



## Wat evaluatoren van kritiek kunnen leren

*Prof. dr. A.F.A. Korsten*

### Aanbevolen citeerwijze bij dit artikel

Prof. dr. A.F.A. Korsten, 'Wat evaluatoren van kritiek kunnen leren', *Beleidsonderzoek Online* december 2013, DOI: 10.5553/Beleidsonderzoek.000030

Is er ook kritiek mogelijk op evaluatieonderzoek van overheidsbeleid, waar evaluatoren, in casu visitatiecommissies of onderzoekscommissies, wat van kunnen leren? Daarop moet worden ingegaan omdat dit weer positief kan inwerken op het 'leren van evalueren' in organisaties. En 'leren' hoort tot de principes uit de 'Code voor goed openbaar bestuur' (kabinet-Balkenende IV).

Eerst dit: waarop komt weinig of geen kritiek? Op enkele belangrijke aspecten van beleidsevaluatie.

1. Opvallend is dat de *onafhankelijkheid* van de evaluator zelden kritiek krijgt. Onafhankelijkheid is overigens wel belangrijk, ook als een visitatiecommissie aantreedt of een parlementaire (evaluatie)commissie. Leden van een parlementaire commissie moeten hun lidmaatschap van een bepaalde partij als het ware terzijde schuiven en ze moeten gaande het onderzoek spelregels in acht nemen, zoals niet lekken (Van der Zwan & Noordegraaf, 1998: 71).
2. De *onderzoeksaanpak* moet te allen tijde helder gestructureerd zijn. De keuze van de opzet is belangrijk omdat dan moet blijken om welk beleid het gaat: om beleid in een één-actorsituatie of in een multi-actorennetwerk?

Waarop bestaat wel kritiek? Ik kom tot de volgende inzichten na het raadplegen van vele evaluatiestudies en op basis van eigen ervaringen als evaluator, lid van visitatiecommissies, opdrachtgever tot evaluatie en voorzitter van begeleidingscommissies bij evaluatiestudies en als toezichthouder gedurende enkele decennia.

### 1 Te weinig bezinning op de opdracht

De eerste kritische waarneming behelst dat onderzoekers *vaak (te) laat bij de evaluatieopdracht betrokken worden*. Deze ervaring baseer ik op tal van opdrachten die ik zelf heb uitgevoerd en op kennisname van studies van anderen.

Waarom? Het proces van opdrachtverlening is vaak sequentieel. Het besluit tot evaluatie is gevallen, er zijn contouren aangegeven over het wat, waarom en hoe van het evaluatieonderzoek en het budget ligt vast. Dan komt de evaluator pas aan bod. Als een externe evaluator gezocht wordt, mag die dan met medestrevers een offerte uitbrengen? Met die offerte gaat hij 'teruggeven' (herformuleren in eigen woorden) aan de opdrachtgever hoe het onderzoek wordt opgevat. Dat is een rem op durf en creativiteit. Juist in de startfase zou eens vrijmoedig gesproken moeten worden, ook aan de kant van de opdrachtverleners over vragen: wat willen we nou echt met het onderzoek?; wat niet?; wat moet het voor soort informatie opleveren? De entreebesprekingen tussen opdrachtgever en onderzoeker(s) en eventueel ook de betrokkenen uit het beleidsnetwerk zijn van groot belang. Bijvoorbeeld omdat blijkt dat een opdracht een compromis is tussen twee departementen met een geheel verschillende invalshoek. Is drugsverslaving een gezondheidsvraagstuk of iets van wetshandhaving of worden beide (on)elegant 'verzoend' in een opdracht tot evaluatie van centra voor verslaafdenopvang? Het belang van bezinning op de startsituatie wordt vaak onderschat.

## 2 Begrip van doelbereiking schiet tekort

Een tweede kritiek raakt de verklaring van output en/of effecten (outcomes).

Tal van evaluaties kijken naar de bereikte prestaties of resultaten, zelfs naar effecten en naar de processen ernaartoe. Maar ze 'leggen geen relatie met de daarvoor verantwoordelijke mechanismen'. Ze gaan niet terug naar het 'waarom' van de effecten.

'De achterliggende beleids- en interventietheorie(en) worden niet geproblematiseerd, gereconstrueerd noch getoetst',

aldus Leeuw en Van Gils (2010: 170) op basis van hun analyse van een steekproef van 30 evaluaties uit een totale populatie van 100 op het terrein van het ministerie van Buitenlandse Zaken. Er ontbreekt iets wat lijkt op – streng gesteld – een *oorzakelijkheidsanalyse*. Als beleid effect heeft of geen effect, moet duidelijk worden of dat ligt aan de beleidsinhoud, de in- of uitvoering of aan iets anders. En als beleid de oorzaak is, hoe kan het dan gebeuren dat de doelstellingen niet gehaald worden? Die vraag moet beantwoord worden.

De implicatie daarvan is duidelijk. De onderzoekers leren de lezers nogal eens te *weinig over het waarom* van het wel of niet bereiken van bepaalde resultaten. Wat nog wel gebeurt, is dat aangegeven wordt of de bereikte resultaten en effecten eventueel overeenstemmen met de beleidsintenties en afspraken.

Evaluaties volstaan vaak met de analyse van doelbereiking. Weinig gebruiksgericte evaluatiestudies zijn *echte* effectiviteitsanalyses. Waarom? Omdat de onderzoekers niet gevraagd wordt te zoeken naar de 'mechanismen', omdat het budget voor de analyse beperkt is en omdat er altijd weer haast is. Onderzoekers worden in feite nogal eens te veel opgejaagd. Dat betekent, anders gesteld, dat de opdrachtgevers niet echt interesse hebben voor degelijke effectiviteitsanalyses.

Veel analyses zijn bovendien slechts procesevaluaties.

Deze aanpak van doelbereikingsgerichtheid heeft gevolgen voor de beleidsrelevantie. Onderzoekers die een bestaande beleidstheorie niet ter discussie kunnen of mogen stellen en ook niet zoeken naar een alternatief of een dynamisering van de beleidstheorie, bereiken dat de beleidsvoerders padafhankelijk blijven opereren. Ze blijven denken in termen van het beleid als potentieel succesvol, terwijl het beleid misschien rigoureuus op de schop moet (zie ook Van Twist, 1993).

Een voorbeeld. Een provinciebestuur zette een omgevingsbeleid op dat moest leiden tot het voldoen aan milieunormen door agrarische ondernemers. De stankoverlast van bedrijven van agrarische ondernemers moest verminderen en de ammoniakcijfers (bijvoorbeeld als gevolg van varkenshouderij) moesten omlaag. Dat leidde tot een provinciaal voornemen om voor bepaalde landbouwbedrijven geen bedrijfsuitbreiding toe te staan. En andere bedrijven zouden moeten worden verplaatst. Daarvoor was een bepaald budget beschikbaar. Na enkele jaren bleek dat amper een varkenshouderij verplaatst was naar een gebied waar de ondernemer binnen aanvaardbare stanknormen bleef. Het beleid was dus volkomen vastgelopen. De doelstellingen werden niet bereikt. Waarom? De centrale veronderstelling achter het verplaatsingsbeleid bleek niet juist. De overgrote meerderheid van agrarische ondernemers was niet bereid om opnieuw fors in een elders gevestigd bedrijf te investeren. De meesten die in aanmerking kwamen, waren al ouder dan 50 jaar en hadden bovendien geen bedrijfsopvolger. Dus zouden de eventuele schulden bij een bedrijfsverplaatsing tot ver na hun 65ste levensjaar om hun nek blijven hangen. Ze redeneerden: 'We zitten het wel uit op de bestaande locatie, ook als geen verdere uitbreiding meer mogelijk is.' Eigenlijk zou in dit geval het roer fors om moeten: een herziening van de achterliggende beleidstheorie diende zich aan. Kwam die er ook? Nee. De evaluator wachtte af of het provinciebestuur het beleid wil stoppen.

### 3 De keuze van de centrale maatstaf

Een derde kritiekpunt slaat op het ontkennen van een aanwezig beleidsnetwerk.

De keuze voor een effectiviteitsevaluatie of een zwakkere variant daarvan impliceert de keuze van doelstellingen die achter beleid *van een centrale overheidsactor* schuilgaan. Dat is een gerechtvaardigde aanpak ingeval deze overheid *volledige* doorzettingsmacht heeft. Dan stelt de centrale actor beleid vast en voert dit ook zelf uit. Bij veel beleid is echter de doorzettingsmacht niet slechts in één hand; er is sprake van *afhankelijkheden* tussen overheden en misschien ook wel in bredere zin (Abma & In 't Veld, 2001; Bekkers & Lips, 1998).

Neem het landschapsbehoud in Zuid-Limburg. Op een in 2013 gehouden conferentie bleek dat heel veel partijen bij elkaar waren met elk eigen taken, bevoegdheden en middelen: de rijksoverheid, het waterschap, de provincie, de dienst landelijk gebied, de gemeenten, landbouwers, recreatieondernemers (pensionhouders, campinghouders, de eigenaar van een skipiste, enz.), chemische bedrijven, de regionale VVV, enzovoort. In dit geval moeten doelen verknoot worden en middelen gebundeld. We zitten daarmee bij netwerkmanagement.

Verdedigers van netwerkmanagement noemen het uitgaan van gerichte doelstellingen van één overheidsactor een eenzijdige en verkeerde *top-down*

benadering (Klijn & Koppenjan, 2001: 189). In een netwerkcontext moet juist worden uitgegaan van een benadering *van onderop*, door de denkwijzen en oriëntaties van *alle* relevante actoren uit het beleidsnetwerk mee te nemen. Wat levert dat op? Erkend moet worden dat zich in het multi-actorennetwerk vaak dynamiek voordoet. Dat doelstellingen stabiel blijven, is een misverstand, aldus Klijn en Koppenjan (2001). Er is sprake van 'de onhoudbaarheid van een eenduidig inhoudelijk ijkpunt'. Bij een evaluatie van beleid dat in een netwerkcontext van van elkaar afhankelijke actoren tot stand kwam, denk maar eens aan de evaluatie van de vernattingspolitiek van De Peel, moeten de doelstellingen van alle actoren meegenomen worden. Al de actoren moeten in feite bij de opdracht tot evaluatie betrokken worden. Zij moeten immers ook de resultaten relevant vinden en benutten.

#### 4 De reikwijdte van het onderzoek

Een vierde kritiekpunt: veel evaluatiestudies hebben een *beperkte geldigheid*. De bevindingen slaan alleen maar op wat is onderzocht en zijn ook begrensd in de tijd. Dat valt in elk geval op te maken uit een studie naar evaluaties op het terrein van het ministerie van Buitenlandse Zaken (Leeuw & Van Gils, 2010: 170). Deze evaluaties zeggen bij wijze van spreken iets over het beleidsprogramma in jaar 1 en eventueel 1 en 2 in de plaatsen a tot en met h. Maar ze hebben niet veel reikwijdte naar jaar 3 en de plaatsen i, j en k. Van generaliseerbaarheid is niet of nauwelijks sprake.

Dat kan te rechtvaardigen zijn omdat *gebruiksgerichte evaluatie* niet tot doel heeft om de kennis in een wetenschapsgebied uit te breiden, maar om bij te dragen aan beter beleid en/of betere besluitvorming of – nog voorzichtigte – 'goede reflectie'. Toch zou meer generaliseerbaarheid hier en daar mogelijk zijn. Denk maar eens aan gebiedsgericht beleid. Het is wel zo nuttig als een analyse van de politiek in De Biesbosch-Noordwaard, Kerngebied Vechtplassen en de Gelderse Poort inzichten oplevert die ook elders van nut zijn (zie Cörvers et al., 2003).

Zoeken naar *synthese* dient zich als opportuniteit aan. Want het is vaak niet zo dat wat overheid x laat onderzoeken zo uniek is dat een andere overheid y dit onderwerp nog nooit bij de kop heeft gepakt. Wie kijkt naar studies van lokale rekenkamers en rekenkamercommissies, ontdekt dat veel van deze organisaties in een bestek van een jaar of vijf een aantal zelfde of sterk vergelijkbare studies ondernemen. Blijkbaar kijkt men bij het opvoeren van mogelijke onderzoekthema's naar wat andere rekenkamers onderzochten.

#### 5 Veldkennis te gering?

Uit het voorgaande vloeit voort dat veel evaluatieonderzoekers wel wat weten van onderzoek verrichten maar gemiddeld genomen zelf *niet uitblinken in kennis van het terrein dat ze onderzoeken* (Bongers & Den Hertog, 2007). Veld- of sector kennis zit niet in het onderzoeksteam dat evalueert. De evaluatoren kijken naar doelstellingen maar laten zich *niet* (sterk) in met de beleidsinhoud. Die wordt als onomstreden vertrekpunt aanvaard.

Dat heeft het voordeel dat de onderzoekers *enige afstand houden* tot de sectorcultuur, waarin 'heilige huisjes' en taboes bestaan. Maar daar staat een nadeel tegenover. Dat zou eigenlijk gecorrigeerd moeten worden. Onderzoekers moeten thuis zijn of verder thuis geraken in de materie, zeker als het om een

complexe casus gaat, zeggen Van der Zwan en Noordegraaf (1998: 71), die het Gemeentelijk Vervoerbedrijf uit Amsterdam hebben doorgelicht. Een ander nadeel is dat wie zich niet grondig verdiept in de beleidsinhoud, ook *moelijk met radicale suggesties* kan komen om beleid over een andere boeg te gooien. Hier dreigt dus iets te ontstaan wat eerder padafhankelijkheid is genoemd. Onderzoekers blijven dan ‘werken’ binnen het bestaande beleidsparadigma (lees: ‘frame’).

Tineke Abma (2001) is een van de evaluatoren die meent dat de onderzoeker een multiële evaluatie kan kiezen. Dat impliceert dat de beleidsinhoud zelf ter discussie kan komen. Dan moet je dus over materiekennis beschikken.

## 6 Te weinig nieuws?

Een zesde kritiekpunt op beleidsevaluaties impliceert dat gebruikers van evaluaties wel eens aangeven dat ze *niet veel nieuws* geleerd hebben van een evaluatiestudie (Leeuw & Van Gils, 2010: 172).

Dat is begrijpelijk als evaluatoren zich niet inlaten met de beleidsinhoud (zie kritiek nr. 5). Evaluatoren die zich niet inlaten met de beleidsinhoud, of die zelfs maar beperkt begrijpen, neigen ertoe om padafhankelijk te blijven denken en dus slechts te komen met suggesties in de vorm van ‘kleine stappen’.

Hier is wat aan te doen door enkele maatregelen.

1. Voor de start van een evaluatie grondiger *nadenken over de evaluatieopzet*. Moet de evaluatie een *multiële of multiperspectivische evaluatie* zijn vanuit meerdere perspectieven in plaats van alleen een doelbereikings- of een effectiviteitsevaluatie? Wat worden maatstaven ter beoordeling van beleid: alleen effectiviteit vanuit een overheid gezien? Is de evaluatie een evaluatie van beleid dat in een netwerkcontext tot stand kwam, waarbij de actoren van elkaar afhankelijk zijn omdat geen actor de volledige doorzettingsmacht heeft?
2. Een tweede suggestie houdt in om *het prospectieve* explicieter onderdeel te laten zijn van de evaluatie. Een evaluator die gevraagd wordt om een wet te evalueren - stel een mestbeperkende wet voor varkenshouders - kan pogen de bevindingen te doordenken vanuit enkele scenario's die in het land de ronde doen, zoals het scenario van radicale deregulering. Die suggestie sluit aan op denken in termen van een multiperspectivische evaluatie (Korsten, 1988; Ringeling, 1993; Abma, 2001: 314). Mijn ervaring bij de (vroegere) Sociale Verzekeringsraad (SVR) is geweest dat, als dit niet bij de opdracht al is meegegeven, evaluatoren hiertoe niet geneigd zijn. Die weerzin kan gevoed zijn door gebrekkige kennis van het maatschappelijke debat rond een thema. Men is niet in staat om scenario's te formuleren of denkt te veel in politiek vaarwater te belanden. Men heeft de durf van het Rathenau-instituut niet.
3. Een derde suggestie is om voor het verschijnen van een definitief rapport één of meer *expertmeetings* te organiseren waarbij juist betrokkenen uit een sector evaluatieresultaten of stellingen voorgelegd krijgen om op te reageren. Dergelijke meetings kunnen een zeefwerking hebben omdat ze uitfilteren wat in de ogen van experts en veldkenners belangrijk of minder belangrijk is. Ze kunnen ook bijdragen aan het begrijpen van resultaten,

argumentaties toevoegen of corrigeren, en additionele aanbevelingen opleveren.

## 7 Evaluaties te weinig prospectief

Nog een zevende kritiekpunt, dat aansluit op de twee genoemde. Menige beleidsevaluatie is *te veel retrospectief* en *te weinig prospectief*. Evaluaties blijven, anders gesteld, te veel in het verleden hangen, in waar het om ging (doelstellingen) en hoe het gegaan is (uitvoering, prestaties). Ze bieden geen of een te geringe denkact over wat er thans en de komende periode aan knelpunten speelt en wat de oplossingen daarvoor zijn. Het ontbreekt bij evaluatie dus vaak aan beantwoording van de vraag 'hoe verder?'. Terwijl bestuurders na een evaluatie juist erop uit zijn om te weten 'wat te doen?', bijvoorbeeld 'hoe voorkomen we uitwassen'. Denk maar aan onderzoek:

- naar de bankencrisis (commissie onder leiding van Jan de Wit) of
- naar fouten bij onderwijsinstelling InHolland (commissie-Leers),
- naar het gemeentelijk optreden bij de asbestaffaire in Utrecht (commissie-Jansen),
- naar het echec van schoolorganisatie Amaranthis (rapport van de commissie onder leiding van Riet de Wit uit 2012 en van de vervolgc commissie-Halsema uit 2013) en
- naar privatisering en verzelfstandiging van overheidsdiensten (commissie-Kuiper).

Reguliere evaluaties, dus *niet* die van commissies welke na een crisis aantreden, zijn dus nogal eens te weinig prospectief.

De kans is groot dat visitatiecommissies ontkomen aan dit verwijt omdat daarin vaak ex-bestuurders naast evaluatoren zitten die weten dat bestuurlijke relevantie van belang is. Het is natuurlijk vaak ook een kwestie van opdrachtformulering.

Denk ook aan de suggestie om de bevindingen te plaatsen in multiperspectief perspectief.

Is het dan een idee om ex ante evaluatie-elementen te koppelen aan de ex post evaluatie? Wie een geëvalueerde wet ineffectief acht, kan een nieuw wetsontwerp doordenken of ontwikkelen en via *spelsimulaties* 'uittesten' (Termeer et al., 1995).

### Wat is nog als reactie op kritiek mogelijk?

Er zijn nog diverse suggesties gedaan. Ik noem vier alternatieve typen evaluatie:

1. responsieve evaluatie (Guba & Lincoln; Abma) en multi-evaluatie (Ringeling);
2. argumentatieve evaluatie (Fischer, Pröpfer);

3. netwerkgericht evalueren (Klijn, Koppenjan);
4. lerend evalueren (Edelenbos).

Ik licht deze toe en bekijk de eventuele meerwaarde.

### ***Alternatief 1 & Responsieve evaluatie: rekening houden met belanghebbenden***

Welk alternatief is er dat enigszins aan de kritiek op bestaande beleidsevaluatie tegemoetkomt? In een historische schets van evalueren maken Egon Guba en Yvonna Lincoln (1989: 22-31) duidelijk dat er volgens hen *vier generaties evaluatie* bestaan hebben (zie ook Abma, 1995, 1997, 2000) en dat de *vierde generatie*, de responsieve evaluatie, een alternatief is voor de andere drie. Welke generaties zouden er dan zijn?

1. De eerste generatie evalueren is te beschouwen als *'meten'*. De beleidsonderzoeker profileert zich als een technicus en richt zich op kwantitatieve gegevens. Tegenwoordig wordt dit wel aangeduid met monitoring (aldus Bongers & Den Hertog, 2007: 32).
2. De tweede generatie evalueren richt zich op *'beschrijven'*. De evaluator analyseert de sterkten en zwakten van beleid en betreft die op de gestelde doeleinden. De *'goal achievement'* komt in beeld.
3. De derde generatie evaluaties richt zich op *'beoordelen'*. De evaluator geeft een soort finaal oordeel als een rechter.

Guba en Lincoln (1989) zijn van mening dat er in deze generaties beleidsevaluatie enkele gemeenschappelijke tekortkomingen zijn aangetroffen.

1. Ten eerste richten de evaluatoren zich in alle drie de generaties evaluatie op de *officiële doelstellingen van het bestuur*. Het is het bestuur dat de opdracht verleent tot evaluatie en de resultaten worden ook weer uitgebracht aan de (machtige) bestuurders alsof er geen belanghebbenden (met minder macht) bestaan die interesse zouden hebben in of een belang bij een evaluatie. Waar staat *'bestuur'*, kan ook worden gelezen *'de leiding van organisaties'*. Beleidsevaluaties zijn in feite *niet machtsneutraal*.
2. Ten tweede gaat er achter de drie generaties evaluatie een gemeenschappelijk geloof schuil, namelijk in *maatschappelijke waardenconsensus*. Terwijl de maatschappij in de kern pluralistisch is geworden op waardengebied. De vraag is daarom *welke waarden* domineren in evaluatieonderzoek. Spelen waarden van andere belanghebbenden dan overheidsbestuurders ook een rol in de evaluatie of zijn deze belanghebbenden alleen geschikt om als informant te fungeren? De vraag stellen is deze beantwoorden.

Beide auteurs, Guba en Lincoln, pleiten voor een vierde generatie evaluatie (*'fourth generation evaluation'*). Ze noemen die *'evaluatie door onderhandelen'*.

Evaluatoren zouden volgens die aanpak moeten kunnen omgaan met divergerende waardeoriëntaties en dus met *meerdere maatstaven* dan alleen

effectiviteit en doelmatigheid. Zeker in omstandigheden dat sprake is van pluriformiteit, en dus van controverses. Tineke Abma (1997, 2000) en Sandra Kensen sluiten hierop in hun werk aan.

Tineke Abma betoogt dat sprake is van eenzijdige rationaliteit in de vorm van instrumentaliteit in het denken volgens de klassieke beleidsevaluatie (volgens Hoogerwerf). Er wordt geen rekening gehouden met perspectieven van andere belanghebbenden dan het overheidsbestuur (Abma, 2001: 313).

Hoe ziet de gewenste responsieve evaluatie er verder uit?

1. De vierde generatie evaluatie moet een *responsieve evaluatie* zijn, volgens Guba en Lincoln (1989). Beide auteurs spreken van '*responsive focusing*'. De vragen in het evaluatieonderzoek moeten in *samenspraak* met belanghebbenden ('stakeholders') tot stand komen. De evaluatiecriteria worden ook ontleend aan de issues waar de belanghebbenden tegen aanlopen, hun 'concerns' en hun 'claims'.
2. De vierde generatie evaluatie wordt ook gekenmerkt door een *specifieke methodologie*. Die heeft te maken met een *bepaalde werkelijkheidsopvatting*. Volgens beide auteurs bestaat er *geen* objectieve werkelijkheid, maar is het maar 'hoe je kijkt' naar de werkelijkheid. Er bestaan meerdere werkelijkheden ('sociale constructies genoemd'). Een evaluatie moet tegemoetkomen aan deze meervoudige wijze van kijken en van de werkelijkheid (de werkelijkheidsconstructies) waarnemen. Hoe dan te werk te gaan? Door te '*Verstehen*': door te proberen om te begrijpen, 'door je in te leven' in de positie van belanghebbenden en door met hen in gesprek te gaan. De werkelijkheidsaanspraken zouden *op elkaar betrokken moeten worden*, wat zou kunnen bijdragen aan meer wederzijds begrip en inzicht in waarover overeenstemming bestaat. Het gaat om zoeken naar consensus tussen de deelnemers aan de evaluatie.

De Rotterdamse hoogleraar Arthur Ringeling (1993) erkent de zin van wat hij noemt een *multipele evaluatie*, waarmee niet gezegd is dat zijn aanpak geheel overeenkomt met die van Abma.

Volgens critici is deze 'vierde generatie evalueren' echter toch te veel een poging om belanghebbenden te laten deelnemen aan beleidsvoorbereiding. Zoals de Twentse hoogleraar Andries Hoogerwerf eens zei: 'Het is geen op waarneming van effecten gerichte beleidsevaluatie meer maar participatie van een onderzoeker in ontwerpprocessen.' Het is een actiegerichte ontwerpaanpak in een context van waardenpluriformiteit.

Toch zijn de opmerkingen van Guba en Lincoln nuttig. Want de neiging bestaat om een beleid te ontwerpen dat vervolgens (bijna) mechanisch 'uitgerold' wordt over de samenleving. Dergelijk beleid, waarin een overheid zich als monocentrisch besturend orgaan situeert, werkt in veel gevallen niet; zeker niet in een netwerkcontext waarin van elkaar afhankelijke actoren zijn opgenomen. Dus beleidsevaluatie zal soms dood kunnen lopen als de netwerkcontext ontkend wordt.

Een waarschuwing is mogelijk? Responsieve evaluatie valt te proberen in een netwerkcontext, maar *pas op voor rolvermenging*. De evaluator dreigt in een rol



te komen van zowel procesregisseur, die partijen bij elkaar brengt in een beleidsontwerp of uitvoeringsproces, als evaluator. Deze responsieve evaluatie is daarom niet zonder problemen. Bijgevolg wordt die ook *niet* vaak toegepast.

Op naar een volgend alternatief.

### ***Alternatief 2 & Evaluatie als argumentatie***

Frank Fischer (1995) is de man die in *Evaluating Public Policy* pleitte voor een verbreding van het gangbare evaluatieonderzoek. De oordeelsvorming moet zijns inziens in evaluatiestudies breder zijn dan dat het alleen zou gaan om effectiviteit of doelmatigheid van beleid. Evalueren moeten zich ook richten op *de betekenis van beleid voor de maatschappij als geheel*.

Fischer pleit voor een *waardengerichte evaluatie*. Evaluatoren moeten minder technocratisch te werk gaan, door voornamelijk of alleen te letten op gegeven beleidsdoelstellingen en daarop aansluitende middelen. Ze evalueren niet of de 'goede dingen gedaan worden', maar meer of 'de dingen goed worden uitgevoerd'. Evaluatoren moeten zijns inziens echter het fundamentele debat bevorderen en *daarbij kan horen* dat ter discussie wordt gesteld of 'de goede dingen gedaan worden'.

Dat betekent dat de beleidswetenschap meer moet zijn dan een rationeel-instrumentalistische wetenschap. Hij bepleit een argumentatieve wending ('the argumentative turn'). Het zou bij evaluatie niet moeten gaan om het beoordelen van instrumenten en het vinden van de goede instrumenten, maar om *overtuigende argumentaties*. En om *dienstbaarheid aan de democratie*. Dit laatste betekent dat *debatanalyse* in beeld komt (zie ook Hoppe & Peterse, 1998).

Is dit alternatief levenskrachtig gebleken, ook ten behoeve van gebruikgerichte evaluaties?

1. Hoppe en Peterse (1998) voelden zich duidelijk geïnspireerd. Zij spreken van de onafwendbaarheid van de argumentatieve wending. Ze menen dat de machthebbers de waarheid verteld moet worden. Hier is weinig tegen in te brengen. Ze stellen ook: 'Politiek als evenwichtskunst vergt argumentatieve beleidsanalyse en netwerkanalyse.'
2. Het analyseren van politieke en maatschappelijke debatten komt weer in beeld.
3. Igor Mayer en Jacques Geurts benadrukken overigens de instrumentele mogelijkheden van de argumentatieve beleidsanalyse (en responsieve evaluatie). Ze noemen een aantal *participatieve methoden*, zoals beleidsexercities (Hoppe & Peterse, 1998).
4. Hoe verdedigbaar ook het werk van Fischer is, in de praktijk blijkt het schema van Fischer met de vier niveaus van beoordeling lastig toepasbaar (Rob Hoppe zei ooit: 'Men gooit de niveaus door elkaar').

Voor een toepassing verwijzen we naar een analyse van Brandenburg (in Hoppe & Peterse, 1998).

## Conclusie:

1. De aanpak van Fischer is te complex voor 'gewone' gebruiksgerichte evaluaties. Hoewel het nagaan van de vraag 'of het beleid het goede beleid is' nuttig is en ook of er sprake is van bredere effecten van beleid op samenleving en democratie.
2. Die verbredingsdiscussie over evaluatiemaatstaven kan meegenomen worden in de hiervoor bepleite entree-onderhandelingen voor aanvang van het evaluatieonderzoek.

Igno Pröpper (2001: 145-165) heeft aan evalueren als argumentatie overigens een eigen uitwerking gegeven. Hij meent dat evalueren zich niet alleen moet richten op (beleids)effectiviteit. Evalueren moet zich volgens hem richten op meer dan instrumentele rationaliteit en ook aandacht hebben voor normatieve, interactieve en expressieve zaken. Hij beoogt via inhoudelijke en procedurele analyses besluit- en oordeelsvorming te beoordelen met behulp van criteria als overtuigingskracht ('overtuigt het beleid?') en deugdelijkheid.

### **Alternatief 3 & Netwerkgericht evalueren**

Het hiervoor genoemde alternatief van responsief evalueren en van evalueren als argumentatie is een vorm van *verbrede evaluatie* ten opzichte van effectiviteitsevaluatie (Abma, 2001: 314). Toch ben ik er hiermee niet. Want is een overheidsbestuur altijd de centrale instantie in de sturing? Een andere brede evaluatie-aanpak is de aanpak waarbij de procesgang in en uitkomsten van een beleidsnetwerk (dus in een multi-actorensetting) wordt beoordeeld. Anders gesteld, het gaat om evaluatie in een 'governance'-setting. Het governance-begrip verwijst naar *beleidsnetwerken* (Hajer et al., 2004). Iets waarvoor Tineke Abma met de door haar bepleite responsieve evaluatie overigens oog had en heeft.

Ook in dit geval komen weer, net als in de zienswijze van Fischer en Pröpper, *andere beoordelingsmaatstaven* naar voren dan effectiviteit of doelmatigheid, namelijk onder meer openheid in het overleg en betrouwbaarheid van overlegpartners (Klijn & Koppenjan, 2001; Cörvers et al., 2003; De Rynck et al., 2010).

### **Bestaan beleidsnetwerken?**

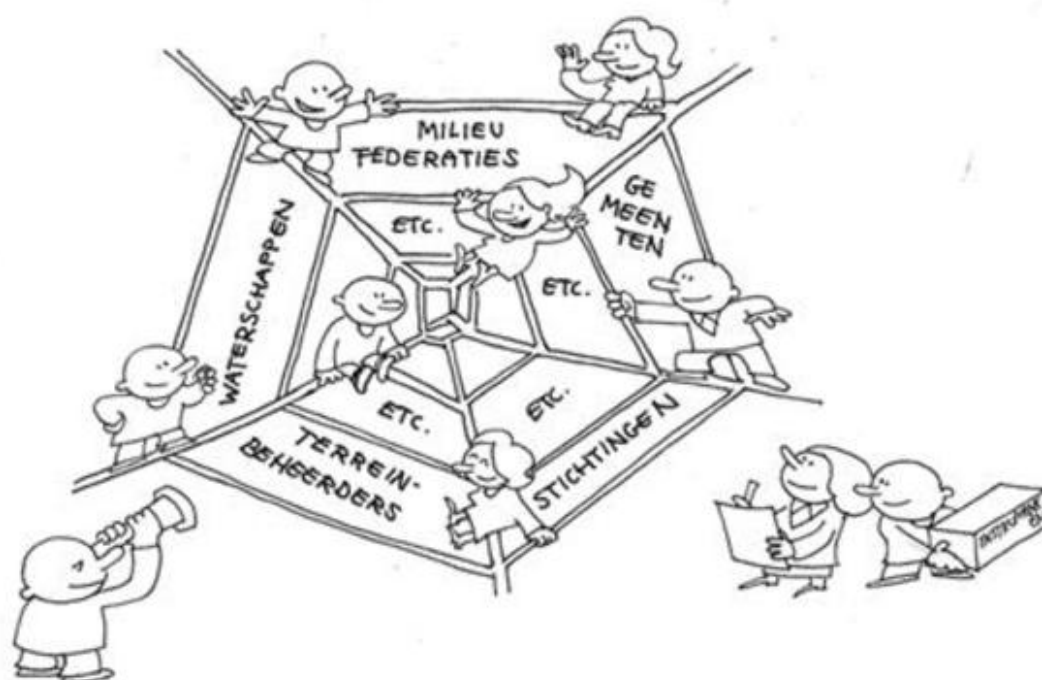
Beleidsnetwerken bestaan. Bij een beleidsnetwerk is sprake van diverse van elkaar afhankelijke actoren, die elk voor zich niet over alle taken, verantwoordelijkheden, bevoegdheden en middelen beschikken. Geen van alle actoren heeft alleen doorzettingsmacht. Het is daarom nodig dat deze actoren bij elkaar komen (netwerkconstitutie) en van gedachten wisselen om te komen tot een gedeeld beeld, en in de uitwerking tot doel- en middelenbundeling. Om dat te bereiken is het nodig dat elk van de actoren in staat is in de huid van de ander te kruipen, er respect voor elkaar bestaat, en beeldvorming die niet productief is gecorrigeerd wordt. Blokkades voor goed overleg moeten verdwijnen.

Het netwerkperspectief benadrukt de informele, gedecentraliseerde en horizontale verhoudingen die kunnen leiden tot bijvoorbeeld *milieuconvenanten*. Denk ook aan *strategische groenprojecten* in de Gelderse

Poort, het *gebiedsgerichte project* Kerngebied Vechtplassen, het *diagonale project* Biesbosch-Noordwaard en de *diagonale benadering* in het beekdalengebied van de Beerze en de Reusel (Cörvers et al., 2003).

Maar niet alle overheidsbeleid voltrekt zich in beleidsnetwerk. Ik noem twee begrenzingsen.

1. Bepaalde problemen zijn *te groot* om in een nationaal, regionaal of lokaal beleidsnetwerk te behandelen, zoals voedselveiligheid of het klimaatprobleem (Hajer et al., 2004: 34). Voedselveiligheid moet ook Europees bekeken worden. Denk aan de gekke koeienziekte/BSE.
2. Of de problemen zijn juist *te klein*, zoals de keuze van een straatnamenplan. Een straatnamenplan is een routinezaak voor de ambtenaren. Daar komt geen netwerk aan te pas.



Evident is dat er beleidsnetwerken bestaan van overheidsactoren en andere organisaties die *van elkaar afhankelijk zijn*. Geen actor beschikt in een netwerk over alle taken, bevoegdheden en middelen en geen enkele actor heeft dus alleen alle doorzettingsmacht (Hajer et al., 2004).

Denk maar eens aan het behoud en de ontwikkeling van een nationaal landschap Zuid-Limburg, dat kwaliteiten heeft om – volgens sommigen – op de lijst van werelderfgoederen te staan.

Hier bewegen zich: een provinciebestuur, waterschap, tal van gemeenten, enkele departementen, woningcorporaties, recreatieondernemers, de land- en tuinbouwbond, landbouwers, VVV, enzovoort. Actoren zijn om iets voor elkaar te krijgen dus *tot elkaar veroordeeld*. Ze moeten taken, doelstellingen en middelen ‘vermengen’ of ‘verknopen’. Het provinciebestuur zal het landschap kwalificeren in het kader van een provinciaal omgevingsplan, maar daarmee worden er nog geen maatschappelijke of bestuurlijke initiatieven genomen. Willen er projecten komen voor biologische landbouw, voor rondleidingen, voor fiets- en wandelroutes, voor wijnbouw en bezoek aan een wijngaard, enzovoort, dan zullen actoren elkaar in gemeenschappelijke initiatieven en

projecten moeten vinden.

De actoren in een netwerk, die beschikken over taken en bevoegdheden en/of middelen, zullen moeten overleggen en het eens moeten worden over wat het probleem is of wat de problemen zijn en hoe die gepercipieerd worden, over een gedeelde probleemaanpak, een kaderstelling en kanalisatie van initiatieven. Meestal zullen dit overheids- of semi-overheidsactoren zijn, maar in het kader van de ontwikkeling van Brainport Eindhoven wordt gesproken van *triple helix*, een verbinding tussen overheden, onderwijsinstellingen en bedrijfsleven. Een bekend ander voorbeeld van netwerksturing is de 'Alders-tafel' over aspecten van de ontwikkeling van Schiphol.

Toegespitst op het vorige voorbeeld: een landschap is er niet om slechts te behouden en conserveren, er moet ook activiteit zijn (misschien wel een verdienmodel). Het denken over beleidsnetwerken benadrukt de *informele, gedecentraliseerde* en *horizontale verhoudingen* tussen actoren in een netwerk, bijvoorbeeld in een landschapsnetwerk of huisvestingsnetwerk.

Hajer et al. (2004) benadrukken dat er meerdere vormen van 'governance' zijn. Een vorm is *multi-level governance*, een type sturing dat beschouwd kan worden als interbestuurlijke (overheids)sturing. Een andere vorm is de *deliberative governance*; deze vorm legt het accent op de kwaliteit van de participatie van burgers en van het beraad. Deze 'deliberative governance', die de 'argumentative turn' (Fischer, Hoppe) als noodzakelijk verwerkt heeft, lijkt opener naar marktpartijen en de 'civil society' dan de 'multi-level governance'.

*Gebiedsgericht beleid* wordt wel gezien als een illustratie van het denken in termen van beleidsnetwerken. Hier vallen ook verschillende aanpakken te onderscheiden. Hajer et al. (2004) spreken hier vanuit een historisch perspectief zelfs al van *drie generaties gebiedsgericht beleid*.

## Netwerkstrategieën

Er bestaan twee netwerkstrategieën om een netwerk verder te brengen:

1. strategieën van netwerkconstituering, en
2. netwerkmanagementstrategieën (Kickert et al., 1997).

De vraag welke actoren met elkaar rond de tafel moeten en op welk moment, is er een van netwerkconstitutie.

De tweede strategie, netwerkmanagement, uit zich in vragen als hoe actoren via het wegwerken van barrières (fixaties) en het scheppen van positieve condities in een netwerk komen tot een gemeenschappelijke koers, bijvoorbeeld in de vorm van gemeenschappelijke beeldvorming, die uitmondt in doelverrijking en middelenbundeling.

## Zijn netwerken zelforganiserend?

Dat wordt wel gesteld (Hajer et al., 2004: 16), maar dat getuigt mijns inziens van overoptimisme. Er is doorgaans in netwerken werk aan de winkel voor een procesregisseur. De bal moet in het spel gebracht worden, daarin gehouden worden en blijven rollen. Het kan ook wat gewichtiger gesteld worden.

Procesmanagement moet de interacties in de ‘beleidsspelen’ die in het netwerk plaatsvinden, bevorderen (De Bruijn & Ten Heuvelhof, 2001; De Bruijn et al., 2002). Mijns inziens gaat het vooral om sturingspogingen om de percepties van de actoren bij elkaar te brengen en over elkaar te leggen. De actoren hebben ieder voor zich wel een zekere ruimte van handelen, verrichten ook wel zelfstandig bepaalde taken (zoals het waterschap in een gebied dat gekwalificeerd is als nationaal landschap), maar slagen er niet in om het probleem of bepaalde uitdagingen alleen op te lossen. Ze moeten zich op elkaar kunnen en willen oriënteren en zijn in feite gedwongen om de eigen positie te relativeren en zich in de ander te verplaatsen om het eens te worden. Hier moet een *procesregisseur* helpen. Dat proces van bij elkaar brengen van actoren, van begripsvorming voor elkaars positie en gedeeld bewustzijn moet uitmonden in arrangementen tussen actoren, zoals een convenant, om verder te kunnen. En conflicten moeten worden uitgepraat of getemperd.

## Hoe kom je tot een gedeeld beeld in een netwerk?

Misschien door ‘reframing’, dus door anders naar een vraagstuk te kijken en verbeeldingskracht te tonen (dus door een ander verhaal). Dit wordt keer op keer door netwerkspecialisten en anderen benadrukt (Korsten, 1988; Schön & Rein, 1994; Klijn & Koppenjan, 2001: 183; Korsten & Leers, 2005).

Netwerkmanagement is al met al geen eenvoudige kost. Het vereist veel competenties, zoals op het vlak van kennis van een netwerk, kennis van wet- en regelgeving en oog voor belangen, en onderhandelingsvaardigheid. Het ontwikkelen en tonen van verbeeldingskracht is ook geen overbodigheid.

## Procesmanager

Wat is een *procesmanager* voor een type manager? Niet iemand die proclameert en verordonneert of geldt als centrale bestuurder, maar eerder iemand die (tijdelijk) stimuleert, faciliteert, verbindt en bemiddelt. Dat kan een van de actoren uit het netwerk zijn of een externe. De rol van regisseur is een tijdelijke. Wie het is, kan afhangen van het vraagstuk, initiatief of project (Klijn et al., 1995; Van Eeten & Termeer, 1996a, 1996b; De Bruijn & Ten Heuvelhof, 2001).

We zien dat in dit soort interacties in netwerken *het proces de inhoud stuurt*. Als het proces erkenning krijgt en partijen niet uitstappen en het vertrouwen groot is en blijft, kan dit de voorwaarden scheppen of bevorderen om het ergens over eens te worden. Is het vertrouwen niet groot (genoeg), dan ontstaat er een kans op een ‘*dialog van doven*’, op *stropigheid* en *proceduralisme*. Dat impliceert dan dat de aandacht voor interactie de aandacht voor de inhoud kan overstijgen, dus dat de inhoud wordt weggedrukt.

Het is mogelijk dat het proces de inhoud stuurt, maar dat de inhoud geen stand houdt. Als partijen het over iets eens zijn, wil dat nog niet zeggen dat het werkt, omdat voor sommige zaken andere kennis van belang is dan een onderhandelingsresultaat (‘negotiated knowledge’) biedt. Een bekend voorbeeld betreft een onderdeel van de ruimteveer Challenger. Dit ruimtevaartuig was veilig volgens de procesbesprekingen, zeg maar het procesmanagement. Toch deugde een bepaalde ring niet en ontplofte de Challenger. Bij hoogtechnologische vraagstukken kun je dus niet volstaan met procesmanagement-oplossingen bij deliberatie.

## Naar netwerkgerichte evaluatie

Wat betekent dit nu voor de beleidsevaluatie vanuit een netwerkperspectief: welke criteria zijn gewenst om uitkomsten van een netwerk te beoordelen? Hoe wordt het onderscheid gemaakt tussen succes en falen?

- Het is *ongewenst* om het verloop en de uitkomsten van netwerkproductiviteit te beoordelen aan de hand van de doelstellingen van één actor, ook al is dat een actor in het overheidsbestel (Klijn & Koppenjan, 2001: 189). Er is *niet* slechts één algemeen belang. Er is niet slechts hiërarchische sturing. In een evaluatie van een beleidsnetwerk, bijvoorbeeld van een huisvestingsnetwerk, moet worden uitgegaan van de visie van al de in het netwerk opgenomen of toegelaten actoren. Actoren zijn autonoom en trekken zich dus maar beperkt wat aan van de ene actor die denkt dat de macht bij hem ligt. Dus luidt het advies: *ga in een beleidsevaluatie uit van meerdere maatstaven*. Dat zien we ook in het proefschrift van Ron Cörvers over regionale netwerkprojecten op het vlak van natuurbeleid (zie ook Cörvers et al., 2003).
- Hou er rekening mee dat actoren in een beleidsnetwerk *doelzoekend gedrag* kennen (Herweijer, 1981). Soms weten ‘ze’ in de fase direct na netwerkconstitutie nog niet precies wat ze willen. Logisch, want alleen kunnen ze maar beperkt wat voor elkaar krijgen, zoals gesteld. Ze hebben andere organisaties nodig om samen dingen na te streven en mogelijk te maken en dat moet verkend worden. *Advies aan de evaluator*: kijk in een netwerkgerichte evaluatie daarom of actoren elkaar opzoeken, en of er sprake is van vertrouwen en of actoren initiatief nemen (Klijn, 2002). Actoren kunnen zelf leerprocessen doormaken. Wat eerst niet lukt, gaat later misschien wel. Percepties van problemen en uitdagingen zijn dynamisch; ze zijn in ontwikkeling. Door uit te gaan van geformuleerde doelstellingen van één actor wordt dit leren van andere actoren ontkend. De netwerkdenkers zijn op dit punt heel uitgesproken. *Verplaats je als evaluator in elk van de betrokken actoren*.
- Het is bovendien gewenst rekening te houden met *belangen* van de verschillende actoren in een netwerk. Leidt samenwerking tot win-winsituaties, meer dan elk alleen zou kunnen bereiken?
- Niet uitgesloten is dat na verloop van tijd de netwerkactoren *samen* besluiten tot een doelstelling van het netwerk. Dat is dan zeker een prominent punt en dus in een evaluatie te betrekken. *Advies aan de evaluator*: *zoek ook naar gezamenlijke netwerkgerichte doelstellingen*. Hierop ga ik nader in.

De Rynck et al. (2010) hebben zich gezet aan een poging om prestaties van netwerken als geheel in een uitgewerkter analysekader te vatten. Zij volstaan dus niet met de evaluator te laten kijken naar hoe afzonderlijke actoren zich opstellen, dus naar wat afzonderlijke actoren procesmatig gezien doen. De auteurs onderscheiden (1) productieprestaties, (2) procesprestaties, en (3) regimeprestaties (zie figuur 1). Ik licht dit toe.

Figuur 1: Beoordelingskader voor netwerkprestaties (vlg. De Rynck et al., 2010)

Dimensies	Criteria	Kosten en baten
Productieprestaties	Doelbereiking en efficiëntie	Kosten en baten verbonden met productieprestaties
Procesprestaties	Legitimiteit Verantwoording <ul style="list-style-type: none"> <li>- verantwoording afleggen</li> <li>- verantwoordelijk gehouden instemming</li> <li>- issues waarover instemming gevraagd wordt</li> <li>- mechanismen voor verlenen instemming</li> <li>- status van instemming</li> </ul>	Kosten en baten verbonden met procesprestaties
Regimeprestaties	Lidmaatschap Netwerkinstitutionalisering <ul style="list-style-type: none"> <li>- netwerkorganisatie</li> </ul> Kwaliteit van relaties <ul style="list-style-type: none"> <li>- dichtheid</li> <li>- veelvormigheid</li> </ul>	Kosten en baten verbonden met regimeprestaties

*Productieprestaties* op het eerste niveau verwijzen naar output (prestaties) van het netwerk. ‘Beleidsnetwerken zijn alleen succesvol als ze effectief zijn’, schrijven de auteurs. Dat wil zeggen dat ze ‘een meerwaarde’ bieden ten opzichte van wat individuele organisaties kunnen en willen. Zonder een zekere effectiviteit verliest een netwerk het vertrouwen van de deelnemers en wordt het netwerk irrelevant. Wat is dan effectiviteit? Effectiviteit heeft van doen met doeleinden. Maar welke?

Op het niveau van de individuele organisaties die deel uitmaken van een netwerk, gaat het om de tevredenheid over het netwerk. De evaluator kan de actoren daar naar vragen. Doelbereiking op het niveau van het netwerk is nagenoeg bereikt als actoren uit organisaties van mening zijn dat wat het netwerk bereikt aan productiviteit, hun doelstellingen dient. Als organisaties de betekenis van het netwerk zien en erkennen, kan dat de netwerkproductiviteit verder ten goede komen.

Het tweede *niveau* van beoordeling van prestaties in en van een netwerk betreft

*procesprestaties*. Procesprestaties in en van een netwerk verwijzen naar democratische kwaliteit. Netwerken kunnen door de kwaliteit van de interacties in processen bijdragen aan het opbouwen van een maatschappelijk en bestuurlijk draagvlak voor beleid. Hierbij worden drie elementen onderscheiden: legitimiteit, verantwoording en goedkeuring of instemming.

Het derde *niveau van analyse* verwijst naar *regimeprestaties*. Hier gaat het erom of en in welke mate het netwerk de capaciteit van de publieke sfeer versterkt: maakt het netwerk de maatschappelijke infrastructuur sterker? Drie criteria ter beoordeling zijn hier te benoemen volgens de auteurs: de aard van het 'lidmaatschap' van het netwerk, de duurzaamheid van de deelname en de kwaliteit van de relaties.

Een lidmaatschap impliceert dat een netwerk toegang biedt tot hulpbronnen die voor een probleemaanpak van betekenis zijn, zoals bevoegdheden, informatie, productiecapaciteit, legitimiteit, financiële middelen. Hoe meer relevante leden, hoe meer hulpbronnen ingebracht worden, of ontstaan, of gemobiliseerd kunnen worden. Bij te veel deelnemers kunnen de coördinatiekosten (of transactiekosten) echter weer gaan oplopen.

Een tweede regimecriterium slaat op de structuur van het netwerk en de manier van organiseren van het netwerk. Stabiliteit van het netwerk kan de capaciteit van de overheid vergroten. Stabiliteit en management van het netwerk kunnen een indirecte prestatie-indicator vormen.

Het derde criterium om de regimeprestaties in het zicht te krijgen betreft de kwaliteit van de relaties tussen actoren in het netwerk. Het aantal relaties, de dichtheid van relaties en veelvormigheid zijn van belang, aldus Philip de Rynck et al. Hoe groter de relatiedichtheid en hoe veelvormiger de relaties, hoe hoger de regimeprestaties van het netwerk (De Rynck et al., 2010: 43). Hiermee komen we weer dicht bij het denken van auteurs als Klijn et al.

Conclusie: we zien met een netwerkgerichte evaluatie dus duidelijk een verbrede evaluatie. Deze vorm van evaluatie komt nog maar schaars voor omdat deze evaluatie relatief duur is. Immers, de evaluatoren moeten zich verplaatsen in afzonderlijke actoren uit het netwerk. Bovendien, zonder een beetje doorleefde kennis van het netwerkdenken kan een evaluator er maar beter niet aan beginnen. Er worden dus ook eisen gesteld aan de evaluator. Is een medewerker die juist heel erg het netwerk kent dan geschikt? Wel als het ook om iemand gaat met competenties op het vlak van evaluatie.

## Vingeroefeningen

Is er een voorbeeld te vinden van een evaluatie op basis van de 'network approach' in 'evaluation'? Ron Cörvers (2001) heeft in zijn proefschrift een aantal gebiedsgebonden netwerkprojecten met een ruimtelijke en natuurcomponent beoordeeld. Deze beleidsvorming in netwerkprojecten werd betiteld als *dialogisch beleid*. Interactie was cruciaal. Uit die evaluatie blijkt onder meer dat in diverse projecten, zoals het project De Gelderse Poort, het project Het Vechtplassengebied en het project De Biesbosch, voor 2001 sprake was van een *heel trage voortgang*; er verschijnen vele plannen op plannen, dat wel. Integrale oplossingen ontbraken vrijwel volledig en van gemeenschappelijke beeldvorming was evenmin sprake. Het bundelen van overheidsmiddelen vond slechts sporadisch plaats. Investeringsmiddelen uit de private sector worden



nauwelijks aangesproken. Win-winuitkomsten waren meer ideologische schijn dan realiteit. Dit zijn teleurstellende bevindingen, maar er waren ook opmerkelijke uitkomsten.

Als er een gemeenschappelijke doelstelling centraal staat, betekent dit meestal toch dat het beleid van een 'hogere' overheid betreft. *Hiërarchische verhoudingen* blijken nog steeds een rol te spelen. Ze blijken hardnekkig. Van nevenschikking tussen overheden bleek niet echt sprake.

Wat ook heel opmerkelijk was, was dat *externe omstandigheden* een cruciale rol speelden bij de voortgang in gebiedsgerichte projecten. Als er een crisis optreedt, zoals een overstroming, blijkt er ineens productie, lees voortgang, mogelijk in het netwerk, zoals dijkversterking (Cörvers et al., 2003).

*Conclusie*: een evaluator moet er altijd op bedacht zijn dat een beleidsnetwerk in een overwegend verticale sturingscontext kan functioneren of weer getrokken kan worden. In zoverre begrijpelijk omdat elke overheidsactor weer functioneert in een democratische context; Gedeputeerde Staten moet terug naar Provinciale Staten, enzovoort. Die contexten kunnen de vrijheid van handelen van dagelijks bestuurders blijkbaar danig verlammen. Dan kan de verticale sturing het op bepaalde momenten gaan winnen. Succes is bij netwerkmanagement dus zeker geen vanzelfsprekendheid.

## Kritiek op netwerkgericht evalueren

Er is ook wel kritiek geuit op de netwerkbenadering bij evaluatie (Pröpper, 1996; Klijn & Koppenjan, 1997).

1. Een kritiekpunt van iemand als de Utrechtse milieukundige Pieter Glasbergen is dat degenen die netwerkproductiviteit beoordelen, te veel aandacht besteden aan de proceskant en te weinig aan wat het netwerk aan productie voortbrengt. Zoals aangegeven maakte promovendus Cörvers (van de promotores Glasbergen en Korsten) zich hieraan niet schuldig.
2. Een tweede kritiekpunt is dat heel goed in de gaten gehouden moet worden hoe de uitkomsten van een netwerkproductieproces zich verhouden tot de verantwoording in politiek-democratische organen waarmee elk van de overheidsactoren te maken heeft. De leden van deze organen kunnen het gevoel hebben dat gepoogd wordt hun een netwerkresultaat 'door de strot te duwen' (Hajer et al., 2004).

Ondanks deze kritiek is er geen alternatief voor netwerkgericht evalueren. Een evaluator kan niet doen alsof er geen netwerkcontext is als de verantwoordelijkheden versnipperd zijn, de afhankelijkheden tussen actoren groot zijn en geen enkele actor de volledige doorzettingsmacht heeft om de eigen beleidsvoorkeur op te leggen of er doorheen te duwen. Dan moet je netwerkgericht evalueren.

Toch is er nog een ander concept gelanceerd, en wel door enkele auteurs die vertrouwd zijn met zowel netwerkdenken als responsief evalueren. Dat voert ons naar alternatief 4. We zullen zien dat dit op zich interessante type in onze ogen niet probleemloos is. Nog niet proberen te doen dus in een context van gebruikgericht evalueren.

## **Alternatief 4 & Lerend evalueren: leren in een beleidsnetwerk**

In een vierde alternatief – *lerend evalueren* genoemd – wordt teruggegrepen op Guba en Lincoln (1989): dat wil zeggen op proberen de belanghebbenden te begrijpen. En op de netwerkbenadering.

Onder het motto van ‘al evaluerend leren’ hebben Jurian Edelenbos en Arwin van Buuren (2005) zich gericht op het verder brengen van evalueren in prospectief opzicht. Het gaat hun om evalueren als leerproces in een netwerkcontext van van elkaar afhankelijke actoren.

Zij laten het inventariseren van succes of falen en het beoordelen als onderdeel van een gebruiksgerichte evaluatie niet vallen, maar doen wel een poging om verder te komen.

Wat is de kern? Evaluator en geëvalueerde moeten volgens hen *gezamenlijk een leerproces organiseren*. Dat stelt uiteenlopende eisen aan evaluatoren en genereert *rolconflicten* voor hen.

Deze aanpak is tot stand gekomen na een doorlichting van het Stimuleringsprogramma Burger en Milieubeleid van het toenmalige ministerie van VROM.

Hun inzichten sluiten aan op die van Guba en Lincoln (fourth generation evaluation; 1989), Abma (1996) en Edelenbos en Van Eeten (2001). Wat betekent dat?

1. Ze betrekken bij hun aanpak het inzicht dat als je klassiek evaluerend te werk gaat, de geëvalueerden of andere gebruikers van een rapport zich niet actief opstellen. Ze zijn te receptief. Ze zoeken niet zelf actief naar wat te doen met de resultaten van een evaluatie. De beide auteurs menen dat het in het belang van het beleid is (‘de zaak zelf’) dat actoren juist wel *participeren*.
2. De evaluator wordt geen controleur meer, maar *samenspeler*. Hij gaat zich verplaatsen in de positie van elk van de ‘partijen’ in een netwerk. Want netwerkconstitutie is nodig (goede partijen aan tafel); en het netwerk moet productief worden (gemeenschappelijke beeldvorming en doelverrijking).
3. De evaluator tracht zijn inzichten *over te brengen* op en *uit te wisselen* met de geëvalueerden om zo tot gezamenlijke kennis en betekenisvolle inzichten te komen. Daarom wordt ook wel gesproken van een ‘*deliberatieve evaluatie*’ (House & Howe, 2000).
4. De evaluator is in hun optiek van een ‘*lerende evaluatie*’ ook een *verandermanager*. Hij neemt verantwoordelijkheid om het programma mee te verbeteren. Deze evaluator wordt een ‘key person’ genoemd die bijdraagt aan ‘a more sophisticated reality’(Guba & Lincoln, 1989: 262).

Conclusie over deze vierde aanpak.

Edelenbos en Van Buuren hebben gepoogd creatief om te gaan met een opdracht om een lerende evaluatieaanpak te introduceren. Ze verdienen lof daarvoor, maar in werkelijkheid levert hun aanpak toch problemen op. De betrokkenen

krijgen te maken met een evaluator die verschillende rollen vervult. Niet steeds is voor hen duidelijk in welke rol de evaluator 'nu weer zit'.

## Literatuur

Abma, T. & R.J. in 't Veld (red.), *Handboek beleidswetenschap*, Boom, Amsterdam, 2001.

Abma, T., *Responsief evalueren*, Rotterdam, 1996 (diss.).

Abma, T., Variaties in beleidsevaluaties: meten, beoordelen en onderhandelen, in: T. Abma & R.J. in 't Veld (red.), *Handboek beleidswetenschap*, Boom, Amsterdam, 2001, pp. 311-321.

Abma, T.A., Beleidsevaluatie zonder beter weten, in: H. van Gunsteren & E. van Ruyven (red.), *Bestuur in de ongekende samenleving*, Sdu, Den Haag, 1995, pp. 163-180.

Abma, T.A., Onderhandelend evalueren, in: *Bestuurskunde*, 2000, nr. 8, pp. 393-403.

Abma, T.A., Variaties in beleidsevaluatie, in: *Bestuurskunde*, 1997, nr. 2, pp. 77-87.

Bekkers, V.J.J.M. & M. Lips, Spierballen, veel stroop en weinig water & Verdrogingsbeleid in de provincie Noord-Brabant, in: F. Hendriks & Th.A.J. Toonen (red.), *Schikken en plooiën*, Van Gorcum, Assen, 1998, pp. 153-168.

Bekkers, V.J.J.M., *Beleid in beweging*, Boom Lemma uitgevers, Den Haag, 2007.

Bemelmans-Videc, M.L. & M. de Vries, Ontwikkelingen in de evaluatie van overheidshandelen, in: Aardema, A. et al. (red.), *Meerwaarde van de bestuurskunde*, Boom Lemma uitgevers, Den Haag, 2010, pp. 127-137.

Berndsen, R., M. Fraanje, A.F.A. Korsten & M. Kort, Beoordeling van provinciale bestuurskracht, in: *Bestuurswetenschappen*, 2007, nr. 6, pp. 11-49.

Boekenoogen, M., P. van Hoesel, Y. Prince & C. Verheijen (red.), *Methoden van beleidsonderzoekers: creatief en oplossingsgericht*, Uitgeverij Lemma, Den Haag, 2009.

Bongers, F.J. & D. Verlet, Indicatoren en beleid: doe meer met ongeveer, in: *Bestuurswetenschappen*, jrg. 67, 2013, nr. 1, pp. 36-55.

Bongers, F.J. & P. den Hertog, Vernieuwing van methoden voor beleidsevaluatie, in: *Bestuurswetenschappen*, 2007, nr. 3, pp. 29-44.

Bouckaert, G. & J. Halligan, *Managing performance. International comparisons*, Routledge, London, 2008.

Bouckaert, G. & W. van Dooren, Performance Measurement and Management in Public Sector Organizations, in: T. Boivard & E. Löffler (eds.), *Public*

*Management and governance*, Routledge, Londen, 2004, pp. 127-137.

Bressers, J. & A. Hoogerwerf (red.), *Beleidsevaluatie*, Samsom, Alphen aan den Rijn, 1995 (3e druk).

Bressers, J.Th.A., P. de Jong, P.J. Klok & A.F.A. Korsten (red.), *Beleidsinstrumenten bestuurskundig beschouwd*, Van Gorcum, Assen, 1993.

Bruijn, H. de, & E. ten Heuvelhof, Procesmanagement, in: T. Abma & R.J. in 't Veld (red.), *Handboek beleidswetenschap*, Boom, Amsterdam, 2001, pp. 193-208.

Bruijn, H. de, *Framing– Over de macht van taal in de politiek*, Atlas, Amsterdam, 2011.

Bruijn, H. de, G.R. Teisman, J. Edelenbos & W. Veeneman (red.), *Meervoudig ruimtegebruik en het management van meerstemmige processen*, Uitgeverij Lemma, Utrecht, 2004.

Bruijn, H. de, *Managers & professionals*, Academic Service, Den Haag, 2008.

Bruijn, H. de, *Managing Performance in the Public Sector*, Routledge, Londen, 2007.

Bruijn, H. de, *Prestatiemeting in de publieke sector*, Uitgeverij Lemma, Utrecht, 2001.

Bruijn, J. de, E. ten Heuvelhof & R.J. in 't Veld, *Procesmanagement*, Academic Service, Schoonhoven, 2002.

Cörvers, R.J.M., *Netwerksturing bij natuurontwikkeling & Bestuurskundige consequentieanalyse bij gebiedsgerichte projecten*, Maastricht, 2001.

Cörvers, R.J.M., P. Glasbergen & A.F.A. Korsten, Netwerksturing bij natuurontwikkeling, in: *Bestuurswetenschappen*, 2003, nr. 3, pp. 205-228.

Derksen, W., A.F.A. Korsten & A.F.M. Bertrand (red.), *De praktijk van onderzoek*, Wolters-Noordhoff, Groningen, 1988.

Dooren, W. van, G. Bouckaert & J. Halligan, *Performance Management in the Public Sector*, Routledge, Londen, 2010.

Dooren, W. van, *Performance Measurement in the Flemish Public Sector*, Instituut voor de Overheid, Leuven, 2006.

Edelenbos, J. & A. van Buuren, Evalueren als leerproces, in: *Bestuurskunde*, jrg. 14, 2005, nr. 6, pp. 2-13.

Edelenbos, J. & M. van Eeten, The Missing Link- Processing Evaluation in Dialogical Evaluation, in: *Evaluation*, vol. 7, 2001, nr. 2, pp. 211-217.

Edelenbos, J., M. van Buuren, M. van As et al., *Milieubeleid meer van, voor en door de burger & Resultaten van de lerende evaluatie van het*

*Stimuleringsprogramma Burger en Milieubeleid*, Erasmus Universiteit Rotterdam, Rotterdam, 2003.

Eeten, M. van & K. Termeer, Het faciliteren van interactieprocessen in beleidsnetwerken, in: *Werken, leren en leven met groepen*, Samsom, Alphen aan den Rijn, okt. 1996a, C3100, pp. 1-30.

Eeten, M. van & K. Termeer, Routebeschrijvingen door het gebied van netwerkmanagement, in: *Werken, leren en leven met groepen*, Samsom, Alphen aan den Rijn, okt. 1996b, E2200, pp. 1-9.

Glasbergen, P.G. (red.), *Milieubeleid*, VUGA, Den Haag, 1994.

Guba, E.G. & Y.S. Lincoln, *Fourth Generation Evaluation*, Sage, Londen, 1989.

Hajer, M.A., J.P.M. van Tatenhove & C. Laurent, *Nieuwe vormen van Governance*, RIVM, Bilthoven, 2004.

Hart, P. 't, P. de Jong & A.F.A. Korsten (red.), *Groepsdenken in het openbaar bestuur*, Samsom, Alphen aan den Rijn, 1991.

Heffen, O. van, Veronderstellingen achter beleid en hun betekenis, in: *Strategie en beleid in de publieke sector*, Samsom, Alphen aan den Rijn, mei 1996, D2020.

Herweijer, M., De dynamiek van doelstellend gedrag: een struikelblok bij evaluatie, in: *Bestuurswetenschappen*, jrg. 35, 1981, nr. 5, pp. 348-366.

Hoesel, P.H.M. van, F.L. Leeuw & J.W.M. Mevissen (red.), *Beleidsonderzoek in Nederland*, Van Gorcum, Assen, 2005.

Hoesel, P. van, *Beleidsonderzoek als professie- Een lang gekoesterd ideaal*, Stichting Economisch Instituut voor het Midden- en Kleinbedrijf, Zoetermeer, 2003 (oratie).

Hoogerwerf, A. & M. Herweijer (red.), *Overheidsbeleid*, Samsom, Alphen aan den Rijn, 1998 (6e druk)(2003, 7e druk).

Hoogerwerf, A., Beleid berust op veronderstellingen: de beleidstheorie, in: *Acta Politica*, jrg. 19, 1984, nr. 4, pp. 493-531.

Hoogerwerf, A., De beleidstheorie uit de beleidspraktijk: een tussenbalans, in: *Beleidswetenschap*, 1989, pp. 320-341.

Hoppe, R. & A. Peterse (red.), *Bouwstenen voor argumentatieve beleidsanalyse*, Elsevier, Den Haag, 1998.

Hoppe, R., *The Governance of Problems & Puzzling, Powering, Participation*, Policy Press, Bristol, 2011.

House, E. & K. Howe, Deliberative Democratic Evaluation, in: *New Directions for Evaluation*, Sage, Londen, 2000, pp. 3-12.

Kerseboom, J., Leren van evalueren, in: *Bestuurskunde*, jrg. 17, 2008, nr. 2, pp.

41-47.

Kickert, W.J.M., E.H. Klijn & J.F.M. Koppenjan (eds.), *Managing Complex Networks: Strategies for the Public Sector*, Sage, Londen, 1997.

Klein Haarhuis, C., M. van Ooyen-Houben, E. Kleemans & F. Leeuw, Rechtshandhaving geëvalueerd; een synthese van 31 onderzoeken, *Justitiële Verkenningen*, jrg. 31, 2005, nr. 8, pp. 54-72.

Klein Haarhuis, C.M. & E. Niemeijer, *Wet en werkelijkheid & Bevindingen uit evaluaties van wetten*, WODC, Boom Juridische uitgevers, Den Haag, 2008.

Klijn, E.H. & G. Teisman, Effective policy making in a multi-actor setting: networks and steering, in: R. in 't Veld et al. (eds.), *Autopoiesis and configuration theory: new approaches to societal steering*, Kluwer, Dordrecht, 1991, pp. 99-113.

Klijn, E.H. & G.R. Teisman, Besluitvorming in beleidsnetwerken, in: *Beleidswetenschap*, 1992, nr. 1, pp. 32-52.

Klijn, E.H. & J. Koppenjan, Besluitvormingen management in netwerken: een multi-actor perspectief op sturing, in: T. Abma & R.J. in 't Veld (red.), *Handboek beleidswetenschap*, Boom, Amsterdam, 2001, pp. 179-193.

Klijn, E.H. & J.F.M. Koppenjan, Beleidsnetwerken als theoretische benadering: een tussenbalans, in: *Beleidswetenschap*, 1997, nr. 2, pp. 143-167.

Klijn, E.H., J.F.M. Koppenjan & C.A.J.M. Termeer, Managing Networks in the Public Sector, in: *Public Administration*, vol. 73, 1995, nr. 3, pp. 437-454.

Klijn, E.H., Netwerken als perspectief op beleid en uitvoering, in: *Beleidswetenschap*, jrg. 19, 2005, nr. 4, pp. 32-55.

Klijn, E.H., Vertrouwen en samenwerking in netwerken, in: *Beleidswetenschap*, jrg. 16, 2002, nr. 3, pp. 259-280.

Knaap, L.M. van der, F. Leeuw, S. Bogaerts & L. Nijssen, Combining Campbell Standards and the Realist Evaluation Approach & The Best of Two Worlds, in: *American Journal of Evaluation*, vol. 29, 2008, pp. 48-57.

Knaap, P. van der & A. Schilder, Resultaatgericht sturen en evalueren, in: Knaap, P. van der, & A. Schilder (red.), *Resultaatgericht sturen en evalueren*, Sdu, Den Haag, 2004, pp. 9-17.

Kool, D. de, *Monitoring in beeld*, Erasmus Universiteit Rotterdam, Rotterdam, 2007.

Koppenjan, J.F.M., Beleidsonderzoek voor het integraal veiligheidsbeleid: lessen over effectrapportages, in: *Beleidsanalyse*, jrg. 24, 1995, nr. 2, pp. 27-37.

Korsten, A. & G. Leers, *Inspirerend leiderschap in de risicomaatschappij*, Uitgeverij Lemma, Utrecht, 2005.

Korsten, A.F.A. & I.M.A.M. Pröpper, Beleidsevaluatie, in: A. Bekke et al. (red.), *Management in overheidsorganisaties*, Samsom, Alphen aan den Rijn, 1993, B1160, pp. 1-14.

Korsten, A.F.A. & P. de Goede (red.), *Bouwen aan vertrouwen in het openbaar bestuur*, Elsevier, Den Haag, 2006.

Korsten, A.F.A. & R. Hoppe, Van beleidswetenschap naar kennissamenleving, in: *Beleidswetenschap*, jrg. 20, 2006, nr. 4, pp. 34-73.

Korsten, A.F.A. & Th.A.J. Toonen (red.), *Bestuurskunde & Hoofdfiguren en kernthema's*, Stenfert Kroese, Leiden, 1988/1993.

Korsten, A.F.A., Beoordeling & De betekenis van beleidsnetwerken voor beoordelen van provinciale visitatiecommissies, in: R. Cörvers & C. Dieperink (red.), *Milieubeleid als sturingsopgave*, Open Universiteit/Universiteit Utrecht, 2011, pp. 201-211.

Korsten, A.F.A., *Bestuurskunde als avontuur*, Kluwer, Deventer, 1988.

Korsten, A.F.A., *Deugdelijk bestuur*, Open Universiteit Nederland, Heerlen, 2010 (afscheidsrede).

Korsten, A.F.A., Effectrapportage, in: *Beleidsanalyse*, 1997, nr. 1/2, pp. 4-8.

Korsten, A.F.A., Evalueren van wetten, in: *Bestuurswetenschappen*, 1981, pp. 462-484.

Korsten, A.F.A., *Grote klasse & Op zoek naar excellente ambtenaren en leiderschap*, Beljon + Westerterp, Heerlen, 2005a.

Korsten, A.F.A., *Maatstaven voor de beoordeling van overheidshandelen*, Open Universiteit, Heerlen, 14 maart 2005b (zie [www.arnokorsten.nl](http://www.arnokorsten.nl)).

Korsten, A.F.A., Visiteren van gemeentebesturen, in: *Bestuurswetenschappen*, jrg. 58, 2004, nr. 4, pp. 305-324.

Leeuw, F., Evaluatie-onderzoek, auditing en controle, in: T. Abma & R.J. in 't Veld (red.), *Handboek beleidswetenschap*, Boom, Amsterdam, 2001, pp. 70-83.

Leeuw, F.L., G. van Gils, M. Gildemyn & L. Inberg, *Leren van evalueren*, Ministerie van Buitenlandse Zaken, Den Haag, 2009.

Leeuw, F.L. & G. van Gils, Over benutting van beleidsevaluatie bij het ministerie van Buitenlandse Zaken, in: A. Aardema et al. (red.), *Meerwaarde van de bestuurskunde*, Boom Lemma uitgevers, Den Haag, 2010, pp. 165-174.

Leeuw, F.L., J.S. Kerseboom & R. Elte (red.), *Turven, tellen, toetsen*, Boom Juridische uitgevers, Den Haag, 2007.

Leeuw, F.L., Reconstructing Program Theories: Methods available and Problems to be Solved, in: *American Journal of Evaluation*, vol. 24, 2003, nr. 1, pp. 5-20.

Meer, F.B. van der & J. Edelenbos, Evaluation in Multi-actor Policy Processes & Accountability, Learning and Co-operation, in: *Evaluation*, vol. 12, 2006, nr. 2, pp. 201-218.

Meer, F.B. van der, G.J. de Vries & G.A.N. Vissers, Evaluatie en leerprocessen bij de overheid: de rol van institutionele condities, in: *Beleidswetenschap*, jrg. 14, 2000, nr. 3, pp. 253-277.

Noordegraaf, M., *Management in het publieke domein*, Uitgeverij Coutinho, Bussum, 2004.

Noordegraaf, M., Prestatiemanagement, in: M. Noordegraaf, K. Geuijen & A. Meijer (red.), *Handboek Publiek Management*, Boom Lemma uitgevers, Den Haag, 2011, pp. 365-397.

Pawson, R. & N. Tilly, *Realistic Evaluation*, Sage, Londen, 2003.

Pawson, R. & C. Klein Haarhuis, Evaluatie van complexe programma's: een theoriegestuurde aanpak, in: *Justitiële Verkenningen*, jrg. 31, 2005, nr. 8, pp. 42-54.

Pawson, R., *Evidence-Based Policy*, Sage, Londen, 2006.

Pröpper, I.M.A.M. & D.-D. Reneman, De reconstructie van de beleidstheorie als argumentatie: Een toepassing op het verkeersveiligheidsbeleid, in: *Beleidswetenschap*, 1993, pp. 238-257.

Pröpper, I.M.A.M., Beleidsevaluatie als argumentatie, in: T. Abma & R.J. in 't Veld (red.), *Handboek beleidswetenschap*, Boom, Amsterdam, 2001, pp. 145-165.

Pröpper, I.M.A.M., Beleidsevaluatie als argumentatie, in: *Beleidswetenschap*, jrg. 1, 1987, nr. 2, pp. 113-136.

Pröpper, I.M.A.M., Succes en falen van sturing in beleidsnetwerken, in: *Beleidswetenschap*, jrg. 10, 1996, pp. 345-365.

Rein, M. & D.A. Schön, Frame-reflective policy discourse, in: *Beleidsanalyse*, 1986, nr. 4, pp. 4-18.

Rein, M., Reframing Problematic Policies, in: M. Moran, M. Rein & R.E. Goodin (eds.), *The Oxford Handbook of Public Policy*, Oxford University Press, 2006, pp. 389-409.

Rhodes, R.A.W., Policy Network Analysis, in: M. Moran, M. Rein & R.E. Goodin (eds.), *The Oxford Handbook of Public Policy*, Oxford University Press, 2006, pp. 425-448.

Ringeling, A.B., *Het imago van de overheid*, VUGA, Den Haag, 1993.

Rossi, P.H., M.W. Lipsey & H.E. Freeman, *Evaluation & A systematic approach*, Sage, Londen, 2004 (7th edition).



Rynck, Ph. De, J. Voet & W. van Dooren, Het beoordelen van de prestaties van netwerken: aanzet tot een analysekader, in: *Bestuurswetenschappen*, jrg. 64, 2010, nr. 1, pp. 35-56.

Schillemans, Th., *Verantwoording in de schaduw van de macht & Horizontale verantwoording bij zelfstandige uitvoeringsorganisaties*, Uitgeverij Lemma, Den Haag, 2007.

Schön, D.A. & M. Rein, *Frame Reflection - Towards the Resolution of Intractable Policy Controversies*, Basic Books, New York, 1994.

Swanborn, P.G., *Evalueren*, Boom, Amsterdam, 1999/2002.

Termeer, C.J.A.M., H. Mastik, R.G. Scalzo & R.J. in 't Veld, Spelsimulatie als methode voor de ex ante evaluatie van wetgeving, in: *Beleidwetenschap*, jrg. 9, 1995, nr. 3, pp. 225-241.

Termeer, C.J.A.M., *Vitale verschillen & Over publiek leiderschap en maatschappelijke innovatie*, Wageningen, 2006 (oratie).

Twist, M.J.W. van, De beleidstheorie vanuit de wetenschapspraktijk: van 'hoe ver?' naar hoe 'verder'?, in: *Beleidswetenschap*, jrg. 7, 1993, nr. 1, pp. 34-48.

Weiss, C.H., *Evaluation research*, Prentice Hall, Englewood Cliffs, 1972.

Zwan, A. van der & M. Noordegraaf, Onderzoek van een complexe casus & Een gemeentevervoerbedrijf als voorbeeld, in: *Bestuurskunde*, jrg. 7, 1998, nr. 2, pp. 62-74.