is in dat verband ook erg belangrijk. Dat gaat ons over het algemeen niet gemakkelijk af. Er is heel lang geoefend met goed verantwoording afleggen. We zijn inmiddels weg uit de fictie dat beleid altijd succesvol zal zijn. Maar naast verantwoording afleggen is het ook van belang dat overheden leren. Ik denk dat de volgende stap is dat je verantwoording zo interpreteert dat verantwoording afgelegd moet worden over het leerproces. Dat betekent ook dat op het moment dat je niet weet of iets werkt of je weet het misschien wel maar in de praktijk blijkt het toch anders te gaan, dat dat dan niet per definitie als een brevet van onvermogen wordt uitgelegd. Bestuurders moeten zich ook verantwoorden over hoe ze ervoor zorgen dat ze dingen die mis gaan tijdig in beeld krijgen. Dat ze niet na vijf jaar volledig verrast worden door een onverwachte boodschap, als ze een evaluatierapport van een extern bureau op hun bureau krijgen.

We zien dat organisaties de neiging hebben om weg te blijven uit een scherpe verantwoording. Dat leidt tot koppensnellen en daar houden we niet van. Een kritische vraag aan een collega: ‘wat een mooi project, maar helpt het nou ook?’ wordt meestal ook niet gesteld. Terwijl je je daarvoor juist open zou moeten stellen vanuit een grondhouding: ‘we weten het niet, maar we proberen het want het is belangrijk dat we er wat aan doen. En als blijkt dat wat we doen niet werkt, dan zijn we daar open in en gaan we het anders proberen. Maar wees er zeker van, beste inwoners van onze provincie: wat we hebben gedaan was wel geïnformeerd, het was niet zomaar iets. Het was *state of the art*.’

Dat is een manier om uit die kramp te komen. Dat vergt ook veel van toezichthouders, rekenkamers en evaluatoren. Die beginnen altijd met iets dat tegenvalt en merken daar zelden bij op dat het ook erg complex is. En dan constateren ze vervolgens dat de informatievoorziening niet op orde was. Als de informatievoorziening wel op orde is en je kunt laten zien: ‘dit is wat we wisten en toen hebben we dat gedaan om te zorgen dat er dit en dit gebeurde’ dan sta je ook veel sterker als provinciaal bestuur. Beleidsevaluatie kan dan helpen om heel gericht verantwoording af te leggen, omdat je kunt laten zien wat je hebt gedaan om ervoor te zorgen dat je zo goed mogelijk bezig bent met het belang van de burger.

Programmeren

Overheden worstelen met het formuleren van langjarige onderzoeksprogramma's. De Amerikanen spreken in dat verband vaak over *monitoring and evaluation*. Verantwoord bestuur begint bij beleidsprogramma's die hun informatiehuishouding op orde hebben: wat is het probleem? hoe groot is het probleem? wie zijn er door geraakt? wat zijn de oorzaken? wat kan ik doen om te zorgen dat ik op die oorzaken ingrijp? wat moet dat dan zijn ten opzichte van de probleemgrootte? hoe ontwikkelt zich dat van kwartaal tot kwartaal en van jaar tot jaar? Op zo'n moment vormt een evaluatie een logisch slotstuk. Met natuurlijk als moeilijkste punt: is de verandering exclusief aan de interventie toe te schrijven?

In de praktijk gaat het meestal anders. Dan blijkt er een evaluatieverplichting te zijn, soms pas op het allerlaatste moment ontdekt. Dan is het helemaal erg als ook nog eens de doelstelling moet worden gereconstrueerd en men tot de ontdekking komt dat allerlei informatie niet voor handen is waar men al lang en breed over had moeten beschikken als ze er gewoon normaal van dag tot dag mee bezig waren geweest. Er worden dan in allerlei interviews met *stakeholders* gehouden, vaak over proceszaken, en dat is het dan.

Een verantwoord bestuurder moet zich niet alleen afvragen: 'doe ik het zo goed mogelijk?', maar ook: 'doe ik het goede?' Wat ik doe, haalt dat wat uit? Je kunt natuurlijk heel goed een subsidieregeling uitvoeren, keurig netjes de beschikkingen op tijd rond krijgen, met een mooie klachtenprocedure erom heen, waardoor iedereen op dat punt tevreden is. Maar als het weinig of niets uithaalt en misschien zelfs contraproductief werkt, dan kun je wel zeggen dat je het goed doet, maar je doet ondertussen wel het verkeerde. En als je dan drie maanden voor het aflopen van de periode nog eens ontdekt, o ja we moeten ook nog evalueren op het nut, dan ben je geen knip voor je neus waard.

Als je wilt programmeren, dan moet je aansluiten bij je bestuurlijke prioriteiten en van daaruit vaststellen wat je in welke volgorde gaat evalueren. Het helpt wel als je een checklist hebt met regelingen die aan de beurt zijn, maar de intrinsieke interesse moet zijn: wat vind ik als provinciebestuur het belangrijkste om te realiseren? Op die terreinen wil ik ook graag laten zien dat ik de middelen zo goed mogelijk inzet, dat ik dat monitor en dat ik daar de effecten van evalueer. Daarmee komt het uit de verplichtings sfeer en past het veel meer bij het verhaal dat je richting burger wilt vertellen.

Uitbesteden

Ik ben een groot voorstander van het in eigen beheer uitvoeren van evaluaties. Ik

heb dat op het ministerie van Financiën zelf gezien. Het interne leereffect is veel groter dan wanneer je ze uitbesteedt. Het is een stimulans voor mensen om dingen zelf te doen. Het geeft een mooie kans voor benutting van de resultaten zowel tijdens als na het onderzoek. Ook de kosten kunnen veel lager zijn. Natuurlijk hangt dat wel af van de manier waarop je het organiseert. Het standaard in huis hebben van allerlei expertise die je eens in de zoveel jaar gebruikt, dat is natuurlijk onzin. Voor bepaalde methodologische aspecten is het verstandig om zaken in te huren of gebruik te maken van kennis en expertise die al beschikbaar is.

Te vaak wordt bij uitbesteding uitgegaan van de fictie dat het onafhankelijk is. Dat klinkt natuurlijk aantrekkelijk. Maar ik heb ook vaak evaluaties uitbesteed en dan krijg je soms het voorstel om maar vast op te schrijven wat er uit moet komen. Hoezo onafhankelijkheid? Wie betaalt, bepaalt dan in hoge mate. Goede onderzoekskantoren zullen dat overigens nooit doen of in ieder geval nooit toegeven. Maar zo groot is Nederland niet en wat dat betreft hebben wij ook wel ervaren dat het lastig is om onderzoeksbureaus te vinden die niet al eerder betrokken zijn geweest bij het beleid dat je wilt evalueren.

Je kunt externe onafhankelijkheid ook organiseren met een begeleidingscommissie of een *second opinion*. Je kunt je prima voorstellen dat je een beleidsdirectie of een organisatie zelf laat evalueren, maar dat je daar een begeleidingscommissie boven op zet, bestaande uit mensen van onbesproken gedrag, kennis, deskundigheid en afstand. Die kijken als eerste of de goede vragen gesteld worden. Als je alleen wilt weten of partners een bepaald beleid waarderen, dan is dat heel wat anders dan wanneer je naar effecten vraagt. Een begeleidingscommissie kan naar de opzet van de evaluatie en de vraagstelling kijken, kan de methode beoordelen, maar kan ook wegen of wat er uit komt als valide beschouwd kan worden en zich uitspreken over evenwichtigheid in de rapportage.

Benutting

Zelf evalueren biedt een goede kans op benutting van de resultaten tijdens en na het onderzoek. Daarbij moet je wel in de gaten houden of het belangrijk is dat de uitkomsten via de volksvertegenwoordiging moeten worden geleid. De directe procesbenutting heb je bij een extern bureau veel minder. Daar wordt vaak afgewacht tot het rapport is verschenen, vervolgens komt het op de agenda en wordt er nog weer later eens een keer naar gekeken. Dan is het niet van jezelf – *not invented here* – en ben je minder geneigd ermee aan de slag te gaan.

Veel reacties op evaluaties, zeker als het kritische evaluaties zijn, zijn afhoudend. Het is niet goed onderzocht, te snel, te laat, te duur, enzovoort. Terwijl als je dan

later naar die organisatie kijkt, dan zie je dat ze sommige lessen wel degelijk hebben geleerd. Dat ze dingen zijn gaan doen die ze in reactie op de evaluatie eerst met de grootst mogelijke stelligheid hebben verworpen. Op het moment dat ze even in de spotlights stonden werd vanuit een verdedigingsstrategie gezegd dat de aanbeveling niet deugde. In het openbaar zeggen dat je niets hebt aan de aanbevelingen, dat past in het patroon dat je je wilt beschermen in het verantwoordingsdebat. Terwijl als de druk er af is, dan kun je wel degelijk je voordeel doen met een evaluatie.

Wat wij als Algemene Rekenkamer steeds meer doen, is niet aan de achterkant een rapport afscheiden en dan daar alle benutting van laten afhangen. We gaan gaandeweg het proces ook tussentijds uitkomsten presenteren en bediscussiëren. Op sommige punten testen we zelfs aanbevelingen. Waarmee je dus degene die geëvalueerd wordt of van wie zijn programma wordt geëvalueerd al in het onderzoeksproces meeneemt. En hem daarmee ‘verleidt’ om, al was het maar voor een deel, probleemeigenaar te worden.

De Ecologische Hoofdstructuur is daar een voorbeeld van. We waren met dat onderzoek begonnen, maar moesten het halverwege stopzetten omdat we een tussentijds verzoek van de Tweede Kamer kregen. Toen is het ministerie met die tussenstand aan de gang gegaan voor het verder verbeteren van het beleid. Als onderzoeker of rekenkamer denk je dan misschien wel even: ‘een beetje jammer, dan kunnen we straks minder in ons rapport kwijt’. Maar als je er vanuit het publieke belang naar kijkt, en dat moet je natuurlijk doen, dan is dat een hartstikke goede zaak.

Wat ook werkt, is het omdraaien van je veronderstelling. Vaak is de gedachte: beleid werkt, tenzij het tegendeel wordt bewezen. Draai het nu eens om: het werkt niet, tenzij bewezen. Dat levert interessante nieuwe perspectieven op. We hebben het meegemaakt bij ons onderzoek naar subsidies. Dan wordt de stelling: als je na vijf jaar subsidies te hebben verstrekt nog steeds niet hebt onderzocht of het werkt, dan moet je er mee stoppen. Dat is misschien een wat minder sympathieke vorm om te zorgen voor benutting, maar het werkt wel. En we zijn er best trots op dat het Kabinet onlangs zo’n horizonbepaling heeft ingesteld voor alle subsidie-regelingen. Dus vertaald naar het provinciaal niveau: als je bij Provinciale Staten autorisatie vraagt voor het uitgeven van publiek geld, dan geldt voortaan de conditie: prima, maar als je na vijf jaar nog niet hebt onderzocht of het ook iets uithaalt, dan stoppen we er mee. Dat zorgt er natuurlijk ook voor dat je er belang bij hebt om de effecten te laten zien en dat je stuurt op benutting.

Controle

Over de rol van de volksvertegenwoordiging. Als de Tweede Kamer zelf een onderzoek wil instellen of dat gaat doen, dan is dat meestal toch een reactie op het feit dat de minister dat niet of niet goed genoeg heeft gedaan. Eigenlijk zou het anders moeten gaan. Het zou natuurlijk zo moeten zijn dat Provinciale Staten, gemeenteraad of Tweede Kamer tegen de uitvoerende macht zeggen: prima om daar geld aan uit te geven, maar leg ons eens uit waarom dat zo’n goed idee is en wat de onderbouwing is voor de aannahme dat het gaat werken. Of, als het een nieuw en ingewikkeld probleem is: laten we maar eens proberen of het kan werken, maar we willen wel geïnformeerd worden of we waar voor ons geld krijgen. Het begint dus bij het autoriseren van het uitgeven van geld, bij de start van nieuw beleid, bij het lastig vallen van burgers en bedrijven. Daar past natuurlijk bij dat je zegt: ‘we verwachten wel dat je duidelijk maakt wat het oplevert’.

Dat moet je ook terugzien in de planning en controldocumenten. Ik weet wel dat onderzoek vaak is ‘los gezongen’ van de formele planning en controlcyclus. Van het meeste beleid kun je geen effecten op korte termijn verwachten. De cyclus kan niet elk jaar gesloten worden. Sowieso zit er altijd een uitvoeringsjaar tussen, daar heb je ook last van. Dus moet je meer in programma’s en meerjarige verantwoordingstermijnen denken als je het hebt over resultaten en effecten.

Dat vraagt om geduld en dat maakt het lastig. We willen heel snel resultaten (van experimenten) zien, zeker tegen de achtergrond van de bestuurlijke cycli van verkiezingen en gekozen worden. Maar als je reëel bent, ontkom je er niet aan om zaken in meerjarig perspectief te plaatsen. Dat kan gewoon niet anders. Ik denk alleen wel dat je dan in de tussentijd kunt monitoren en de vinger aan de pols houden. De basisregel blijft: bij het autoriseren van beleidsuitgaven hoort altijd interesse voor de achterkant.

Tot slot

Afsluitend, is er ook een vraag die we onszelf als evaluatoren mogen stellen: wat is onze toegevoegde waarde? Niet alleen in termen van wat we adviseren of wat de uitkomsten zijn van onze onderzoeken en wat er mee gebeurt, maar heel direct ook over ons eigen presteren: wat kunnen we zeggen over de kosten-baten-verhouding van onze inspanningen? Dat is een hele interessante vraag, waar we nog te weinig aan toe komen. Een vraag die wel belangrijk is, al was het maar om onszelf scherp te houden.

Mijn beeld is dat evaluaties te vaak nog een ondergeschoven kindje zijn, waar niet iedereen het nut van inziet. Volgens mij komt dat doordat de insteek van de meeste evaluaties vooral gericht is op verantwoording. Een evaluatie wordt te veel gezien als verplicht nummer en te weinig als een kans om te leren van de beleidsuitvoering. Dan is het natuurlijk ook niet zo verwonderlijk dat mensen niet staan te springen om een evaluatie prominent op de agenda te krijgen. Het heeft ook te maken met timing. Vaak wordt er alleen pas achteraf iets over het beleid gezegd. Tegen de tijd dat dat gebeurt - omdat er een evaluatierapport op tafel ligt en men daar iets van moet vinden - is het beleid zelf al afgerond en spelen er weer allerlei andere onderwerpen. En voor een deel ligt het ook aan de kwaliteit van evaluaties. Die bovendien meestal ook nog eens alleen ingaan op één specifiek beleidsprogramma, waardoor er nauwelijks lessen uit getrokken worden voor andere onderdelen van een organisatie. Met dat soort evaluaties slaag je er niet in om de materie te 'overstijgen' en op een hoger niveau conclusies te trekken.

Vaak zie je dat de aandacht voor een beleidsevaluatie is neergelegd bij een aparte eenheid in de organisatie. Dat draagt een afbreukrisico in zich. Als het onderwerp meer geïntegreerd wordt in de beleidscyclus en programmamanagers er zelf verantwoordelijk voor zijn en op aangesproken kunnen worden, dan ontstaat een inhoudelijk 'eigenaarschap'. Dan leer je ook dat als je een evaluatie op een bepaalde manier organiseert, het onderwerp meer gaat leven en je echt iets aan de uitkomsten kunt hebben.

Dat pleit er volgens mij voor om een onderzoekseenheid binnen een ministerie, provincie of grote gemeente vooral een faciliterende rol te geven. Een eenheid die anderen in de organisatie helpt om beleidsevaluaties op te zetten of uit te voeren. Zoals ik het zie, moet de uiteindelijke verantwoordelijkheid in de lijn worden neergelegd. Bij de medewerkers die aan het eind van de rit ook met de uitkomsten aan de slag moeten.

Verplichting

De evaluatie van de Reconstructiewet (gericht op herinrichting van het landelijk gebied) is op dat punt een interessant voorbeeld. De aanleiding was heel traditioneel. Er stond nu eenmaal een bepaling in de wet dat er geëvalueerd moest worden. En zo werd het ook gevoeld door alle partijen. Dat gaf reacties in de trant van: 'Wat moeten we daar nou mee?' en: 'De inkt van de reconstructieplannen is nauwelijks droog en nu moeten we alweer kijken hoe het er voor staat.' Dat zorgde voor een negatieve, bijna vijandige sfeer. Dat was één van de redenen dat men op zoek ging naar een andere manier van evalueren, waarbij gebruikers en betrokken partijen overtuigd konden worden van de meerwaarde van de evaluatie.

Het ministerie heeft er vervolgens voor gekozen om de evaluatie niet zelf uit te voeren, maar uit te zetten bij Alterra, onderdeel van Wageningen University & Research centre. Het opdrachtgeverschap werd neergelegd bij de lijnmanagers die de reconstructie vormgaven en niet bij de directie van het ministerie die zich met kennis en evaluatie bezighoudt. Dat was een hele bewuste keuze en vormde uiteindelijk een belangrijke succesfactor.

Het had vooral met de invulling van de opdrachtgeversrol te maken. De medewerker in kwestie wist van de hoed en de rand wat betreft de reconstructie, maar had tegelijkertijd feeling met de onderzoekskant en een lerende houding. Hij was zich er terdege van bewust dat als het ministerie, provincies en andere stakeholders iets aan de evaluatie wilde hebben, dat hij dan voldoende afstand moest houden tot de onderzoekers om hen in intellectuele vrijheid hun werk te kunnen laten doen. Juist daardoor is het goed gegaan.

Een verplichting om te evalueren is geen garantie voor succes. Je kunt het onderwerp er wel mee op de agenda krijgen en ervoor zorgen dat mensen er aandacht aan besteden. Maar of dat dan ook op een goede manier gebeurt, zodat je er echt iets aan hebt, dat is nog maar de vraag.

Moet echt alle beleid geëvalueerd worden? En op welk niveau dan en hoe ver ga je daarin? Dat je prestaties en voortgang meet is voor alle beleid een goede zaak. Wat mij betreft is daar weinig discussie over nodig. Maar het optuigen en uitvoeren van doe-evaluaties voor alle typen beleid, dat is een wel erg zwaar en kostbaar middel. Ik betwijfel sterk of dat altijd nut heeft. Ik kan me wel voorstellen dat je centrale beleidsprogramma's consequent met beleidsevaluaties belegt. Waarbij je natuurlijk ook weer een discussie kunt krijgen wat centrale beleidsprogramma's zijn en wat niet. Ik denk dat de volksvertegenwoordiging daar een belangrijke rol in kan en moet spelen.

Veel onderzoek wordt trouwens als beleidsevaluatie aangemerkt terwijl het dat, als je er goed naar kijkt, niet is. Waar het in veel gevallen op vastloopt, is dat er in het begin geen duidelijke beleidsdoelen zijn geformuleerd. Dan is het sowieso al heel lastig om daar achteraf criteria voor te ontwikkelen. Ook zie je nogal eens dat evaluatoren niet duidelijk maken aan de hand van welke criteria zij evalueren. Dan is er geen duidelijke norm waartegen de beleidsuitvoering wordt afgezet. Ook vind ik dat veel beleidsevaluaties te kort schieten op het punt van verklaringen: je krijgt dan geen zicht op waarom beleid wel of niet heeft gewerkt, hoe het complexe spel van beleidsuitvoering in elkaar zit. Wat vervolgens leidt tot simplistische conclusies als 'er moet meer geld bij' of 'de regels moeten strenger'.

Als lid van de Provinciale Raad voor het Omgevingsbeleid in Gelderland kreeg ik laatst uitkomsten van een uitgebreide monitoring van het streekplan voorgeschoteld. Dat bevat een hele hoop indicatoren die lang geleden zijn ontwikkeld. Maar veel van de doelen die in het beleidsplan staan, zijn inmiddels niet meer relevant door alle beleidswijzigingen die er hebben plaatsgevonden. Zijn dit dan nog de zaken die we willen weten? En bovendien, misschien nog wel belangrijker: Wat betekenen al die gegevens? Wordt de ontwikkeling die uit de monitoring naar voren komt inderdaad veroorzaakt door het provinciaal beleid, of is er sprake van een autonome trend?

Als evaluator kun je proberen om daar zelf een antwoord op te formuleren, maar ik denk dat je beter de dialoog daarover kunt organiseren met mensen uit het veld. Zij hebben verstand van het onderwerp. Dan gaat het over vragen als: Wat betekenen deze trends die we zien? Is dit een wenselijke ontwikkeling of niet? Wat is het effect van provinciaal beleid op deze ontwikkeling? De antwoorden hierop zijn niet altijd eenduidig en daarom is het belangrijker dat je daar een gesprek met elkaar over hebt want dat bevordert het ‘beleidsleren’.

Programmeren

Het is ingewikkeld om in lange termijnperspectief een evaluatieprogramma te ontwikkelen. Bij ministeries zie je dat alle onderzoek meerjarig wordt geprogrammeerd, niet alleen evaluatieonderzoek. Op het oog zit zo’n systeem heel goed in elkaar. Het begint met een agenderingsfase, daarna wordt het verfijnd en uitgezet en vervolgens brengen onderzoeksbureaus offertes uit voor de uitvoering. Dat is een jaarlijks terugkerend proces. En dan heb je nog een orgaan dat in de gaten houdt of zo’n kennisprogramma een beetje op koers ligt of dat er bijstellingen nodig zijn. Waarbij als het goed is ook nog geprobeerd wordt om in te spelen op zaken die op wat langere termijn spelen.

Maar de keerzijde van een dergelijke, op papier, fraaie aanpak is dat je ongemerkt steeds verder van de beleidspraktijk komt te staan. Dat andere directies denken: ‘Er is een aparte club die het organiseert, dus wij kunnen ons met de lopende zaken bezighouden’. Dat is een spanning waar je in evaluatieonderzoek ook mee te maken hebt. Beleidsambtenaren, de goede niet te na gesproken, houden zich over het algemeen vooral bezig met de vragen van vandaag en niet met die van morgen. Ze hebben wel vragen, ook evaluatieve vragen, maar die bestrijken zelden een langere periode. Terwijl je met je onderzoek en je evaluatie juist iets wilt neerleggen waar het beleid op langere termijn ook wat aan heeft.

Uitbesteden

Het grote risico van alleen uitbesteden is dat het onderwerp niet binnen de organisatie beklijft. Daar kun je als extern onderzoeker overigens wel het nodige aan doen. Ik denk juist dat het voordelen kan hebben om een evaluatie door een externe te laten doen. Je hebt dan in ieder geval een ander perspectief van waaruit je naar het onderwerp kijkt. Belangrijk is wel dat het onderzoek zelf goed ingebed is in de organisatie. Dat betekent dat de externe onderzoeker samen met de opdrachtgever op zijn minst heel scherp kijkt naar de probleemstelling en vraagstelling van het onderzoek. Ook tussentijdse resultaten moeten met mensen in de organisatie worden doorgesproken. Als dat niet goed georganiseerd is, dan dreigt het een ver-van-mijn-bed-show te worden.

Zeker in de beginfase moeten opdrachtgevers dus flink investeren in de relatie met de onderzoekers. Zij moeten hen meenemen in het onderwerp, de context schetsen en zo helder mogelijk aangeven wat de oorsprong van de vraag is. Vaak gaat het daar fout. Dan heeft de organisatie daar te weinig tijd voor ingeruimd. Het onderzoek is immers uitbesteed, zodat de lijnorganisatie het ‘eigenlijke’ werk kan blijven doen.

Een andere, eveneens foute gedachte is: de externe onderzoeker moet zijn werk onafhankelijk kunnen doen, dus dan moet ik hem ook maar niet al te veel vertellen over de dingen die ons bezighouden, want dat zou hem beïnvloeden. Terwijl de onderzoeker die inside-informatie juist nodig heeft om scherp te krijgen waar de vraag van de opdrachtgever zit. Dat betekent dat de opdrachtgever ook bereid moet zijn om de onderzoeker een kijkje in de keuken te gunnen.

Het begint natuurlijk al met de keuze van een goed onderzoeksbureau en een goede onderzoeker. Maar ook dat je waardering kunt opbrengen voor een kritisch geluid. Als je te krampachtig bent als opdrachtgever en de vraagstelling zo beperkt houdt dat je alleen maar iets op de vierkante centimeter kunt doen, dan gaat het natuurlijk nooit goed. Dan heb je het eigenlijk over het verschil tussen de vraag: ‘Doen we de dingen goed?’ en de vraag: ‘Doen we de goede dingen?’. Vaak zie je dat men zich beperkt tot het eerste, maar dat men de onderliggende vraag, op het niveau van het beoogd maatschappelijk effect, uit de weg gaat. Aan die diepere vraag komt men dan niet toe. En daar kunnen juist externen of meer wetenschappelijk georiënteerde instituten mee aan de slag. Puur het technische meten van prestaties en die langs een maatstaf leggen, dat kan een ministerie of provincie heel goed zelf. Maar als het meer de strategische kant opgaat en er ook een lange termijnperspectief gevraagd wordt, dan is het zinnig om de buitenwereld daarbij te betrekken. Goed opdrachtgeverschap houdt dus investeren in tijd

in, maar ook erkennen dat er mensen zijn die vaker met dat bijltje hebben gehakt.

Wat ik nogal eens zie, is dat men vooral op basis van de vraagstelling handelt. Die wordt dan van alle kanten bekeken en helemaal doorgeakkerd, terwijl men geen goed beeld heeft wat er uit een onderzoek zou kunnen komen. Dat kan tot verrassende discussies aan het eind van een onderzoek leiden. Dan heeft de opdrachtgever ineens allerlei extra vragen die hij onvoldoende in het onderzoek beantwoord ziet. Het blijkt voor opdrachtgevers lastig zich puur op basis van de vraagstelling een beeld te vormen wat er uit kan komen. Het is dus zaak de opdrachtgever gedurende het onderzoek goed te blijven informeren over keuzes en tussenresultaten.

Breder verband

Op het punt van de benutting van onderzoek is nog veel winst te boeken. Vaak wordt alleen gekeken naar de één-op-één relatie met de opdrachtgever. En dat is dan meestal ook nog maar eens één persoon binnen een bepaalde directie. En dan moet je maar hopen dat die ene ambtenaar binnen de kolom waarin hij werkzaam is zijn relaties op orde heeft en verbindingen weet te leggen.

Ik vind dat je daar als onderzoeker en onderzoeksbureau ook een verantwoordelijkheid in hebt om dat te organiseren. Of, als dat een brug te ver is en als te bemoeizuchtig overkomt, om het in ieder geval bespreekbaar te maken. Mijn ervaring als onderzoeker is dat wanneer jij het niet doet, het meestal helemaal niet gebeurt. Daarom moet je eigenlijk in je onderzoeksplan al aandacht vragen voor het grotere geheel. En dus bewust breder gaan dan die ene afdeling waar je het onderzoek voor doet. Dan moet je er natuurlijk wel in slagen om een goede aansluiting te vinden bij datgene waar die andere eenheden mee bezig zijn.

Bij het onderzoek naar de reconstructie in Noord-Brabant was de vraag vanuit Provinciale Staten (PS) ook duidelijk: Wat kunnen we vanuit de reconstructie leren voor andere grote transitieprocessen? Dat ging dus verder dan alleen kijken naar het onderwerp. Het onderzoek is later om andere redenen helemaal de mist ingegaan, maar dat laat onverlet dat de vraagstelling op zich goed was. Een brede invalshoek dus, die het voor de hele organisatie en de politiek relevanter en interessanter maakt.

Overigens lukt het ons als onderzoekers niet altijd om die brede scope erin te krijgen. Om zaken praktisch toepasbaar te houden, moet je het vaak juist klein houden. De truc is dan om te proberen om over het onderwerp bijvoorbeeld een gesprek met de gedeputeerde te hebben. Die heeft veel meer overzicht. Op die

manier kun je soms toch verbreden en onderwerpen naar een hoger niveau tillen. Als er een onderzoekseenheid is en deze voldoende statuur heeft, zou zij die rol ook uitstekend op zich kunnen nemen. Maar je zult volgens mij altijd ook een bestuurlijke hoeder moeten hebben.

Leren

Waar wij steeds meer op uitkomen, is het combineren van verantwoorden en leren. Dat gaat dan onder de noemer van lerende evaluaties. Daar zijn natuurlijk allerlei gradaties in. Essentieel is dat je gebruikers mee laat denken, sowieso over je onderzoeksopzet, maar ook dat je regelmatig terugkoppelt met tussenresultaten en dat je probeert de organisatie te betrekken bij het geven van betekenis aan de resultaten.

Het nadeel van klassieke evaluaties is: je voert je onderzoek uit, presenteert je resultaten en dan hoop je maar dat het werking gaat hebben. Met lerende evaluaties probeer je het zo in te richten dat die werking al veel eerder begint. En overigens kunnen die tussentijdse momenten ook als effect hebben dat het onderzoek wordt bijgesteld. Mensen gaan dingen eerder in de praktijk toepassen, waardoor ook jouw object van onderzoek gedurende de rit verandert. Iemand die het denken daarover in het bijzonder heeft gevoed is Jurian Edelenbos, hoogleraar Bestuurskunde aan de Erasmus Universiteit. In zijn visie zijn gebruikers geen mensen die pas aan het eind van de rit in beeld komen. Je praat dan over echt *ex durante* onderzoek.

Wat er ook erg in zit, is het benadrukken van het doorlopen van leercycli: het gezamenlijk betekenis geven aan bevindingen zodat mensen daarvan kunnen leren en ze kunnen toepassen in hun praktijk. Idealiter maak je in een lerende evaluatie verschillende leercycli door. Bij *ex post* onderzoek is het eigenlijk eenmalig. Als de focus op leren is gericht, ben je langere tijd betrokken, zodat je zo'n leercyclus na verloop van tijd ook kunt herhalen. Het is niet een scherp omschreven methodiek, maar meer een inspiratiebron. En afhankelijk van de situatie probeer je daar elementen van toe te passen.

Als je ervan uitgaat dat de werkelijkheid complex is en er verschillende perspectieven zijn, geef dat dan ook een plek in je evaluatie. Probeer dat niet plat te slaan in één kader. Geef betrokkenen een stem in het normenkader. Dat is altijd een heel cruciaal moment. Waar ga je op evalueren? Soms kom je tegen dat de beleidsdoelen niet helder zijn of dat er nog geen overeenstemming is over de doelen. Ook kun je te maken krijgen met meerdere perspectieven of met doelen die in de loop van de tijd zijn veranderd. Of de provincie heeft geen zeggenschap over de

doelen, iets wat je bij gedecentraliseerd beleid nogal eens tegenkomt. Probeer daar dan op zijn minst verschillende partijen over te horen. Dan krijgt de evaluatie bijna een soort beleidsontwikkende rol. Want dan ga je je nog eens een keer heroriënteren op de vraag: Waarom deden we dit eigenlijk ook alweer? Het kan invloed hebben op de beleidstheorie achter het beleid. Is die nog wel zo passend? Dan ga je schuiven. Ik vind dat een interessante ontwikkeling.

Controle

Over de rol van de volksvertegenwoordiging. Ik denk dat evaluaties zowel voor Gedeputeerde Staten (GS) als uitvoerende partij, als voor de controlerende rol van PS belangrijk zijn. Overigens kunnen PS in dat verband hele andere vragen stellen dan GS. Juist voor PS zijn vragen op meta-niveau belangrijker. Dus niet of het allemaal volgens de planning verloopt, maar meer: Wat zijn de maatschappelijke effecten en wat betekent dat nu allemaal voor de wat langere termijn, voor de komende beleidsontwikkeling?

Het onderzoek in Brabant naar de reconstructie was daar een goed voorbeeld van. Daar stelde de commissie Beleidsevaluatie de goede vragen. Nogmaals, dan heb ik het niet over wat er daarna allemaal fout ging. Zelf deed ik een andere ervaring op in Overijssel, eveneens rond de reconstructie. Daar vraag ik me in alle oprechtheid af of men iets aan het onderzoek heeft gehad. PS hadden een tijdelijke constructie voor dat onderzoek opgetuigd en vervolgens hebben ze lang onderhandeld over de precieze vraagstelling. Het onderwerp was van meet af aan heel politiek. Toen wij als onderzoekers vervolgens aan de slag gingen, was in ieder geval één ding helder: wat er ook gebeurde, aan de vraagstelling kon niet worden getornd. Die was immers de uitkomst van een langdurig en uiterst delicaat onderhandelingsproces. Maar als onderzoeker vond ik die vraagstelling behoorlijk beperkt. De lange termijn vragen zaten daar onvoldoende in. Het onderzoek had meer op kunnen leveren.

Beleidsevaluatie is ook echt een vak, daar hoeft je niet omheen te draaien. Dat doe je er niet zo maar even bij, daar moet je je echt in verdiepen wil je er uit halen wat er in zit. Als PS zou je een soort ‘kop’ bovenop de door GS uit te voeren evaluaties kunnen zetten met specifieke aandacht voor vragen die samenhangen met de kaderstellende en controlerende rol.

Een voorbeeld is Gelderland waar men nu bezig is met een nieuwe structuurvisie. Daar wordt een heel interactief proces rondom heen ingericht, waar de politiek overigens op onderdelen moeite mee heeft omdat zij zich soms buitenspel gezet voelt. Maar als je het vanuit een andere invalshoek bekijkt, is het heel logisch om te zeggen: ‘We weten eigenlijk nog helemaal niet wat het effect van het vorige

streekplan is geweest. We kunnen nu natuurlijk wel een nieuw proces in gang zetten, maar het is toch niet verkeerd om eerst eens stil te staan bij wat het eerdere proces heeft opgeleverd en wat we daaruit kunnen leren. Waar blijft de beleidsevaluatie?’ Dat zou ik typisch een vraag vanuit PS vinden om op een soort meta-niveau de vinger aan de pols te houden of dat soort dingen gebeuren. Als je zoals PS voor een lange reeks van jaren grote sommen geld moet autoriseren voor een beleid, dan moet je toch willen weten, om een goede afweging te kunnen maken, wat er van het vorige beleid op hoofdlijnen terecht is gekomen. En daarmee GS dus ook controleren op wat zij gerealiseerd hebben. In die zin is er voor PS een taak.

Partijen

Beleid komt meestal tot stand in een samenspel van meerdere partijen. Dan is het moeilijk om de relatie tussen de interventie en het effect te leggen. Dat levert methodische problemen op. Het probleem doet zich met name voor bij de verantwoording, want daar is het juist die ene actor die zich moet verantwoorden tegenover de volksvertegenwoordiging.

Voor het leren is het minder een probleem, dan kun je aangeven dat je een gezamenlijke verantwoordelijkheid hebt. Natuurlijk, ieder vanuit een eigen rol en verantwoordelijkheid, maar uiteindelijk is het toch een samenspel waarbij het er om gaat hoe je je processen zo goed mogelijk kunt inrichten. En daar kun je over leren, hoe je dat beter kunt doen. Zeker als je van meet af aan de partijen die ermee te maken hebben, bij een dergelijk evaluatieonderzoek betreft. Dan kun je dat ook onderwerp laten zijn van die evaluatie. En biedt de evaluatie als je die goed aanpakt een uitstekende kans aan partijen om even uit de arena te treden waarin zij elkaar steeds tegenkomen en op een andere manier met elkaar te praten over waar ze mee bezig zijn. Als onderzoeker geeft dat ook veel voldoening als je dat gesprek kunt organiseren en partijen niet in hun ‘normale’ vechthouding zitten.

Er zijn helaas geen gemakkelijke antwoorden. Maar dat wil niet zeggen dat je helemaal niks kunt zeggen. Je kunt verschillende vormen van onderzoek inzetten om meer zicht te krijgen op wat bepaalde interventies teweeg brengen. Alleen kun je het niet altijd rechtstreeks en heel eenduidig ‘terugredeneren’ naar één actor en één interventie. Je moet af van het sturingsmodel dat uitgaat van één centrale actor. Dat betekent ook dat de evaluatie dat niet moet suggereren. Dan kun je overigens nog steeds uitspraken doen over wat meer of minder effectief is. De gesprekken die je daarvoor moet hebben, kun je voeden met goed onderzoek. Waarbij je laat zien hoe dingen werken, welke patronen en mechanismen er zijn. Dat soort onderzoek is ook geschikt om de heersende beleidstheorie nog eens

tegen het licht te houden. Op zo'n manier kan ook het beleid weer een stukje verder gebracht worden.

Ontwikkeling

Een belangrijke trend in evaluatieonderzoek, zoals hiervoor al opgemerkt, is het steeds meer combineren van evalueren en leren. De meest uitgesproken vorm daarvan is de *responsieve evaluatie*, waarin nauwelijks verantwoordingselementen zitten en die vooral dient als ondersteuning van het leerproces.

Een andere stroming gaat meer uit van de *kritische theorie*. Daarbij wordt beleid niet zozeer langs de meetlat gelegd van de eigen beleidsdoelen, maar wordt vooral gezocht naar het antwoord op de vraag wat de functie van het onderzochte beleid is. Lost het de maatschappelijke problemen op? Welke machtsrelaties spelen een rol? Wie wint, wie verliest?

Een andere trend heeft te maken met het monitoren van allerlei ontwikkelingen. Er wordt heel veel gemeten, terwijl het evaluatieve aspect vaak achterwege blijft. Dat is het zoeken naar de verklaring achter al die gegevens. Licht het aan het beleid? Licht het aan andere dingen? Als je niet weet waardoor de dingen komen, dan kun je het wel van allerlei kanten bekijken, maar dan weet je nog steeds niet wat je moet doen. Dat zie ik ook wel steeds meer ontstaan. Dus een beweging weg van het pure monitoren, meer in de richting van het zoeken naar verklaringen. Een continue zoektocht naar inzicht in hoe beleid werkt. En daarmee draag je misschien ook een beetje bij aan een meer bescheiden rol van de overheid. Die overheid is niet almachtig.

Op die manier kan er van onderzoek een opvoedende rol richting overheid uitgaan. 'Trek niet een te grote broek aan, het zit allemaal heel ingewikkeld in elkaar. Je bent maar één van de radertjes in het grotere geheel.'

Frans Leeuw

31 mei 2012

Als ik gevraagd wordt om trends op het gebied van evaluatiestudies aan te geven, wil ik allereerst een zeer recente ervaring delen. Gisteren woonde ik de viering bij van het 35-jarig bestaan van de IOB, de Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie. Ik ben voorzitter van het panel dat de kwaliteit, bruikbaarheid en gebruik van de evaluaties van de IOB bekijkt. Kees Vendrik van de Algemene Rekenkamer en Marjolein van Asselt van de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) spraken daar ook. De WRR heeft een groot project 'Lessen van evaluatie'. Eén van de kernpunten uit hun bijdragen was: vroeger keek men bij evaluaties *ex post* naar beleid en dan werd er over resultaten gerapporteerd. Die resultaten werden vervolgens vrij snel in verband gebracht met het beleid en als er geen resultaten waren dan lag de verklaring voor de hand: dat kwam door het slechte beleid. Methodologisch allemaal niet erg geavanceerd. De kwaliteit van effectonderzoeken is de laatste tien jaar ten gevolge van de discussie over het belang om robuuste effectstudies te doen, waarbij al snel gedacht wordt aan *experimentele* en *quasi experimentele* onderzoeken, steeds meer op de agenda aan het komen. Het is interessant dat ook de IOB op dat gebied vooruitgang boekt.

Serius

Dat thema was dan ook één van de rode lijnen in dit jubileumcongres van de IOB. Met name bij ontwikkelingssamenwerking speelt al een aantal jaren de vraag hoe op robuuste wijze tot uitspraken te komen over de effecten van ontwikkelingsinterventies, programma's en beleid. De discussie daarover is vooral aangejaagd door de publicatie van het rapport *When will we ever learn?*¹ uit 2006. Daarin wordt de balans opgemaakt van het vele evaluatiewerk en komt men tot de vernietigende conclusie: 'We weten het nog steeds niet'. Dat heeft een serieuze boost gegeven om, net zoals de medische wetenschap dat doet, effectstudies veel serieuzer op te pakken. Het heeft trouwens ook geleid tot uiterst felle discussies tussen methodologen. Met aan de ene kant de 'randomistas' die *randomized trials* bepleiten (de 'experimentele effectevaluaties') en aan de andere kant de meer rekkelijken, soms wel 'relativistas' genoemd. Die discussie is inmiddels enigszins beslecht. Ik heb daar een bescheiden bijdrage aan geleverd met een handboek dat ik samen met een Vlaamse collega voor de Wereldbank heb geschreven en waarin de voor- en nadelen van verschillende methoden van onderzoek op een rij worden gezet en wordt aangegeven wanneer welk evaluatiedesign het beste wel en niet kan worden ingezet.²

¹ Center for Global Development: *When will we ever learn? Improving lives through impact evaluation*. Rapport van de Evaluation GAP Working Group, mei 2006.

² Frans Leeuw en Jos Vaessen, *Impact Evaluations and Development - NONIE Guidance on Impact Evaluation*, Network of Networks on Impact Evaluation, 2009.

Ik beschouw dat als de *eerste trend*: dat we in een aantal maatschappelijke sectoren afstappen van de slordige, door iedereen op een achternamiddag te verrichten ex post studietjes die resultaten te snel toe wijzen aan beleid. Dat toewijzen kan alleen als dat methodologisch verankerd en zorgvuldig gebeurt. Dat betekent dus ook dat niet iedereen zomaar in het wilde weg kan roepen dat hij bezig is met effectonderzoek.

Steeds meer dringt het besef door dat als we willen weten of instrumenten werken, we er dan niet mee weg komen door wat tevredenheidsmetinkjes te doen en aan mensen in een twintig-seconden-survey op internet te vragen: 'Vindt u ook niet dat enzovoort.....?'.

Ex ante

Trend twee, het WODC, het Wetenschappelijk Onderzoek- en Documentatiecentrum van het ministerie van Veiligheid en Justitie, brengt daar over enige tijd een rapport over uit: de sterke groei van *ex ante* evaluaties. Het zijn soms een soort van uitvoeringstoetsen gericht op de vraag: 'Zal dit gaan werken?'. Die trend is in de eerste plaats een gevolg van het feit dat ex post evaluaties, of ze nu slecht of goed zijn, doorgaans veel tijd kosten. Een bestuurder kan daar niet op wachten, die moet morgen of overmorgen al wat weten. Dat geeft een druk richting ex ante evaluatie. En op dat vlak begint er ook steeds meer serieuze methodologie te komen.

Kennischatkamers

Een derde trend is het werken met zogenaamde *repositories*. Dat zijn kennischatkamers, portals waar systematische onderzoeksbevindingen naar de werking van interventies maar ook van andere zaken te vinden zijn. De organisaties die deze 'systematic reviews' en syntheses maken doen vaak niet zelf evaluaties, maar beoordelen de kwaliteit van de aangetroffen evaluaties en andere studies, trekken daar conclusies uit en – als het kan – rekenen opnieuw de 'effect maten' uit door gebruik te maken van de – vaak – tientallen onderzoeken die ze samenvatten. Die samenvattingen worden per onderwerp bij elkaar gebracht. Wetenschappers fileren honderden evaluaties, scheiden het kaf van het koren en maken vervolgens deze 'systematic reviews' van enkele tientallen pagina's. En daarbij brengen ze, als het kan en wetenschappelijk verantwoord is, ook de effecten van verschillende interventies in beeld. Je ziet steeds meer van die repositories komen, bijna altijd zijn ze openbaar en gratis toegankelijk. Deze ontwikkeling komt veertig jaar nadat de medische wetenschap daarmee is begonnen. De eerste repository was die van Archie Cochrane. In zijn 'bibliotheek' zitten resultaten naar honderdduizenden effectstudies van therapieën, pillen, organisatieveranderingen in medische 'edifices' enzovoort. Als je bijvoorbeeld wilt weten wat de beste manier is om nierstenen te

verwijderen, dan zie je in tientallen studies samengevat dat als je het op de ene manier doet je dit effect hebt en doe je het op een andere manier dan zie je die en die effecten optreden. Deze manier van werken slaat nu sinds zo'n tien/vijftien jaar over naar de wereld van overheidsinterventies en – ook – interventies door de private sector.

Ritualisering

De vierde trend zou ik willen omschrijven als: 'Vrienden, waar zijn wij mee bezig?'. Het is een trend van ritualisering. Of, om het anders te zeggen: geef iemand een hamer en alles moet gespijkerd worden. Er is een ongelooflijk grote hoeveelheid organisaties die allemaal bezig zijn met M&E, *monitoring and evaluation*. Het houdt niet op. Ik zeg wel eens tegen mijn studenten of in een spreekbeurt: 'Wanneer is bij u het lokale café toe aan een sous-chef Monitoring and Evaluation?'. Let wel, het idee van M&E is natuurlijk helemaal niet slecht. Maar als de hele wereld bezig is met het opzetten van allerlei monitoringsystemen, dan moet je natuurlijk niet raar staan te kijken als het op een gegeven moment niet meer serieus wordt genomen. Hadden we vroeger de situatie dat er niets was – Hoogerwerf vroeg daar aandacht voor in zijn Twentse oratie in 1977, ik heb het zelf in 1993 in mijn Utrechtse oratie laten zien – nu zijn we een paar decennia verder en is het een *mer à boire* geworden met over het algemeen een beperkte kwaliteit. Peter Dahler-Larsen, oud-voorzitter van de European Evaluation Society, sprak in dat verband over *evaluation machines*.

Verankering

Wat kun je als ministerie of provincie doen om niet in die valkuil van ritualisering te trappen? Met andere woorden: hoe veranker je op een goede manier de aandacht voor evaluatie? Ik zie een aantal antwoorden op die vraag. Allereerst moet je een klein clubje hebben dat in de gaten houdt of, met name bij de grote programma-uitgaven, de mensen die daarvoor verantwoordelijk zijn zelf bijhouden wat de doelen zijn en informatie over die doelen vergaren. Zo'n club moet je niet te groot maken, het moet geen 'directie' of zo worden. Wat mij betreft moet je het ook zoveel mogelijk integreren met *financial audit*, ook al is dit echt ander werk.

Als je zo'n club hebt, dan moet je nog steeds oppassen voor ritualisering. Want voor je het weet, worden ze 'ingepakt' door beleidsdirecties of hoe die mogen heten. Officieel zeggen die hun medewerking toe, maar als de deur dicht is, haalt men de schouders op en gaat men vaak over tot de orde van de dag. Je zult je dus heel slim moeten organiseren.

Echter, vrienden, robuuste effectstudies voor de grote programma's, dat is te hoog

gegrepen voor zo'n klein clubje. Die kunnen een tevredenheidsonderzoek doen of een quick scan. Serieuze effectstudies moet je niet zelf doen – denk aan de eerste trend – die moet je uitbesteden en dan niet zomaar aan de eerste de beste, maar aan absolute specialisten. Dat moet je alleen incidenteel doen, je hebt niet elke dag een groot beleidsprogramma dat op die manier tegen het licht gehouden moet worden.

Het eigen 'clubje' kan wel de selectie van het onderzoeksbureau voor zijn rekening nemen en de aanbesteding organiseren. Mits er natuurlijk één of enkele mensen in zitten die een stevige verankering in dat werk hadden en nog hebben. Anders wordt het weer amateurisme. Door zo te werken ben je ook van het probleem af dat je geen impact hebt, want in deze situatie heb je op jaarbasis misschien vijf of zes grote rapporten en die zullen zeker niet ongemerkt in de la verdwijnen. De organisatie zal zich er wel degelijk om bekreunen. Het kost veel geld, maar je hebt wel uitstekend onderzoek met uiterst relevante resultaten in handen.

Daar staat tegenover dat er mensen zijn die zeggen dat je het onderzoek juist wel in huis moet uitvoeren, omdat je er alleen dan van verzekerd bent dat het onderwerp beklijft. Omdat je de ambtelijke organisatie vanaf het begin meeneemt in wat er daarna moet gebeuren. Ik ben het daar niet mee eens. In mijn visie zorgt de kleine onderzoeksclub voor de agendering en voor tijdigheid van rapportages. Bovendien beschikt het over enige expertise om bij de aanbesteding de goede keuzes te kunnen maken. Maar het onderzoek zelf doen ze niet. Daarvoor hebben ze ook geen materiedeskundigen of methodologen in huis. Zij zijn alleen de 'lopers' die het organiseren.

Grote organisaties beschikken wel over eigen onderzoekscapaciteit. Zoals het SCP (het Sociaal en Cultureel Planbureau), het WODC bij het ministerie van Veiligheid en Justitie en het KIM, het Kennisinstituut voor Mobiliteitsbeleid van het ministerie van Infrastructuur en Milieu. Bij het WODC hebben we 100 fte en een budget van € 17 miljoen. Daar werken hoogleraren en methodologen. Dat is ook zo bij de paar andere 'Haagse' onderzoeksinstituten/planbureaus. Maar provincies ... dat is een andere schaal, die kunnen dat niet aan. Als een provincie toch een SCP-achtige constructie wil maken, dan ben je er niet met 3,9 fte en een beetje geld. Dan moet je daar fors in investeren. Ik raad dat af. Kennis kan ook gedeeld worden (via de repositories bijvoorbeeld of in samenwerkingsarrangementen).

Kwaliteit

Een belangrijke vraag is vervolgens hoe je de kwaliteit van evaluaties borgt. Ik vind dat in dat kleine clubje – waar ik dus voorstander van ben – enkele fte's aanwezig moet zijn van mensen die uit de onderzoekswereld komen. Die zijn te vinden.

Zij beschikken over een netwerk en weten de weg als ze op zoek moeten naar kwalitatief hoogstaande onderzoekers en onderzoeksbureaus. Als het iemand is die dat niet in zich heeft en in zijn eentje op de 17e verdieping van het hoofdgebouw zit te 'studeren', dan wordt het niks.

Over de samenstelling van zo'n clubje is nog meer op te merken. Je zult er een paar 'piskijkers' bij moeten hebben. Eigenlijk zoals mijn rol bij de ontwikkelings-samenwerking. De IOB is een intern verzelfstandigde eenheid van het ministerie, met een hoogleraar als directeur die met zijn positie en bevoegdheden buiten het schootveld van de minister ligt. Maar toch is – op verzoek van toenmalig minister Koenders – een extern adviespanel georganiseerd, bestaande uit drie Nederlanders en een buitenlander, dat vier keer per jaar scherp kijkt naar wat zij aan het doen zijn. Niet als een grote toezichthouder, daarvoor is dat panel te klein, maar als een buitenboordmotor die de organisatie mede op koers houdt en toetert als het onzes inziens de verkeerde kant op gaat.

Als provincie kun je evaluatieonderzoek en zeker dat naar effecten maar beter niet in eigen beheer houden, want dan word je de slager die zijn eigen vlees keurt. Het kost je nauwelijks iets zo'n panel van mee- en piskijkers, want experts op dit gebied vinden het doorgaans spannend om het te doen en met een beetje vacatiegeld ben je rond. Mensen moeten er ook niet al te veel geld voor krijgen want anders worden ze belanghebbend. Er is zoveel evaluatieonderzoek in Nederland en er zijn zoveel uitmuntende onderzoekers die daar al jaren mee bezig zijn, dat je zo een aantal geschikte mensen daarvoor kunt vinden. Wat in ieder geval nooit mag gebeuren, is dat de organisatie een paar mensen, die men elders liever kwijt dan rijk is, in zo'n clubje parachuteert. Overheden hebben nog wel eens de neiging zo hun personele problemen op te lossen. Ook daar moet gekeken worden naar kwaliteit.

En er is nog een belangrijk punt om je kwaliteit te borgen. Elke beslissing rondom aanbesteding komt wat mij betreft in een databestand. Ik geef een voorbeeld hoe dat bij het WODC gaat. Wij werken al 25 jaar met aanbestedingsprocedures en beschikken over een bestand waar 220 instituten in zitten. Dat zijn bureaus die allemaal op enig moment geoffreerd hebben en waarvan er een aantal in de loop der jaren ook producten hebben afgeleverd. In dat bestand staat precies hoe goed ze waren, wat voor narigheid ze bezorgd hebben, of er geld terug is gevorderd en nog veel meer. Dat is de basis van je lerend vermogen. Wat je bij veel provincies en andere overheden ziet, is dat die informatie alleen in de hoofden van sommige medewerkers is opgeslagen. En dus van de ene op de andere dag verloren kan gaan. Dat werkt niet. Je moet dat externaliseren, buiten jezelf plaatsen en transparant maken. Je kunt dat gewoon met open source software organiseren. Je moet

alleen 0,4 fte hebben om het in te kloppen. De eerste twee jaar zullen mensen zeggen: 'Wat hebben we hier in hemelsnaam aan?' Maar dat geluid zal verstommen, als men er achter komt dat je door hierin te investeren je eigen kwaliteit op niveau brengt.

Een klein kwalitatief hoogstaand clubje dus, lerend vermogen en een paar grote 'klappers' van onderzoeken. Onderzoeken die allemaal in de openbaarheid komen. Je kunt evaluatieonderzoek niet achter de gordijnen uitvoeren. Dat wil niet zeggen dat je het aan de betreffende geleerde overlaat om het tijdstip te bepalen waarop het rapport openbaar wordt. Je kunt daar hele goede afspraken over maken. Bijvoorbeeld: het rapport wordt door de externe onderzoeker afgerond, een begeleidingscommissie met in meerderheid onderzoekers/geleerden geeft haar fiat en vervolgens krijgt de provincie drie maanden de tijd om te bepalen wat ze ermee wil gaan doen. En na die periode wordt het openbaar gemaakt.

Kwaliteit heeft ook te maken met je informatiehuishouding, met de manier waarop je met data omgaat. Ik geef een voorbeeld op het terrein van Veiligheid en Justitie. De politie heeft tal van bestanden waar ze mee werkt: bedrijfsvoeringssystemen, bestanden met gegevens over verdachten die worden opgepakt, etc. Het Openbaar Ministerie heeft weer andere bestanden met gegevens over bijvoorbeeld: hoeveel verdachten van de politie bij het OM komen, wat daar mee gebeurt en hoe lang zaken duren. En daarnaast zijn er nog veel meer organisaties met allereerste gegevens: de rechterlijke macht, de Raad voor de Kinderbescherming, Slachtofferhulp Nederland, Bureau HALT. Deze informatie bestaat al jaren en wordt beheerd door de Justitiële Informatiedienst van het ministerie van Veiligheid en Justitie in Almelo.

Aan het begin van deze eeuw kwamen we er achter dat het met moderne technieken als *dataspaces* en relatiemanagers heel goed mogelijk is om data uit die verschillende bestanden met elkaar te verbinden.

Vanuit evaluatieoogpunt kunnen wij nu dus laten zien hoe het zit met doorlooptijden, wachttijden, verdwenen zaken, verdwenen respondenten, doden in politieregisters – er zitten in Nederland 8.000 mensen in politieregisters die al lang dood zijn –, enzovoort. We kunnen zo op gezette tijden de *flow* en *stock* van het hele terrein van Veiligheid en Justitie (in dit geval betreffende criminaliteit en strafrecht) laten zien. Dan denk je misschien: 'Dat zal wel gigantische hoeveelheden geld kosten'. Helemaal niet, die data zijn er al en met een paar fte (en relevante software) kom je een heel eind.

Ergens in het provinciehuis moet een eenheid zijn – vroeger zou je het misschien

registratuur en statistiek hebben genoemd – die zich daarmee bezighoudt. Het gaat om het ontsluiten van akten, prestaties, activiteiten enzovoort in een digitale wereld. Bij provincies kun je dan denken aan gegevens over: subsidies, vergunningen, inspecties, heffingen, onderhoud van wegen, incidenten, meldingen, maar ook uitspraken van bestuurders en toezeggingen die ze hebben gedaan.

Je hoeft dus zelf geen onderzoek te starten om die gegevens boven water te krijgen. Mijn advies voor de verankering van evaluatieonderzoek is dan ook om een programma te ontwikkelen en een plan te maken voor aanwezige registerdata, hoe je die kunt opschonen en veredelen. Je kunt dan ook *early warnings* gaan formuleren. Niemand kan dan meer aankomen met het argument dat je opmerkingen gebaseerd zijn op een veel te kleine steekproef. Nee, ze zijn gebaseerd op duizenden data over een reeks van jaren: een *compelling force of evidence*.

Een laatste manier om de kwaliteit te verankeren is het openbaar maken van de veronderstellingen die aan het beleid ten grondslag liggen. Er is vandaag de dag heel veel informatie over beleidstheorieën – zeker een paar honderd gezaghebbende studies inmiddels – en als onderzoekseenheid binnen een provincie zul je de beleidsmensen moeten 'verleiden', misschien zelfs dwingen om zich daartoe te verhouden en de assumpties die aan het beleid ten grondslag liggen openbaar te maken.

In dat verband herinner ik mij mijn eerste contact ooit met een lid van de Tweede Kamer, dat was Gerd Leers, eind jaren 80/begin jaren 90. Naar aanleiding van het centraal thema onderzoek 1991 van de Algemene Rekenkamer naar de stand van de beleidsevaluatie wilde hij een motie opstellen en indienen waarin, analoog aan de Memorie van Toelichting bij wetgeving, er ook een Memorie van Toelichting zou moeten komen bij elk nieuw beleid. Alle nieuw beleid van enige omvang zou in zijn gedachte de beleidstheorie, waar de voorstellen op gestoeld waren, moeten articuleren. De motie heeft het op een paar stemmen na niet gehaald maar is volgens mij nog steeds heel relevant. Je zou bij provinciaal beleid dus een oplegnotitie over de beleidstheorie kunnen afdwingen voordat de politieke besluitvorming erover plaatsvindt. Ik ben ervan overtuigd dat dat het debat aanscherpt en de kwaliteit van het beleid ten goede komt.

Programmering

Het formuleren van meerjarige onderzoeksprogramma's is heel ingewikkeld. We zijn als Panel van Advies voor de IOB daar recent mee bezig gegaan. Waar wij tegen aanliepen is dat er wel een heel keurig interactief proces voor ontwikkeld is en dat er aan het eind van zo'n rit een fraai programma uitrolt, maar een echt

intelligente onderbouwing zit daar nog niet in. We wilden weten hoe we dat op een betere manier kunnen doen en hoe dat in andere landen gaat. Overigens geldt deze bevinding zeker niet alleen voor deze inspectie, maar breder.

Eén van de gedachten is dat je, los van het programma dat je *bottom up* in co-creatie met andere partijen maakt, als onderzoeksinstituut moet zeggen: wij willen in ieder geval dat deze twee of drie projecten hoe dan ook altijd plaatsvinden. Die zijn van zo'n vitaal belang dat ze buiten de programmering staan en altijd doorgaan. Voor het WODC is dat bijvoorbeeld de recidivemonitor. Wij weten, vanaf begin jaren 90, van alle Nederlanders hoe vaak ze met politie en justitie in aanraking zijn gekomen. We volgen hen door de jaren heen. Als iemand zou zeggen: we moeten maar ophouden met die monitor dan is dat hetzelfde als wanneer je tegen het KNMI zou zeggen: de komende vijf jaar verzamelen we geen gegevens meer over het weer. Een ander onderzoek dat we altijd blijven doen is onderzoek naar de manier waarop Nederlanders hun geschillen oplossen. Dat is maatschappelijk zo belangrijk dat het moet worden uitgevoerd. Essentiële onderwerpen moet je dus altijd buiten de programmering houden. Doe je dat niet, dan kunnen ze onderdeel worden van allerlei nut- en noodzaakdiscussies, met alle risico's van dien. Je moet wel heel zorgvuldig kiezen voor welke onderwerpen je dat wilt doen. Als je 25 onderwerpen zo wilt behandelen, dan kan dat niet.

Wat we ook wel hebben geprobeerd is om de programmering in één plaatje samen te brengen, gerangschikt naar de kerndoelen en subdoelen van het beleid. Zodat je kunt zien of je spreiding van onderzoeksonderwerpen enigszins op orde is. Dat geeft enig houvast, maar grosso modo blijft een meerjarige onderzoeks-programmering best ingewikkeld.

En dan moet je daarbij ook nog eens bedenken dat Nederland een geschiedenis heeft in programmering. In de jaren 70 is het geprobeerd met behulp van de zogenaamde nationale programma's. Die waren bepalend voor de onderzoekswereld. Vervolgens is het co-creatie geworden, waar je partijen met elkaar het gesprek liet voeren over wat er moest gebeuren en een gezamenlijke verantwoordelijkheid gaf voor het eindresultaat. Als je me vraagt: Is er ook nog een derde weg? Ik zie hem niet.

Burgers

Wat de rol van de volksvertegenwoordiging is bij evaluaties? Allereerst wil ik daarbij opmerken dat een flink deel van de onderzoeken die wij bij het WODC uitvoeren, rechtstreekse verzoeken zijn van de Tweede Kamer. Dus de link met de volksvertegenwoordiging is duidelijk.

Maar misschien moet je ook nog wel een andere, meer principiële vraag stellen: Is de volksvertegenwoordiging het volk? In dat verband is het goed om stil te staan bij iets dat *civic evaluation* wordt genoemd. Daarbij gaat de burger zelf als een evaluator te werk: hij bestudeert de doelen van het beleid, verzamelt en analyseert gegevens die hij ophaalt uit bronnen van allerlei instituten en confronteert daarna de overheid met zijn bevindingen. Via het open databeleid dat in Europa is afgesproken komen steeds meer evaluaties van onderzoeksinstituten ter beschikking. Dat geldt ook voor statistische gegevens van de nationale bureaus. Eén van de effecten daarvan is dat burgers, al dan niet georganiseerd, elkaar steeds vaker vinden rond thema's die voor hen belangrijk zijn en in de meest uiteenlopende netwerkconstructies voor hen relevante onderwerpen bij de kop pakken en uitspitten. Het is het begin van een nieuwe traditie aan het worden in Europa.

Ik wil in dat verband ook wijzen op het essay van Marc Chavannes in de publicatie *'Turven, tellen en toetsen'*, het jubileumboek uit 2007 van Vide, de beroepsvereniging voor toezichthouders, inspecteurs, handhavers en evaluatoren. Chavannes laat daarin met bijzonder fraaie voorbeelden zien hoe het establishment in de Verenigde Staten niet meer om de burger heen kan. Het zijn juist de burgers die zich roeren en in toenemende mate onrechtmatigheden aan het licht brengen.

In de App-wereld zie je de ene na de andere toepassing op de markt komen waarmee gegevens op alle mogelijke terreinen verzameld kunnen worden. Ik ken schitterende studies bijvoorbeeld waarin via burgers het beleid op het gebied van geluidsoverlast wordt getoetst. Met hun smartphones verrichten ze geluidmetingen die *realtime* worden doorgestuurd en opgeslagen in een centraal gegevensbestand, van waaruit weer analyses gemaakt kunnen worden. Dat is een volstrekt nieuwe benadering. Nu is het meestal zo dat een gemeente na klachten van burgers een buurtonderzoek naar geluidsbeleving houdt ('Hebt u veel of weinig last van de disco?') en als dat niet voldoende gegevens oplevert misschien één of twee keer met een geluidswagen uitrukt om de decibellen te meten. In zo'n nieuwe setting brengen honderden burgers geheel vrijwillig duizenden metingen bij elkaar en zorgen daarmee voor een gigantische hoeveelheid gegevens waar de gemeente niet omheen kan.

Niemand weet op dit moment nog precies hoe je dit soort acties van de burger moet incorporeren in de formele evaluatieprocessen, maar dat het een ontwikkeling is waar enorme mogelijkheden liggen, is voor mij duidelijk. Wat ook duidelijk is, is dat je op twee manieren met zo'n ontwikkeling kunt omgaan: of je laat de burger links liggen of betreft hem erbij. Als je echt wilt innoveren, vind ik dat je in deze laatste richting verder moet zoeken. Ik vind dat een fascinerend iets. Ook de

volksvertegenwoordiging zal moeten aantonen dat zij echt de burger representeert. Want slaagt ze daar niet in, dan zal de burger steeds meer zijn eigen weg gaan en met nieuwe media het bestuur gaan bestoken.

Actoren

Provincies voeren hun beleid uit in samenspel met andere overheden, bedrijfsleven en maatschappelijke organisaties. Bij evaluaties loop je dan tegen het probleem aan hoe je kunt laten zien wat de interventie van die ene actor te midden van al die andere actoren heeft opgeleverd. Met andere woorden: Hoe doeltreffend en doelmatig was de euro die de provincie heeft ingezet, naast de euro's van die andere partijen?

Daar is eigenlijk maar één oplossing voor. Je moet daarvoor de cruciale actoren redelijk vroeg in het traject bevragen: 'Wat is uw verwachting van dit programma? Waarom doet u mee? Wat is uw bijdrage in termen van middelen en resultaten?'. Je brengt dus in kaart hoe de actoren die erover gaan het gewicht van hun eigen bijdrage wegen. Dan krijg je bijvoorbeeld van de zeven belangrijkste actoren een soort overzichtsplaatje. Je kunt door middel van *infographics* laten zien hoe partijen zich tot elkaar verhouden. En wanneer je in een *netwerkevaluatie* naar de resultaten van een programma hebt gekeken, dan ga je de partijen confronteren met wat zij in het begin hierover hebben gezegd. Dan wordt hun bijdrage aan het geheel ook in het goede perspectief geplaatst. En dan kan de muis misschien wel roepen dat hij zo lekker samen met de olifant heeft gestampt, maar de andere deelnemers zullen hem dan toch tot de orde roepen en zijn verhaal in perspectief plaatsen.

Wanneer het alleen gaat om uitvoering, dan is het wel mogelijk om ieders bijdrage scherp in beeld te krijgen. Er is namelijk altijd iemand verantwoordelijk voor (een deel van) de uitvoering. En als het daar mis is gegaan, dan kun je in ieder geval uitspraken doen over de aansturing.

De enige andere manier om de mate van doeltreffendheid en doelmatigheid in kaart te brengen, is om elke afzonderlijke bijdrage te kwantificeren in termen van zijn betekenis. Ik denk niet dat dat haalbaar is. Met een project van € 600 miljoen kun je het misschien proberen, maar als je niet aan die schaal komt – en voor de provincie is dat doorgaans het geval is – dan is het een heel inefficiënte manier van onderzoek doen. Meer aandacht voor de *cost and benefit* van evaluaties is in dat verband ook een belangrijke zaak.

Colofon

Interviews: Zuidelijke Rekenkamer

Landschapsfotografie: Henk van den Berg, Oss

Vormgeving: Henk van den Berg Creatieve Communicatie, Oss

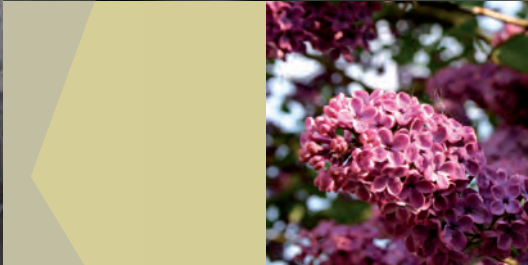
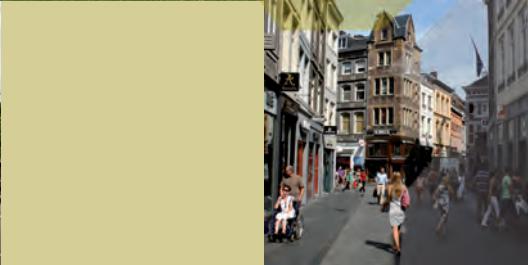
Drukwerk: Drukkerij Wihabo, Geffen

Artikelen of delen daarvan mogen alleen uit deze publicatie worden overgenomen met bronvermelding en goedkeuring van de Zuidelijke Rekenkamer.

ISBN 978-90-8768-035-0



9 789087 680350



Postbus 2000
5600 CA Eindhoven
Keizersgracht 5
5611 GB Eindhoven
tel. 040 - 232 93 38
www.zuidelijkerekenkamer.nl