

SVR - Studie

2008 / 2

# **Over beleidsevaluatie: van theorie naar praktijk en terug**

**Dries Verlet & Carl Devos (red.)**

Studiedienst van de Vlaamse Regering

Vlaamse overheid



**Samenstelling**

Diensten voor het Algemeen Regeringsbeleid  
Studiedienst van de Vlaamse Regering

**Verantwoordelijke uitgever**

Josée Lemaître  
Administrateur-generaal  
Boudewijnlaan 30  
1000 Brussel

**Review**

Ann Carton, Adélien Decramer, Luc Deschamps, Josée Lemaître, Alex Vanderstraeten, Eddy Van Loocke

**Lay-out cover**

Diensten voor het Algemeen Regeringsbeleid  
Communicatie  
Patricia Van Dichel

**Druk en lay-out**

Drukkerij Gazelle NV

**Depotnummer**

D/2008/3241/293

**Bestellingen**

<http://publicaties.vlaanderen.be>

Delen uit deze uitgave mogen worden verveelvoudigd en openbaar gemaakt indien de auteur van de bijdrage en de titel van deze studie uitdrukkelijk als bron worden vermeld.

Elke auteur is verantwoordelijk voor de inhoud van haar of zijn deel.

Vlaamse overheid



# **Over beleidsevaluatie: van theorie naar praktijk en terug**

Dries Verlet & Carl Devos (red.)

# Korte inhoud

Beleidsvaluatie: een inleidende situering .....	7
<i>Dries Verlet &amp; Carl Devos</i>	
Hoofdstuk 1: Beleidsvaluatie: gelijk hebben ≠ gelijk krijgen .....	20
<i>Henk Klaassen &amp; Frans van Nispen</i>	
Hoofdstuk 2: Prestatiemeting: doel of middel? Over de niet-evidente meting van prestaties binnen de publieke sector .....	40
<i>Dries Verlet &amp; Carl Devos</i>	
Hoofdstuk 3: Preventief gedrag bij beleidsvaluaties .....	62
<i>Frank Bongers</i>	
Hoofdstuk 4: Risico's op perverse effecten bij beleidsvaluatie? Een vergelijking vanuit mogelijke evaluatiemotieven tegen de achtergrond van de Vlaamse hervormingscontext .....	77
<i>Bart De Peuter, Valérie Pattyn &amp; Marleen Brans</i>	
Hoofdstuk 5: De plaats en betekenis van beleidsvaluaties in de Vlaamse overheid –de domeinen Binnenlands Bestuur en Stedenbeleid verkend ...	106
<i>Ellen Wayenberg</i>	
Hoofdstuk 6: Evaluatie van antiradicaliseringsprojecten. Hoe meet je of radicalisering wordt tegengegaan? .....	122
<i>Amy-Jane Gielen &amp; Gerd Junne</i>	
Hoofdstuk 7: Doelmatige publieke diensten in een democratische samenleving: een geïntegreerde benadering en praktische evaluatie-instrumenten .....	152
<i>Walter Van Dongen</i>	



Hoofdstuk 8: Als meten het doel wordt: de vertaling van gender mainstreaming naar de beleidspraktijk .....	177
<i>Petra Meier &amp; Karen Celis</i>	
Hoofdstuk 9: Een Dorpshuis als Ontmoetingsplaats: Ontwerp van Beleidsevaluatie .....	207
<i>Steven Lannoo &amp; Dries Verlet</i>	
Epiloog: Beleidsevaluatie is een evidentie, beleid evalueren is dat niet .....	230
<i>Dries Verlet &amp; Carl Devos</i>	
Over de auteurs .....	232
Samenvattingen .....	235
Lijst met figuren .....	239
Lijst met tabellen .....	241

# Gedetailleerde inhoudsopgave

<b>Beleidsvaluatie: een inleidende situering .....</b>	<b>7</b>
1. <i>Wat is beleidsvaluatie?</i> .....	9
2. <i>Beleidsvaluatie als een onderdeel van een veelzijdig en –meerlagig totaalproces</i> .....	11
3. <i>Motieven tot beleidsvaluatie</i> .....	12
4. <i>Types van evaluaties</i> .....	13
5. <i>Wat volgt?</i> .....	16
<i>Bibliografie</i> .....	18
<b>Hoofdstuk 1: Beleidsvaluatie: gelijk hebben ≠ gelijk krijgen .....</b>	<b>20</b>
1. <i>Inleiding</i> .....	21
2. <i>Beleidsanalyse, een plaatsbepaling</i> .....	22
3. <i>De institutionalisering van beleidsanalyse in Nederland</i> .....	24
4. <i>Beleidsanalyse als kosten-batenanalyse (kba): “input” versus “outcome” ..</i>	25
5. <i>Beleidsanalyse als effectrapportage: “output” versus “outcome” .....</i>	27
6. <i>Beleidsanalyse als performance measurement: “input” versus “output/ outcome” .....</i>	30
7. <i>De veranderende rol van de beleidsanalyse, i.c. de beleidsanalist .....</i>	32
8. <i>Vergelijking</i> .....	33
9. <i>Conclusie</i> .....	35
<i>Bibliografie</i> .....	36
<b>Hoofdstuk 2: Prestatiemeting: doel of middel? Over de niet-evidente meting van prestaties binnen de publieke sector .....</b>	<b>40</b>
1. <i>Inleiding</i> .....	41
2. <i>Meten van prestaties in de publieke sector</i> .....	44
2.1. <i>Op welk niveau meten we de prestaties van de overheid?</i> .....	44
2.2. <i>Meten we perceptie of realiteit?</i> .....	45

2.3. Welke waarde-indicatoren zijn er voor dienstverlening in de publieke sector? .....	46
2.4. Welke kwaliteitsnormen kunnen we gebruiken? .....	46
2.5. Zijn alle nodige gegevens beschikbaar? .....	46
<b>3. Effecten van prestatiemeting .....</b>	<b>47</b>
3.1. Situering en het goede aan prestatiemeting .....	48
3.2. Negatieve effecten van prestatiemeting .....	49
3.3. Een algemene verklaring van negatieve effecten en algemene strategieën om ze te voorkomen .....	55
<b>4. Besluit .....</b>	<b>57</b>
<i>Bibliografie .....</i>	<i>59</i>
<b>Hoofdstuk 3: Preventief gedrag bij beleidsevaluaties .....</b>	<b>62</b>
1. Inleiding .....	63
2. De rol van beleidsevaluaties .....	63
3. Belemmerende factoren .....	65
3.1. Institutionele factoren .....	65
3.2. Beleidsmatige factoren .....	67
3.3. Persoonlijke factoren .....	71
4. Effect van beleidsevaluaties op preventief gedrag en op het beleidsproces ..	72
5. Tot besluit .....	74
<i>Bibliografie .....</i>	<i>75</i>
<b>Hoofdstuk 4: Risico's op perverse effecten bij beleidsevaluatie? Een vergelijking vanuit mogelijke evaluatiemotieven tegen de achtergrond van de Vlaamse hervormingscontext .....</b>	<b>77</b>
1. Inleiding .....	78
2. Evaluatiemotieven .....	78
3. Evaluatiemotieven en risico's op meetproblemen .....	80
4. Evaluatiemotieven bij de Vlaamse overheid .....	81
4.1. Evaluatiemotieven in een hervormingscontext .....	81
4.2. Drijfveren .....	83
4.3. Casestudy: Vlaanderen .....	84
5. Risico's verkleinen .....	91
<i>Bibliografie .....</i>	<i>104</i>

## **Hoofdstuk 5: De plaats en betekenis van beleidsevaluaties in de Vlaamse overheid –de domeinen Binnenlands Bestuur en Stedenbeleid verkend ... 106**

1. Inleiding .....	107
2. <i>Evalueren van het Vlaamse beleid inzake binnenlands bestuur en steden: de puntjes op de i</i> .....	107
2.1. Het kernstuk van beleidsevaluatie .....	108
2.2. Het voorwerp van beleidsevaluatie .....	108
2.3. Evalueren van “eigen” of andermans beleid .....	109
2.4. Evalueren: zelf doen of laten doen .....	110
2.5. Evalueren: samen doen of toch alleen .....	110
2.6. Evalueren niet enkel als sluitstuk van beleidsvoering .....	112
2.7. Kwaliteitsvol evalueren .....	113
3. <i>Evalueren: ook om te leren?</i> .....	114
4. <i>Besluit</i> .....	118
<i>Bibliografie</i> .....	120

## **Hoofdstuk 6: Evaluatie van antiradicaliseringsprojecten. Hoe meet je of radicalisering wordt tegengegaan? ..... 122**

1. <i>Introductie</i> .....	123
2. <i>Projecten om radicalisering tegen te gaan</i> .....	124
2.1. Het actieplan ‘Wij Amsterdammers’ .....	124
2.2. Antiradicalisering nader uitgewerkt in Stadsdeel Slotervaart .....	125
3. <i>Traditionele evaluatiemethoden in vogelvlucht</i> .....	127
3.1. Meten is weten .....	127
3.2. Pragmatische en beschrijvende evaluatie .....	129
3.3. Interpretieren .....	129
3.4. Onderhandelen .....	129
3.5. Realistische evaluatie .....	130
4. <i>Evaluatiemethoden van internationale onderzoeksprogramma’s</i> .....	131
4.1. Het “Local Capacities for Peace Project” (LCPP) .....	131
4.2. Peace and Conflict Assessment (PCA) .....	137
4.3. Outcome mapping .....	139
5. <i>Aanzet tot een evaluatie</i> .....	141
6. <i>Het trainingsproject ‘Identiteit en weerbaarheid’</i> .....	144
7. <i>Conclusie</i> .....	149
<i>Bibliografie</i> .....	150

## **Hoofdstuk 7: Doelmatige publieke diensten in een democratische samenleving: een geïntegreerde benadering en praktische evaluatie-instrumenten ..... 152**

<i>Inleiding</i> .....	153
<i>1. Dienstverlening in de dagelijkse leefsituatie</i> .....	154
1.1. De leefsituatie en levensloop van mannen en vrouwen .....	154
1.2. Tijdsverdeling binnen het marktsysteem .....	155
<i>2. Vraag en aanbod van publieke diensten in het marktsysteem</i> .....	157
2.1. Actoren en aspecten van de marktwerking .....	157
2.2. Positie van de aanbieders en de ontvangers .....	158
2.3. Mate van gelijke beschikbaarheid en collectieve financiering .....	161
<i>3. Doelmatigheid van publieke diensten</i> .....	162
3.1. Betekenis .....	162
3.2. Meten van effectiviteit en efficiëntie .....	164
3.3. Ontwikkeling van evaluatie-instrumenten voor publieke diensten ...	164
<i>4. Toepassing: evolutie v/d dagopvang voor jonge kinderen in het Vlaamse Gewest</i> .....	167
4.1. Beschikbaarheid .....	167
4.2. Participatie .....	168
4.3. Omvang van het gebruik .....	169
4.4. Bezetting .....	170
4.5. Productiviteit .....	170
4.6. Kostenefficiëntie .....	171
<i>5. Besluit</i> .....	174
<i>Bibliografie</i> .....	175

## **Hoofdstuk 8: Als meten het doel wordt: de vertaling van gender mainstreaming naar de beleidspraktijk ..... 177**

<i>1. Inleiding</i> .....	178
<i>2. Gender mainstreaming: een intentioneel en rationeel beleidsproces</i> .....	179
2.1. Intentionaliteit en rationaliteit van het beleidsproces .....	179
2.2. Intentionaliteit en rationaliteit van gender mainstreaming .....	180
2.3. Een gestuurd beleidsproces als instrument voor het realiseren van gelijkheid? .....	183
<i>3. De vertaling van gender mainstreaming naar de Belgische beleidspraktijk</i> ..	185
3.1. De Pekingwet van 1996 (en het Vlaamse Pekingdecreet van 1997) ...	186
3.2. De Vlaamse emancipatie-effect-rapportage van 1997 en 2000 .....	187
3.3. De federale Cel Gender Mainstreaming van 2001 en 2002 .....	188
3.4. De Pekingdecreten en –ordonantie van 2002 en 2006 .....	191
3.5. De Vlaamse Open Coördinatie Methode (2005) .....	193
3.6. De wet gender mainstreaming van 2007 .....	194

4. Verklaringen voor gender mainstreaming als een proces van meten en weten .....	197
5. Gender mainstreaming: een meerwaarde voor het realiseren van gendergelijkheid? .....	200
Bibliografie .....	203
<b>Hoofdstuk 9: Een Dorpshuis als Ontmoetingsplaats: Ontwerp van Beleidsevaluatie .....</b>	<b>207</b>
<i>Inleiding</i> .....	208
1. <i>De Interventie</i> .....	208
1.1. Een dorpshuis als ontmoetingsplaats .....	208
1.2. De stakeholders .....	209
1.3. De beleidstheorie: Sociale Verbondenheid en Lokale Faciliteiten ..	210
2. <i>Methodologisch Raamwerk</i> .....	213
2.1. Types beleidsevaluatieonderzoek .....	213
2.2. Ex-post evaluatie .....	216
2.3. Methoden: kwalitatieve versus kwantitatieve methoden: twee werelden van waarheden? .....	218
2.4. Interviews en Focusgroepen .....	219
3. <i>Het Concrete Beleidsevaluatieontwerp</i> .....	223
3.1. Input .....	223
3.2. Throughput .....	223
3.3. Output .....	224
3.4. Outcome .....	224
4. <i>Uitleiding</i> .....	226
<i>Bibliografie</i> .....	227
<b>Epiloog: Beleidsevaluatie is een evidentie, beleid evalueren is dat niet .....</b>	<b>230</b>
<b>Over de auteurs .....</b>	<b>232</b>
<b>Samenvattingen .....</b>	<b>235</b>
<b>Lijst met figuren .....</b>	<b>239</b>
<b>Lijst met tabellen .....</b>	<b>241</b>

# ***Beleidsvaluatie: een inleidende situering***

Dries Verlet & Carl Devos

Studiedienst van de Vlaamse Regering & Universiteit Gent

<b>1. Wat is beleidsvaluatie? .....</b>	<b>9</b>
<b>2. Beleidsvaluatie als een onderdeel van een veelzijdig en –meerlagig totaalproces .....</b>	<b>11</b>
<b>3. Motieven tot beleidsvaluatie .....</b>	<b>12</b>
<b>4. Types van evaluaties .....</b>	<b>13</b>
<b>5. Wat volgt? .....</b>	<b>16</b>
<b>Bibliografie .....</b>	<b>18</b>

Beleed beoogt het realiseren van gewenste effecten. Beleid wordt geëvalueerd om na te gaan of en hoe die effecten bereikt worden. Naast beoogde beleidseffecten kunnen er zich ook neveneffecten voordoen. Dat is alom bekend.

Het evalueren van beleid is geen evidentie. Over de vraag wat we hoe kunnen evalueren bestaat discussie. Evalueren is keuzes maken. Het evalueren van beleid is geen simpele zaak van objectief meten. Er komt heel veel bij kijken. Bovendien kennen ook de beleidsactoren de methoden waarmee hun activiteiten – de beleidsvoering – geëvalueerd worden of kunnen worden. Ze kunnen daarop reageren, bv. door officiële beleidsdoelstellingen – een cruciale standaard in het evaluatieproces – aan te passen, rekening houdend met de methoden van evaluatie. Er is dus sprake van strategisch gedrag van de beleidsactoren (Swanborn, 1999).

Over deze en andere neveneffecten werd in 2008 een workshop georganiseerd op het Politicologenetmaal.<sup>1</sup> Het centrale thema van de workshop was: “beleidsevaluatie: doel of middel?”. Dit initiatief werd gedragen door de Studiedienst van de Vlaamse Regering en de Vakgroep Politieke Wetenschappen van de Universiteit Gent. De workshop vormde de aanleiding, de papers die tijdens dit Vlaams-Nederlands internationaal congres werden voorgesteld vormen de basis van dit boek.

Met beleidsevaluatie wordt er in deze SVR-studie een thema aangesneden dat meer en meer aan belang wint binnen de publieke sector. De voorbije jaren zijn we getuige van een ware explosie aan beleidsevaluatie. Beleidsevaluatie krijgt vorm als een afzonderlijk onderzoeksveld met eigen professionele verenigingen ter zake (bijv. European Evaluation Society en recentelijk Vlaams Evaluatie Platform), tijdschriften, congressen, opleidingen, consultancy bedrijven en andere instellingen. In belangrijke mate organiseert de publieke sector overigens zelf de evaluatie. Het was dan ook te verwachten dat al die evaluatie ook aan evaluatie onderworpen zou worden.

De voorbije decennia is het denken over en doen van publieke dienstverlening sterk veranderd. Dit gebeurde vooral onder impuls van evoluties in het publieke management. Cornforth (2003, o.c. in Spanhove & Verhoest, 2007) stelt dat twee hervormingen daarvoor cruciaal zijn. Vooreerst staat een toenemend aantal (quasi-)autonome overheidsbedrijven in voor publieke dienstverlening. Ten tweede is er de introductie en het stevig belijden van marktmechanismen bij publieke dienstverlening en dus ook van begrippen als concurrentie, efficiëntie en effectiviteit (Verhoest & Spanhove, 2007). Ook onder impuls van *Reinventing Government* van Osborne en Gaebler, komt prestatie meting in de publieke sector meer op de voorgrond. Het idee is verleidelijk eenvoudig: een overheidsorganisatie definieert haar “producten” (bv. dienstverlening) en ontwikkelt indicatoren om de productie ervan meetbaar te maken. Dat stelt een organisatie in staat om dankzij een planning- en controlecyclus te werken aan een goed presterende organisatie (De Bruijn, 2002). Zo werkt de overheid optimaal.

---

<sup>1</sup> Het politicologenetmaal is een jaarlijks congres op initiatief van de Nederlandse Kring voor Wetenschap der Politiek en het Politologisch Instituut, die academici en praktijkpolitologen uit respectievelijk Nederland en Vlaanderen samenbrengt.



We moeten er op wijzen dat “evaluatie” veel meer omvat dan enkel en alleen het in kaart brengen van prestaties van de overheid in termen van output en outcome, zoals deze centraal worden gesteld in de New Public Management (NPM)-benadering. Daarenboven is opgang van beleidsevaluatie niet overal gelijk. Zo zijn er belangrijke verschillen tussen beleidsdomeinen en landen. Wat de beleidsdomeinen betreft is het zo dat er bijvoorbeeld meer evaluatie is van het onderwijs dan van de defensie. Daarnaast heeft bijvoorbeeld de VS een meer uitgebouwde traditie qua evaluatie dan bijvoorbeeld België (Dahler-Larsen, 2005, 615-622). In de opgang die beleidsevaluatie maakt zien we een toenemende integratie in prestatie-meetsystemen en andere vormen van de NPM-geïnspireerde monitoring. Overigens is de boost in het beleidsevaluatieonderzoek niet alleen toe te schrijven aan de toenemende aandacht voor efficiëntie en effectiviteit zoals dit tot uiting komt in het NPM-denken. Sinds de jaren zestig is er immers een toename van evaluatieonderzoek, terwijl de het efficiëntie- en effectiviteitsdenken pas sinds de jaren tachtig opgeld maakt. De opgang van beleidsevaluatie wordt dan ook gekoppeld aan een bredere culturele trend die zich laat vatten onder de globale noemer van modernisering (Dahler-Larsen, 2005, 619-620).

In dit inleidende hoofdstuk willen we ingaan op het concept beleidsevaluatie. We pogen geen centrale definitie te formuleren die nadien consequent moet aangehouden worden. We kozen ervoor om verschillende aspecten en dimensies van beleidsevaluatie te belichten, die in hiernavolgende hoofdstukken van het boek verder uitgewerkt zullen worden.

## 1. Wat is beleidsevaluatie?

Evalueren is een dagelijkse, vaak triviale activiteit. In essentie is evalueren waardebeoordeling (Hall e.a., 2004), bedoeld om gedrag te sturen. Goed en juist evalueren en vooral evaluatieonderzoek is wel bijzonder. Onderzoek moet immers wetenschappelijk gebeuren. Vereist is o.a. een duidelijk geformuleerde probleemstelling, een doordacht design voor het onderzoek, een systematische dataverzameling en -analyse en rapportering die aangeeft in welke mate de probleemstelling te beantwoorden valt en hoe dat precies is gebeurd (Swanborn, 1999).

Door al dat studiewerk bestaan er ondertussen veel en veel verschillende definities van beleidsevaluatie. Soms wordt naar dat proces ook verwezen onder de noemer beleidsanalyse, wat breder kan zijn. Dye (1976, 1) omschrijft *'policy analysis'* als *'finding out what governments do, why they do it, and what difference it makes. More elaborate definitions of policy analysis can be found in the academic literature. But they really boil down to the same thing - the description and explanation of the causes and consequences of government activity'*. Beleidsanalyse is hier ruimer dan strikte beleidsevaluatie.

Wanneer we het bij een meer beperkte omschrijving van evaluatie houden, blijkt het doorgaans in essentie te gaan om het systematisch verzamelen van informatie en de beoordeling daarvan (Trochim, 2008). Deze essentie vinden we ook terug in de definitie bij Scriven (1980, 2007): *'Evaluation is the systematic and objective*

*determination of the worth or merit of an object*.<sup>2</sup> Hiermee raakt hij de kern van evaluatie aan: een waarde toekennen. Evalueren kan dus moeilijk los gezien worden van waarden, waarbij waardenschalen cruciaal zijn. Wie evalueert waardeert. Dat maakt van evaluatie een belangrijke activiteit, want aan die waardetoekenning zijn doorgaans heel wat gevolgen verbonden. Indien niet is ze nutteloos en, gelet op de nodige kosten die gepaard gaan met evaluatie, zelfs te veroordelen omdat evaluatie zonder gevolgen elders bestuurscapaciteit ontnemt.

Beleidsvaluatie staat dan voor een systematische analyse van doelstellingen die men wil bereiken, alternatieven (middelencombinaties) om deze doelstellingen te bereiken en de relaties tussen doelstellingen en de alternatieven, teneinde te beslissen welke de beste middelencombinatie is om de gestelde doelen te bereiken (Nagel, 2002). Dat laatste is niet evident binnen een steeds veranderende context. Beleid is nooit af, maar ook de maatschappelijke realiteit waarbinnen dit beleid vorm moet krijgen is voortdurend in beweging. Daarom moet ook het evaluatieproces evolueren.

Conceptueel is evaluatie te omschrijven als *'a systematic or careful assessment of the merit, worth, and value of administration, output, and outcome of government interventions, which is intended to play a role in the future, practical action situation'* (Vedung, 1997, 3, o.c. in Dahler-Larsen, 2005, 616). Met deze definitie wordt er nog een element aan toegevoegd, met name de praktische bruikbaarheid van beleidsvaluatie.

Er zijn veel begrippen die aanleunen bij beleidsvaluatie. Evaluatie en monitoring worden vaak door elkaar gebruikt. Ze zijn nochtans van elkaar te onderscheiden. Monitoring omvat het verzamelen van informatie in de context van het beleid en het beheer, over geselecteerde aspecten of factoren. Het is een proces dat op een systematische manier moet plaatsvinden en gebeurt op grond van periodieke metingen of registraties, bijvoorbeeld via een managementinformatiesysteem (MIS). Monitoring kan gebruikt worden om informatie aan te leveren voor de verschillende componenten in de beleids- en beheerscyclus.

Belangrijk is dat monitoring vooral beschrijvende informatie oplevert. Het biedt een antwoord op de wat-vragen. De evaluatie van prestaties is voornamelijk gericht op het beantwoorden van hoe- en waarom-vragen. Andere verschillpunten vinden we bij McDavid & Hawthorn (2006). Zo is evaluatie meer *ad hoc* en onderwerpspecifiek, terwijl monitoring een permanent gegeven is en opgezet wordt uit meer algemene overwegingen. Kortom, monitoring en evaluatie verschillen wat betreft focus en gebruik (Dunn, 1994, De Peuter, De Smedt, Van Dooren & Bouckaert, 2007; UNDP, 2002). Er zijn ook gelijkenissen tussen monitoring en evaluatie. In hun analyse komen Nielsen en Ejler (2008) tot de vaststelling dat beiden gelijkaardig zijn qua structurering en planning alsook het verkrijgen, analyseren en beoordelen van data. De verschillen zien ze vooral wat betreft de scope, de diepgang, de veelheid en de frequentie waarmee de taken moeten worden gedaan. Ze zien beiden dan ook als methodologisch complementair.

---

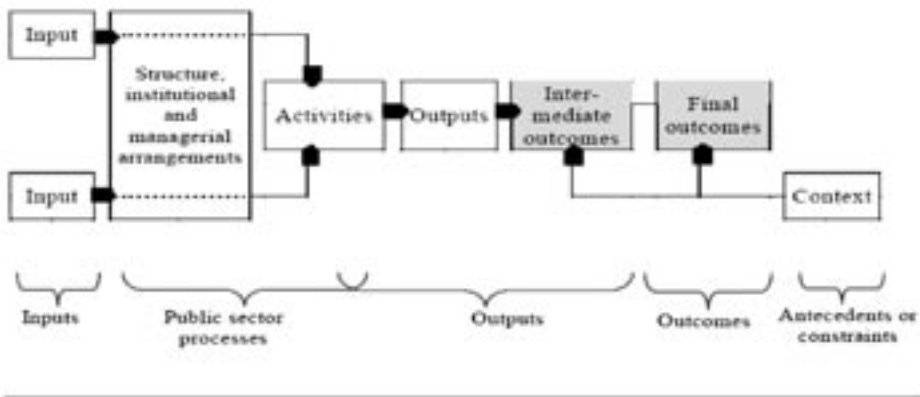
<sup>2</sup> Scriven, M. (2007). The logic of evaluation. Nota te vinden op : <http://homepages.wmich.edu/~mscriven/>  
Van dezelfde auteur (Scriven, 1997) vinden we ook een iets andere definitie terug: *'Evaluation is primarily concerned with determining the merit, worth or value of an established policy or a planned intervention'*. Citaat overgenomen uit Hall, e.a. 2004, 28.

Het voorwerp van evaluatie, het zogenaamde evaluandum, kan heel divers zijn. Toch worden vooral de prestaties in de publieke sector geëvalueerd. Burgers verwachten veel van de overheid, en er is geen “markt” die zoals in de private sector gedragssturend werkt. Dus wordt vooral de “*performance*” van de overheid expliciet geëvalueerd. Wat is *performance*? Alvast een veelzijdig begrip dat verschillende elementen omvat. Dat maakt het lastig om *performance* samen te vatten in één enkele indicator. Ook de verhouding tussen proces en de uitkomst (“*outcome*”) is daarbij van tel (Van De Walle & Bouckaert, 2007 en Tyler, 1990). Van de Walle stelt dat we *performance* en effectiviteit van de overheid niet kunnen meten enkel door het afwegen van *outputs* en *outcomes* t.o.v. bepaalde doelstellingen, omdat de doelstellingen van overheden over het algemeen vaag en soms contradictorisch zijn (Van de Walle, 2008). De overheid is een huis met vele kamers. Gezien beleidsdoelen voor meerdere interpretaties vatbaar zijn, zijn meervoudige indicatoren vereist. Wie effecten wil meten moet over de juiste indicatoren of meetinstrumenten beschikken.

## 2. Beleidsevaluatie als een onderdeel van een veelzijdig en –meerlagig totaalproces

Beleidsevaluatie is onderdeel van een ruimer beleidsproces. De evaluatie kan pas goed gebeuren als er voldoende aandacht is voor de complexiteit van het hele beleidsproces. Het productieproces in de publieke sector kunnen we als volgt weergeven (OECD, 2007, 16):

**Figuur 1: Het productieproces in de publieke sector**



Efficiëntie duidt op de relatie tussen de ingezette middelen (input) en de afgeleverde producten of dienstverlening (output). Productiviteit is dan weer de keerzijde van efficiëntie. Efficiëntie wijst op de hoeveelheid input die nodig is per eenheid output, terwijl productiviteit een maatstaf is voor de hoeveelheid output die men kan realiseren per eenheid input. Effectiviteit verwijst naar de oorzaak- en gevolgrelatie tussen output en effect. Heeft het uitgevoerde beleid het beoogde effect gehad (binnen de vooropgestelde termijn)?

Langs de inputzijde is het voor beleidsevaluatie essentieel om zicht te krijgen op de middenstromen. Langs de outputzijde stelt zich voor de publieke sector het probleem dat haar diensten doorgaans niet op de vrije markt worden verhandeld en hun (markt)waarde meestal moeilijk te berekenen is. Daarom wordt vaak gebruik gemaakt van fysieke productindicatoren die een (in)directe maat zijn voor de productie.

In tegenstelling tot de *output*, kunnen *outcomes* vaak niet simpelweg worden toegeschreven aan handelingen van de overheid (Hatry e.a., 1994). Diverse andere (f)actoren, buiten de controle van een overheid om, kunnen een rol spelen. In welke mate is de overheid door eigen output verantwoordelijk voor veranderingen in de outcome?<sup>3</sup>

Een bijkomend probleem is dat het niet altijd even evident is om een onderscheid te maken tussen de verschillende onderdelen van het productieproces in de publieke sector. Zo is het soms moeilijk om het onderscheid te maken tussen de output en outcome. Het komt alvast de transparantie niet ten goede en kan problematisch zijn voor de interpretatie in het kader van beleidsevaluatie. Bouckaert en Auwers (1999) zien het dan ook als één van de (mogelijke) algemene gebreken van meetsystemen. In deze context heeft Van Dooren (2006) het over de pollutie van het meetobject. Daarnaast is het eigen aan overheidsorganisaties dat diverse onderdelen opereren in een bestuurlijke keten. Dit maakt dat wat voor het ene onderdeel een output is, dan weer als een input wordt gezien voor een ander onderdeel.

Efficiëntie en effectiviteit zijn specifieke aspecten van het ruimere “*good governance*”. Hoewel vandaag bijzonder veel nadruk wordt gelegd op efficiëntie en effectiviteit is “*good governance*” meer dan dat (Verlet, 2008). Ook andere zijn belangrijk: rechtszekerheid, rechtmatigheid, *equity*, transparantie, *accountability*, ... (Smith, 1995). Het overbenadrukken van efficiëntie en effectiviteit ontnemt echter het zicht op andere waarden en criteria. In deze context kunnen we verwijzen naar de drie types waarden die Hood onderscheidt: *sigma*-, *theta*- en *lambda*-waarden. Bij de *sigma*-waarden staan efficiëntie en effectiviteit centraal, met als kern ‘*keep it lean and purposeful*’. *Theta*-waarden hebben dan weer het volgende als centraal kenmerk: ‘*keep it honest and fair*’, terwijl bij de *lambda*-waarden de nadruk komt te liggen op duurzaamheid onder het motto ‘*keep it robust and resilient*’ (Dijkstra, 2008). In zijn analyse benadrukt Dijkstra het belang om rekening te houden met culturele en institutionele factoren teneinde prestatiemeetsystemen op te zetten en toe te passen. Hij wijst er ons ook op dat de perverse effecten van prestatie meting zich vooral manifesteren op het niveau van de *sigma*-waarden zelf (efficiëntie en effectiviteit). Hij merkt op dat ook constitutionele waarden (zoals democratie en rechtsstaat) relevant zijn.

---

<sup>3</sup> Bijvoorbeeld kan men de stijging van de levensverwachting in goede gezondheid (effect) in relatie brengen met het aantal uitgevoerde vaccinaties, het aantal bereikte personen met de preventiecampagnes (output) of met de hoeveelheid overheidsmiddelen (input). Het is niet omdat de indicatoren stijgen in dezelfde richting en met dezelfde snelheid dat de levensverwachting enkel en alleen kan verklaard worden door het preventieve beleid. Ook de vooruitgang van de geneeskunde (waaraan we bijdragen via onderwijs en O&O), de hogere levensstandaard (waaraan we bijdragen via economische stimuli) waardoor men zich een gezondere voeding en woning kan permitteren en medische behandeling niet langer moet uitstellen, kunnen een invloed hebben.

Samen met Dijkstra (2008) leggen we hier het verband naar de drie niveaus die Kiser en Ostrom benoemen in hun *'Three worlds of action'*: het operationele, het collectieve en het constitutionele keuzeniveau. Op het operationele niveau staat het gedrag van individuele ambtenaren centraal, het collectieve niveau handelt over managementsystemen (prestatie metingen), terwijl het constitutionele niveau waarden en de institutionele inbedding omvat. In dezelfde lijn ligt het onderscheid tussen het micro-, meso- en macroniveau.

### 3. Motieven tot beleidsevaluatie

De beweegredenen tot evaluatie zijn talrijk en divers. Voor een overzicht daarvan verwijzen we naar het werk van De Peuter e.a. (2007). Samenvattend onderscheidt men vier beweegredenen die aan de basis liggen van het evalueren van het beleid.

Vooreerst is er het ondersteunen van de beleidsplanning. In deze context kunnen we beleidsevaluaties situeren die tot doel hebben om na te gaan in welke mate vooropgesteld beleid de vooropgestelde noden kan lenigen. Ook de evaluaties waarbij diverse alternatieve oplossingen worden geschetst en vergeleken kunnen we onder deze noemer plaatsen. De beleidsevaluatie met het doel de beleidsplanning te ondersteunen kan eveneens een evaluatie inhouden van de coherentie en consistentie van het geplande beleid.

Ten tweede kan de evaluatie ook als grondslag het ondersteunen van de beleidsuitvoering hebben. De nadruk komt dan te liggen op de optimalisering van de uitvoering en de werking van de entiteiten die instaan voor deze uitvoering.

Een derde beweegreden is te vinden in het verschaffen van verantwoording omtrent de werking en geleverde dienstverlening door overheidsdiensten.

Een vierde en laatste beweegreden is het beleidsleren. Immers, ook binnen de publieke sector wil men de nodige lessen trekken uit het gevoerde beleid teneinde dit –zo nodig– bij te sturen.

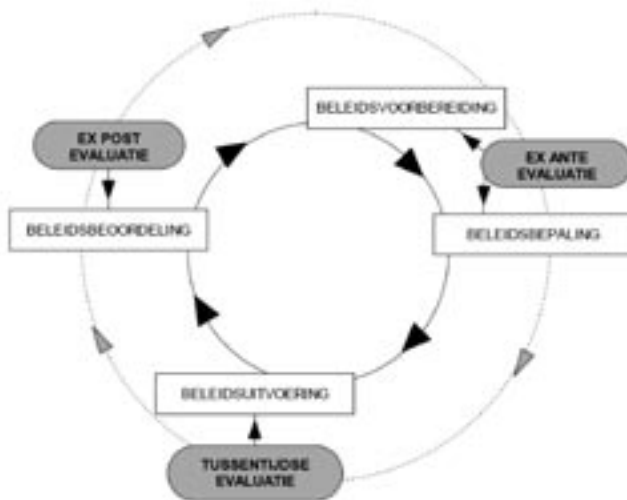
### 4. Types van evaluaties

Eigen aan de theorie en de praktijk van beleidsevaluatie is dat er tal van evaluatietypes en –typologieën zijn te onderscheiden. Hoewel verschillende van deze types nog verder aan bod komen in het boek, willen we in deze inleiding een bondig overzicht opmaken. We verwijzen opnieuw naar het basiswerk van De Peuter e.a. (2007).

In een eerste typologie maakt men een onderscheid naargelang van het tijdstip waarop de evaluatie in de beleidscyclus plaatsvindt. De drie klassieke types die in deze context naar voren komen zijn de ex ante, de tussentijdse en de ex post evaluatie. Een ex ante-evaluatie vindt plaats tijdens de beleidsplanning, een tussentijdse evaluatie gebeurt gedurende de beleidsuitvoering, terwijl een ex post-

evaluatie gebeurt gedurende de beleidsbeoordeling (Wollmann, 2007). In figuur 2 worden deze types weergegeven.

**Figuur 2: Beleidsevaluatie geplaatst binnen de beleidscyclus**  
(De Peuter, e.a. 2007 , 36)

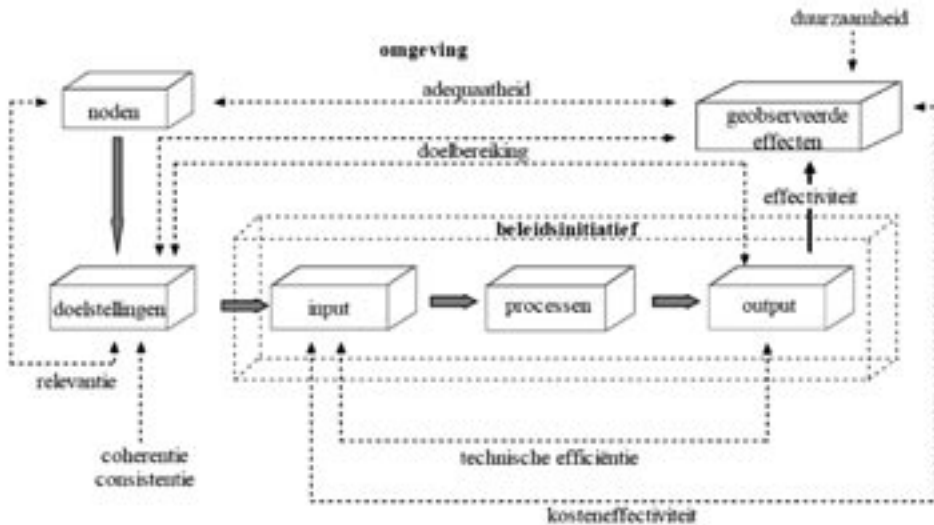


Een tweede typologie focust zich op de inhoudelijke kern van de evaluatie. Hierbij maakt men gebruik van het onderscheid dat Hoogerwerf (2003) maakt tussen de inhoud van het beleid enerzijds en de beleidsimpact anderzijds. Het is duidelijk dat de beleidsinhoud op zich een veelomvattende term is. Men doelt er mee op het concept, de structuur, de instrumenten, de processen en de acties van het beleid. De term beleidsimpact is alvast scherper te omlijnen, het betreft de veranderingen of effecten die het beleid teweegbrengt in de maatschappelijke omgeving.<sup>4</sup> We zien overigens een duidelijke link met de vorige typologie. Ex ante evaluaties leggen de nadruk op de beleidsinhoud, terwijl ex post evaluaties eerder de focus leggen op de beleidsimpact.

Een derde evaluatietypologie maakt het onderscheid op basis van de evaluatiecriteria die men vooropstelt. De evaluatiecriteria om een oordeel te vellen over het functioneren van de publieke sector kunnen we afleiden uit de verschillende componenten uit het open systeemmodel alsook de relaties tussen deze componenten.

<sup>4</sup> Hoogerwerf, A., (2003). Beleid, processen en effecten. In: Hoogerwerf, A. & Herweijer, M. Overheidsbeleid. Een inleiding in de beleidswetenschap. Kluwer, Alphen aan den Rijn, pp.17-35.

**Figuur 3: Evaluatiecriteria afgeleid uit de beleids- en beheerscyclus als open systeemmodel**  
(De Peuter, e.a. 2007 , 38)



Op grond van dit schema kunnen we duiding geven met betrekking tot de volgende aspecten die essentieel zijn in de evaluatie van de “performance” van de overheid.

- Beleidsrelevantie: de mate waarin doelstellingen tegemoet komen aan de maatschappelijke noden en behoeften.
- Interne coherentie: mate waarin strategische en operationele doelstellingen een samenhangend geheel vormen.
- Externe consistentie: de mate waarin de doelstellingen compatibel zijn met doelstellingen binnen andere beleidsvelden.
- Technische efficiëntie: de verhouding tussen input en output.
- Doelbereiking: de mate waarin de doelstellingen en meer bepaald de ambitiesniveaus of normen worden gehaald.
- Effectiviteit: de mate waarin de geobserveerde effecten of veranderingen toe te schrijven zijn aan de beleidsoutput.
- Kosteneffectiviteit: de verhouding tussen input en effect.
- Duurzaamheid van effecten: de mate waarin effecten blijven bestaan na de beëindiging van het beleid.
- Adequaatheid van effecten: de mate waarin de effecten of veranderingen tegemoet komen aan de maatschappelijke noden en behoeften.

Daarnaast zijn er nog andere bijkomende belangrijke evaluatiecriteria, die niet door middel van dit model worden weergegeven. Zo is er het draagvlak voor het beleid, te omschrijven als de mate waarin de doelstellingen, maatregelen en instrumenten van het beleid aanvaard worden door de *stakeholders*. Een ander criterium kan zijn: de sterkte van de achterliggende beleidslogica. Het gaat over de mate waarin de



assumpties waarop een beleidsinitiatief gestoeld is, doordacht en onderbouwd zijn.

Een vierde en laatste evaluatietypologie baseert zich op de organisatorische positie van de evaluator. Meer specifiek handelt deze typologie over het onderscheid tussen een interne dan wel externe evaluatie (Wollmann, 2007). Men spreekt van een interne evaluatie indien de evaluator deel uitmaakt van de organisatie waarvan het beleid wordt geëvalueerd. Bij een externe evaluatie, staat de evaluator los van de organisatie die men onder de loep neemt.<sup>5</sup>

Daarnaast kan deze algemene typologie nog worden verfijnd. Meer specifiek zien we onder meer in het werk van De Peuter e.a. (2007) dat men een onderscheid kan maken tussen verschillende subtypes wat ex ante evaluatie betreft. Deze subtypes worden dan gemaakt op grond van de klemtonen die worden gelegd in de evaluatie, zoals de relevantie of geldigheid van het beleid, de interne logica, het afwegen van alternatieven voor beleid, de coherentie en consistentie van het beleid, de verwachte effecten en impact van het beleid en de implementatie van het beleid.

Ook wat ex post-evaluaties betreft, maakt men een onderscheid tussen verschillende subtypes. Zo heeft men het over een experimenteel en quasi-experimenteel design. Eigen aan experimentele designs is dat men een vergelijking maakt tussen “policy-on”- en de “policy-off”-groepen. Het idee is dat men in experimentele design een vergelijking maakt tussen enerzijds één of meerdere groepen waarbij men beleid(salternatieven) toepast (de zogenaamde beleidsgroep) en anderzijds een groep waarbij het beleid niet wordt toegepast (de controlegroep). Het is duidelijk dat in de context van beleidsevaluatieonderzoek het veelal niet wenselijk/mogelijk is dergelijke experimentele designs op te zetten. Als alternatief kan een quasi-experimenteel onderzoeksdesign dienen. Het essentiële verschil tussen een experimenteel en quasi-experimenteel onderzoek is dat in een quasi-experimenteel onderzoek men de vergelijkingsbasis niet door toeval of ad random kan vastleggen. Veelal zijn praktische dan wel ethische overwegingen hiervoor de reden. Indien ook een quasi-experimenteel ex post-onderzoeksopzet niet mogelijk is, kan men zich richten tot een kwalitatief design, waarbij door middel van hoofdzakelijk kwalitatieve methoden toch een beeld kan ontstaan van de oorzaak-gevolgrelaties.

Binnen de vele types van beleidsevaluatie zijn er heel veel verschillende methoden en technieken mogelijk. Denken we bijvoorbeeld aan technieken van ex ante evaluatie, zoals een kostenbatenanalyse, een kosteneffectiviteitsanalyse of een multicriteria-analyse bij het onderling afwegen van verschillende beleidsalternatieven. Verderop in dit boek wordt dieper ingegaan op deze technieken.

Dit overzicht van verschillende types evaluatieonderzoek heeft geenzins de pretentie om volledig te zijn. In de literatuur over beleidsevaluatieonderzoek zijn nog diverse andere mogelijke typologieën te vinden. Een aantal daarvan gaan terug tot basisindelingen binnen het wetenschappelijk onderzoek. Zo maakt Wollman

---

<sup>5</sup> Voor een overzicht van de voor- en nadelen van een interne dan wel externe evaluatie kunnen we verwijzen naar het werk van De Peuter, e.a. (2007: 40-42).



(2007) het onderscheid tussen fundamenteel onderzoek, toegepast onderzoek en contractonderzoek. Nog anderen maken dan weer een onderscheid naargelang van de specifieke doelstelling van de evaluatie. Zo hebben Hall en Hall (2004) het in hun overzichtswerk over het onderscheid tussen drie types doelstellingen: “*accountability*”, ontwikkeling en kennis. Nog een andere mogelijke typologie maakt het onderscheid op basis van de globale benadering van het evaluatieonderzoek. Zo maakt Dahler-Larsen (2005) het onderscheid tussen drie benaderingen in het beleidsevaluatieonderzoek: een doelgerichte benadering, een theoriegebaseerde benadering en een responsieve/participatorische benadering. Verderop in het boek gaan Lannoo en Verlet dieper in op het onderscheid tussen deze benaderingen. Ook het onderscheid tussen kwantitatieve dan wel kwalitatieve methoden bij evaluatieonderzoek komen aan bod. In diverse basiswerken is ook de gebruikte methode een criterium voor het oplist van diverse types onderzoek.

Daarnaast kunnen we een onderscheid maken tussen vier generaties inzake beleidsevaluatieonderzoek. Het onderscheid tussen die vier generaties evaluatie hebben we te danken aan Guba en Lincoln (1989). In de eerste generatie ligt de nadruk op het meten als dusdanig (cf. monitoring). Het beschrijven staat dan weer voorop in wat men de tweede generatie noemt. In de derde generatie staat het beoordelen centraal. Bij evaluatieonderzoek uit de vierde generatie geven Guba en Lincoln onderhandelen een centrale plaats. In deze benadering zijn pluraliteit, dialoog en participatie centrale waarden zijn bij het voeren van beleidsevaluatieonderzoek. In de bijdragen van Bongers (hoofdstuk 3) en Gielen en Junne (hoofdstuk 6) wordt er dieper ingegaan op dit onderscheid.

## 5. Wat volgt?

Ruw omschreven bevatten de bijdragen in dit boek een kritische analyse van beleidsevaluatie en het onderzoek ter zake, met een bijzondere aandacht voor de invloed van beleidsevaluatie op het beleidsproces. Beleidsevaluatie kan leiden tot beleidsleren en tot betere verantwoordingsprocessen. Maar daarnaast kan beleidsevaluatie op een onvoorziene/oneigenlijke manier het beleidsproces beïnvloeden.

De bijdragen kunnen opgedeeld worden in eerder theoretische bijdragen en bijdragen die de nadruk leggen op zo genaamde case-studies.<sup>6</sup> In beide groepen zijn er hoofdstukken uit Nederland en uit Vlaanderen.

In het eerste hoofdstuk breken Klaassen en van Nispen een lans voor integrale beleidsevaluatie. Ze analyseren verschillende types van beleidsanalyse en gaan in op de vraag hoe veranderingen ter zake de rol van beleidsevaluatie hebben beïnvloed. Verlet en Devos staan in hun bijdrage stil bij een *hot issue* in het beleids-evaluatieonderzoek: welke positieve en negatieve effecten kunnen prestatie-meting

---

<sup>6</sup> Hiermee willen we natuurlijk geen afbreuk doen aan de praktijktoets binnen de theoretische bijdragen, noch aan de theoretische basis van de case-studies.

met zich mee brengen? De bijdrage van Bongers is daar een logisch vervolg op: hij stelt zich de vraag of er, bij beleidsevaluatie, sprake is van preventief gedrag binnen overheidsorganisaties. Hij wijst ons onder meer op de cruciale rol die institutionele, beleidsmatige, maar ook persoonlijke factoren kunnen spelen in deze. In de daaropvolgende bijdrage staan De Peuter, Pattyn & Brans stil bij een risicoanalyse van de effecten van beleidsevaluatie. Bijzondere aandacht gaat hierbij ook naar de Vlaamse situatie.

In haar bijdrage gaat Wayenberg dieper in op de plaats en betekenis die aan beleidsevaluatie wordt gegeven binnen de Vlaamse overheid, meer specifiek in de beleidsdomeinen Binnenlands Bestuur en Stedenbeleid. We begeven ons in de Nederlandse context met de bijdrage van Gielen en Junne. Zij bekijken een specifiek onderwerp van evaluatie, met name de evaluatie van anti-radicalisering sprocessen. In deze bijdrage worden minder traditionele evaluatiemethoden voor het licht gehouden. Van Dongen werkt op grond van een algemeen analysemodel voor overheidsdiensten een evaluatie uit voor de dagopvang voor jonge kinderen in het Vlaams gewest. In het voorlaatste hoofdstuk van dit boek vinden we een goed voorbeeld van een situatie waarbij meten het doel dreigt te worden. Meier en Celis gaan meer specifiek na in welke mate de *gender mainstreaming* vertaald wordt naar de beleidspraktijk. In de verschillende bijdragen zijn kritische geluiden over de beleidsevaluatie te horen. Dit doet geen afbreuk aan het belang van beleidsevaluatie. Alleen is, zo blijkt, de vertaling naar de praktijk niet altijd even evident. Daarom wordt in het laatste hoofdstuk, van Lannoo en Verlet, een voorbeeld gegeven over hoe de theorie in de praktijk van beleidsevaluatie kan worden omgezet. Afsluiten doen we met een epiloog.

## Bibliografie

Bouckaert, G. & Auwers, T. (1999). *Prestaties meten in de overheid*. Brugge: Die Keure.

Dahler-Larsen, P. (2005). Evaluation and public management. In: Ferlie, E., Lynn, L. & Pollitt, C. (ed.) *The Oxford handbook of public management*. New York: Oxford University Press, 615-639.

De Bruijn, H. (2002). Prestatiemeting in de publieke sector. Strategieën om perverse effecten te neutraliseren. In: *Bestuurswetenschappen*, 56 (2), 139-159.

De Peuter, B., De Smedt, J. Van Dooren, W. & Bouckaert, G. (2007). *Handleiding beleidsevaluatie. Deel 2: monitoring van beleid*. Leuven, Onderzoeksrapport.

Dijkstra, G. (2008). Outputmeting van overheidspersoneel: een simpele prestatie? In: *Bestuurswetenschappen*, 62 (1), 54-68.

Dunn, W. (1994). *Public Policy Analysis. An introduction* (2<sup>nd</sup> ed.). London: Prentice-Hall International.

Dye, T. (1976). *Policy analysis. What governments do, why they do it, and what difference it makes*. Alabama: The University of Alabama Press.

Guba, E.G. & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park: Sage Publications.

Hall, I., Hall, D. & Campling, J. (2004). *Evaluation and social research: introducing small-scale practice*. Houndmills: Palgrave.

Hatry, H., Gerhart, C. & Marshall, M. (1994). Eleven Ways to Make Performance Management More Useful to Public Managers. In: *Public Management*, September 1994, 15-18.

Nagel, S. (2002). *Handbook of public Policy Evaluation*. Thousand Oaks/London/New Delhi: Sage Publication.

Nielsen, S.B. & Ejler, N. (2008). Improving performance? Exploring the complementarities between evaluation and performance management. In: *Evaluation*, 14 (2), 171-192.

OECD (2007). *Issues in Output Measurement for "Government at a Glance"* OECD project on Management in Government: Comparative Country Data. OECD GOV Technical Paper 2 (Second Draft).

Scriven, M. (2007). *The logic of evaluation*. Nota te vinden op : <http://homepages.wmich.edu/~mscriven/>, laatst geraadpleegd op 10 oktober 2008.

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. In: *International Journal of Public Administration*, 18 (2/3), 277-310.

Spanhove, J. & Verhoest, K. (2007). *Corporate governance vs. Government governance: translation or adoption?* Paper, EIASM, 4<sup>th</sup> workshop on corporate governance, Brussel, 15-16 november 2007.

Swanborn, P. G. (1999). *Evaluëren: het ontwerpen, begeleiden en evalueren van interventies: een methodische basis voor evaluatie-onderzoek*. Amsterdam, Boom.

Trochim, W. (2008). *Introduction to evaluation*. <http://www.socialresearchmethods.net/kb/intreval.html>, laatst geraadpleegd op 8 oktober 2008.

UNDP, (2002). *Handbook on Monitoring and Evaluating for Results*. New York: United Nations Development Programme, Evaluation Office.

Van de Walle S. & Bouckaert G. (2007). Perceptions of productivity and performance in Europe and the United States. In: *International Journal of Public Administration*, 30 (11), 1123-1140.

Van de Walle, S. (2008). Comparing the performance of national public sectors: conceptual problems. *International Journal of Productivity and Performance Management*, 57 (4), 329-338.

Van Dooren, W. (2006). *Performance measurement in the Flemish public sector: A supply and demand approach*. K.U. Leuven, gepubliceerd doctoraal proefschrift.

Verlet, D., (2008). *Good governance, corporate governance, government governance: what's in a name? Een theoretische situering van Beter Bestuurlijk Beleid in Vlaanderen*. Brussel, Studiedienst van de Vlaamse Regering, SVR-nota 2008/4 (in druk).

Wollmann, H. (2007). Policy Evaluation and Evaluation Research. In: Fischer, F., Miller, G. & Sidney, M. *Handbook of Public Policy Analysis*. Boca Raton/London/New York, CRC Press, 393-402.

# Hoofdstuk 1: Beleidsevaluatie: gelijk hebben ≠ gelijk krijgen

Henk L. Klaassen & Frank K.M. van Nispen<sup>7</sup>

Faculteit der Sociale Wetenschappen, Erasmus Universiteit Rotterdam

1. Inleiding .....	21
2. Beleidsanalyse, een plaatsbepaling .....	22
3. De institutionalisering van beleidsanalyse in Nederland .....	24
4. Beleidsanalyse als kosten-batenanalyse (kba): “input” versus “outcome” □ .....	25
5. Beleidsanalyse als effectrapportage: “output” versus “outcome” .....	27
6. Beleidsanalyse als performance measurement: “input” versus “output/ outcome” .....	30
7. De veranderende rol van de beleidsanalyse, i.c. de beleidsanalist .....	32
8. Vergelijking .....	33
9. Conclusie .....	35
Bibliografie .....	36

---

<sup>7</sup> De auteurs danken de deelnemers aan de workshop “Beleidsevaluatie: doel of middel?” op het Politicologenetmaal voor hun commentaar op een eerdere versie van deze bijdrage.

## 1. Inleiding

In de afgelopen decennia heeft zich, aldus Osbourne & Gaebler, een ontwikkeling voorgedaan van “government” naar “governance”, te weten ‘... *the process by which we collectively solve our problems and meet our society’s needs*’ (Osbourne & Gaebler, 1993, 24). Kersbergen & Van Waarden (2004) constateren daarnaast dat zich in de loop van de tijd heeft een aantal “shifts in governance” voorgedaan. In dit verband dient met name de ontwikkeling van de klassieke top-down benadering naar de netwerkbenadering te worden genoemd (Klijn, 1996; Börzel 1997, 5; Rhodes 1997). Ten opzichte van een hiërarchie (unicentrisch) en de markt (multicentrisch) kan een netwerk worden gekarakteriseerd als een pluricentrische vorm van governance (Van Kersbergen & Van Waarden, 2004)<sup>8</sup>. In de netwerkbenadering heeft de overheid haar monopolie verloren en wordt beleid opgevat als de co-productie van meerdere – deels samenwerkende, deels tegenwerkende – actoren die wederzijds van elkaar afhankelijk zijn voor de realisering van hun doelstellingen (Börzel, 1997).

In dezelfde periode heeft zich een verschuiving voorgedaan van beleid naar management. Het sterkst is deze ontwikkeling herkenbaar in de brede omarming van het gedachtegoed van New Public Management (NPM). Beide ontwikkelingen hebben sterke invloed op de rol die beleidsanalyse speelt. De netwerkbenadering heeft bijvoorbeeld belangrijke consequenties voor de rol van de beleidsanalist. Hij is zijn monopolie kwijt (Radin, 2000) en moet de concurrentie aan met andere beleidsanalisten die in opdracht werken van actoren die deel uitmaken van dat netwerk. De opkomst van NPM en het daarmee samenhangende streven naar van performance management heeft ook invloed op de rol en de betekenis van de beleidsanalist: niet meer de “of- vraag” staat centraal, maar meer en meer de “hoe-vraag”

In deze bijdrage staan wij stil bij de veranderende positie van beleidsanalyse (waaronder beleidsevaluatie) en de veranderde positie van de beleidsanalist. Wij vatten dit samen in de volgende vraag:

*In hoeverre is de inhoud en daarmee de betekenis van beleidsanalyse als analytisch instrument ten behoeve van beleidsvorming veranderd in de afgelopen decennia en wat impliceert dit voor de functie die beleidsanalyse?*

In de literatuur zijn vele opvattingen van beleidsanalyse en –evaluatie te onderscheiden. Wij gaan na ons gepositioneerd te hebben (paragraaf 2), kort in op de institutionalisering van beleidsanalyse in Nederland waarbij met name de initiërende rol van de Commissie voor de Ontwikkeling van Beleidsanalyse (COBA) dient te worden genoemd. De uitvoering van beleidsanalyse is daarna gedecentraliseerd en verspreid over allerlei instituten (paragraaf 3). Wij onderkennen drie stadia in

---

<sup>8</sup> Wij abstraheren hierbij van de vraag of de netwerkbenadering een theorie vormt of “slechts” een metafoor is (Konig, 1998, 387; Peterson, 2003, 12).

de ontwikkeling van beleidsanalyse<sup>9</sup> die zich min of meer chronologisch hebben voorgedaan. Het vertrekpunt is de ontmanteling van het Planning Programming Budgeting System (PPBS), waarover Aaron Wildavsky een fraai essay heeft geschreven: *'rescuing policy analysis from PPBS'* (Wildavsky, 1969). In die tijd werd beleidsanalyse min of meer gelijk gesteld aan een kosten-batenanalyse (kba) welke zich kenmerkt door een confrontatie tussen "inputs" en "outcomes". De methodische problemen die inherent zijn aan het waarderen van effecten op geld, heeft onder andere geleid tot de ontwikkeling van niet-monetaire methoden zoals multicriteria-analyse en de "score-card" (paragraaf 4). De institutionalisering van belangen in de samenleving heeft vervolgens geleid tot een proliferatie van beleidsanalyse resulterend in een "wildgroei" van over elkaar heen buitelandse effectrapportages, waarin wordt getracht een relatie te leggen tussen "outputs" en "outcomes" (paragraaf 5). Het gedachtegoed van beleidsanalyse heeft een nieuwe impuls gekregen door de zgn. "performance movement" welke onder andere is gericht sturing op resultaten: *'funding outcomes, not inputs'* (Osbourne & Gaebler, 1993, 138). Het streven naar sturing op resultaten heeft in ons land gestalte gekregen in de VBTB (Van Beleidsbegroting tot Beleidsverantwoording), i.e. de drie www-vragen (wat willen we bereiken, wat moeten we daarvoor doen en wat mag het kosten) en de bijbehorende drie hhh-vragen (hebben we bereikt wat we wilden bereiken, hebben we gedaan wat we moesten doen en heeft het gekost wat het mocht kosten)<sup>10</sup>, waarmee wordt beoogd een directe relatie te leggen tussen inputs (middelen) en outputs (prestaties) en zelfs outcomes (effecten) (paragraaf 6). We besluiten deze bijdrage met de conclusie dat er verschuivingen hebben plaatsgevonden in het domein van beleidsanalyse. De betekenis van beleidsanalyse zelf is een andere geworden en daarmee heeft ook de rol van de beleidsanalist een verandering ondergaan.

## 2. Beleidsanalyse, een plaatsbepaling

Beleidsanalyse, waaronder beleidsevaluatie, heeft als doelstelling het genereren van informatie die kan worden gebruikt bij positiebepaling en oordeelsvorming van actoren bij beleidskwesties. Gezien het feit dat een beleidsprobleem geen rekening houdt met de traditionele, disciplinaire grenzen is beleidsanalyse een multi- of interdisciplinaire activiteit (Van Nispen, 1993): *'... policy analysis is a process of multidisciplinary inquiry designed to create, critical assess, and communicate information that is useful in understanding and improving policies'* (Dunn, 2004, 2).

In de literatuur is sprake van een amalgaam van benaderingen, opvattingen en stijlen van beleidsanalyse (Parsons, 1995; Mayer 2002; Van Nispen, 2003). In een poging om deze in kaart te brengen, kan allereerst onderscheid worden gemaakt tussen een theoriegerichte benadering van beleidsanalyse ("theory-driven") die primair is gericht op het leveren van een bijdrage aan de "body of knowledge" op een bepaald gebied

---

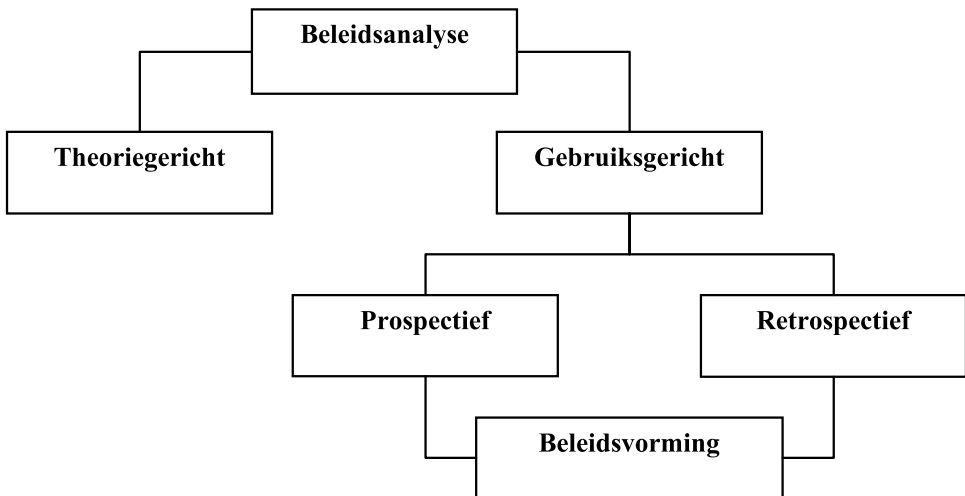
<sup>9</sup> Wij abstraheren hiermee van de ontwikkeling van de neo-constructivistische beleidsanalyse waarin argumentatie en discourse centraal staan (zie bijv. Fischer, 2003; Hajer & Wagenaar, 2003; Stone, 2002).

<sup>10</sup> De beleidsanalytische grondslag van de VBTB wordt uit de doeken gedaan door Boorsma e.a. in hun historische vergelijking.

en een meer gebruiksgerichte benadering van beleidsanalyse (“utilization-focused”) die beoogt een bijdrage te leveren aan het oplossen van maatschappelijke problemen. In deze bijdrage beperken wij ons voornamelijk tot de laatste vorm van beleidsanalyse (Dunn, 2004; Radin, 2000; Weimer & Vining, 1999).

In navolging van Dunn kan nader onderscheid worden gemaakt tussen in prospectief en retrospectief onderzoek (Dunn, 2004). Waar prospectief onderzoek toekomstgericht is, heeft retrospectief onderzoek een terugblikkend karakter.<sup>11</sup> Zij vormen tezamen de integratieve beleidsanalyse, dat ‘... has all the methodological advantages of prospective and retrospective analysis but none of their weaknesses. Integrated (policy) analysis provides for the continuous monitoring and evaluation of policies over time’ (Dunn, 1994, 79). Een nuancering is overigens op zijn plaats. De mogelijkheid bestaat dat de resultaten van theoriegericht onderzoek bruikbaar zijn voor het oplossen van maatschappelijke problemen, terwijl de resultaten van gebruiksgericht onderzoek van nul en generlei waarde blijken te hebben.

**Figuur 1.1: Plaatsing van beleidsanalyse**



Uit de figuur mag niet worden afgeleid dat de gemaakte analysevormen zich in de praktijk altijd duidelijk van elkaar laten onderscheiden. In een analyse kunnen zowel prospectieve als retrospectieve componenten voorkomen. Het is goed mogelijk dat een ex post analyse een essentieel onderdeel uitmaakt van een ex ante analyse. Zo geldt voor het ontwerpen van overheidsbeleid dat de beleidsopties verschillend kunnen scoren in termen van de bijdragen die zij leveren aan doelstellingen, aan onbedoelde negatieve effecten of aan het budgettaire beslag dat realisatie met zich brengt. Om die reden dienen beleidsopties tegen elkaar te worden afgewogen. Onderzoek en analyse die in dit verband worden verricht, hebben tot doel de beleidsopties van een zeker argumentatiegehalte te voorzien om ze op die manier te kunnen laten concurreren

<sup>11</sup> In dat verband wordt ook wel onderscheid gemaakt tussen analyse van beleid en analyse voor beleid (zie bijv. Parsons, 1995, 54-57). In de praktijk loopt dat onderscheid soms door elkaar. Het is goed mogelijk dat een analyse van beleid een essentieel onderdeel uitmaakt van een analyse voor beleid.



met andere voorstellen. Het is niet zo dat slechts ex ante analyse hieraan een bijdrage kan leveren. Ex post analysevormen kunnen hierbij eveneens waardevol zijn. Daarbij geldt dan overigens wel de veronderstelling dat er een zekere continuïteit bestaat in het beleid en dat de omstandigheden waaronder het beleid ten uitvoer wordt gebracht geen drastische wijzigingen ondergaan. Discontinuïteit of sterk veranderende omstandigheden belemmeren dat oordelen, bijvoorbeeld over effectiviteit van beleid, doorgetrokken kunnen worden naar de toekomst. In de praktijk vervaagt het onderscheid dan ook dikwijls tussen zuivere vormen van ex post evaluatie en ex ante evaluatie. De heroverweging van overheidsuitgaven die in de jaren tachtig van de vorige eeuw heeft plaatsgevonden is daarvan een goede illustratie (Van Nispen, 1993).<sup>12</sup>

Beleidsanalyse met een directe relatie tot beleidsvorming kan spanning opleveren bij het objectief uitvoeren van een analyse. Zo wijst Lynn op het gevaar van beïnvloeding door de beleidsanalist dat bestaat, indien een beleidsanalyse een relatie krijgt met het beleidsvormingsproces. Hij ziet liever dat beleidsanalyse een meer afstandelijke aangelegenheid blijft: *'Because of their prescriptive orientation, policy analysts tend to become advocates for their own views. Acting as advocates may lead policy analysts, consciously or unconsciously, to present arguments and evidence in a biased or even deceptive manner, which raises questions as to whether their advice can be trusted'* (Lynn, 1987, 166). De consequentie van deze erkenning dat sociaal-wetenschappelijk onderzoek niet waarde vrij is, is dat minder gehecht kan en mag worden aan de brugfunctie die met name door overheden geëntameerd onderzoek heeft. Het heeft meer een beleidsonderbouwende dan beleidsondersteunende functie, hetgeen overigens niet per se slecht hoeft te zijn: *'The purpose of policy analysis is not to eliminate advocacy, but to raise the level of argument among contesting interests ... The end result, hopefully, would be a higher quality debate and perhaps eventually public choice among better known alternatives'* (Wildavsky, 1969, 190).

### 3. De institutionalisering van beleidsanalyse in Nederland

Een belangrijk moment voor institutionalisering van beleidsanalyse in Nederland was de oprichting van de Commissie voor de Ontwikkeling van Beleidsanalyse (COBA) in 1971. Een maatschappelijke kosten-batenanalyse, als de meest vergaande en geavanceerde vorm van beleidsanalyse, werd beschouwd als het middel om de effectiviteit en doelmatigheid van overheidshandelen te toetsen. Door middel van een kosten-batenanalyse kon dan worden bepaald welke oplossing maatschappelijke gezien het meest gunstig was. Door overheidsuitgaven te bezien in het licht van hun functionaliteit in termen van kosten en baten werd eveneens verwacht dat beleidsanalyse een grote rol zou kunnen vervullen in het beheersen van de overheidsfinanciën. Dat was kortweg de missie van de COBA.<sup>13</sup>

---

<sup>12</sup> De heroverweging van overheidsuitgaven heeft later een vervolg gekregen in de interdepartementele beleidsonderzoeken en zgn. beleidsdoorlichtingen.

<sup>13</sup> Een uitgebreide beschrijving van de geschiedenis van de COBA is te vinden in Obbink & Van Zwieten (1993). Zie tevens Klaassen & Van Nispen (1995).

Het beleidsanalytische bolwerk begon al ver voor de voltooiing scheuren te vertonen. De inspanning om voor alle overheidsuitgaven beleidsanalytische studies te verrichten bleek te groot. De variëteit aan activiteiten die met de overheidsuitgaven samenhangen was bovendien dermate groot dat men met het beleidsanalytisch instrumentarium daarop onvoldoende kon inspelen. Ook methodologische problemen werden steeds duidelijker, zoals de juiste hoogte van de disconteringsvoet, het meenemen van indirecte effecten, het ontbreken of onduidelijk zijn van beleidsdoelen, het ontbreken van duidelijke causaliteiten tussen beleidsinstrumenten en de doelen waarvoor ze worden ingezet, of het optreden van beleidsinterferentie in positieve of negatieve zin. Het hoge ambitieniveau van beleidsanalyse werd daarom al vroegtijdig bijgesteld. Beleidsanalyse zou worden tot het instrument ter facilitering van beleidsvoorbereiding en beleidsbepaling (beleidsvorming).

In de jaren '90 zien we een steeds verdere afkalving van beleidsanalytische instituties. Organisatieonderdelen met een expliciete taak op het terrein van beleidsanalyse worden opgeheven en ondergebracht bij bestaande organisaties die zich direct bezighouden met beleidsvorming. Achterliggende redenering was dat de expertise op het terrein van de bij organisatieonderdelen in de lijn van voldoende niveau was geworden. En dat de methodologische ontwikkeling en de taken inzake de verspreiding van beleidsanalytische vaardigheden (de cursussen beleidsanalyse) door organisaties buiten de rijksdienst opgepakt zouden kunnen worden. Deze insteek heeft voor- en nadelen. Zo is enerzijds waarneembaar dat de ontwikkeling in methodologisch opzicht doorgang heeft gevonden, getuige de ontwikkeling van OEEI (onderzoek economische effecten infrastructuur) en later OEI (onderzoek effecten infrastructuur)<sup>14</sup>. De afhankelijkheid van professionele instituten bij het uitvoeren van complexe beleidsanalytische studies is anderzijds groter geworden. Op sommige gebieden is expertise op het terrein van de toepassing zo geconcentreerd dat men zou kunnen spreken van ongewenste concentratie (Buck, 2002).<sup>15</sup>

#### **4. Beleidsanalyse als kosten-batenanalyse (kba): “input” versus “outcome”**

In de beginjaren was beleidsanalyse min of meer gelijk gesteld aan een kosten-batenanalyse (Jenkins-Smith, 1990; Parsons, 1995). Op zich is dat niet vreemd, omdat het (nog) steeds kan worden beschouwd als één van de krachtigste beleidsanalytische instrumenten. In principe is dit instrument bruikbaar voor de afweging van zowel projecten als programma's: de kosten-batenanalyse is een evaluatiemethode waarbij een zoveel mogelijk gekwantificeerd overzicht wordt gegeven van de voor- en nadelen over een langere periode van een project of programma. Bij een maatschappelijke kosten-batenanalyse wordt getracht te bepalen of de maatschappij in haar totaliteit een netto voordeel heeft bij het uitvoeren van het

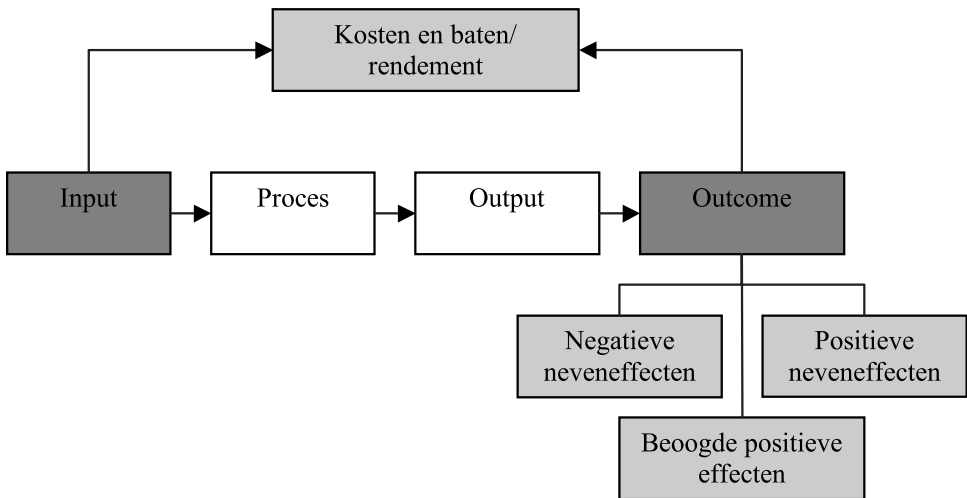
---

<sup>14</sup> Sinds 2000 wordt in Nederland voor (grote) infrastructuurprojecten een Overzicht Effecten Infrastructuur (OEI) opgesteld. Voor het opstellen van een OEI is een procedurebeschrijving ontworpen die nagenoeg overeenkomt met het opstellen van een kosten-batenanalyse.

<sup>15</sup> Buck Consultants constateerde bij een evaluatie van de OEI (2002) dat de expertise op het terrein van kosten-batenanalyse nagenoeg geheel is geconcentreerd bij het Nederlands Economisch Instituut (tegenwoordig Ecorys)

project of de maatregel. Uitgaande van de input-outputssystemetiek die gebruikelijk is bij de toepassing van kosten-batenanalyse, kan de opzet als volgt in beeld worden gebracht:

**Figuur 1.2: Het concept kosten-batenanalyse**



De inputs worden gevormd door de direct voor het project of maatregel op te offeren productiefactoren, uitgedrukt in geldtermen. De outputs zijn alle gevolgen (inclusief doorwerkingen) van het project of de maatregel. Uiteraard gaat het om de beoogde positieve effecten, maar de positieve en negatieve neveneffecten dienen ook te worden meegerekend. Uit de figuur blijkt de verbluffende eenvoud van de kosten-batenanalyse als concept. In de praktijk van de kosten-batenanalyse doen zich echter lastige problemen voor:

- *Referentiesituatie.* In de vorm van de nulsituatie is deze soms niet of slechts onder extreme veronderstellingen voorstelbaar. Het uitblijven van additionele maatregelen voor waterberging in Nederland levert een referentiesituatie (nulsituatie) op die maatschappelijk gezien dermate onwenselijk is (ondergelopen steden e.d.) dat het zinniger is alternatieven niet ten opzichte van deze referentiesituatie te analyseren, maar ten opzichte van elkaar. Daarmee wordt wel afbreuk gedaan aan het uitgangspunt dat de kosten-batenanalyse een verschilanalyse is.
- *Afbakening van het relevante domein.* Het domein voor maatschappelijke kosten-batenanalyse wordt uit praktisch oogpunt meestal afgebakend via de landsgrenzen. De gevolgen van maatregelen of projecten die zich voordoen buiten die grenzen worden niet expliciet opgenomen in de analyse.
- *Afbakening in termen van doorwerking.* Het te analyseren project of beleidsmaatregel heeft een aantal directe consequenties tot gevolg. Dit zijn de gevolgen die zo goed mogelijk gemonetariseerd in de analyse worden opgenomen. De consequenties van het project of beleidsmaatregel werken echter ook weer door (de indirecte effecten). En zo ontstaat een keten aan consequenties. Hoe verder in de keten des te groter is de kans op interveniërende variabelen en des

te lastiger om een gevolg eenduidig aan het project of maatregel te verbinden. De stelregel om de afkapping voor het meenemen van gevolgen daar te leggen, waar een gevolg niet meer in staat is de rangorde van alternatieven te beïnvloeden, lijkt een heldere, maar negeert de (toenemende) onzekerheden in de analyse.

- *Rubricering van baten en kosten.* De wijze van rubricering van baten en kosten kan invloed hebben op de uitkomst van de analyse. Wordt gerekend met de baten-kostenverhouding als selectiecriteria, dan maakt het uit of men de met het project of maatregel negatieve externe effecten rekent tot de offers die men zich moet getroosten om het project of de maatregel te realiseren, of dat men die opvat als de logisch met het project of maatregel samenhangende consequenties.
- *Waarderingsproblemen.* De toepassingsmogelijkheden van de kosten-batenanalyse worden grotendeels bepaald door de mate waarin de effecten op bevredigende wijze in geld uit te drukken zijn. De niet in geld te waarderen effecten kunnen immers niet in een selectiecriteria als de baten-kostenverhouding tot uitdrukking worden gebracht. Er werd en wordt naarstig gezocht om moeilijk te waarderen effecten toch expliciet in de analyse te betrekken. Voorbeelden daarvan zijn de schaduwprojectmethode en de contingent valuation method.
- *Beschikbaarheid en tijdigheid van informatie.* Ten behoeve van de opstelling van een kosten-batenanalyse is, zelfs voor betrekkelijk geringe complexe vraagstukken, veel informatie nodig. Het betreft informatie voor het in beeld brengen van die aspecten van de nul-situatie die door het project of de maatregel worden beïnvloed, gegevens voor de inhoud en vorm van de alternatieve oplossingen, gegevens voor het in beeld brengen van de doorwerkingen, et cetera.
- *Volgtijdelijkheid van potentiële oplossingen (alternatieven).* De kosten-batenanalyse steunt op de gedachte dat alternatieven zich tegelijkertijd aandienen. In de praktijk is dit, zeker in wat complexere situaties, een minder realistische veronderstelling. De besluitvorming rond de Oosterschelde afsluiting is een goed voorbeeld van zich volgtijdelijk aandienende alternatieven. Ook hier zijn alternatieve methodische benaderingen ontworpen. Een voorbeeld hiervan is de sequentiële investeringsanalyse (SIA), waarbij een stapsgewijze analyse van alternatieven wordt gehanteerd (Hakvoort & Klaassen, 2004).

Bovenstaande problemen moeten bij een concrete toepassing van een kosten-batenanalyse (ad hoc) worden opgelost. Dat betekent dat in de praktijk lastige keuzes moeten worden gemaakt. Een niet onbelangrijk neveneffect is dat daarmee ook de methode zelf in discussie kan komen, juist vanwege de gemaakte keuzes. Voor- en tegenstanders van het project of de maatregel kunnen daar op strategische manieren gebruik van maken. Dit is onvermijdelijk, maar wetenschappelijk gezien onvoldoende om in alle gevallen afstand te nemen van het instrument kosten-batenanalyse.

De gebreken van de kosten-batenanalyse hebben geleid tot twee ontwikkelingen. Enerzijds is geprobeerd om de gebreken te repareren met behoud van de kosten-batenanalyse, anderzijds heeft men allerlei nieuwe methoden en technieken ontwikkeld, waarvoor die bezwaren niet of in mindere mate gelden. Bij het eerste moet bijvoorbeeld gedacht worden aan methodisch-technische verfijningen om

milieu-effecten beter in een kosten-batenanalyse op te nemen.<sup>16</sup> Bij het tweede moet men denken aan de ontwikkeling van niet-monetaire methoden zoals de scorecard methode of de multicriteriatechnieken. Bij de scorecard methode worden alle effecten in hun natuurlijke eenheden op overzichtelijke wijze in een tabel weergegeven. Op die manier ontstaat een beeld van de sterke en minder sterke kanten van een alternatief. Bij multicriteriamethoden worden de effecten ook in natuurlijke eenheden in een tabel weergegeven. Door de effecten te wegen met gewichten ontstaan scores voor de alternatieven. Op basis van de scores kan een rangordening van de alternatieven worden gemaakt. Behalve dat de methodische problemen die inherent zijn aan het waarderen van effecten op geld hebben geleid tot de ontwikkeling van de scorecard methode en de verschillende multicriteriatechnieken, geldt ook dat ze een positieve impuls hebben gegeven tot de ontwikkeling van andere vormen van beleidsgericht onderzoek zoals effectenonderzoek (policy impact analysis) waaronder de milieu-effectrapportage.<sup>17</sup> Op het fenomeen effectrapportages wordt in de volgende paragraaf nader ingegaan.

## 5. Beleidsanalyse als effectrapportage: “output” versus “outcome”

Een kosten-batenanalyse is gebaseerd op het waarderen van effecten op geld. Alleen wanneer effecten op bevredigende wijze in geld kunnen worden uitgedrukt, kunnen ze worden meegenomen in een kosten-batenanalyse. Er is dus duidelijk sprake van een economische oriëntatie. Het is vooral deze oriëntatie geweest, die er toe heeft geleid dat bepaalde categorieën effecten in evaluatiestudies onderbelicht bleven. De noodzaak om de effecten in geld te vertalen, leverden dikwijls moeilijk oplosbare problemen op met betrekking tot het kwantificeren en waarderen van effecten. De analyse werd om die reden veelal (impliciet) beperkt tot de gemakkelijk op te nemen effecten. De categorie niet of moeilijk te waarderen effecten behoeft echter niet samen te vallen met de categorie minst relevante effecten. Zo bleven milieu-effecten – anders dan via de bekende “pro-memorie post (p.m.-post)” – voor een belangrijk deel onzichtbaar in evaluatiestudies, terwijl men tegelijkertijd beseftte dat deze van wezenlijk belang zijn voor een deugdelijke afweging.

De perceptie dat in evaluatiestudies te veel de nadruk was komen te liggen op de op geld waardeerbare effecten, heeft geleid tot wat we eerder noemden ‘*een emancipatiebeweging binnen de evaluatiecultuur*’ (Klaassen & Van Nispen, 1996). Door separate behandeling van moeilijk in evaluatiestudies op te nemen effecten of effecten die anderszins buiten beeld waren gebleven, is getracht het evenwicht in de behandeling van de gevolgen van beleidsmaatregelen te herstellen. Men kan dit opvatten als de bestaansreden voor effectrapportages. Methodisch gezien kan de

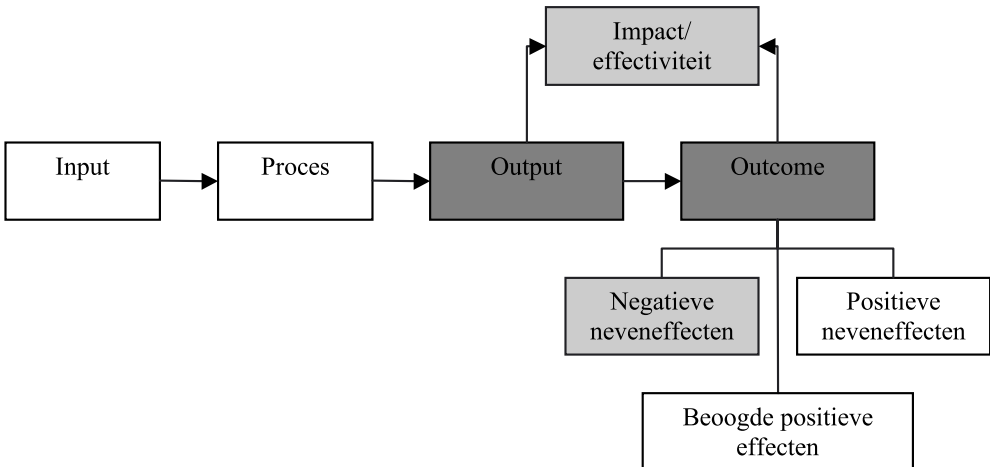
---

<sup>16</sup> Een goed voorbeeld hiervan is de schaduwprojectmethode. Dit behelst dat men ter compensatie van milieu-aantasting door een project een nieuw (milieu)project ontwerpt dat de negatieve milieugevolgen van het eerste project teniet doet. Het verlies van zelfreinigend vermogen van open water door bijvoorbeeld een inpoldering, kan men gedeeltelijk compenseren door de aanleg van een waterzuiveringsinstallatie

<sup>17</sup> Het onderscheid tussen een effectrapportage en effectenonderzoek hangt samen met het aantal effecten dat wordt meegenomen. Een effectrapportage beperkt zich tot één bepaald gevolg van een of meer beleidsmaatregelen richt, bijvoorbeeld de gevolgen voor het milieu of de veiligheid. Een effectenonderzoek is breder en richt zich in beginsel op alle gevolgen van beleidsmaatregelen (Klaassen & Van Nispen, 1995).

werkwijze bij effectrapportage als volgt in beeld worden gebracht:

**Figuur 1.3: Het concept effectrapportage**



Het belang van effectrapportages is met name gelegen in het in kaart brengen van de mogelijke gevolgen van beleidsmaatregelen, en dan wel in het bijzonder gericht op het in beeld brengen van één categorie effecten. Methodisch gezien maakt het uit of er sprake is van één alternatief dan wel meerdere alternatieven. Elke opzet waarbij meer dan één alternatief in beschouwing wordt genomen, biedt de mogelijkheid tot onderlinge vergelijking. De effectscores van een alternatief kunnen dienen als referentiepunt voor de beoordeling van de andere. Kwalificaties als meer/minder, beter/slechter hebben betekenis bij dergelijke analyses in de zin dat de alternatieven daardoor ten opzichte van elkaar kunnen worden gepositioneerd. Uiteraard vervalt deze mogelijkheid indien nog slechts één alternatief in het onderzoek wordt betrokken. De effectscores dienen dan in “absolute” zin te worden gewaardeerd, dat wil zeggen, er moet worden geoordeeld over waarden die het alternatief bewerkstelligt.

Een effectenonderzoek kan een opstap vormen voor een maatschappelijke kosten-batenanalyse, een multicriteria-analyse of een “score-card” (en vormen dan samen een evaluatie ex ante)<sup>18</sup>. De koppeling van effectenonderzoek en ex ante evaluatiemethoden is recent onderstreept door Tennekes en Hornis (2008) in hun leidraad voor ruimtelijke effectenanalyse van EU-beleid.

Een nuancering is wel op zijn plaats. In een aantal gevallen kan worden volstaan met een partiële analyse, omdat een integraal effectenonderzoek niet relevant is voor het onderhavige probleem en/of de baten niet tegen de kosten van een dergelijk effectenonderzoek opwegen. Wij zullen ons voortdurend moeten afvragen of een

<sup>18</sup> De score-card is ontwikkeld door de Rand Corporation in Santa Monica ten behoeve van de Oosterscheldestudie. Hij dient zorgvuldig te worden onderscheiden van de balanced score card. De balanced scorecard is een methode die gericht is op de bedrijfsvoering van publieke of private ondernemingen. Nieuwe initiatieven (projecten) van een organisatie worden beoordeeld op 4 perspectieven: het financieel perspectief, het klantperspectief, het perspectief interne bedrijfsvoering en het perspectief lerend vermogen.

integraal effectenonderzoek wel nodig is, hetgeen –naar analogie– wel de “kosten-batenanalyse van de kosten-batenanalyse” wordt genoemd. Verder dient te worden bedacht dat een volledig integraal effectenonderzoek niet mogelijk is, al was het maar omdat de gevolgen van een maatregel nu eenmaal niet allemaal van tevoren zijn te voorspellen.

Methodische problemen bij effectrapportages en effectenonderzoek doen zich met name voor bij het bepalen van de causaliteit. Het optreden van een effect na het realiseren van een project of na het nemen van een beleidsmaatregel behoeft niet oorzakelijk samen te hangen. Ook het niet optreden van een effect kan impliceren dat het project of de maatregel ineffectief is. Immers gelijktijdig veranderende omstandigheden of gelijktijdig optredende effecten van andere maatregelen of projecten kunnen het oorspronkelijke effect mitigeren.

Bij een eerdere gelegenheid hebben wij al gesproken over de ongebreidelde groei die zich heeft voorgedaan in de diverse verschijningsvormen van effectrapportages. Behalve de milieueffectrapportage spreken we nu over bestuurseffectrapportage, consumenteneffect-rapportage, economische effectrapportage, rechtsplegingseffectrapportage, veiligheidseffect-rapportage, vrouweneffectrapportage en wapenbeheersings-effectrapportage. In dat verband spraken we over een wildgroei van effectrapportages (Klaassen & van Nispen, 1996). Elke belangengroep verzekert zich van analytische ondersteuning, zodat elke expertise met contra-expertise kan worden bestreden (Goemans, 1985). Zo heeft de aandacht van economische overwegingen geleid tot een emancipatie van de niet-economische overwegingen in de vorm van milieu-effectrapportage. Het resultaat is een amalgaam van effectrapportages, hetgeen een zware wissel trekt op het integratievermogen van de overheid, omdat niet alleen wordt uitgegaan van andere probleemdefinities en daardoor andere oplossingsrichtingen, maar soms ook andere uitgangspunten worden gehanteerd. Zij maken daarnaast gebruik van verschillende van andere methoden van onderzoek en meetinstrumenten (Van Hoesel & Mevissen, 2005). De kunst is om die vervolgens weer op een noemer te krijgen.

De proliferatie van beleidsanalyse heeft daarnaast geleid tot een ingrijpende verandering van het speelveld (Lindblom & Cohen, 1979). De ervaring leert dat een effectrapportage alleen wordt opgesteld wanneer er sprake is van een duidelijk aanwijsbare groep, die de voor- dan wel de nadelen van een maatregel ondervindt. Het is bovendien niet ongebruikelijk om gebruik te maken van de afwezigheid van een duidelijk aanwijsbare groep, bijvoorbeeld omdat de effecten zijn gespreid over de gehele bevolking of omdat er sprake is van een toekomstige generatie, om de handen voor een beleid op elkaar te krijgen. Het is dan ook niet verwonderlijk dat de resultaten van een effectrapportage zich dikwijls kenmerken door een zekere vooringenomenheid. Zo wordt in een milieu-effectrapportage zelden aandacht besteed aan de positieve gevolgen voor het milieu. De negatieve gevolgen worden echter breed uitgemeten. Een effectrapportage krijgt hierdoor het karakter van “policy advocacy”, hetgeen overigens –zoals Wildavsky terecht heeft opgemerkt– niet per se negatief behoeft te zijn.

Ten slotte bestaat het gevaar dat de proliferatie van beleidsonderzoek leidt ook tot een politisering van het gebruik van de resultaten van effectenonderzoek. De kans



mag niet worden uitgesloten dat zij leiden tot een *'battle of analyses'* (Goemans, 1988) ook wel *'dueling swords'* genoemd (Radin, 2000), waarbij de partiële analyses als wapen worden gebruikt. *'We are a field with multiple languages, values, and forms and with multiple individuals and groups as clients for the analyst. We confront conflict between analysts who strive for objectivity and those who begin with an advocacy posture. Often our work involves dueling between analysts'* (Radin, 2000, 51). De consequenties van de proliferatie van beleidsanalyse laten zich goed illustreren aan de hand van de gang van zaken rond de uitbreiding van de luchthaven Schiphol die door Van Eeten eerder is getypeerd als een *'dialogue of the deaf'* (Van Eeten, 1999). In de loop der tijd zijn talloze rapporten uitgebracht over de mogelijke consequenties van een uitbreiding van Schiphol voor het milieu. De probleemdefinitie en daarmee de mogelijke oplossingen van de verschillende belanghebbenden liepen niet alleen uiteen, maar zij hanteerden ook andere veronderstellingen. De strijdende partijen hadden bovendien weinig vertrouwen in "elkaars" cijfers en gegevens waardoor veel negatieve energie is gestoken in het bestrijden van elkaars standpunten (Van Nispen, 2000).

## 6. Beleidsanalyse als performance measurement: "input" versus "output/outcome"

Het gedachtegoed van beleidsanalyse heeft een nieuwe impuls gekregen door de zgn. "performance movement" welke onder andere is gericht sturing op resultaten: *'funding outcomes, not inputs'* (Osbourne & Gaebler, 1993, 138). De kern van sturing op resultaten is gelegen in een koppeling van middelen aan prestaties (producten) en aan de gewenste maatschappelijke effecten. De ratio tussen input en output (maat voor efficiëntie), respectievelijk tussen output en outcome (maat voor effectiviteit), vormt het ijkpunt voor sturen op resultaat.

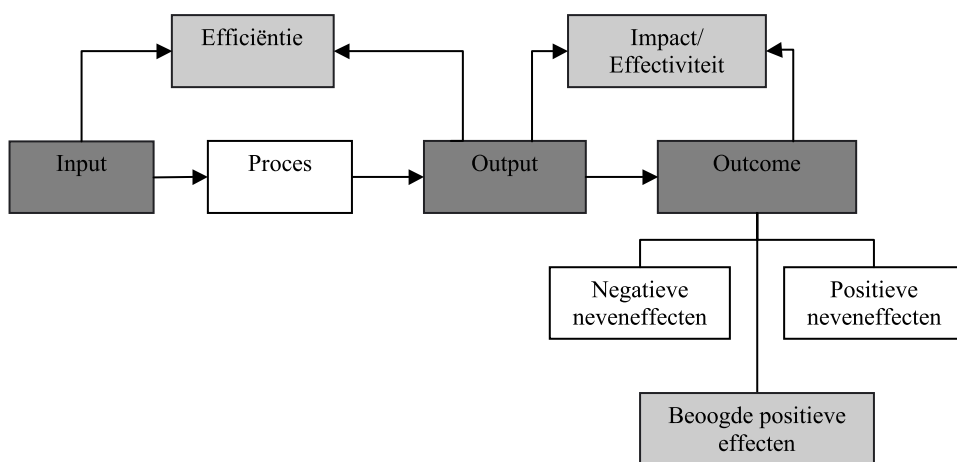
De focus op prestaties en effecten is, aldus Bouckaert & Halligan, internationaal gezien al zo'n honderd jaar oud. Zij onderscheiden drie typen van performance. In de beginperiode had performance management de vorm van productiviteitsmeting. Zij noemen dat *'performance administration'*. Kern van productiviteitsmeting is het kunnen leggen van een verbinding tussen een toename in output en een toename in input. Productiviteitsmetingen waren zeker in de beginperiode te kenschetsen als een mechanistische benadering (cf. scientific management denken). In de loop van de tijd is het mechanistische model vervangen door een model met een lossere koppeling tussen in management van de organisatie en het bereiken van prestaties. Kwaliteit, ambiguïteit, aandacht voor outputs zonder de directe link te leggen met inputs zijn kenmerkend voor wat Boeckeaert & Halligan *'management of performance'* hebben genoemd. Ten slotte onderscheiden zij nog *'performance management'* waarbij een organisatie open staat voor ijkpunten van derden (benchmarks), ratio's tussen input, proces, output en outcome en voor indicatoren naast technische, ook functionele en intern gelegitimeerde standaards hanteert (Bouckaert & Halligan, 2008).

Methodische problemen op het terrein van performance management zijn, als het



gaat om de relatie tussen middelen en outcome overeenkomstig de problemen bij effectrapportages. Het gaat met name om de causaliteitsvraag tussen middelen en effecten. In figuur 1.4 is de gebruikelijke voortbrengingsketen (input, proces, output) weergegeven van producerende eenheden. De relatie tussen indicatoren over het proces en de inputs voor dat proces geeft aan hoe zuinig men met de inputs omgaat. De relatie tussen de inputs en de gewenste output geeft een indicatie van de beoogde efficiëntie. De kosteneffectiviteit wordt gevormd door de verhouding tussen de output en de effecten die uiteindelijk worden beoogd.<sup>19</sup>

**Figuur 1.4: Het concept performance management**



In een vergelijking van prestatiebesturing tussen landen plaatsen Boukaert en Halligan de Anglo-Saksische landen –Australië, Canada, Groot-Brittannië en de Verenigde Staten– in de categorie performance management. Nederland en Zweden plaatsen zij in de categorie management van performance, hetgeen overeenkomt met een in 2006 verricht vergelijkend onderzoek onder 17 noordwest Europese gemeenten (Klaassen, 2007). Een dergelijke conclusie trekken Van Nispen en Posseth voor Nederland in hun onderzoek naar budgetary reform op rijksniveau.<sup>20</sup> Er is in hun ogen veeleer sprake is van ‘*performance-informed budgeting*’ dan van ‘*performance budgeting*’ (Van Nispen & Posseth, 2006; Curristine, 2007). Als argument wordt aangevoerd dat het niet realistisch is om te verwachten dat een directe relatie kan worden gelegd tussen inputs (middelen) en outputs (prestaties) omdat deze veelal heterogeen zijn. Het verband tussen inputs en outcomes (effecten) is nog lastig vast

<sup>19</sup> Uiteraard is het denkbaar dit figuur ook te gebruiken voor ex post doeleinden (verantwoordingsdoeleinden). Het gaat dan om daadwerkelijk gebruikte inputs, gerealiseerde output en gerealiseerde outcome. In dit artikel beperken wij ons tot de ex ante varianten van de verschillende analysevormen.

<sup>20</sup> Sturing op resultaat is op dit moment uiterst populair. In Nederland vindt men dit op alle niveaus overheidsbemoedening terug. Op departementaal niveau moet bij begrotingen zo concreet mogelijk worden aangegeven wat er wordt nagestreefd, welke middelen men daarvoor inzet, wat de kosten van deze inzet zijn, en welke concrete beleidsprestaties moeten worden geleverd (de huidige VBTB). Op provinciaal en gemeentelijk niveau is hetzelfde gedachtegoed ook omarmd. Ook de publieke uitvoeringsorganisaties besteden op diverse niveaus aandacht aan resultaatsturing (Van Lindenberg e.a, 2008).

te stellen omdat allerlei interveniërende variabelen een rol spelen. Een conclusie van hun onderzoek luidt dan ook dat “performance budgeting” zich enkel leent voor dat deel van de begroting waarvoor geldt dat sprake is van homogene activiteiten. In het geval van heterogene activiteiten lijkt “performance informed budgeting” het maximaal haalbare, hetgeen overigens een majeure stap voorwaarts is ten opzichte van de traditionele “line-item” begroting (Van Nispen & Posseth, 2006).

De ervaring die met de VBTB is opgedaan leert dat het vrijwel onmogelijk is om de doelmatigheid van het beleid te bepalen, omdat de effecten niet of nauwelijks zijn te meten.<sup>21</sup> De commissie die een tussentijdse evaluatie heeft uitgevoerd, pleit onder andere voor een andere strategie, waarin sprake is van (de programmering van) beleidsanalyse, respectievelijk evaluatieonderzoek. Zij merkt daarbij op dat een beleidsanalyse zich niet dient te beperken tot een beschrijving van bestaand beleid, maar ook een synthese dient te bevatten van het beschikbare evaluatieonderzoek naar de effecten van het beleid. Een evaluatieonderzoek maakt derhalve deel uit van een beleidsanalyse. Een beleidsanalyse gaat over in een interdepartementaal beleidsonderzoek wanneer op grond daarvan mogelijke alternatieven worden ontwikkeld voor het tot dan toe gevoerde beleid (TK 29 949, nr. 1).<sup>22</sup>

## 7. De veranderende rol van de beleidsanalyse, i.c. de beleidsanalist

In het voorafgaande is beleidsanalyse omschreven als beleidsgericht onderzoek. Het streven moet er op zijn gericht beleidsrelevante informatie te genereren. De doorwerking van de resultaten van beleidsgericht onderzoek is, aldus Hoppe in navolging van Lindblom & Weiss,<sup>23</sup> beperkt: *‘De invloed van professionele beleidsanalyse is gering en voegt weinig toe aan de praktijkkennis van politici, bestuurders en ambtenaren. Het enige wat beleidsanalisten rest, is hun onderzoek en adviezen zo goed mogelijk afstemmen op de praktijk (‘research as argument’), af en toe – bij toeval gekruid – met een briljant idee (‘enlightment’)’* (Hoppe, 1988, 25).

De doorwerking van de resultaten van beleidsgericht onderzoek geschiedt meestal indirect en vindt plaats op de lange termijn. Zij verhelderen de situatie of worden gebruikt als munitie in de beleidsarena, zoals Weiss reeds lang geleden heeft vastgesteld: *‘It becomes ammunition for the side that finds its conclusions congenial and supportive. Partisanans flourish the evidence in an attempt to neutralize opponents, convince waverers, and bolsters supporters. Even if conclusions have to be ripped out of context (with suppression of qualifications and of the evidence ‘on the other*

---

<sup>21</sup> In het kader van de VBTB wordt een ruime definitie van doelmatigheid, i.c. de ratio tussen input en outcome, gehanteerd.

<sup>22</sup> De doorlichting van beleid geschiedt aan de hand van een vragenlijst van 10 vragen (TK 29 949, nr. 1, 42-43), welke vergeleken kan worden met de beleidsanalytische toets die de Raad van State hanteert (RvS 2003, Bijlage 4)

<sup>23</sup> De toon is gezet door Carol Weiss. Zij heeft zes vormen van gebruik van (de resultaten van) onderzoek onderscheiden (Weiss 1979). Zij komt tot de conclusie dat knowledge creeps (Weiss, 1979; Weiss, 1980). Zie ook de recente literatuur op het gebied van ‘boundary arrangements’ (Wittrock, 1991; Hoppe, 2005).

*hand'*) *research becomes grist to the mill'* (Weiss, 1979, 429). In dat verband dient te worden bedacht dat de uitkomst van een beleidsanalyse moet concurreren met informatie uit andere bronnen. De informatie fungeert voornamelijk als argument, dan wel als contra-argument in de politieke arena (Van Nispen, 2003).

In de rol van de beleidsanalist treedt als gevolg van de proliferatie van beleidsanalyse een verandering op. Hij is niet langer de objectieve, neutrale onderzoeker, maar een geëngageerd meedenker die als “hired gun” fungeert om iedere gewenste positie te verdedigen (bijvoorbeeld het belang van de opdrachtgever). De resultaten van onderzoek worden gebruikt voor de onderbouwing van een reeds vooringenomen standpunt, dan wel de ondermijning van het standpunt van een tegenstander. De vraag in hoeverre hij er in slaagt om een bijdrage te leveren aan de oplossing voor een maatschappelijk probleem is moeilijk te beantwoorden (Post & Schmidt, 2005). Actoren proberen de resultaten van onderzoeken te gebruiken om hun eigen machtspositie te versterken. De adviezen zijn onderdeel van de politieke strijd, en staan er niet “boven”, zoals min of meer wordt verondersteld in het rationele model (Bekkers e.a., 2004).

## 8. Vergelijking

In het voorafgaande hebben wij aandacht besteed aan verschillende invalshoeken voor beleidsanalyses: de klassieke kosten-batenanalyse, effectrapportages en de performance-meting. Hoewel er methodologisch grote verschillen zijn aan te wijzen, zoals ook al bij de behandeling van de verschillende vormen naar voren is gekomen, beschouwen wij ze toch alle drie lid van dezelfde familie. Zij kunnen als volgt kort worden getypeerd:

Benadering	Domein	Focus
Kosten-batenanalyse	Input vs outcome	Monetaire effecten
Effectrapportages	Output vs outcome	Negatieve neveneffecten
Performance management	Input vs output/outcome	Positive beoogde effecten

Een onderlinge vergelijking leert dat zij alle drie hun eigen vertekening kennen. Een kosten-batenanalyse bestrijkt in beginsel het hele spectrum van input tot outcome, maar beperkt zich tot de monetaire effecten. Niet op geld te waarden effecten worden als “p.m.-post” opgevoerd. Een “p.m.-post” is echter niet gelijk aan nul. In het gunstigste geval wordt een schaduwprijs opgevoerd. Een effectrapportage beperkt voornamelijk tot de relatie tussen output en outcome –het rechter gedeelte van het basisschema– en in het bijzonder tot de negatieve neveneffecten. Het streven van performance management is gericht op een allocatie van middelen die zo gunstig mogelijk de gewenste effecten genereren. De neveneffecten blijven buiten beschouwing. Performance management richt zich echter voornamelijk tot de relatie tussen input en output. Het risico bestaat daarnaast dat performance management

perverse effecten oproept bijvoorbeeld omdat alleen wordt gemeten wat er valt te meten. Wat niet valt te meten wordt buiten beschouwing gelaten. In dat verband wordt ook wel gesproken over de “performance paradox” (De Bruijn, 2001; Van Thiel & Leeuw, 2002).

### *Kosten-batenanalyse versus effectrapportages*

Het belangrijkste verschil tussen kosten-batenanalyse en een effectrapportage is dat een kosten-batenanalyse een integrale analyse is en een effectrapportage een partiële analyse. Vanuit de optiek van facilitering van beleidsvormingsprocessen, hebben beide een eigenstandige functie. Met een kosten-batenanalyse wordt gepoogd een bijdrage te leveren aan het proces van afweging van alternatieven: welke optie werkt maatschappelijk gezien het beste uit. Bedacht moet worden het een economische studie is. Verdelingsaspecten (hoe werkt het uit voor verschillende landsdelen, regio's, steden, wijken, bevolkingsgroepen) blijven buiten beschouwing. Budgettaire consequenties maken geen onderdeel uit van een kosten-batenanalyse. En strikt genomen is ook het probleemoplossend vermogen geen aspect in een kosten-batenanalyse. Het gaat bij een kosten-batenanalyse immers om het bepalen van een alternatief dat maatschappelijk gezien het meeste bijdraagt. Deze kanttekeningen maken duidelijk dat een goed uitgevoerde kosten-batenanalyse bij kan dragen aan een afweging, maar niet in de plaats kan treden van de afweging zelf. Dit lijkt nog al eens te worden vergeten in de debatten over de aanleg van nieuwe infrastructuur, zoals blijkt uit de verhitte debatten die in Nederland zijn gevoerd over de aanleg van nieuwe railinfrastructuur naar het noorden van het land. Een maatschappelijke kosten-batenanalyse voor de railverbinding tussen de Randstad en het noorden van het land toonde een laag rendement, Tegenstanders betoogden daarop dat de railverbinding niet moest worden aangelegd. Voorstanders betoogden dat verdelingsaspecten ten onrechte niet in de analyse waren meegenomen. En daarmee bleven de ontwikkelingseffecten voor het noorden van het land in hun ogen ten onrechte buiten de afweging

In tegenstelling tot een kosten-batenanalyse wordt met een effectrapportage getracht de aandacht te vestigen op negatieve neveneffecten van projecten of maatregelen. Een groot voordeel van een effectrapportage kan zijn dat het leidt tot diepgaande inzichten in de gevolgen van een project of maatregel. Dit voordeel is tegelijkertijd het nadeel. De neiging zal bestaan om vanuit de optiek van de in beschouwing genomen effecten een oordeel te vellen, hetgeen andere effectrapportages kan uitlokken. Het is vanuit deze waarneming dat wij in het verleden een pleidooi hebben gehouden voor integrale analyses.

### *Kosten-batenanalyse versus performance management*

Bij performance management ligt de nadruk op de te bereiken prestaties en beoogde effecten. De betekenis en vorm die performance management heeft gekregen komt met name tot uiting in de allocatiefunctie van begrotingen (ex ante) en later bij de verantwoording (ex post). Ontegenzeggelijk heeft performance management daarmee een zeer belangrijke plaats verworven in het proces van beleidsvorming. En performance management heeft in dat opzicht ook veel bijgedragen aan het verhelderen van begrotingdebatten. Daarin schuilt tevens het grote gevaar. Immers,

performance management is primair gericht op de beoogde effecten. Op zich is dat nog geen probleem, maar het lijkt er in veel gevallen op dat er simplificatie optreedt: lineaire verbanden worden verondersteld tussen de ingezette middelen en de gevolgen. Kortom de beleidstheorie ontbreekt in veel gevallen. Nog wezenlijker is dat negatieve neveneffecten buiten beschouwing blijven. Daarvoor is simpelweg geen plaats bij performance management. Ten slotte is er een verschil in tijdsdimensie. Een kosten-batenanalyse is gericht op een weging van offers en effecten over een bepaalde periode als gevolg van een alternatief. Bij performance management gaat het om min of meer routinematig handelen van een organisatie teneinde bepaalde effecten te bereiken, en die op basis van indicatoren over het voortbrengingsproces te optimaliseren.

### *Effectrapportages versus performancemeting*

Methodisch gezien is ligt bij zowel de effectrapportage als bij performancemeting de nadruk op output en outcome. Het verschil is voornamelijk gelegen in de effecten in beschouwing worden betrokken. In een effectrapportage ligt de nadruk op negatieve effecten als gevolg van handelen, terwijl het bij prestatiemeting juist gaat om de beoogde effecten. Methodisch gezien is dit verschil echter niet principieel. Het gaat in beide gevallen om effecten als gevolg van overheidsoptreden. In beide gevallen speelt de causaliteitsvraag. Bij performancemeting wordt de link gelegd tussen de inputs en de output en outcome, zodat zicht ontstaat op de efficiëntie respectievelijk de kosteneffectiviteit van de voortbrenging. Bij effectrapportages gaat het veeleer om de verbinding tussen een alternatief en de (mogelijk negatieve) gevolgen daarvan. De causaliteitsvraag speelt bij beide: is de outcome een direct gevolg van het handelen.

## **9. Conclusie**

In deze bijdrage is duidelijk gemaakt dat het beleidsanalytisch instrumentarium in de loop van de tijd is uitgebreid. En daarmee is ook betekenis en reikwijdte veranderd. Ook is duidelijk gemaakt dat de rol van de beleidsanalist een andere is geworden. Zoals gesteld zien we twee veranderingen in de setting waarin beleidsanalyse plaatsvindt. Ten eerste de herkenning en erkenning dat de verschillende overheden deel zijn van een netwerk. De overheden zijn wat betreft de realisering van de ambities afhankelijk van andere spelers. Het is niet zo dat die afhankelijkheid noodzakelijkerwijs groter is geworden in de afgelopen decennia (dat vereist een afzonderlijk onderzoek), maar vormt wel een verklaring waarom er in de bloeitijd van de beleidsanalyse grote discrepanties konden ontstaan tussen de uitkomsten van beleidsanalytische studies (in het bijzonder de kosten-batenanalyse's) en de uitkomsten van het feitelijke besluitvormingsproces. De rol van de beleidsanalist verandert mee. Hij verricht nog steeds onderzoek, maar hij treedt meer op als belangenbehartiger dan als waarheidsvinder en neutraal dienaar van het algemeen belang. De tweede verandering die bij beleidsanalyse plaatsvindt, is het verschuiven van de inhoud van de informatie. De informatie die wordt gegenereerd heeft niet alleen meer betrekking op het keuzeprocess aangaande alternatieve beleidsmaatregelen

of projecten, maar gaat nu ook om het in beeld brengen van de performance van organisaties (de rol van performancemeting). Ook hier zien we een nieuwe rol voor de beleidsanalist.

Ten slotte dient te worden gewezen op het politiek gebruik van de resultaten van beleidsanalyse. Een kosten-batenanalyse kan bijvoorbeeld uitmonden in een dominant alternatief, dat desondanks niet wordt gerealiseerd omdat daarvoor onvoldoende politieke en maatschappelijke steun bestaat. In dat verband dient te worden bedacht dat de overheid slechts een van de vele spelers is in een netwerk van actoren. Zij zal de andere actoren moeten zien te overtuigen van haar gelijk –“speaking truth to power”– maar zelfs dan behoeft zij nog geen gelijk te krijgen. De uitslag wordt niet zozeer bepaald door de macht van de argumentatie, als wel door de macht van de sterkste. Een actor kan gelijk krijgen bij de gratie van ongelijk hebben (voor zover al valt vast te stellen of iemand gelijk heeft). De strijd is daarmee overigens nog niet gestreden, omdat daarmee voorbij zou worden gegaan aan eventuele perverse effecten. Actoren leren om te gaan met onwelgevallig beleid, waardoor het beleid als het ware wordt uitgehoud.<sup>24</sup> Wie niet sterk is, moet slim zijn!

---

<sup>24</sup> Het optreden van perverse effecten is door In 't Veld de “wet van de afnemende effectiviteit” gedoopt. Hij speelt haasje over met de “wet van de beleidsaccumulatie”, die er van uitgaat dat beleidsmakers op eenzelfde manier reageren op effectiviteit als op een gebrek aan effectiviteit.

## Bibliografie

Bekkers, V.J., H.J.M. Fengers, V.M. Homburg & Puffers, K. (2004). *Doorwerking van strategische beleidsadvisering*. Erasmus Universiteit Rotterdam/Universiteit Tilburg.

Börzel, Tanja A. (1997). What's So Special About Policy Networks? An Exploration of the Concept and Its Usefulness in Studying European Governance. In: *European Integration Online Papers*, Vol. 1, Issue 16.

Bouckaert, G. & Halligan, J. (2008). *Managing Performance. International Comparisons*. London/New York: Routledge.

Boorsma, P. B., Maessen, F.C. & Schild, J.A. (1999). Van beleidsbegroting tot beleidsverantwoording in relatie tot beleidsanalyse en prestatiebegroting: een historische vergelijking vanuit een beleidsanalytisch perspectief. In: *Beleidsanalyse*, 28 (1-2), 15-28.

Bruijn, H. de (2001). *Prestatiemeting in de publieke sector: tussen professie en verantwoording*. Utrecht: Lemma.

Buck Consultants International (2002). *Evaluatie OEEI-leidraad*. Nijmegen.

Curristine, T. (ed.) (2007). *Performance Budgeting in OECD Countries*. Paris: OECD.

Dunn, W.N. (1994, 2004). *Public Policy Analysis. An Introduction*. Upper Saddle River/New Jersey: Pearson Prentice Hall.

Eeten, M. van (1999). *Dialogues of the Deaf. Defining New Agendas for Environmental Deadlocks*. Delft: Eburon.

Fischer, F. (2003). *Reframing Public Policy. Discursive Politics and Deliberate Practices*. Oxford/New York: Oxford University Press.

Goemans, T. (1988). Beleidsanalyse op weg naar volwassenheid. In: Korsten, A.F. & Toonen, T.A. (red.). *Bestuurskunde: hoofdfiguren en kernthema's*. Leiden: Stenfert Kroese.

Hakvoort, J.L.M. & Klaassen, H.L. (2004). *Bedrijfsvoeringstechnieken voor overheid en non-profitorganisaties*. Den Haag: Sdu.

Hajer, M. & Wagenaar, H. (ed.) (2003). *Deliberate Policy Analysis. Understanding Governance in the Network Society*. Cambridge: Cambridge University Press.

Hoesel, P. van & Mevissen, J. (2005). Beleidsonderzoek onder vuur. In: van Hoesel, P.H., Leeuw, F.L. & Mevissen, J.W. (red.), *Beleidsonderzoek in Nederland. Kennis voor beleid; ontwikkeling van een professie*. Assen: Van Gorcum.

Hoppe, R. (1988). *De broosheid van debat en argumentatieve beleidsanalyse*. Inaugurale rede Enschede: Universiteit Twente.

Hoppe, R. (2005). Rethinking the science-policy nexus: From Knowledge Utilization and Science Technology Studies to Types of Boundary Arrangement, Poiesis & Praxis. In: *International Journal of Technology Assessment and Ethics of Science*, 3 (3), 199-215.

Jenkins-Smith, H. (1990). *Democratic Politics and Policy Analysis*. Pacific Grove: Brooks/Cole.

Kersbergen, K. van & Waarden, F. van (2004). Governance as a bridge between disciplines: Cross-disciplinary inspiration regarding shifts in governance and problems of governability, accountability and legitimacy. In: *European Journal of Political Research*, 43 (2), 143–171.

Klaassen, H.L. (2007). Begrotingen vergeleken. Een vergelijking van vorderingen, In: *Tijdschrift voor Public Governance (TPC), Audit & Control*, 5 (4), 12-17.

Klaassen, H. L. & van Nispen, F. K. (1994). *Drowned as Icarus? The Flight of Policy Analysis*. Paper at a conference on the Transformation of Values in a Pluriform Society, Erasmus Universiteit Rotterdam, Rotterdam, October 19, 1994.

Klaassen, H. L. & van Nispen, F. K. (1995). *Policy Analysis in the Netherlands*. George Mason University: Policy Analysis Center.

Klaassen, H. L. & van Nispen, F. K. (1996). De wildgroei van effectrapportages: een rapportage. In: *Bestuurskunde*, 5 (7), 308-316.

Klijn, E.H. (1996). *Regels en Sturing in Netwerken*. Delft:Eburon.

König, T. (1998). Introduction: Modeling Policy Networks. In: *Journal of Theoretical Politics*, 10 (4), 387-388.

Lindblom, C. E. & Cohen, D.K. (1979). *Usable Knowledge. Social Science and Social Problem Solving*. New Haven/London: Yale University Press.

Lindenberg, P. van, van den Burg, J. van 't Hoog, B., Yu, J. Kleijn, A. & Eising, B. (2008). *Sturen op resultaat. Op weg naar meer doelmatigheid en doeltreffendheid bij publieke uitvoeringsorganisaties*. Den Haag: Sdu Uitgevers.

Lynn, L. E. (1987). *Managing Public Policy*. Boston/Toronto: Little Brown and Company.

Mayer, I.S., van Daalen, C.E. & Bots, P.W. (2002). De verbeelding van de beleidsanalyse. In: *Beleidswetenschap*, 16 (2), 95-113.

Minister van Financien (2004). Evaluatie VBTB, *Tweede Kamer* 29 949, nr. 1.



Nispen, F. K. van (1993). *Het dossier Heroverweging*. Delft: Eburon.

Nispen, F.K. van (2003). Beleidsanalyse tussen idealisme en realisme. In: Bekkers, V. & Ringeling, A. (red.). *Vragen over beleid. Perspectieven op waardering*. Utrecht: Lemma.

Nispen, F.K van, & Posseth, J.J. (2006). Performance Budgeting in the Netherlands: Beyond Arithmetic. In: *OECD Journal on Budgeting*, 6(4), 1-26.

Obbink, N.S. & van Zwieten, M.T. (1993). *Institutionalisering van beleidsanalyse. Opkomst en ondergang van de COBA*. Rotterdam: Erasmus Universiteit Rotterdam.

Osbourne, D. & Gaebler, T. (1993). *Reinventing Government. How the Entrepreneurial Spirit is Transforming the Public Sector*. Harmondsworth (UK): Penguin Books.

Parsons, W. (1995). *Public Policy. An Introduction to the Theory and Practice of Policy Analysis*. Cheltenham: Edward Elgar.

Peterson, J. (2003). *Policy Networks*. Vienna: Institute for Advanced Studies.

Pollitt, C. & Bouckaert, G. (2004). *Public Management Reform. A Comparative Analysis* (2<sup>nd</sup> ed.). Oxford: Oxford University Press.

Post, H. M. & Schmidt, J. (2005). Geschiedenis van het beleidsonderzoek in de Verenigde Staten. In: van Hoesel, P.H., Leeuw, F.L. & Mevissen, J.W. (red.), *Beleidsonderzoek in Nederland. Kennis voor beleid; ontwikkeling van een professie*. Assen, Van Gorcum, 14-34.

Raad van State (2003). *Jaarverslag 2003*. 's-Gravenhage.

Radin, B. A. (2000). *Beyond Machiavelli. Policy Analysis Comes of Age*. Washington D.C.: George Washington Press.

Rhodes, R.A. (1997). *Understanding Governance*. Buckingham: Open University Press.

Stone, D. A. (2002). *Policy Paradox. The Art of Political Decision Making*. New York: Norton.

Tennekes, J & Hornis, W. (2008). *Ruimtelijke effectanalyse van EU-beleid: een Leidraad*. Beleidsadviezen 08. Den Haag: Ruimtelijk Planbureau.

Thiel, S. van & Leeuw, F. L. (2002). The Performance Paradox in the Public Sector. In: *Public Performance and Management Review*, 25 (3), 267-281.

Veld, R. J. in 't (1998). The Dynamics of Policy Instruments. In: Peters, B.G. & van Nispen, F.K. (ed.). *Public Policy Instruments. Evaluation the Tools of Public*

*Administration*. Cheltenham/Northampton: Edward Elgar, 153-161.

Weimer, D. L. & Vining, A.R. (1999). *Policy Analysis. Concepts and Practice* (3<sup>rd</sup> edition). Upper Saddle River/New Jersey: Prentice Hall.

Weiss, C. H. (1979). The many meanings of research utilization. In: *Public Administration Review*, 39 (5), 426-431.

Weiss, C. H. (1980). Knowledge creep and decision accretion. In: *Knowledge: Creation, Diffusion, Utilization*, 1 (3), 381-404.

White House (2008). *The President's Management Agenda: The Scorecard*, [www.whitehouse.gov/results/agenda/scorecard.html](http://www.whitehouse.gov/results/agenda/scorecard.html), laatst geraadpleegd op 24 oktober 2008.

Wildavsky, A. (1969). Rescuing Policy Analysis from PPBS. In: *Public Administration Review*, 29 (2), 189-202.

Wildavsky, A. (1979). *The Art and Craft of Policy Analysis: Speaking Truth to Power*. Boston: Little, Brown and Co.

Wittrock, B. (1991). Social Knowledge and Public Policy: Eight Models of Interaction. In: Wagner, P. (ed.). *Social Sciences and Modern States: National Experiences and Theoretical Crossroads*. Cambridge: Cambridge University Press, 28-85.

# **Hoofdstuk 2: Prestatiemeting: doel of middel? Over de niet-evidente meting van prestaties binnen de publieke sector**

Dries Verlet & Carl Devos

Studiedienst van de Vlaamse Regering & Universiteit Gent

<b>1. Inleiding</b> .....	<b>41</b>
<b>2. Meten van prestaties in de publieke sector</b> .....	<b>44</b>
2.1. Op welk niveau meten we de prestaties van de overheid? .....	44
2.2. Meten we perceptie of realiteit? .....	45
2.3. Welke waarde-indicatoren zijn er voor dienstverlening in de publieke sector? ...	46
2.4. Welke kwaliteitsnormen kunnen we gebruiken? .....	46
2.5. Zijn alle nodige gegevens beschikbaar? .....	46
<b>3. Effecten van prestatiemeting</b> .....	<b>47</b>
3.1. Situering en het goede aan prestatiemeting .....	48
3.2. Negatieve effecten van prestatiemeting .....	49
3.3. Een algemene verklaring van negatieve effecten en algemene strategieën om ze te voorkomen .....	55
<b>4. Besluit</b> .....	<b>57</b>
<b>Bibliografie</b> .....	<b>59</b>

## 1. Inleiding

Hoewel beleidsevaluatie veel meer omvat dan het in kaart brengen van de prestaties binnen de publieke sector, is deze prestatiemeting een centraal gegeven binnen het onderzoeks- en toepassingsveld. Met deze bijdrage willen we stilstaan bij de prestatiemeting in de publieke sector. Bijzondere aandacht gaat uit naar de potentiële onbedoelde effecten die een dergelijke meting kan hebben op het functioneren van een organisatie zelf. Hiermee willen we geenszins een afbreuk doen aan het belang van prestatiemeting. We willen enkel bijdragen tot een bedachtzame meting, die onontbeerlijk is indien we op een correcte manier o.a. de efficiëntie en effectiviteit van het gevoerde beleid in kaart willen brengen.

De grotere aandacht voor het evalueren van het overheidsoptreden past in de algemene evolutie van de publieke sector. Brans e.a. (2008) detecteren drie brede trends in de publieke sector: organisaties krijgen meer ruimte om prestaties te verbeteren omdat de overheid steeds verder terugtreedt, door de toegenomen autonomie is ook de aansprakelijkheid voor kwaliteit van de dienstverlening toegenomen en ten slotte is er de opkomende marktwerking. Daardoor is (informatie over) de prestatie van publieke dienstverleners steeds belangrijker.

De legitimering van deze toenemende zorg voor de kwaliteit van de publieke dienstverlening kreeg vooral een *boost* vanuit het NPM-denken. Er bestaan veel omschrijvingen van het Nieuw Publiek Management. Doorgaans staat het voor ‘*a management culture that emphasizes the centrality of the citizen or customer, as well as accountability for results. It also suggests structural or organizational choices that promote decentralized control through a wide variety of alternative service delivery mechanisms, including quasi-markets with public and private service providers competing for resources from policymakers and donors*’ (Manning, 2000). Een essentiële component van NPM zijn ‘*explicit standards and measures of performance*’ (Hood, 1996, 271). Gezien in regel bij overheidsprestaties geen directe marktprikkels spelen – waardoor slechte of te dure prestaties via dalende verkoop of inkomsten gesanctioneerd worden en bijsturing onvermijdelijk is – moeten de prestaties van de overheid onophoudelijk en uitvoerig geëvalueerd worden. Zo kunnen slechte of te dure prestaties bijgestuurd worden. De publieke sector kan daarbij zoveel mogelijk beroep doen op methoden van de private sectoren, al moet daarbij rekening gehouden worden met de eigenheid van de publieke sector.

Binnen het NPM kreeg prestatiemeting, vooral de meting van efficiëntie en effectiviteit van de dienstverlening, soms mythische proporties. Onder meer via de ‘*Balanced Scorecard*’ (cf. Kaplan & Norton, 1996), konden publieke sectoren leren om strategisch en doelgericht te werken. De toepassing binnen de publieke sector verloopt evenwel niet altijd rimpelloos (Modell, 2004).

Vanuit NMP wordt aangestuurd op een meer resultaatgericht overheidsmanagement. Daardoor krijgt de efficiëntie en effectiviteit van maatregelen veel meer aandacht. Dit vereist anderzijds dat doelstellingen duidelijk zijn en duidelijk vooropgesteld worden, dat de prestaties en de effectieve inzet van middelen met het oog op het bereiken van de vooropgestelde doelstellingen precies te meten zijn en dat er nadien gevolgen verbonden zijn aan de uitkomst van die evaluatie. Dat alles moet een organisatie in

staat stellen om primaire processen te optimaliseren, de niet-waardetoevoegende activiteiten te reduceren en de beschikbare capaciteit beter te benutten. In deze context is er vaak nood aan een monitoringsysteem met kwantitatieve en kwalitatieve indicatoren. De ontwikkeling van een dergelijk systeem moet bovendien bij voorkeur geïntegreerd te worden in de beleidscyclus.

NPM wordt stilaan gebruikt als een verzamelnaam voor opvattingen die ervan uitgaan dat de kwaliteit van de publieke dienstverlening gebaat is bij heldere doelen en afspraken, bij een scheiding van verantwoordelijkheden, meetbare prestaties en het afleggen van verantwoording.<sup>25</sup> Hoewel het niet zo aandoet, en zelfs precies het tegenovergestelde suggereert, is ook dit discours ideologisch geladen. Nogal vaak ligt het streven eraan ten grondslag om de publieke sector zoveel mogelijk als de private te laten functioneren. Vandaag is daarop een correctie gekomen, omdat velen beseffen dat de publieke sector haar eigenheid heeft en dat bij de blinde toepassing van het private evaluatiedenken allerlei ongewenste neveneffecten opduiken. Bovendien is het ook vrij moeilijk om de kwaliteit van publieke dienstverlening in kaart te brengen via efficiëntie- en effectiviteitsmetingen. In zijn analyse –op basis van de studie van prestatie-sturing, het publiceren van prestatiegegevens en persoonsgebonden budgetten– stelt Pestman evenwel vast dat die neveneffecten in belangrijke mate bestreden kunnen worden. Hij roept op om het ideologische debat te begraven en te focussen op de vraag in welke situaties en onder welke voorwaarden meetinstrumenten zinvol toe te passen zijn (Pestman, 2007).<sup>26</sup>

Er zijn vele redenen waarom men, los van NPM, ervoor kan pleiten om het overheidsbeleid meer en beter te evalueren. Een ervan is omdat beter overheidsbeleid ook het vertrouwen in de publieke dienstverlening zal versterken. Hoewel het empirisch materiaal ter zake slechts fragmentair is, zijn er belangrijke aanwijzingen dat het vooropgestelde doel van een verhoogd vertrouwen niet wordt bereikt, zelfs soms in tegendeel, indien niet zorgvuldig wordt omgegaan met de publicatie van prestatie-metingen (Hayes & Pidd, 2005).

Andere redenen voor meer prestatie-metingen liggen voor de hand. Het schaarse belastingsgeld moet bij een uitdeinende overheidsopdracht zo zinvol mogelijk aangewend worden, burgers hebben recht op de beste dienstverlening. De aandacht voor efficiëntie en effectiviteit van de publieke dienstverlening staat al een tijdje bovenaan de politieke agenda. Burgers en hun politieke vertegenwoordigers eisen daarom een maximale “*return on investment*”. Daarom is er politieke druk om meer aandacht te besteden aan het meten van overheidsbeleid. De burger-consument heeft recht op kwalitatieve dienstverlening.

Het meten van overheidsprestaties, waarover velen zich graag uitspreken, is evenwel geen evidente opdracht. Wat is, bijvoorbeeld, effectiviteit? Ruw en eenvoudig gesteld is effectiviteit de mate waarin de beleidsoutput doelstellingen –gewenste effecten (outcome)– bereikt, los van de manier waarop die effecten bereikt worden.

---

<sup>25</sup> Deze definitie is gebaseerd op het werk van Hood, 1991; Pollitt, 1993 en Kickert, 1993. Referentie gevonden in: Pestman (2007,104).

<sup>26</sup> Dit is natuurlijk op zich een ideologisch statement.

Dat betekent dat veel termen ingevuld en geïnterpreteerd moeten worden. Daardoor wordt “effectiviteit” een soort superwaarde, die meerdere andere waarden en indicatoren omvat (Jorgensen, 2006). Ook het streven naar “deugdelijk bestuur” laat vele invullingen toe, die niet los kunnen staan van normatieve kwesties (Verlet, 2008). Deze invullingen worden mede bepaald door het heersende politieke klimaat en economische inzichten, maar ook de bredere culturele setting.<sup>27</sup> “Good governance” is een sociale constructie (Edwards & Clough, 2005) die geen sterke basis heeft in empirisch onderzoek. Indicatoren voor *governance* lijken –aldus Van Roosbroek– in dit opzicht vooral beleidsinstrumenten, eerder dan academische oefeningen. Het is vooral een socio-politiek fenomeen (Van Roosbroek, 2007).

Dat neemt niet weg dat er veel studies gepubliceerd worden over overheidsprestaties, waaruit beleidsmakers gevolgen willen trekken. Daarom zijn er ook allerlei indicatoren en *rankings* ontstaan, die de prestaties van de ene overheidsdienst t.o.v. de andere stellen. *Benchmarking* is dan het logisch gevolg. Evalueren gaat immers over de vraag of er een gewenst verschil vast te stellen is als gevolg van een overheidsingreep. Dus moet bij voorkeur vergeleken worden tussen een oude (zonder de te evalueren overheidsingreep) en nieuwe situatie (met de te evalueren overheidsingreep). Vaak is het onmogelijk om gegevens over de situatie voorafgaand aan de overheidsingreep te verzamelen. Daarom wordt de nieuwe situatie vaak vergeleken met een vergelijkbare situatie elders, al is dat vergelijken vaak geen sinecure.

Hoe dergelijke internationale en binnenlandse *rankings* worden samengesteld is echter vaak onduidelijk. Van de Walle e.a. analyseerden vergelijkende studies. Hun verdict is duidelijk en ongenadig: de gebruikte indicatoren meten meestal slechts een vrij beperkt deel van het overheidsfunctioneren, percepties over dit functioneren moeten veelal doorgaan voor objectieve metingen van prestaties. De versnippering van de verantwoordelijkheid voor het verzamelen van data is een belangrijke reden voor het kwaliteitsdeficit van bestaande indicatoren. Daardoor is vergelijken problematisch. Ze wijzen dan ook op de nood aan goede databanken om de publieke sector in Europa te vergelijken en te monitoren (Van de Walle, Sterck, Van Dooren & Bouckaert, 2004). Problemen met algemene *rankings* is dat ze vaak al te veel subjectieve indicatoren bevatten, dat er weinig garanties zijn over de kwaliteit van de steekproeven en dat er te vaak ongepaste aggregaties plaatsvinden (Van de Walle, 2006).

Hoewel dergelijke vergelijkingen zinvol kunnen zijn, zeggen ze dus vaak weinig over de eigenlijke performantie van de publieke sector. Doelstellingen en context zijn vaak heel verschillend. Soms overbenadrukken ze enkele parameters, zoals het aantal ambtenaren, en kunnen ze de (kwaliteit van de) output/outcome van overheidsdiensten niet voldoende operationaliseren en dus vergelijken. De discussie over de prestaties van de publieke sector is evenwel een onvermijdelijke internationale discussie, die o.a. werd versterkt door de Lissabon Agenda. Tegen 2010 moet de EU een van de meest concurrerende economische gebieden zijn (Kuhry, 2004).

---

<sup>27</sup> Zo heeft Zondergeld-Hamer het over de rol van godsdienstige verwachtingen (Zondergeld-Hamer, 2007).

Een “performante overheid” is daar een belangrijk instrument toe. In Vlaanderen is het streven naar een efficiëntere overheid ook deel van de doelstellingen van “Vlaanderen in Actie”.

Beleid gaat om het veroorzaken van gewenste effecten, daar is niets nieuws aan. Maar recent is er wel veel meer aandacht voor de evaluatie van het overheidsbeleid. De aandacht voor de gevolgen van het meten van overheidsbeleid zijn eveneens niet nieuw. Over de perverse en onbedoelde effecten van het meten van overheidsprestaties schreef Ridgway reeds in 1956. Ook recenter werd daarover gerapporteerd. Smith (1995) toonde aan dat er consensus is over het feit dat prestatiemetingen ook ongewenste effecten kunnen hebben. Die ongewenste effecten hebben bovendien ook een kost, die bij het instellen van meetsystemen vaak over het hoofd worden gezien (Pidd, 2005). Maar de aandacht voor de onvoorziene gevolgen van beleidsevaluaties blijft beperkt. Verwacht wordt dat dit de komende jaren, precies door de toegenomen evaluering, zal veranderen. Het evaluatieproces zal zelf steeds meer geëvalueerd worden.

Strategische en operationele doelen helder vastleggen en communiceren, uitvoeren en registreren, periodiek rapporteren, prestaties evalueren en indien nodig bijsturen... Het klinkt goed en niemand zal deze algemene beginselen tegenspreken. Maar evaluatie vergt niet enkel heel wat energie, het kan ook ongewenste effecten opwekken. Zoals diensten die de meest begeleidbare doelgroepen naar gewenste effecten begeleiden en de moeilijkere groepen “opgeven”. Zo halen ze wel hun vereiste prestaties maar kan dit vanuit globaal beleidsoogpunt niet als optimaal worden omschreven.

In deze probleemschetstende bijdrage willen we ingaan op het fenomeen van het meten van overheidsprestaties. We richten de aandacht op geïventariseerde negatieve effecten van prestatiemeting en -evaluaties in de publieke sector. Dit is niet te lezen als een pleidooi om overheidsbeleid niet langer of minder te evalueren, maar wel als een kritische analyse van de analyse van overheidsbeleid. Doel van deze analyse is om, door rekening te houden met negatieve, ongewenste (neven)effecten van de prestatiemeting, de evaluatie van overheidsbeleid te verbeteren om zodoende tot beter overheidsbeleid te komen.

Deze bijdrage bestaat verder uit twee delen. In de tweede paragraaf staan we stil bij de uitdagingen van het meten van overheidsbeleid of -prestaties. In de derde en laatste paragraaf komt het hoofdonderwerp aan bod: welke positieve en negatieve effecten roept publieke prestatiemeting op? Hoe kunnen we de negatieve effecten verklaren en voorkomen?

## 2. Meten van prestaties in de publieke sector

Vooraleer in te gaan op de vraag welke ongewenste effecten het evalueren van overheidsbeleid of meten van overheidsprestaties kan hebben, gaan we eerst in op de bijzondere vragen die bij een dergelijke beleidsevaluatie moeten beantwoord worden. De antwoorden zijn immers niet zonder betekenis. Ze bepalen samen met het substantiële overheidsbeleid mee het evaluatieresultaat.

## 2.1. Op welk niveau meten we de prestaties van de overheid?

In essentie kunnen we het onderscheid maken tussen analyses van de prestaties van de overheid op drie niveaus: macro-, meso- of microniveau (Callens, 2007). Dit heeft betrekking op doelstelling van de analyse: willen we het productieproces van de overheid in haar geheel (macro) analyseren, in een specifieke overheidssector (sectoriëel) of een specifieke dienstverlening aan eindgebruikers (micro) in kaart brengen.

Het idee van een globale index lijkt heel interessant. Een dergelijke index laat immers toe om bijvoorbeeld de positie van Vlaanderen te vergelijken met een aantal landen met het oog op een onderlinge rangschikking. Callens (2007) maakt in zijn analyse melding van vier voorbeelden van dergelijke globale indexen: de Europese Centrale Bank, het Institute for Management Development, het World Economic Forum en de World Bank. Twee van deze “globale” studies werden al gerepliceerd voor Vlaanderen: ECB (Moesen, 2004) en WEF (Bowen, 2007 in voorlopige versie). Voor een kritische analyse van deze en andere indexen kunnen we eveneens verwijzen naar het werk van Luts e.a. (2008) in het kader van een onderzoek van het Steunpunt Bestuurlijke Organisatie Vlaanderen (SBOV).

Het probleem bij het maken van dergelijke algemene performantie-indicatoren is de aggregatie. De complexiteit van een overheid is veelal niet te herleiden in één enkele indicator. Zo'n algemene indicator laat ons bijvoorbeeld niet of onvoldoende toe om rekening te houden met de administratieve cultuur, de verschillen in de staatsstructuren, et cetera. Een specifiek probleem voor de Vlaamse regio is daarenboven dat vrij weinig data op het niveau van de regio's beschikbaar zijn. Enige haalbare manier vandaag is het gebruiken van indicatoren van gepercipieerde performantie teneinde verschillende overheden te vergelijken (Van de Walle & Bouckaert, 2007).

Bij een zogenaamde sectoriele studie vergelijkt men bijvoorbeeld de efficiëntie en de effectiviteit van een bepaalde sector in een land of regio met deze van die sector in andere landen of regio's. Schoolvoorbeeld in deze is het onderzoek van het Sociaal Cultureel Planbureau omtrent '*Prestaties van de Publieke Sector*' (Kuhry, 2004). Deze studie heeft als doel om verschillen in productiviteit, kwaliteit en doeltreffendheid van de door de overheid georganiseerde dienstverlening tussen Nederland en andere ontwikkelde landen op te sporen. Vier beleidsterreinen worden daarbij afzonderlijk bestudeerd: onderwijs, gezondheidszorg, criminaliteitsbestrijding en openbaar bestuur. Daarnaast is ook de OESO actief op het vlak van sectoriële studies. Dit is onder meer het geval wat betreft onderwijs, gezondheidszorg, openbare orde en veiligheid.<sup>28</sup>

In een micro-benadering worden specifieke publieke diensten binnen een land onderling vergeleken (Callens, 2007). Zo bestaan er o.m. vergelijkende studies

---

<sup>28</sup> Voor een overzicht kunnen we verwijzen naar (OECD, 2007, 38).



van brandweerdiensten, diensten burgerlijke stand, ziekenhuizen, scholen, gevangenissen en rechtbanken (bv. Bouckaert 1992; Bouckaert 1993; Blank 2000; referenties uit Callens, 2007). Hier kunnen we ook wijzen naar de studie van Van Dongen over kinderopvang die aan bod komt in hoofdstuk 7 van het boek.

## **2.2. Meten we perceptie of realiteit?**

Het probleem met meetindicatoren is dat ze vaak geen recht doen aan de complexiteit van maatschappelijke keuzes die ten grondslag liggen aan het beleid. Zo kan een doel van de publieke dienstverlening erin bestaan om aanwezig te zijn, om de dienst aan te bieden aan groepen of in gebieden waar private dienstverlening te kort schiet. De overheid heeft een voorbeeld- of alvast referentiefunctie in haar eigen organisatievorm en -beleid en moet bijvoorbeeld groepen tewerk stellen die elders weinig kansen krijgen. Zelfs wanneer dat vanuit efficiëntie- of effectiviteits oogmerk niet rationeel is. Anderzijds moet de gebruiker aan gunstige voorwaarden goed bediend worden. Het verzoenen van al deze uiteenlopende criteria is geen evidentie. De publieke dienstverlening moet immers legitiem bevonden worden. Vandaar dat velen subjectieve indicatoren van overheidsprestaties gebruiken. Het gaat daarbij vaak niet zozeer om performantie als dusdanig, wel de gepercipieerde performantie.

Toch is het belangrijk om performantie en perceptie van performantie gescheiden te houden. Helaas zijn er tot op de dag van vandaag geen goede algemene meetinstrumenten van de performantie van de publieke sector die ons toelaten om zinvolle vergelijkingen te maken tussen overheden (Van de Walle & Bouckaert, 2007). Daarnaast moeten we voorzichtig omspringen met subjectieve indicatoren. Immers, een negatieve attitude van de bevolking ten opzichte van de overheid, kan leiden tot een negatieve perceptie van de prestaties van die overheid. Deze attitude kan meer te maken hebben met de algemene culturele context, dan wel met de overheid in kwestie (Van de Walle, Sterck, Van Dooren & Bouckaert, 2004). Daarnaast moet er ook rekening worden gehouden met het feit dat verwachtingspatronen de perceptie in belangrijke mate kunnen beïnvloeden.

## **2.3. Welke waarde-indicatoren zijn er voor dienstverlening in de publieke sector?**

Het productievolume kan in de marktsector gemakkelijk afgeleid worden uit de marktwaarde van de betreffende goederen. Tijdreeksen kunnen op die manier worden geconstrueerd, rekening houdend met de prijsindex. Op die manier kan men "waarde-indicatoren" ontwikkelen. Diensten door de publieke sector geproduceerd, worden doorgaans niet op de vrije markt verhandeld. Dit maakt dat hun marktwaarde meestal niet bekend is. De waarde van deze productie kan men op die manier niet in geld uitdrukken. Daarom wordt in de meeste gevallen gebruik gemaakt van fysieke productindicatoren (Kuhry, 2004).

Dit is een verzamelterm, die betrekking heeft op verschillende types indicatoren, die als directe of indirecte maat voor de productie kunnen gelden. Hierbij kan een onderscheid worden gemaakt tussen: prestatie-, gebruiks- en procesindicatoren. Prestatie-indicatoren hebben betrekking op de geleverde eindproducten, bijvoorbeeld

het aantal diploma's dat een onderwijsinstelling afleverde. Gebruiksindicatoren van hun kant handelen over de afnemers van de diensten, zoals het aantal leerlingen/studenten. Procesindicatoren zijn dan weer indicatoren die slaan op de verrichte werkzaamheden of tussenproducten, bijvoorbeeld het aantal gegeven lesuren.

Probleem blijft dat het moeilijk is om zuiver collectieve goederen/diensten op die manier te meten. Een alternatieve manier om deze goederen/diensten te meten bestaat erin om de ingezette middelen te plaatsen tegenover het BBP (Kuhry, 2004).

Niet enkel het volume, vooral de kwaliteit van het overheidsbeleid valt soms moeilijk te meten. Wat is de kwaliteit van defensie? Er is grote onenigheid over de definitie van kwaliteit (Eggink en Blank, 2002). Die auteurs zien een eventuele "trade-off" tussen kwaliteit en doelmatigheid. Als kwaliteit niet voldoende tot uitdrukking komt in de productiematen, dan kunnen lage kwaliteitsinstellingen zich als zeer doelmatig kwalificeren. Deze *trade-off* is overigens geen wetmatigheid, efficiëntie en kwaliteit kunnen ook hand in hand gaan (van Thiel & Leeuw, 2003).

#### **2.4. Welke kwaliteitsnormen kunnen we gebruiken?**

Met welke, goed omschreven kwaliteitsnormen zullen we de data analyseren die we in het evaluatieonderzoek opnemen? Voorbeelden van dergelijke criteria vinden we bij de meting die het Sociaal Cultureel Planbureau uitvoerde, of bij Eurostat of de OESO. In ieder geval is het essentieel om kwalitatief goede data op te nemen opdat een meetinstrument betrouwbaar en valide zou zijn. De kwaliteit van de data is overigens een cruciale factor als het gaat meer specifiek om de effecten van prestatie-meetsystemen (cf. infra).

#### **2.5. Zijn alle nodige gegevens beschikbaar?**

Essentieel is het verkrijgen van voldoende en correcte informatie over het overheidsfunctioneren. Wat zijn de kosten, de processen, de eindproducten, wat willen en vinden de gebruikers ervan, wie is waarvoor verantwoordelijk, ... Een analytische boekhouding helpt alvast bij de doelmatigheidsanalyse. Binnen een analytische boekhouding zijn kosten en opbrengsten immers niet alleen naar soort, maar ook naar plaats en verantwoordelijkheid gekend. In de publieke sector (net als in de private) is het vaak niet evident om alle kosten toe te wijzen aan bepaalde activiteiten. Vooral de kosten voor algemeen ondersteunende activiteiten zijn moeilijk toewijsbaar. Dit kan ondermeer op basis van tijdobservaties die een beeld geven van de tijd die personeelsleden besteden aan diverse taken en de tijd die activiteiten in beslag nemen. Deze registratie gebeurt bij voorkeur gespreid in de tijd om vertekeningen eigen aan specifieke periodes in het werkjaar te vermijden. Een dergelijke registratie is vaak statisch en tijdrovend. Men kan eventueel gebruik maken van de techniek van normbepalingstijd waarbij men vooropstelt hoeveel tijd mag/moet worden besteed aan een activiteit. In dat geval werkt men taakstellend, wat nuttig is bij nieuwe activiteiten. Deze aanpak laat toe om te rekenen of alle beschikbare personeelscapaciteit volledig wordt benut met de huidige activiteiten.

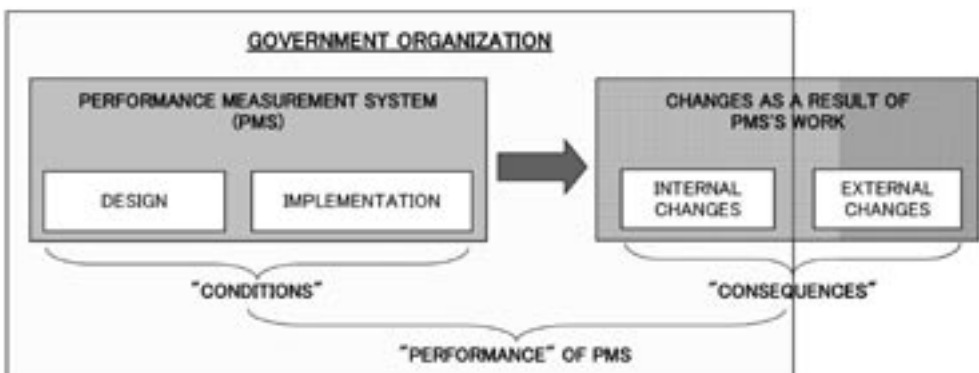
Daarnaast is er nood aan een prestatie-meetsysteem dat de organisatiedoelstellingen koppelt aan middelen, activiteiten, prestaties en effecten (MAPE). Deze opdracht kadert in de opvolging van de beheersovereenkomst en ondernemingsplannen die bijvoorbeeld de agentschappen in de Vlaamse overheid moeten uitvoeren. Er moet zoveel mogelijk geharmoniseerd worden naar definities en categorieën zodat vergelijkingen over gelijkaardige doelgroepen en/of prestaties mogelijk wordt.

Bij het verzamelen van gegevens kunnen we een onderscheid maken tussen primaire en secundaire data. Secundaire data vinden we onder meer terug in de interne meetsystemen waarmee het management beheers- en beleidsinformatie verzamelt. Daarnaast kan er behoefte zijn aan additionele primaire dataverzameling. Bij primaire dataverzameling kan men zich richten tot verschillende categorieën *stakeholders*: ontvangers van het beleid/de doelgroep, beleidsmakers en beleidsverstrekkers (politici en ambtenaren), derden die kennis hebben over de beleidsontvangers en de bredere beleidscontext (zoals experts, wetenschappers, ouders van leerlingen, et cetera). Deze primaire dataverzameling kan ook peilen naar aspecten van kwaliteit van de dienstverlening (bijvoorbeeld tevredenheid over de geleverde diensten, de wachttijden).

### 3. Effecten van prestatie-meting

Bij beleids- of prestatie-meetsysteem moeten we het onderscheid maken tussen twee dimensies: de condities en de consequenties. De condities hebben betrekking op het design en de implementatie van het meetsysteem, terwijl de consequenties betrekking hebben op de resultaten van de werking van een dergelijk systeem. De gevolgen kunnen intern en extern zijn. Interne gevolgen zijn bijvoorbeeld veranderingen in attitudes van werknemers, verhoging van de efficiëntie en veranderingen in de toewijzing van middelen. Externe veranderingen situeren zich buiten de organisatiegrenzen en hebben bv. betrekking op veranderingen in percepties van burgers en veranderingen in de maatschappelijke omgeving. Deze begrippen worden samengebracht in figuur 2.1 (Hiraki, 2007, 5).

**Figuur 2.1.: Dimensies bij beleids- of prestatie-meetsystemen**



### 3.1. Situering en het goede aan prestatiemeting

In deze paragraaf komt de kern van voorliggende bijdrage aan bod en gaan we vooral in op de onbedoelde en perverse of negatieve<sup>29</sup> effecten van beleidsevaluatie. Meer in het bijzonder gaat onze aandacht uit naar de effecten van het meten van prestaties. Bij het meten van prestaties zien we een overwegende output-oriëntatie, hoewel het vanuit beleidsoogpunt interessanter is om te kijken naar het uiteindelijke effect van het overheidsoptreden (outcome). Zoals aangegeven is outcome evenwel moeilijk te meten. Vandaar de aandacht voor output, die wel meetbaar is en op grond waarvan bijsturing wel mogelijk is (De Bruijn, 2002). In de hierna volgende uiteenzetting gaat het dus, zoals in de literatuur ter zake, voornamelijk over outputmeting.

Binnen de indrukwekkende literatuur over prestatiemeting en prestatie-meetsystemen is het aantal bijdragen die specifiek handelen over de negatieve of perverse effecten van prestatiemeting relatief beperkt. Niettemin werd er al vrij vroeg aandacht aan geschonken, denk aan de analyse van Ridgway (1956). We delen de indruk van Pidd (2005) dat net zoals prestatiemeting zelf, het bestaan van dergelijke perverse effecten niet controversieel blijkt te zijn. Het lijkt een inherent en aanvaard gegeven te zijn. Toch hebben die perverse effecten ook direct en indirect een kostprijs. Ze worden nochtans vaak over het hoofd gezien bij het opzetten van prestatie-meetsystemen, wellicht vanuit de idee dat het meetsysteem er geen last van heeft.

Binnen een aantal specifieke beleidsdomeinen, zoals de zorgsector, vinden we wel relatief veel aandacht voor de ongewilde effecten van meetsystemen (Brans e.a., 2008). Voor een meer globale analyse ter zake kunnen we verwijzen naar het werk van De Bruin (2002, 2006) en Smith (1995).

Het spreekt voor zich dat de communicatie over prestatiemeting en over de resultaten ervan een cruciale rol speelt in de effecten die deze prestatiemeting hebben. Het is dan ook essentieel om zowel oog te hebben voor interne (bv. t.a.v. medewerkers) als voor externe communicatie (bv. t.a.v. het algemene publiek) bij de rapportage n.a.v. prestatiemeting (Garnet e.a., 2008). Die communicatie kan op zich effecten genereren die aan de prestatiemeting verbonden zijn.

Hoewel onze aandacht vooral gaat naar de negatieve effecten van prestatiemeting en naar de strategieën om deze tegen te gaan (resp. §3.2. en §3.3.), is het evident dat prestatiemeting ook positieve effecten heeft. Zoals aangegeven is deze bijdrage geen pleidooi tegen de evaluatie van prestaties omdat die negatieve effecten kunnen hebben.

Evaluatieonderzoek kan immers bijdragen tot de sturing van het gedrag van beleidsinstanties (Swanborn, 1999). In opklimmende mate van dwang maakt de Bruijn (2002) melding van vier functies van prestatiemeting. Vooreerst draagt prestatiemeting bij tot transparantie. Het laat immers toe dat organisaties duidelijkheid bieden over de “producten” of dienstverlening die ze aanbieden en de middelen die ze daarvoor

---

<sup>29</sup> We beschouwen beide bijvoeglijke naamwoorden als synoniemen.

inzetten. Ten tweede kan een organisatie op grond van de prestatiemeting ook leren wat goed is en wat voor verbetering vatbaar is. Ten derde laat een dergelijke meting eveneens toe om (bestuurlijk) te oordelen over het functioneren van een organisatie, wat bijdraagt tot een betere want objectievere verantwoording. Op grond van het resultaat van de prestatiemeting kan men immers positief dan wel negatief gaan sanctioneren.

Bij zijn analyse van effecten van prestatiemeting stelt Smith (1995) vast dat de effecten zich vooral laten voelen op het interne management van organisaties binnen de publieke sector, ook wanneer evaluaties duidelijk gericht zijn op externe stakeholders (bv. burgers). Vandaar dat we bij de analyse van effecten van prestatiemeting ons richten op het gedrag van de organisatie zelf, en niet zozeer op bijvoorbeeld wijzigende attitudes van burgers t.a.v. die organisaties.

### 3.2. Negatieve effecten van prestatiemeting

De negatieve effecten van prestatiemeting, krijgen in dit tijdperk van meten en wegen vaak minder aandacht. In zijn analyse detecteert Smith (1995) acht onbedoelde (negatieve) effecten of dysfuncties van prestatiemeting.<sup>30</sup> Deze zijn volgens hem het resultaat van een gebrek aan congruentie tussen de doelstellingen van de *agent* en de eigenlijke doelstellingen van de *principal*. Een zelfde redenering vinden we terug bij De Bruijn (2006), die het respectievelijk heeft over de professional (degenen die in het primaire proces actief is) en de bestuurder/manager (die uiteindelijk op grond van de prestatiemeting wil sturen). Ook De Bruijn heeft in zijn analyse oog voor negatieve effecten in deze context. In wat volgt geven we een overzicht van de meest genoemde negatieve effecten van prestatiemetingen terzake. Dit sluit natuurlijk niet uit dat we in deze beschouwing ook oog hebben voor perverse effecten die niet onmiddellijk voortvloeien uit het spanningveld tussen de professional en manager. De reden hiervoor is dat ook los van dit potentiële spanningsveld zich negatieve effecten kunnen manifesteren. Daarbij komt nog dat –zeker in kleinere– organisaties dit onderscheid tussen professional en manager hooguit een theoretisch gegeven is, gezien beide taken soms door dezelfde personen worden waargenomen.

De mate waarin specifieke (positieve en negatieve) effecten zich manifesteren is sterk afhankelijk van de structuur en cultuur van een specifieke organisatie. Ook de kwaliteit van de indicatoren waarop prestatiegegevens worden gebaseerd is essentieel (Brans e.a., 2008). Daarnaast is het belangrijk om te benadrukken dat we het hier enkel hebben over effecten die zich manifesteren indien er überhaupt een meetsysteem is. Het is evenwel niet ondenkbaar dat dergelijke meetsystemen er niet komen omwille van houdingen van verantwoordelijken om deze te voorzien (De Peuter & Pattyn, 2008). Zo kan men van oordeel zijn dat er geen nood is aan dergelijke meetsystemen, dat het onmogelijk is om geschikte indicatoren te vinden of dat een dergelijk meetsysteem geen nut heeft (Bouckaert & Auwers, 1999). Ook dit kunnen we beschouwen als negatieve effecten van het meten, maar valt niet binnen de scope van deze bijdrage. We focussen ons op de effecten indien er gemeten

---

<sup>30</sup> Meer specifiek heeft hij het over : tunnervisie, suboptimalisatie, myopia, measure fixation, misrepresentation, misinterpretation, gaming en ossification. We gaan uiteraard verder in de tekst in op de invulling van deze concepten.

wordt. Hoewel de strategieën om om te gaan met negatieve effecten op zich het onderwerp vormt van §3.3., wordt hier voor een aantal negatieve effecten al meteen een aanzet tot remediëring gegeven.

### *Een te sterke nadruk op het gemakkelijk kwantificeerbare*

In prestatie-meetsystemen ligt de nadruk op kwantificeerbare data. Als gevolg daarvan zal ook het management vooral aandacht hebben voor kwantificeerbare processen, ten nadele van niet of moeilijker kwantificeerde aspecten van het overheidsbeleid. Een oorzaak hiervan is de moeilijkheid en onenigheid over de definitie van kwaliteit en/of verandering van de invulling ervan (Eggink & Blank, 2002).

Smith (1995) heeft het in deze context over “*tunnel vision*”. Hij geeft ook een voorbeeld uit de gezondheidszorg in Groot-Brittannië. De sterke nadruk op perinatale sterfcijfers leidde er tot veranderingen in de aard van de dienstverlening op materniteiten, ten nadele van niet-kwantificeerbare doelstellingen.

Ook De Bruijn verwijst naar dit probleem door ons erop te wijzen dat prestatie-meting potentieel de professionele habitus verdrift, door kwantiteiten en dus vooral meetbare en goed te definiëren aspecten van een prestatie centraal te stellen. Het voorbeeld dat hij aanhaalt is dat van musea, waar een te sterke focus op makkelijk meetbare data –zoals het bezoekersaantal– andere indicatoren en dus overwegingen, zoals de artistieke waarde van een collectie, naar achteren duwen.

Dit probleem is te verklaren door de divergentie tussen de doelstelling van een organisatie en van het meetsysteem. Het is specifiek voor de publieke sector. Eigen aan de sector is dat deze een hele reeks verschillende doelen moet realiseren, en dat veel belangrijke doelstellingen in die sector moeilijk te kwantificeren zijn. Daarbij komt nog dat de doelstellingen van organisaties binnen de publieke sector vaak veel verder reiken dan het directe doel van de dienstverlening. Zo moet onderwijs niet enkel kennis en vaardigheden overbrengen, maar ook attitudes aanleren, normen en waarden overbrengen, enz.

Veelal is het zeer moeilijk om alle activiteiten en doelstellingen te inventariseren. Als gevolg wordt het belang dat men aan prestatie-meting hecht verminderd en worden waarden meer in de verf gezet.

### *Te weinig aandacht voor de doelstellingen van het ruimer geheel*

Dit tweede probleem is wat Smith (1995) suboptimalisatie noemt. Verantwoordelijken voor een onderdeel van een ruimere organisatie zijn vooral geconcentreerd op hun particuliere doelstellingen. Dit gebeurt onder impuls van de prestatie-meting en gaat ten koste van de doelstellingen van de organisatie als geheel. Ook dit probleem komt vooral in de publieke sector voor, waar er vaak vele deelorganisaties betrokken worden bij de realisatie van doelstellingen en bevoegdheidsregels een belangrijk gegeven zijn.

De Bruijn (2002) verwijst naar dit probleem als hij stelt dat prestatie-meting “bestel-

rantwoordelijkheid” in de weg kan staan. Met deze term duidt hij op het onderling ter beschikking stellen van expertise tussen (onderdelen van) organisaties. Zo had het invoeren van prestatiemeting in scholen een slechte invloed op de samenwerking en onderlinge verstandhouding tussen de scholen in kwestie.

Veel is afhankelijk van het soort activiteiten van een organisatie binnen de publieke sector. Daarenboven kan de centrale overheid dit probleem enigszins vermijden door een goede afstemming tussen de verschillende geledingen, bv. via algemene dienstverleningscharters en globale handvesten die vertaald worden in meer operationele charters en handvesten (Verlet, 2008).

#### *Te veel aandacht voor doelstellingen op korte termijn*

Ook dit probleem is er één dat voorkomt uit het eventuele verschil tussen de doelstellingen van een organisatie en wat een prestatie-meetsysteem in kaart kan brengen. Smith spreekt over *myopia*, de geneeskundige term voor bijziendheid. Het gaat om kortzichtigheid, in die zin dat men doelstellingen op korte termijn nastreeft ten nadelen van legitieme doelstellingen op langere termijn.

Prestatiemeting is veelal slechts een momentopname. Activiteiten kunnen op lange termijn een groot voordeel opleveren, maar dat komt niet steeds tot uiting in de meetsystemen. Veel prestatie-meetsystemen geven immers geen beeld van beleid over een langere periode, noch van toekomstige gevolgen van huidige managementacties.

Dit effect wordt versterkt wanneer leidinggevenden en medewerkers functies waarnemen op kortere termijn. Natuurlijk is ook de mate waarin dit probleem zich voordoet sterk afhankelijk van het soort dienstverlening, de cultuur en de structuur van een organisatie. Een manier om dit specifieke probleem te voorkomen is door in de prestatiemeting ook aandacht te hebben voor processen met betrekking tot thema's op langere termijn, eerder dan enkel output te meten.

#### *Een te sterke nadruk op maatstaven voor succes*

Dit effect is wat Smith (1995, 290) *measure fixation* noemt. Onder impuls van prestatiemeting heeft een organisatie de neiging om maar al te sterk de nadruk te leggen op de criteria op grond waarvan ze worden afgerekend, eerder dan het onderliggende objectief. In deze context wijzen Brans e.a. (2008) op het feit dat prestatiemeting kan leiden tot ritualisering. Dit komt erop neer dat men hoofdzakelijk goed tracht te scoren op de kengetallen om belanghebbenden tevreden te stellen. Men wijst er ook op het concept van Power (1999) die het heeft over “ontkoppeling”, een voorstelling van zaken die niet meer overeenstemt met de realiteit.

Een te sterke nadruk op maatstaven voor succes komt voort uit het onvermogen van vele meetsystemen om complexe fenomenen in kaart te brengen. Smith geeft in zijn analyse een voorbeeld van het reduceren van de wachttijden in de gezondheidszorg, meer specifiek de doelstelling dat patiënten niet langer dan twee jaar zouden mogen wachten op een heelkundige ingreep. Dit had als onvoorzien effect dat het



aantal patiënten dat één jaar diende te wachten op een ingreep toenam en dat de “intake” van patiënten op een later moment gebeurde. Patiënten kwamen later op de wachtlijst terecht.

Een mogelijke oplossing voor dit probleem is het opdrijven van het aantal criteria op grond waarvan het functioneren van een dienst wordt beoordeeld, hoewel dit natuurlijk op zich de focus kan vervagen en kan leiden tot “demoralisering”. Een alternatieve oplossing die Smith naar voor schuift, is het erkennen dat de meeste metingen proxies zijn voor output en dat de ultieme arbiters van de kwaliteit van de output de “cliënten” van de organisatie zijn. Om dit te verwezenlijken hebben we een duidelijk beeld nodig van wie die “klanten” zijn en van hun verwachtingen en behoeften. Daarnaast moeten we indachtig zijn dat de perceptie van het functioneren van een dienst niet noodzakelijk een goede indicator is voor het feitelijk functioneren voor deze dienst (Verlet, Reynaert & Devos, 2005).

### *Een foute voorstelling van de prestaties*

Dit effect verwijst naar het opzettelijk manipuleren van data. Daardoor stemt het gerapporteerde gedrag niet overeen met het feitelijke gedrag. Het spreekt voor zich dat de incentive om deze laakbare praktijken te hanteren het grootst is indien de nadruk sterk ligt op prestatie-indicatoren. De mogelijkheden tot een foute weergave van de prestaties is volgens Smith (1995) in de publieke sector dikwijls hoog, omdat de desbetreffende organisaties vaak zelf instaan voor het toeleveren van data en indicatoren over hun presteren (of het gebrek eraan). Ook hier kunnen we verwijzen naar de moeilijkheid om ingewikkelde fenomenen met precisie en betrouwbaarheid in kaart te brengen. Smith maakt een onderscheid tussen twee vormen van wat hij noemt *misrepresentation*: “creatieve” rapportering en fraude. Het onderscheid tussen beide is soms moeilijk te maken. In deze context heeft Van Dooren (2006) het specifiek over problemen die zich kunnen voordoen bij (des)aggregatie van data bij het gebruik van indexen. Zo bestaat de mogelijkheid dat men een minder goede score gaat verdoezelen door te aggregeren, terwijl men door desaggregatie de mogelijkheid heeft om gunstige resultaten extra in de verf te zetten. Het euvel van een foute weergave van de prestaties kan worden tegengegaan door te voorzien in (interne en externe) audit en in de mogelijkheid tot sancties wanneer men een foute representatie van prestaties op het spoor komt.

### *Gebrekkige validiteit en betrouwbaarheid*

Onder deze noemer kunnen we de meetproblemen plaatsen die Bouckaert en Auwers (1999, 77) beschouwden als ‘*pathologieën*’ die ‘*verwijzen naar foutieve percepties van volumes en aantallen*’. Meer specifiek heeft het men in deze context over convexe en concave meetinstrumenten omdat die respectievelijk hogere en lagere waarden laten optekenen dan er in de werkelijkheid bestaan. In essentie gaat het hier over een over- en onderschatting van datgene dat men wil meten, omwille van minder goede indicatoren. In dezelfde context stelde men dat de output abnormaal kan dalen of stijgen door het het feit dat er gemeten wordt. Ten slotte merkt men op dat door het meten van een bepaald fenomeen de frequentie ervan ogenschijnlijk kan toenemen. Dit is bijvoorbeeld het geval indien een type misdrijf voor het eerst



wordt gemeten. Alleen al door de bekendheid dat een misdrijf toeneemt, zijn er meer registraties, wat nog niet per definitie wil zeggen dat dergelijke voorvallen zich in de werkelijkheid meer voordoen. Deze kunnen op hun beurt leiden tot wat Bouckaert en Auwen (1999) waanbeelden noemen, verkeerde percepties van wat men meet. De genoemde pathologieën houden allen verband met de betrouwbaarheid en validiteit van het meetinstrumenten. Het is duidelijk dat deze problemen niet zozeer voortkomen uit het spanningsveld tussen managers/professionals, maar eigen zijn aan het meten als dusdanig.

### *Foute interpretaties*

Het productieproces van publieke dienstverlening is veelal een heel complex gegeven. Daarbij komt nog dat deze actoren zich in een complexe omgeving bevinden. Dus ook al kunnen we de prestaties perfect in kaart brengen, dan nog is het niet evident om de signalen die in de data te onderkennen zijn te vertalen. Het spreekt voor zich dat deze foute interpretaties een reëel probleem vormen bij het hanteren van prestatie-indicatoren. Daarnaast stellen we vast dat de prestaties van diverse organisaties vaak met elkaar vergeleken worden. We wezen er eerder op dat dit geen evidentie is, rekening houdend met het feit dat ze verschillende doelstellingen, middelen, context, enz. hebben. Het correct omgaan met gegevens uit prestatie-meting is een kunst. Door het aantal indicatoren beperkt te houden kan het probleem van de foute interpretatie van data enigszins worden tegengewerkt, al kan dit natuurlijk op zich weer bijdragen tot andere perverse effecten.

### *Gaming*

Dit effect van prestatie-meting betreft het opzettelijk manipuleren van gedrag om strategisch voordeel veilig te stellen. Terwijl we bij "*misrepresentation*" vertekeningen zien in het gerapporteerde gedrag, gaat het bij *gaming* om een manipulatie van het feitelijke gedrag. In het werk van De Bruijn (2006) vinden we een voorbeeld van hoe prestatie-meting kan leiden tot strategisch gedrag. Het gaat om de prestatie-meting bij een dienst binnen het Australische leger die moet zorgen voor huisvesting van militairen die ver van huis zijn gelegerd. De dienst biedt deze militairen huisvesting aan. De prestatie-indicator die men hanteerde was het aantal militairen dat na maximaal drie aanbiedingen instemde met de huisvesting. Na het invoeren van deze indicator, stelde men al gauw vast dat de volle 100% van de militairen instemde na maximaal drie aanbiedingen. De verklaring hiervoor was eenvoudig. De dienst bood informeel huisvesting aan. Pas wanneer de betrokkenen op de dienst vrij zeker waren dat de militair met het aanbod zou instemmen werd het aanbod formeel gedaan. Hier is sprake van strategisch gedrag: de prestaties zijn er enkel op papier, de maatschappelijke betekenis ervan is beperkt.

Hoe kan *gaming* worden tegengegaan? In eerste instantie kan men, aldus Smith (1995), het tegenwerken door een brede waaier aan prestatie-indicatoren op te nemen. Een tweede mogelijkheid is *benchmarking*, of kunnen er voor managers carrièreperspectieven op kortere termijn geboden worden. Dit kan anderzijds leiden tot *myopia*.

## *Fossilisering*

De “verstening” of fossilisering, verwijst naar het ontraden van innovatie door een te rigide prestatiemeting. Vele prestatie-meetsystemen hebben de neiging om het voortdurend reproduceren van het bestaande te belonen. De noodzaak om op voorhand prestatie-indicatoren te selecteren en doelstellingen voorop te stellen, kan er toe bijdragen dat men geen oog heeft voor nieuwe bedreigingen en mogelijkheden. Het is in diezelfde omgeving dat *gaming* kan kiemen. Het gevaar van fossilisering komt voort uit de onvermijdelijke *time lag* tussen het opzetten van prestatie-metingen en de mogelijkheid/moeite om het meetinstrument bij te sturen. Dit maakt dat men een soort van meta-controlemechanisme nodig heeft om toe te zien op het adequaat zijn van de prestatie-indicatoren.

Deze verstening van organisaties kan worden tegengewerkt door het voorzien van *incentives* voor het anticiperen op nieuwe uitdagingen en innovatief gedrag, ook al dragen deze activiteiten niet rechtstreeks bij tot de huidige prestatie-indicatoren.

## *Versterkt interne bureaucratie*

Een ander negatief effect van prestatiemeting, is dat deze de interne bureaucratie in sterke mate kan doen toenemen. Men moet immers meer en meer tijd en middelen besteden aan het verzamelen en ontsluiten van informatie. Een degelijke prestatiemeting vergt namelijk een nauwkeurige registratie van input, processen, output, outcome en houdt daarenboven rekening met steeds veranderende omgevingsfactoren (cf. supra). Het spreekt voor zich dat een dergelijke meting ook extra middelen en mensen vergt. Wanneer men dus een meetsysteem opzet, onder meer vanuit het idee om op termijn kostenbesparend te kunnen optreden, dient rekening gehouden te worden met het feit dat meten kost.

## *Het afremmen van ambities/cherry picking*

Dat prestatiemeting een rem kan vormen op de ambities van een organisatie, komt voort uit het gegeven dat organisaties hun prestaties, bijvoorbeeld in termen van output, kunnen opdrijven door de input te optimaliseren. Meer specifiek komt het er op neer dat men ervoor kan opteren om de input zo te selecteren zodat deze een zo min mogelijke *throughput* vereist. In deze context heeft men het over “*cherry picking*”. Bij wijze van voorbeeld kunnen we verwijzen naar het onderwijs, waarbij een school haar output (bv. in termen van percentage geslaagde studenten), kan verhogen door strenge selectiecriteria te hanteren bij het toelaten van studenten (De Bruijn, 2006).

Dit neveneffect wordt ook “*cream-skimming*” genoemd, het afremmen van de doelgroep door zich vooral te richten tot subgroepen die het gemakkelijkst te bereiken zijn. Swanborn (1999) merkt overigens op dat dit effect zich ook kan manifesteren door zelfselectie vanuit de doelgroep. Onder deze noemer kunnen we ook het negatieve effect “polarisatie” vermelden (Van Dooren, 2006). De situatie die hij hierbij schetst is deze waarin bepaalde vormen van dienstverlening of dossiers worden afgestoten omdat ze als hopeloos worden beschouwd met het oog op de gestelde normen.

In de plaats van te investeren in deze probleemgevallen kan men desgevallend opteren om deze af te stoten.

### *Kan goed presteren bestraffen*

Hoewel het opzetten van een prestatie-meetsysteem onder meer als doelstelling heeft om het goed presteren in kaart te brengen en te bevorderen, kan datzelfde meetsysteem er ook toe bijdragen dat datzelfde goed presteren wordt afgestraft. Bij het beschrijven van dit effect verwijst De Bruijn (2002) naar het werk van Bordewijk en Klaassen (2000) die ons er op hun beurt op wijzen dat het investeren in transparantie en efficiëntie niet zonder risico's is. Organisaties die investeren in transparantie kunnen daar eventueel voor gesanctioneerd worden via een budgetvermindering. Immers, controle leert dat gelijkblijvende prestaties ook met minder middelen kunnen gebeuren. Een gelijkaardige organisatie die niet investeert in transparantie en efficiëntie wordt beloond met eenzelfde budget voor gelijkblijvende prestaties.

### **3.3. Een algemene verklaring van negatieve effecten en algemene strategieën om ze te voorkomen**

Zoals aangegeven vindt het optreden van diverse negatieve effecten zijn oorsprong bij een mismatch tussen de doelstellingen van *principals*/het management enerzijds en *agents*/de professionals anderzijds.

De Bruijn (2002, 2006) ziet twee algemene verklaringen voor de negatieve effecten van prestatie-meting. Vooreerst stelt hij dat professionals het prestatie-meetsysteem kunnen perverteren en dat zij zich hiertoe gelegitimeerd achten. Deze legitimatie kent op zijn beurt diverse gronden. Zo beschouwt men de prestatie-meting als gebrekkig gezien prestaties –zeker in de publieke sector– een *trade-off* zijn tussen (concurrerende) waarden. Publieke prestaties zijn meervoudig en dit komt niet altijd tot uiting in de systemen om deze te meten. Daarnaast beschouwen vele professionals prestatie-meting als “unfair”, gezien het vaak onvoldoende rekenschap geeft van het feit dat prestaties veelal het resultaat zijn van coproductie. Een derde en laatste legitimatiegrond is gestoeld op de mening dat prestatie-meting op zich veelal een statisch gegeven is, terwijl prestaties dynamisch van aard zijn.

Een tweede algemene verklaring is te vinden in de wetmatigheid dat hoe meer bestuurders willen sturen door prestatie-meting, des te minder effectief de prestatie-meting zal zijn. De Bruijn heeft het in deze context over de paradox van toenemende perverse effecten: hoe meer een bestuurder door prestatie-meting invloed wil uitoefenen op het primaire proces, hoe meer perverse effecten zich manifesteren. De verklaring voor deze paradox is tweeledig. Vooreerst zullen professionals zich trachten af te scherm van prestatie-meting. Ten tweede stelt hij dat hoe sterker de werking van prestatie-meting voelbaar is, hoe minder recht wordt gedaan aan het meervoudig, co-productief en dynamisch karakter van de prestaties. De Bruijn stelt overigens dat deze paradox bijzonder hardnekkig is. Stel dat professionals zich niet conformeren aan het meetsysteem en zich afschermen, dan kan dit aanleiding geven tot bijvoorbeeld strategisch gedrag waardoor de prestatie-meting op zich niet effectief is. Stel dat men wel wil conformeren, dan nog kunnen er zich effecten voordoen, bijvoorbeeld een te sterke nadruk op maatstaven voor succes, waardoor

meting eveneens niet effectief is.

Zowel Smith (1995) als De Bruijn (2002, 2006) geven in hun analyse strategieën weer om de negatieve effecten die kunnen optreden bij het gebruik van prestatiemeting zoveel als mogelijk tegen te gaan. Een element dat in beide studies naar voren komt is het hanteren van subjectieve indicatoren in de vorm van tevredenheidsmetingen. De Bruijn beschouwt dit evenwel als een alternatief beoordelingssysteem, naast de prestatiemeting. Hierin kunnen we hem niet geheel volgen, in die zin dat ook deze beoordelingen inherent deel kunnen uitmaken van het prestatie-meetsysteem. De perceptie van het functioneren van een dienst kan immers een wezenlijke invloed uitoefenen op het functioneren van deze dienst (Stipak, 1979). Dat dergelijke subjectieve indicatoren niet alleenzalmakend zijn, betwijfelen weinigen. Niettemin lijken dergelijke indicatoren ons cruciaal te zijn in de meting, hoewel we de nodige voorzichtigheid in acht moeten nemen bij het hanteren ervan. Zo is het gebruik van subjectieve indicatoren bij de beleidsevaluatie vaak gebaseerd op de simplistische veronderstelling dat dergelijke indicatoren ook een goed meetinstrument zijn voor de kwaliteit van de dienstverlening. Naast vele andere factoren kan het gebrek aan kennis of zichtbaarheid van een dienst de subjectieve evaluatie van die dienst systematisch vertekenen (Trent e.a., 1984). Daarom vereisen ook deze elementen aandacht, vooraleer de informatie die subjectieve indicatoren ons opleveren gebruikt kan worden in de beleidsevaluatie (Anderson e.a., 1984). Subjectieve indicatoren kunnen dus diverse types van beleidsrelevante informatie geven aan de beleidsmakers. Of deze subjectieve indicatoren op zich volstaan om de kwaliteit van de dienstverlening te beoordelen, is een ander paar mouwen (Stipak, 1979).

Daarnaast zijn er nog een aantal andere strategieën die de prestatiemeting ten goede komen. Zo stelt De Bruijn dat een tolerantie ten aanzien van een diversiteit aan (zelfs concurrerende) productdefinities aangewezen is. Dergelijke verschillende definities kunnen perfect naast elkaar bestaan. Het hanteren van een variëteit aan productdefinities biedt overigens een aantal voordelen: het werkt conflictreducerend, biedt een rijker beeld van de geleverde prestaties, het matigt perverterend gedrag en kan ook voor het management interessant zijn. Wat dit laatste betreft kan de diversiteit aan productdefinities immers bevorderlijk zijn voor het gezag van de uitkomsten. Immers, als een organisatie bijvoorbeeld op grond van alle productdefinities slechts/ goed scoort, heeft men een stevig argument.

Een andere gouden raad van De Bruijn is het verbod op een monopolie van betekenisgeving. Prestatiemeting kan een verhullende betekenis hebben. Deze neemt toe naarmate de afstand tussen de producent en de ontvanger van de cijfers groter is. Hoe groter deze afstand, hoe moeilijker ook om betekenis te geven en hoe hoger ook de vermeende "hardheid" van de gegevens. Dit verbod op een monopolie van betekenisgeving kan gerealiseerd worden door duidelijke afspraken te maken tussen bijvoorbeeld de bestuurders en de professionals.

Ook het beperken van de functies en de fora waartoe de prestatiemeting wordt gebruikt, komt de prestatiemeting vaak ten goede. Immers, hoe hoger het aantal functies en fora, hoe groter de kans op de paradox van toenemende perverse effecten. Ook hieromtrent zijn dus duidelijke afspraken essentieel voor het welslagen van de prestatiemeting en vooral om te vermijden dat perverse effecten zich manifesteren.

Een gelijkaardige aanbeveling vinden we in het werk van Smith (1995), die van oordeel is dat perverse effecten kunnen worden tegengewerkt door het betrekken van medewerkers op alle niveaus in de ontwikkeling en de implementatie van prestatie-meetsystemen. Verder pleit hij voor het flexibel hanteren van prestatie-indicatoren en om ze niet enkel en alleen te hanteren als controlemechanisme.

De Bruin (2002) breekt ook een lans voor een strategische selectie van de “producten” die via het meetsysteem in beeld komen. Op zich kan men opteren voor een zware dan wel lichte meting.<sup>31</sup> De selectie van de “producten” is een strategische keuze, veelal gemotiveerd door het streven naar volledigheid (al leidt dit vaak tot een *overload* aan informatie en is een dergelijke meting niet kosteneffectief), terwijl door een intelligente selectie van een beperkter aantal “producten” toch invloed kan worden uitgeoefend op de organisatie als geheel. Verder kan men een onderscheid maken tussen een operationalisering van meervoudige producten (bijvoorbeeld het geven van onderwijs) versus enkelvoudige producten (bijvoorbeeld het aantal zittenblijvers). In het werk van De Bruin vinden we een pleidooi om te kiezen voor hoofdzakelijk enkelvoudige producten en slechts een beperkt aantal meervoudige producten.<sup>32</sup> Enigszins strijdig hiermee wordt er in het werk van Smith (1995) gepleit voor een kwantificering van elke doelstelling. Smith pleit voor een kritische houding tegenover het design van het prestatie-meetsysteem en wat er al dan niet wordt door gemeten.

Een andere strategie van De Bruijn is deze van het management van de concurrerende benaderingen “product” en proces. Hiermee bedoelt hij dat het essentieel is om bij prestatie-meting niet enkel te kijken naar de output, maar dat het eveneens essentieel is om een beeld te krijgen van de processen of de *throughput*. Bij prestatie-meting is het in deze context dan ook essentieel om voldoende aandacht te besteden aan de zogenaamde *black box* waarbij input wordt vertaald in output (Pawson & Tilley, 1997). Overigens is het risico op perverse effecten groter indien productindicatoren worden gebruikt (Brans e.a., 2008), een reden te meer om ook procesindicatoren op te nemen in de evaluatie.

De essentie van strategieën die De Bruijn (2002) naar voor schuift is een terughoudend gebruik, ruimte en vertrouwen enerzijds en het afspreken van spelregels anderzijds. Daarnaast vinden we in het werk van Smith (1995) een aantal suggesties terug die vooral relevant zijn wanneer doelstellingen vrij vaag zijn en de meting van de output problematisch is. In dat geval benadrukt hij het opnemen van subjectieve indicatoren, door het meten van de tevredenheid van “klanten”. In diezelfde context is het aangewezen om de interpretatie van prestatie-indicatoren over te laten aan experts en een zorgvuldige audit te voorzien.

---

<sup>31</sup> Het onderscheid tussen beiden wordt gemaakt door het al dan niet verbinden van een vorm van sanctionering aan de beoordeling op grond van het prestatie-meetsysteem.

<sup>32</sup> Het is aantrekkelijk voor de professionals, het heeft hen ruimte en het is voor bestuurders ook makkelijker om af te rekenen op enkelvoudige producten. Daarnaast kan een focus op enkelvoudige producten positieve effecten hebben op meervoudige producten. Ten slotte kan het goede gedrag van professionals worden beloond door een beperkte selectie van producten. Voor een meer gedetailleerde argumentatie kunnen we verwijzen naar De Bruin (2002, 152-154).

## 4. Besluit

Niemand betwist dat in de publieke sector beleidsevaluatie via zo objectief mogelijke indicatoren steeds belangrijker wordt. In het kader van de groeiende complexiteit van de beleidsomgeving en in het licht van toenemende verantwoordingsvragen streven beleidsmakers naar *evidence-based policy* dat een (indruk van) hoge *return on public investment* moet garanderen.

Hoe deze beleidsevaluatie in het algemeen en het meten van prestaties in het bijzonder moet gebeuren, welke functies men eraan mag of moet toedichten en op welke fora ze hoe met wie besproken moeten worden, dat en meer is voorwerp van discussie. Deze discussies houden verband met de maatstaven van goed bestuur en van morele en democratische legitimiteit. Maar er zijn nog andere vragen te stellen. Gezien beleidsevaluatie in essentie nagaat of en hoe doelstellingen, gelet op de middeleninzet, gehaald worden, lijkt het ons logisch dat de keuzes van doelstellingen en middelen als het ware primeren op, of alvast extern zijn en vooraf gaan aan, de evaluatie. Beleidsevaluatie kan leiden tot beleidsleren en tot de versterking van verantwoordingsprocessen en is dus in essentie voordelig. Maar, als het denken over de hoge nood aan geobjectiveerde, gekwantificeerde evaluatie overheersend wordt, kan dit invloed hebben op de keuze van beleidsdoelstellingen en dus op de probleemformulering. Op die manier kan evaluatie op een onvoorziene, oneigenlijke manier het beleidsproces beïnvloeden. Doelstellingen (probleemoplossingen) worden soms zodanig geformuleerd of gekozen dat ze nadien objectief geëvalueerd kunnen worden en bovendien tot een goede score leiden.

De negatieve effecten van prestatiemeting komen in belangrijke mate voort uit het spanningsveld tussen bestuurders/managers en de uitvoerders/professionals. Deze laatste kunnen zich in hun autonomie aangetast voelen gezien bij beleidsevaluatie ze veelal niet langer zelf bepalen in welke mate hun werk goed of minder goed is (Swanborn, 1999). Daarnaast kunnen er zich eveneens spanningen voordoen met andere actoren, zoals bijvoorbeeld de evaluatoren zelf, naast de bestuurders en professionals. Natuurlijk zijn er ook negatieve effecten die te wijten zijn aan het meten op zich.

In deze bijdrage stonden we stil bij diverse mogelijke negatieve of perverse effecten. In het bijzonder concentreerden we ons op de interne gevolgen. Een algemene bemerking is dat de mate waarin perverse effecten zich manifesteren sterk afhankelijk is van het soort organisatie en activiteit, de structuur en de heersende cultuur.

Al die negatieve effecten spelen een rol in het debat over prestatiemeting. Met de woorden van Ridgway: *'The cure is sometimes worse than the disease.'* (Ridgway, 1956, 240). Hij had het over de vraag of het überhaupt aangewezen is om gebruik te maken van kwantitatieve meetinstrumenten om prestaties in de publieke sector in kaart te brengen. Vandaag lijkt het nut van het kwantificeren van de prestaties binnen de publieke sector niet het onderwerp van discussie te zijn. Al kunnen we – ruim 52 jaar na zijn analyse – ons nog steeds aansluiten bij zijn conclusie dat de perverse effecten van prestatiemeting onvoldoende onderkend worden. Daarnaast dienen we er ons bewust van te zijn dat strategieën om specifieke perverse effecten

tegen te gaan, op zich weer aanleiding kunnen geven tot het genereren van andere perverse effecten. Omgaan met perverse effecten is dus op zich een moeilijke evenwichtsoefening. Niettemin, dat bewustzijn is op zich al een goed startpunt.

Beleidsvaluatie dient een nobel streven waarin een “goed” bestuur voorop staat en is in deze een middel. In een aantal gevallen dreigt dit middel evenwel een doel op zichzelf te vormen. Dit is meteen een hardnekkig probleem waarmee beleidsvaluatie in het algemeen en prestatie­meting in het bijzonder te kampen heeft. Zeker als enkele beroeps­categorieën zich volop op prestatie­meting concentreren en dat hun voornaamste broodwinning is. Dit probleem verdient de nodige aandacht, willen we vermijden dat het doel van beleidsvaluatie ondergesneeuwd raakt.



## Bibliografie

Anderson, B., Ryan, V. & Goudy, W. (1984). Consistency in subjective evaluations of community attributes. In: *Social Indicators Research*, 14 (2), 165-175.

Bouckaert G. & Auwers T. (1999). *Prestaties meten in de overheid*. Brugge: Die Keure.

Brans, M., Giesbers, S. & Meijer, A. (2008). Alle ogen op de ziekenhuizen gericht? De effecten van openbaarheid van prestatiegegevens. In: *Bestuurswetenschappen*, 62 (2), 32-52.

Callens, M., (2007). *Efficiëntie en effectiviteit in de publieke sector*. Draft Nota. Brussel: Studiedienst van de Vlaamse Regering.

De Bruijn, H. (2002). Prestatiemeting in de publieke sector. Strategieën om perverse effecten te neutraliseren. In: *Bestuurswetenschappen*, 56 (2), 139-159.

De Bruijn, H. (2006). *Prestatiemeting in de publieke sector. Tussen professie en verantwoording* (2<sup>e</sup> druk). Den Haag: Lemma.

De Peuter, B., Pattyn, V. & Brans, M. (2008). *Perverse effecten bij beleidsevaluatie? Een risicoanalyse van de praktijk in Vlaanderen in het licht van mogelijke evaluatiemotieven*. Paper voor het Politicologenetmaal, Nijmegen, 29-30 mei 2008.

Edwards, M. & Clough, R. (2005). *Corporate Governance and Performance: An Exploration of the Connection in a Public Sector Context*. Issues paper 1. Gevonden op: [http://www.canberra.edu.au/corpgov-aps/pub/IssuesPaperNo.1\\_GovernancePerformance\\_Issues.pdf](http://www.canberra.edu.au/corpgov-aps/pub/IssuesPaperNo.1_GovernancePerformance_Issues.pdf), laatst geraadpleegd op 10 oktober 2008

Eggink, E. & Blank, J.L. (2002). Efficiëntie in de publieke sector. In: *Beleidswetenschap*, 16 (2), 144-161.

Garnet, J., Marlowe, J. & Pandey, S. (2008). Penetrating the Performance Predicament: Communication as a Mediator or Moderator of Organizational Culture's Impact on Public Organizational Performance. In: *Public Administration Review*, 68 (2), 266-281.

Hayes, M. & Pidd, M. (2005). *Public Announcement of Performance Ratings: implications for Trust Relationships*. <http://www.lancs.ac.uk/staff/smamp/working%20paper%20on%20trust.pdf>, laatst geraadpleegd op 10 oktober 2008.

Hiraki, T. (2007). *Factors affecting the "performance" of performance measurement systems: an analysis of the Japanese Practice*. Paper presented at the EGPA conference, EGPA Study Group on productivity and quality in the public sector, Madrid.



[http://www.vlaandereninactie.be/documenten/septemberverklaring\\_2007.pdf](http://www.vlaandereninactie.be/documenten/septemberverklaring_2007.pdf),  
laatst geraadpleegd op 10 oktober 2008.

Hood, C. (1996). Exploring variations in public management reform of the 1980's. In: Bekke, H., Perry, J. & Toonen, T. (ed.). *Civil service systems in comparative perspective*. Bloomington: Indianapolis.

Jorgensen, T. (2006). Value consciousness and public management. In: *International journal of organization theory and behavior*, 9 (4), 510-536.

Kaplan, R. & Norton, D. (1996). *The balanced scorecard: translating strategy into action*. Boston (Mass.): Harvard business school press.

Kuhry, B. (ed.) (2004). *Prestaties van de publieke sector*. Den Haag: SCP Nederland.

Luts, M., Van Dooren, W. & Bouckaert, G. (2008). *Internationale rangschikkingen gerangschikt. Een meta-analyse van rangschikkingen van publieke sectoren*. Leuven: Steunpunt beleidrelevant onderzoek-bestuurlijke organisatie Vlaanderen.

Manning, N. (2000). The New Public Management & its legacy. [www.worldbank.org/publicsector/civilservice/debate1.htm](http://www.worldbank.org/publicsector/civilservice/debate1.htm)

Modell, S. (2004). Performance measurement myths in the public sector: a research note. In: *Financial Accountability & Management*, 20 (1), 39-55.

OECD (2007). *Issues in Output Measurement for "Government at a Glance"* OECD project on Management in Government: Comparative Country Data. OECD GOV Technical Paper 2 (Second Draft).

Pestman, P. (2007). NMP-instrumenten in de praktijk. Van ideologisch debat naar kritische toepassing. In: *Bestuurskunde*, 61 (3), 104-114.

Pidd, M. (2005). Perversity in Public Service Performance Measurement. In: *International Journal of Productivity and Performance Management*, 54 (5/6), 482-493.

Ridgway, V. (1956). Dysfunctional Consequences of Performance Measurements. In: *Administrative Science Quarterly*, 1 (2), 240-247.

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. In: *International Journal of Public Administration*, 18 (2/3), 277-310.

Stipak, B. (1979). Are there sensible ways to analyse and use subjective indicators of urban service quality? In: *Social Indicators Research*, 6 (4), 421-438.

Swanborn, P. G. (1999). *Evaluëren: het ontwerpen, begeleiden en evalueren van interventies: een methodische basis voor evaluatie-onderzoek*. Amsterdam, Boom.

Trent, R., Stout-Wiegand, N. & Furbee, P. (1984). The nature of the connection between life course and satisfaction with community services. In: *Social Indicators Research*, 15 (4), 417-429.

Van de Walle S. & Bouckaert G. (2007). Perceptions of productivity and performance in Europe and the United States. In: *International Journal of Public Administration*, 30 (11), 1123-1140.

Van de Walle, S. (2006). The State of the World's Bureaucracies . In: *Journal of Comparative Policy Analysis*, 8 (4), 439-450.

Van Dooren, W. (2006). *Performance measurement in the Flemish public sector: A supply and demand approach*. Leuven, gepubliceerd doctoraal proefschrift.

Van Roosbroek, S. (2007). *Rethinking governance indicators: What can quality management tell us about the debate on governance indicators?* Paper EGPA annual conference. Study Group: Performance and Quality, Madrid, 19-22 september 2007.

van Thiel, S. & Leeuw, F. (2003). De prestatieparadox in de publieke sector. In: *Beleidswetenschap*, 17 (2), 123-144.

Van de walle, S., Sterck, M., van Dooren, W. & Bouckaert, G. (2004). *What you see is not necessarily what you get. Een verkenning van de mogelijkheden en moeilijkheden van internationale vergelijkingen van publieke sectoren op basis van indicatoren*. Leuven, Onderzoeksrapport.

Verlet, D. (2008). *Good governance, corporate governance, government governance: what's in a name? Een theoretische situering van Beter Bestuurlijk Beleid in Vlaanderen*. Brussel, Studiedienst van de Vlaamse Regering, SVR-nota 2008/4 (in druk).

Verlet, D., Reynaert, H. & Devos, C. (2005). *Burgers in Vlaamse grootsteden. Tevredenheid, vertrouwen, veiligheidsgevoel en participatie in Gent, Brugge en Antwerpen*. Brugge, Vanden Broele.

Zondergeld-Hamer, A. (2007). Een kwestie van goed bestuur. In: *Openbaar Bestuur*, 17 (10), 5-10.

# Hoofdstuk 3: Preventief gedrag bij beleidsevaluaties

Frank Bongers  
Provincie Noord-Brabant, 's-Hertogenbosch<sup>33</sup>

<b>1. Inleiding</b> .....	<b>63</b>
<b>2. De rol van beleidsevaluaties</b> .....	<b>63</b>
<b>3. Belemmerende factoren</b> .....	<b>65</b>
3.1. Institutionele factoren .....	65
3.2. Beleidsmatige factoren .....	67
3.3. Persoonlijke factoren .....	71
<b>4. Effect van beleidsevaluaties op preventief gedrag en op het beleidsproces</b> □ .....	<b>72</b>
<b>5. Tot besluit</b> .....	<b>74</b>
<b>Bibliografie</b> .....	<b>75</b>

---

<sup>33</sup> Dit hoofdstuk is op persoonlijke titel geschreven.

## 1. Inleiding

In sociaal- en gedragswetenschappelijk onderzoek speelt de uitdaging dat mensen hun gedrag kunnen aanpassen, omdat ze weten dat ze onderzocht (gaan) worden. Veel psychologische verklaringen voor menselijk gedrag zijn bijvoorbeeld gebaseerd op experimenteel onderzoek in laboratoria, waarbij de subjecten (vaak studenten) weten dat zij deelnemen aan een experiment. De vraag is dan aan de orde of de uitkomsten van dergelijk onderzoek ook buiten het laboratorium gelden. Het is bijvoorbeeld denkbaar dat onderzochte personen extra hun best gaan doen tijdens het experiment. In de sociologie en politicologie speelt dit probleem iets minder, want deze disciplines maken meer gebruik van veldonderzoek dan van experimenteel onderzoek, maar ook hier geldt dat verklaringen van menselijk gedrag invloed kunnen hebben op de gedragingen van individuen die onderzocht worden. Of zoals Trusted vaststelt '*[a]ny explanation offered [in human sciences, FB] can affect the behaviour of those whose behaviour the theory was devised to explain, that is the theory (and any other explanatory theory) has a direct effect on subsequent behaviour and hence may be self-defeating or self-justifying*' (1987, 32). De wetenschap dat men onderworpen wordt aan een (wetenschappelijk) onderzoek komt ook in beleid en politiek voor. Wanneer een beleidsvoerder (dat is de politiek verantwoordelijke samen met zijn ambtelijke staf; ofwel de uitvoerende macht) weet dat zijn beleid op een later moment geëvalueerd gaat worden, is de kans aanwezig dat hij (of zij) de inhoud en vormgeving van dat beleid aanpast met het oog op de evaluatie; zeker wanneer de "meetlat" voor het beoordelen van dit beleid bekend is. Hij kan ook politiek-strategisch gedrag ontwikkelen om een evaluatie van de agenda af te voeren. Gesteld kan worden dat het hier om een ongewenst effect gaat van beleidsevaluaties. Evaluaties hebben immers als doel om de kwaliteit van het openbaar bestuur en beleidsprocessen te verbeteren. Het is echter ook mogelijk dat de beleidsvoerder –omdat hij weet van de geagendeerde evaluatie– extra inspanningen gaat verrichten om doeleinden te halen. Dat is te bestempelen als gewenst preventief gedrag.

Ongewenst preventief gedrag kan leiden tot inflatie van beleid en van evaluaties. De kans neemt toe dat evaluaties rapporteren over beleid en beleidsvoerders die zich al "ingedekt" hebben tegen evaluaties. De vraag is wat de omvang van dit preventief gedrag is en of dit gedrag schadelijk is. Complicerende factor hierbij is dat wat de één als ongewenst preventief gedrag beschouwt, de ander juist als gewenst preventief gedrag ziet. Een analyse van het effect van evaluaties op preventief gedrag heeft dus veel haken en ogen. Het gaat immers ook over de integriteit van het openbaar bestuur. Mocht er sprake zijn van preventief gedrag is het onduidelijk of dit gedrag alleen door evaluaties wordt veroorzaakt (attributieprobleem) of dat andere factoren in het geding zijn.

## 2. De rol van beleidsevaluaties

Beleidsevaluatie –in dit hoofdstuk gedefinieerd als '*het beoordelen van de waargenomen inhoud, processen of de effecten van een beleid dan de hand van bepaalde criteria*' (Bressers & Hoogerwerf, 1991, 22)– draagt bij aan de *politieke*

*verantwoording* over het gevoerde beleid (ex post evaluatie) of over het nog te voeren beleid (ex ante evaluatie). In het –voor de beleidsvoerder gunstige geval– onderbouwt een evaluatie het gewenste of gesteunde beleid of ondergraaft de evaluatie een beleid dat de beleidsvoerder liever ziet eindigen. Ten tweede zetten evaluaties idealiter een *leerproces* in gang over de werking en effecten van een bepaald beleid. Ten derde kunnen beleidsevaluaties een bijdrage leveren aan een cultuur- en structuuromslag bij de overheid waarin *prestaties en effecten van overheidsorganisaties* meer centraal komen te staan.

Deze functies van beleidsevaluaties kunnen alleen gerealiseerd worden wanneer evaluaties voldoen aan criteria als onafhankelijkheid, betrouwbaarheid, geldigheid en bruikbaarheid. In bestuurskundige kring bestaat echter kritiek op de rol van beleidsevaluaties in het beleidsproces (Verlet & Devos, 2008). Zo valt te beluisteren dat een evaluatie het beleidproces op een onvoorziene en oneigenlijke manier kan beïnvloeden. Een voorbeeld daarvan is dat doelstellingen zodanig gekozen worden dat ze na afloop van het beleid relatief eenvoudig en objectief geëvalueerd kunnen worden. Bovendien kunnen doelstellingen minder ambitieus worden geformuleerd waardoor het beleid achteraf sneller als succesvol kan worden bestempeld. Evaluaties lijken daardoor aan te zetten tot preventief gedrag. Beleidsvoerders stemmen hun doelstellingen, doelgroepen, werkwijze en inspanningen af op wat haalbaar is (en de kans tot een positief eindoordeel vergroot) in plaats van wat (politiek) wenselijk is (en de kans tot een positief oordeel verkleint). Aanvullend valt er ook kritiek te bespeuren op beleidsonderzoek waarvan evaluaties een onderdeel zijn. Van Hoesel en Mevissen (2005) hebben een tiental kritiekpunten verzameld op beleidsonderzoek, in te delen naar de kwaliteit van het onderzoek zelf (betrouwbaarheid, gehanteerde methoden), het uitvoerende bureau (onafhankelijkheid, expertise) of het verschijnsel beleidsonderzoek in het algemeen (hoeveelheid, bruikbaarheid, benutting). Het gaat hier echter om afgeronde onderzoeken, terwijl preventief gedrag uitgelokt wordt door beleidsonderzoek dat nog moet plaatsvinden, gelijk een ex post evaluatie.

In de beleidspraktijk komt het regelmatig voor om doelstellingen vooraf relatief open en onduidelijk te formuleren, zodat er voldoende “speelruimte” tijdens het beleid resteert. Dat gebeurt ook om achteraf minder gemakkelijk vastgepind te kunnen worden op het niet behalen van doelen. Het gaat echter nogal ver om het vooruitzicht op een beleidsevaluatie één op één te verbinden met het formuleren van vage of eenvoudig haalbare doelstellingen of andere vormen van preventief gedrag. Deze vooronderstelling geeft een nog niet uitgevoerde evaluatie een wel erg grote betekenis in een beleidsproces. Zelfs van afgerond evaluatieonderzoek is maar de vraag welke invloed zij heeft op beleid. Pröpper (1998) onderscheidt daarbij het vervangen, bijstellen, continueren, opschorten en bijstellen van beleid, maar het is ook mogelijk dat er helemaal geen consequentie is. De rol van beleidsonderzoek - en dus beleidsevaluatie - in beleid is voor sommige wetenschappers sowieso minder groot dan vaak wordt aangenomen. Vooral vanuit de hoek van de *Utilization of Knowledge School* zijn hier kanttekeningen bij geplaatst. Caplan verwoordt het als volgt: *‘the level of knowledge utilization is not so much the result of a slow flow of relevant and valid knowledge from knowledge producers to policymakers, but is due more to factors involving values, ideology, and decision-making styles’* (geciteerd in Dunn, 1994, 56). Aanvullend is er kritiek op beleidsevaluaties, omdat zij niet

ondubbelzinnig antwoord kunnen geven op de vraag of een programma succesvol is. Abma stelt daarover vast dat evaluatoren ‘*onheilsboodschappers*’ worden ‘*die niet in staat zijn om ondubbelzinnige antwoorden te geven op de vraag waarom programma’s falen of succesvol zijn. Resultaten uit evaluaties hebben nauwelijks direct effect op de besluitvorming*’ (1996, 7). Het is tevens onduidelijk of een mogelijke, latere beleidsevaluatie op het netvlies staat van de beleidsvoerder en tot ongewenst preventief gedrag zal leiden.<sup>34</sup> Deze zienswijze impliceert ook dat een beleidsevaluatie de officiële beleidsdoeleinden als referentiepunt neemt. Dat is in de meeste studies ook aan de orde, maar er zijn – zoals later wordt behandeld – ook andere vormen van evaluatieonderzoek die niet uitgaan van de officiële doelstellingen (Huppe, 2007; Bressers, 1998; Bressers & Hoogerwerf, 1991) of die deze doelstellingen ter discussie stellen, zoals in de maatschappijkritische theorie gebeurt (zie daarvoor Abma, 1996, 11 e.v.).

### 3. Belemmerende factoren

Er bestaan factoren die zorgen dat beleidsvoerders geen of slechts beperkt ongewenst preventief gedrag vertonen. Alvorens deze factoren te benoemen, wordt eerst stilgestaan bij de term “preventief gedrag”. Daarvan is –in een beleidscontext– sprake wanneer een beleidsvoerder zijn gedrag vooruitlopend afstemt op een bepaalde gebeurtenis, bijvoorbeeld een beleidsevaluatie. Deze afstemming heeft tot doel om zijn gedrag in een positief daglicht van deze gebeurtenis te kunnen plaatsen. Dit preventief gedrag kan gewenst of ongewenst zijn. Ongewenst preventief is aan de orde wanneer de beleidsvoerder door zijn gedrag de doelmatigheid en doeltreffendheid van een beleid onder druk zet, bijvoorbeeld door het formuleren van minder ambitieuze doelstellingen of de kaderstellende of controlerende taken van een volksvertegenwoordiging bemoeilijkt met dat gedrag. De factoren die van invloed zijn op de mate waarin een beleidsvoerder preventief gedrag vertoont, spelen op drie onderling verbonden niveaus: (1) de institutie, (2) het beleid en (3) de beleidsvoerder.

#### 3.1. Institutionele factoren

Institutionele factoren hebben betrekking op instituties als waarden, normen en regels die het preventief gedrag van beleidsvoerders mee sturen en bepalen. Op dit niveau worden twee dominante factoren onderscheiden: de organisatorische inbedding van beleidsevaluaties en uiteenlopende corrigerende, institutionele mechanismen die ongewenst preventief gedrag tegengaan.

---

<sup>34</sup> Een ander treffend voorbeeld van ongewenst preventief gedrag is het beperkt opzetten en bijhouden van administratieve gegevens die een belangrijke basis zijn voor een evaluatie. In de praktijk blijkt nogal eens dat het ontbreken van dergelijke administratieve gegevens, bijvoorbeeld over deelnemers van een beleidsprogramma, voor problemen zorgt in een evaluatie: het steekproefkader is niet compleet, het verspreiden van een vragenlijst onder deelnemers wordt vertraagd, maar ook het ontbreken van voortgangsrapportages valt hieronder. Overigens is hier ook vaak sprake van onbewust gedrag of heeft het opzetten van een dergelijk bestand geen prioriteit. Het tegenovergestelde kan ook. Het uitzicht op een evaluatie leidt tot nauwkeurige registratie van administratieve gegevens door de beleidsvoerder.

## *Organisatorische inbedding van beleidsevaluatie*

Ondanks dat een beleidsvoerder vaak een grote stem heeft in het besluit om een evaluatie uit te voeren, betekent dat geenszins dat hij de opzet en uitvoering van een evaluatie alleen kan sturen of bepalen. Ten eerste zal de beleidsvoerder –zeker wanneer het beleid een resultaat is van interactieve processen waarbij uiteenlopende partijen binnen en buiten de overheid betrokken zijn (zie paragraaf 3.2)– rekening moeten houden met opvattingen van derden over de evaluatie. Ten tweede worden evaluaties steeds vaker door externe partijen geagendeerd en/of uitgevoerd, bijvoorbeeld door het Belgische Rekenhof of de Nederlandse Algemene Rekenkamer. Ook wetenschappers en maatschappelijke organisaties voeren evaluaties uit. Rekenkamers hebben het voordeel dat zij zelf bepalen welk beleid zij evalueren (en hoe). Hierop heeft de beleidsvoerder nauwelijks invloed. Hij heeft wel meer invloed op evaluatieonderzoek dat hij laat uitvoeren door externe bureaus. Hij is dan vaak betrokken bij de formulering van de offerteaanvraag, bij de keuze van het bureau, en de rapportage van de evaluatie. Soms is hij ook betrokken als lid of voorzitter van de begeleidingscommissie van het onderzoek. Vooral dat laatste is vanuit een onafhankelijke en objectieve evaluatie minder gewenst. Beleidsvoerders die lid zijn van een begeleidingscommissie hebben de neiging om negatieve resultaten van het beleid minder op te nemen in het rapport en positieve resultaten juist meer aandacht te geven. Het is dan aan het externe bureau of de andere leden van de begeleidingscommissie om ongewenste sturing tegen te gaan. Overigens kan een kritische houding van de opdrachtgever over inhoud en methode ook ingegeven zijn door de rol die het onderzoek nog gaat spelen in een beleidsproces. Steeds vaker besluiten overheidsorganisaties om de verantwoordelijke beleidsmedewerkers geen lid te maken van een begeleidingscommissie of hen een ondergeschikte rol te geven, maar de ervaring leert dat het nog geen gemeengoed is. Een uitzondering op deze praktijk is een ministerie dat de begeleidingscommissie van een evaluatieonderzoek altijd laat voorzitten door een medewerker die niet bij het dossier betrokken is. De commissie wordt aangevuld met externe deskundigen en belanghebbenden, terwijl het secretariaat en het aanleveren van relevante informatie wel neergelegd wordt bij de beleidsafdeling. Een ander voorbeeld betreft een provincie die externe deskundigen begeleidingscommissies van evaluatieonderzoeken laat voorzitten.

Wanneer een beleidsvoerder de (latere) opzet en uitvoering van een evaluatie niet onder controle heeft, kan dat leiden tot het formuleren van minder ambitieuze en minder concrete doelstellingen. Dit kan dus (ongewenst) preventief gedrag bevorderen, want de beleidsvoerder kan op deze situatie anticiperen door de toekomstige evaluatie te sturen met voor hem gunstige doelen. Daar staat echter tegenover dat een onafhankelijke evaluatieonderzoeker met open vizier een evaluatie zal uitvoeren waarbij niet alleen de vooraf vastgestelde doeleinden als referentie zullen dienen. Daarnaast is het –zeker bij grote en langlopende beleidsprogramma's– de vraag of de beleidsvoerder na enkele jaren nog betrokken is bij het dossier. Naarmate de evaluatie verder terug in de tijd ligt, neemt de kans dat de huidige beleidsvoerder nog direct betrokken is bij de evaluatie af.<sup>35</sup>

---

<sup>35</sup> Veel evaluatieonderzoeken strekken zich uit over enkele jaren van een beleid. In deze periode zijn er bijna altijd wel personele wisselingen geweest. Een evaluatie heeft daardoor bijna altijd te maken met meerdere beleidsvoerders die achtereenvolgens betrokken waren bij een beleid en daar een eigen invulling aan gaven.



## *Corrigerende mechanismen*

Er bestaan binnen en buiten het openbaar bestuur mechanismen die helpen te voorkomen dat beleidsvoerders ongewenst preventief gedrag vertonen. Ten eerste draagt de volksvertegenwoordiging (wetgevende macht) bij aan de inhoudelijke koers van een beleidsvoerder (uitvoerende macht). De ambities van de beleidsvoerder zullen aan moeten sluiten bij de wensen van de volksvertegenwoordiging. Uiteraard bestaat er beleidsvrijheid om een en ander naar eigen inzicht uit te werken, maar het primaat van kaderstelling en controle ligt bij de volksvertegenwoordiging. Ten tweede onderhouden beleidsvoerders contacten met collega's, doelgroepen en experts. Deze contacten dragen bij aan de invulling en concretisering van doeleinden, maar maken ook vergelijkingen tussen doelen van vrijwel identieke beleidsprogramma's mogelijk. Zo zijn er honderden gemeenten die op identieke terreinen beleid voeren. Onderlinge vergelijking zowel in doeleinden, aanpak als prestaties zijn dan eenvoudig mogelijk en dragen bij aan beleidsontwikkeling en –evaluatie in eigen huis. Beleidsvergelijkingen kunnen tekortkomingen in het beleid bloot leggen.

### **3.2. Beleidsmatige factoren**

Beleidsmatige factoren hebben betrekking op karakteristieken van beleidsprocessen die ertoe leiden dat ongewenst preventief gedrag minder opportuun is. In dit verband spelen de volgende eigenschappen een rol: (1) model van beleidsprocessen; (2) agendering van evaluaties; (3) interactieve doelformulering en evaluaties; (4) dynamiek van beleidsprocessen; en (5) de beleidsinhoud.

#### *Model van beleidsprocessen*

Het verloop van een beleidsproces wordt in de beleidssociologische literatuur vaak als een lineair en rationeel model voorgesteld (Hoogerwerf & Herweijer, 1998; Dunn, 1994). In dat model handelt een centrale partij, meestal de beleidsvoerder, als een alleswetende en allesbepalende speler die een aantal logische stappen zet waarmee een beleidscyclus (van agendering tot en met evaluatie) voltooid wordt. Deze werkelijkheid bestaat hooguit in de vakliteratuur en fungeert daar vooral als een nuttig en ordenend analysekader. De idee dat een beleidsvoerder vooraf met het beïnvloeden van doelstellingen de latere beleidsevaluatie op een oneigenlijke manier kan sturen, past wellicht in dit model maar niet in de beleidspraktijk. Ten eerste veronderstelt dit namelijk dat er sprake is van een echt "startschot" van een beleid. Nu zal dat gedeeltelijk wel het geval zijn, maar beleid bouwt bijna altijd voort op eerder beleid en bovendien baseert beleid zich continu op voortschrijdend inzicht. Het formele startschot is dus niet altijd zichtbaar. Ten tweede veronderstelt deze visie dat beleid pas aan het einde wordt geëvalueerd. Ex post evaluaties zijn een gangbare praktijk, maar zij veronderstellen wel dat een beleid eindigt. Niets is minder waar. Veel beleid continueert ongeacht uitkomsten van een evaluatie. Bovendien lijkt deze visie zich vooral uit te spreken voor inhoudelijke evaluaties, terwijl veel evaluaties zich richten op het proces of juist voorafgaand aan een beleid worden uitgevoerd (ex ante evaluaties). Vooral ex ante evaluaties bieden de mogelijkheid om vooraf te toetsen of beleidsdoelstellingen haalbaar zijn.



## *Agendering van evaluaties*

Anno 2008 kan worden vastgesteld dat beleidsevaluaties steeds hoger op de bestuurlijke en politieke agenda staan. Het is empirisch echter lastig vast te stellen wat de totale omvang van evaluatieonderzoek in Nederland en Vlaanderen is (in mensjaren en budgetten) en wat het aandeel van evaluatieonderzoek in het totale beleidsonderzoek is. Het Nederlandse Evaluatie Overzicht Rijksoverheid registreert ongeveer 250 evaluaties per jaar (De Jong & Kabel, 2005). De aandacht voor beleidsresultaten en –effecten is één van de redenen waarom overheden steeds vaker evalueren. Evaluaties komen ook voort uit wettelijke verplichtingen en uit de externe agendering van evaluatie, bijvoorbeeld door rekenkamercommissies die met de invoering van de nieuwe Gemeentewet in Nederland verplicht zijn. Bovendien neemt de belangstelling voor de evaluatie van overheidsbeleid ook toe onder betrokkenen die niet direct tot de harde kern van het politiek-bestuurlijke systeem worden gerekend, zoals wetenschappers, “gewone” burgers en maatschappelijke organisaties.

Ondanks deze grote aandacht geldt nog steeds dat bij veel dossiers vooraf niet bekend is of een beleid geëvalueerd zal gaan worden en hoe. Nu is het ontegenzeggelijk zo dat veel nieuw beleid gepaard gaat met de mededeling dat het later geëvalueerd gaat worden, zoals bijvoorbeeld in begrotingen gebeurt.<sup>36</sup> Echter, voor veel beleid geldt dat deze informatie vooraf niet bekend is. Er zijn zelfs evaluaties die in het geheim worden uitgevoerd. Waarom zouden beleidsvoerders zich vooraf dan gaan indekken tegen een evaluatie? Dit zou hooguit onderbouwd kunnen worden met een “no regret”-scenario: voor het geval dat er een evaluatie komt, gaat de beleidsvoerder over tot preventief gedrag. Of dat een reële optie is, valt te betwijfelen. De volksvertegenwoordiging heeft immers een kaderstellende en controlerende rol en kijkt daardoor mee over de schouder van de beleidsvoerder; zeker waar het een belangrijk dossier betreft. De volksvertegenwoordiging in Nederland kan er in dit verband ook voor kiezen om bijvoorbeeld achteraf een parlementair onderzoek in te stellen, zoals het Parlementair Onderzoek Onderwijsvernieuwingen of zelf het zware middel van de parlementaire enquête, zoals de Parlementaire Enquête Bijlmerramp. Deze middelen zijn niet noodzakelijkerwijs evaluatieonderzoek, maar zij komen wel voort uit de maatschappelijke en politieke behoefte om beleidsterreinen waar langere tijd problemen spelen of gebeurtenissen die tot de nodige maatschappelijke en politieke commotie leiden een keer goed door te lichten. Deze onderzoeken – die vaak wel een evaluatief karakter hebben – richten zich op ontwikkelingen en beleid die zich over meerdere jaren of zelfs decennia uitstrekken. Preventief handelen kan hier gewoon niet aan de orde zijn geweest. Het onderzoek komt immers onverwacht en vaak pas jaren nadat beleid loopt (en tussentijds vaak is aangepast).

---

<sup>36</sup> In dit verband is de opkomst van zogenaamde “beleidsdoorlichtingen” bij de Nederlandse rijksoverheid relevant. Dit nieuwe evaluatie-instrument verplicht departementen periodiek, bijvoorbeeld eens in de vijf jaar, hun beleid door te lichten en aandacht te geven aan een aantal terugkerende vragen. Deze systematiek is ontwikkeld, omdat een systematische programmering van beleidsevaluaties ontbrak en de meeste evaluaties zich bovendien op één beleidsinstrument richtten. Een eerste evaluatie van deze doorlichtingen toont dat de bruikbaarheid en kwaliteit nog tegenvallen, maar de doorlichtingen zorgen wel dat er meer aandacht is gekomen voor systematische programmering van evaluatieonderzoek (Pestman & Van Wijk, 2008).

Een relatief nieuwe impuls voor beleidsevaluaties komt vanuit eerder genoemde lokale en provinciale rekenkamer(commissie)s in Nederland. Deze onafhankelijke organen zijn in de afgelopen jaren als gevolg van invoering van dualisme opgericht. Aangezien deze organen hun onderzoeksagenda zelf bepalen, verrichten zij relatief veel onderzoek naar beleid waarvan vooraf helemaal niet vaststond dat het geëvalueerd zou gaan worden. Onlangs is een eerste evaluatie naar het functioneren van rekenkamer(commissie)s verschenen (B&A Groep, 2007). Een resultaat van deze evaluatie is dat rekenkamer(commissie)s over voldoende randvoorwaarden (budget, ondersteuning) beschikken om hun taken uit te voeren. Ook is er voldoende deskundigheid en wordt er geïnvesteerd in het verbeteren van deskundigheid. De aanbevelingen van de rekenkamer(commissie)s worden over het algemeen wel overgenomen door de raad, maar slechts een minderheid van de rapporten wordt in de raad behandeld. Het evaluatierapport ontvouwt twee perspectieven op rekenkamers, namelijk een perspectief dat de rekenkamer ziet als een ondersteuning van politiek-bestuurlijke vernieuwing en reflectie bij de beleidsontwikkeling. Het andere perspectief stelt het principe van rekenschap centraal. Hierbij wordt het inzichtelijk maken van de kwaliteit van het bestuurlijk handelen bedoeld inclusief het afleggen van verantwoording daarover. De bijdrage van rekenkameronderzoek aan beleidsontwikkeling geeft aan dat evaluatieonderzoek wel degelijk invloed kan hebben op het beleid (en dus ook preventief gedrag in de hand kan werken). Voor zover er sprake is van doorwerking lijkt zich dat vooral te voltrekken rondom het SMART maken van doelstellingen (Specifiek, Meetbaar, Afgesproken, Realistisch en Tijdsgebonden). De meeste rekenkamercommissies hangen hun onderzoek immers sterk op aan het zogenaamde SMART model. Beleidsdoelstellingen worden dan beoordeeld op de mate waarin zij specifiek, meetbaar, acceptabel, realistisch en tijdgebonden zijn geformuleerd. Het is echter de vraag of de kwaliteit van beleid en beleidsevaluaties gediend is met de toepassing van dit strakke kader, maar het effect is wel dat beleidsvoerders steeds meer aandacht geven aan het SMART formuleren van doelstellingen. Dit is echter een magere invloed van evaluaties op het beleidsproces.

### *Interactieve doelformulering en evaluaties*

Aanvullend op het vorige thema is er steeds meer sprake van interactieve beleidsprocessen waarbij doelgroepen van (nieuw) beleid in samenspraak met de beleidsvoerder spreken over de inhoud en de vorm van het te voeren beleid. Daaraan gekoppeld zit vaak ook de vraag welke doelstellingen een beleid moet bereiken. Kortom, het formuleren van doelstellingen is dan niet alleen een zaak van de beleidsvoerder, maar ook van derden zoals burgers, bedrijven en maatschappelijke instellingen. Uiteraard betekent deze betrokkenheid niet dat er dan automatisch meer concrete en ambitieuze doelstellingen worden vastgesteld. Meer partijen, betekent ook meer belangen en meer complexiteit. Dit kan er juist toe leiden dat er doelstellingen worden vastgesteld waarin de meeste partijen zich kunnen vinden. Deze neiging tot consensus kan daardoor leiden tot minder ambitieuze en conflictueuze doelstellingen. Daar heeft een potentiële evaluatie niks mee te maken. Overigens kan zich hier een interessant twistpunt ontwikkelen. Het is denkbaar dat de beleidsvoerder doelstellingen prefereert waarvan hij het bereiken beter onder controle heeft. Hij zou dan kunnen aansturen - en dat is wel preventief gedrag - op

doelstellingen die gerelateerd zijn aan “output” van een beleidsafdeling, zoals het verstrekken van een “x” aantal cultuursubsidies of het verstrekken van een “x” aantal adviezen over het gebruik van computers en internet aan kleine ondernemingen. De betrokken maatschappelijke organisaties kunnen echter aansturen op doelstellingen die de “outcomes” (ofwel maatschappelijke effecten) van beleidsinstrumenten en –programma’s die overheidsorganisaties centraal stellen. Het gaat dan – in het licht van de genoemde voorbeelden – om toename van cultuurparticipatie onder jongeren of kostenbesparingen dankzij computer bij bedrijven. Dat zijn echter indicatoren die moeilijker te meten zijn en die beleidsvoerders minder makkelijk “onder controle” hebben. Ook omdat andere factoren deze effecten bepalen, het zogenaamde attributieprobleem (Bongers, 2008).

De idee om andere partijen te betrekken bij de beleidsontwikkeling en -evaluatie past bij wat in beleidswetenschappelijke kringen bekend staat als vierde generatie evaluatie (Bongers & Den Hertog, 2007; Abma, 1996). In een historische schets van evaluatieonderzoek onderscheiden Guba en Lincoln vier generaties van evalueren (1989, 22-31). De eerste generatie duiden zij aan met de term “meten”: de beleidsonderzoeker stelt zich op als een technicus en verzamelt alleen kwantitatieve gegevens (“monitoring”) aangeduid. In de tweede generatie staat “beschrijven” centraal. De beleidsonderzoeker beschrijft de sterke en zwakke kanten van het beleid in relatie tot de gestelde doeleinden. In de derde generatie staat “beoordelen” voorop. De beleidsonderzoeker krijgt een rol als rechter omdat hij een finaal oordeel velt over het beleid. Ondanks dat elke generatie een verbetering is ten opzichte van de voorgaande generatie stellen Guba en Lincoln (1989) enkele gemeenschappelijke tekortkomingen van deze generaties vast. Ten eerste richten deze drie generaties zich op de doeleinden van de leiding die een evaluatie laat uitvoeren. De opvatting dat beleidsvoerders anticiperen op een mogelijke evaluatie door beleidsdoelstellingen en dus evaluatiecriteria aan te passen, past bij deze eerste drie generaties. Ten tweede kenmerken deze generaties zich door een geloof in een maatschappelijke overeenstemming over geldende waarden. Deze consensus ontbreekt in een pluriforme samenleving. Het is dan ook de vraag welke waarden domineren in een evaluatieonderzoek. Ook komen zij tot de conclusie dat veel beleidsevaluaties niet gebruikt worden in de besluitvorming en belanghebbenden worden in een evaluatieonderzoek doorgaans alleen benaderd als informatieverstrekker. Guba en Lincoln (1989) pleiten dan ook voor een vierde generatie evaluatieonderzoek die zij “onderhandelen” noemen. In deze generatie komen de onderzoeksvragen en afbakening van een evaluatie in samenspraak met belanghebbenden tot stand (dus niet alleen in overleg tussen evaluatieonderzoeker en de formele opdrachtgever). De evaluatiecriteria in een onderzoek worden samengesteld op basis van opvattingen (“claims”, “concerns” en “issues”) van uiteenlopende belanghebbenden en niet alleen op basis van de officiële beleidsdoeleinden. Uiteraard kunnen deze wel overlappen; zeker wanneer doeleinden op basis van een interactief proces tot stand zijn gekomen. Dat betekent dat de aanpak van beleidsevaluaties recht moet doen aan deze interactieve kenmerken van beleidsprocessen, bijvoorbeeld door meer nadruk op communicatie en onderhandelen. Met andere woorden: wie derden betreft bij het ontwikkelen van beleid zal hen ook moeten betrekken bij de vraag op welke wijze dit beleid geëvalueerd kan worden. Zelfs als deze externe partijen niet betrokken waren bij de beleidsontwikkeling geldt dat een vierde generatie evaluatie

normen kan hanteren die de beleidsvoerder bij het opstellen van het beleid nooit voor ogen had. Deze aanpak impliceert dat een evaluatie niet alleen de officiële beleidsdoeleinden als referentiepunt zal nemen. In de praktijk komen zogenaamde vierde generatie beleidsevaluaties nog niet veel voor.

### *Dynamiek binnen en buiten beleidsprocessen*

Preventief gedrag heeft wellicht meer zin in een statische omgeving, omdat dit gedrag immers veronderstelt dat een toekomstige evaluatie en de gehanteerde methoden en criteria vastliggen. Deze werkelijkheid bestaat alleen bij beleidsvoerders die ooit een aantal doelstellingen hebben geformuleerd en politiek laten accorderen en die vervolgens gedurende het beleidsproces blijven vasthouden aan deze doelstellingen. Formeel gesproken is dat correct, maar beleid, politiek en samenleving zijn complex en dynamisch. Doelstellingen kunnen achterhaald raken door nieuwe ontwikkelingen. Dat hoeft overigens niet voor alle doelstellingen te gelden. Denk aan meer fundamentele doelstellingen over inkomensverdeling, economische ontwikkeling of het handhaven van de rechtstaat. Ook doelstellingen op deze terreinen worden uitgewerkt in meer concrete doelstellingen en beleidsprogramma's en zijn dus onderhevig aan maatschappelijke, economische, politieke en technologische ontwikkelingen. Een voorbeeld van een beleidsdoelstelling die door de realiteit wordt ingehaald betreft de doelstelling van het beleids- en ICT-experiment Kenniswijk in de regio Eindhoven/Helmond. Dit experiment had als doelstelling om binnen twee jaar een consumentenmarkt van de toekomst te creëren met een innovatief karakter en internationale uitstraling, waarbinnen op basis van hoogwaardige infrastructuren de ontwikkeling en toepassing van interactieve en elektronische dienstverlening voor de consument op open wijze zouden plaatsvinden. Deze ambitieuze doelstelling werd geformuleerd in een tijd van economische groei en de internethype. Na aanvang van het experiment bleek echter al snel dat economische groei afnam en dat marktpartijen minder geneigd waren te investeren in infrastructuur en diensten. Deze ontwikkelingen hebben de vaart uit het project gehaald en de doelstelling eigenlijk ook gedeeltelijk achterhaald gemaakt. Overigens was Kenniswijk een experiment dus het niet behalen van de doelstelling doet niets af aan het nut en het leereffect van dit middel (Dialogic, 2006). Ook geheel andere ontwikkelingen zoals het aantreden van een nieuw kabinet of gemeentebestuur met een andere politieke signatuur kunnen doelstellingen van het voorgaande beleidsvoerders ongewenst of verouderd maken.

Bovenstaande voorbeelden illustreren de dynamiek zowel binnen als buiten beleidsprocessen. Deze dynamiek kan er in veel gevallen toe leiden dat doelstellingen in een ander daglicht komen staan of zelfs aanpassing vergen. Dat laatste is overigens niet eenvoudig, omdat dat vaak ook een politieke goedkeuring vergt. Het vooraf bedenken van relatief brede of onduidelijke doeleinden heeft dan als voordeel dat de beleidsvoerder flexibiliteit creëert om effectief te kunnen handelen onder veranderende omstandigheden. Het anticiperende gedrag wordt dan dus geprikkeld door andere factoren dan een mogelijke beleidsevaluatie. En hier wordt een fundamenteel punt geraakt: in hoeverre is het mogelijk om preventief gedrag te herleiden tot een evaluatie en is de invloed van een mogelijke evaluatie op dit gedrag te isoleren tegenover andere factoren?

## *Beleidsinhoud*

De hiervoor geschetste factoren hangen sterk samen met methodologische, procedurele en politieke kanten van beleidsprocessen. De beleidsinhoud is vooralsnog buiten beschouwing gebleven. Vanuit deze inhoud kan echter ook een preventieve werking uitgaan. Overheidsbeleid richt zich op uiteenlopende terreinen als economie, zorg, onderwijs en veiligheid. De inhoudelijke betrokkenheid van beleidsmakers bij deze onderwerpen voedt hun nieuwsgierigheid over de vraag of beleid eventuele knelpunten in deze terreinen vermindert. Dat vereist effecten- en evaluatieonderzoek.

### **3.3. Persoonlijke factoren**

Het derde niveau dat preventief gedrag kan afremmen, betreft persoonlijke factoren. Het gaat daarbij om persoonlijke eigenschappen van de beleidsvoerder die van invloed zijn op de ontvankelijkheid voor preventief gedrag. Denk hierbij aan de ervaring van de beleidsvoerder met eerdere evaluaties, zijn positie in de organisatie en het politieke krachtenveld, zijn welwillendheid en ambities en zijn eventuele risicomijdende gedrag. De beleidsvoerder die langere tijd meedraait in het politieke circuit kan beter inschatten welke effecten een evaluatie op het beleid of op zijn positie kunnen hebben. Niet elke evaluatie is “politieke springstof”. Er bestaat nogal een verschil tussen evaluaties die voor de vorm worden uitgevoerd, evaluaties die voeding geven aan een actueel, dominant of controversieel politiek onderwerp of evaluaties die tot grote problemen kunnen leiden. Uiteraard is dit vooraf niet allemaal in te schatten, maar de ervaren beleidsvoerder zal met enige nuance naar een evaluatieonderzoek kijken. Een en ander hangt ook af van de positie van een beleidsvoerder: ligt er nog een carrièrepad open in het openbaar bestuur (waarbij geen negatieve berichtgeving past) of gaat de beleidsvoerder bijna met pensioen. Bovendien zijn er voldoende ambitieuze en enthousiaste beleidsvoerders die niets ophebben met preventief gedrag en het risico durven te nemen om later ook “afgerekend” te worden op basis van het gevoerde beleid. Sterker nog: de wetenschap dat er een evaluatie aankomt, kan de beleidsvoerder uitnodigen om extra inspanningen te plegen om van het beleid een succes te maken. Dit lijkt op het Hawthorne-effect.<sup>37</sup> Een goed evaluatieonderzoek houdt rekening met factoren die de beleidsvoerder niet onder controle heeft, maar die wel van invloed zijn op het welslagen van het beleid. We weten allemaal dat een economische recessie “altijd” veroorzaakt wordt door internationale ontwikkelingen, terwijl een economische voorspoed “altijd” te danken is aan het kabinetsbeleid.

---

<sup>37</sup> Het Hawthorne-effect is het effect van waarnemers op de productiviteit van groepen die zij onderzoeken. Dit effect is in de jaren twintig en dertig van de vorige eeuw vastgesteld toen onderzoekers op allerlei manieren probeerden de productiviteit van arbeiders in een fabriek te vergroten. Wat zij ook deden, de productiviteit bleef toenemen. Deze stijging had kennelijk te maken met de aanwezigheid van onderzoekers, hetgeen leidde tot vooraf niet voorziene groepsprocessen (Morgan, 1986).

## 4. Effect van beleidsevaluaties op preventief gedrag en op het beleidsproces

Ondanks de belemmerende factoren is preventief gedrag niet helemaal uit te sluiten. Het probleem is echter hoe preventief vastgesteld kan worden. Daarbij speelt ook de vraag of dat gedrag op grote schaal voorkomt en of het schadelijk is en hoe het kan worden voorkomen. Kortom, laten beleidsvoerders zich bij het bepalen en inrichten van een beleid leiden door specifieke evaluatiemethoden- en –criteria en is dat gedrag dan ook onwenselijk.

Allereerst speelt het probleem van het empirisch vaststellen van de effecten op ongewenst gedrag dat voortkomt uit de agendering van een beleidsevaluatie. Om dat goed vast te stellen zouden er drie vergelijkbare groepen beleidsvoerders moeten worden samengesteld die een (vrijwel) identiek beleidsprobleem gaan aanpakken. De experimentele groep beleidsvoerders krijgt te horen dat er na afloop van het beleid een evaluatie zal worden uitgevoerd. Een controlegroep krijgt expliciet te horen dat er geen evaluatie plaats zal vinden, terwijl de andere controlegroep hierover in het ongewisse blijft. Idealiter weten de deelnemers aan dit veldexperiment niet dat zij deelnemen aan een onderzoek. Dit ook om sociaal wenselijk gedrag te voorkomen. Desondanks zal het lastig blijven het effect precies vast te stellen. Het is immers nauwelijks denkbaar dat de omstandigheden voor deze drie groepen gelijk zijn. Bovendien is het lastig om drie vrijwel identieke groepen beleidsvoerders te selecteren die ook nog eens dezelfde problemen gaan aanpakken (elk probleem heeft een voorgeschiedenis, etc.). Een mogelijke ingang zouden vergelijkbare gemeenten met identieke problemen kunnen zijn. Uiteraard zal bij een dergelijk veldexperiment vooraf ook bepaald moeten worden wat er onder “ongewenst, preventief gedrag” moet worden verstaan. De onderzochte beleidsvoerder kan daar heel anders over denken dan de onderzoeker. Denk aan een situatie waarin de beleidsvoerder –in dit geval een beleidsmedewerker van een gemeente– gedrag vertoont dat gericht is op “damage control” voor de wethouder. Onderzoek naar preventief gedrag vergt dus ook inlevingsvermogen (“Verstehen”) van de onderzoeker. Wanneer de evaluatie de kans vergroot op een beschadiging van de wethouder (schepen) kan de medewerker proberen de latere evaluatie te beïnvloeden. Hij is immers ook aangesteld om de wethouder te dienen. Mochten de groepen en omstandigheden in het veldexperiment goed vergelijkbaar zijn, dan is het mogelijk om eventuele verschillen in de wijze waarop het beleid tot stand komt en de inhoud van dat beleid te herleiden tot de kennis dat er (geen) evaluatie aankomt.

Al zou de aankondiging dat er een evaluatie op komst is, leiden tot onwenselijk gedrag is natuurlijk de vraag aan de orde wat dan de omvang van dit ongewenst gedrag is en welke schade dat ongewenste gedrag veroorzaakt. Kunnen de resultaten van een veldexperiment zoals hierboven beschreven worden veralgemeniseerd naar *alle* beleidsvoerders, wetende dat er allerlei condities zijn die ongewenst gedrag tegengaan. Met andere woorden: is het mogelijk om vast te stellen dat in x% van alle aangekondigde beleidsevaluaties de aankondiging tot ongewenst gedrag leidt. Daaraan gekoppeld is de vraag of dit ongewenste gedrag dan ook schadelijk is. Het is schadelijk wanneer het gedrag ertoe leidt dat de kaderstellende en controlerende taak van de volksvertegenwoordiging in het geding is. Of wanneer de doeltreffendheid en doelmatigheid van het beleid hierdoor aangetast dreigen te



worden. Dat is natuurlijk aan de orde wanneer beleidsdoelstellingen bewust (naar beneden) worden bijgesteld om achteraf makkelijker beter te kunnen scoren op de lat van de evaluatiecriteria of wanneer administratieve gegevens die nodig zijn voor een beleidsevaluatie onvoldoende worden geregistreerd. Dat is natuurlijk makkelijker te realiseren wanneer de evaluatiecriteria bij aanvang van het beleid bekend zijn. De beleidsvoerder en de beleidsuitvoerder kunnen dan naar deze norm gaan werken. Dit is een van de minder gewenste effecten bij sturing op kwantificeerbare normen. Wanneer het evalueren van beleid en het meten van effecten vervallen tot het louter meten van scores op bepaalde normen beweegt de evaluatie zich echter op het vlak van monitoring. Hiervoor is aangetoond dat dit een beperkte en smalle vorm van evalueren is die geenszins voldoet aan de wensen die er in het openbaar bestuur (en daarbuiten) leven ten aanzien van politieke oordeelsvorming op basis van een gedegen beleidsevaluatie. Het punt bij beleidsevaluaties is daarbij wel dat zij doorgaans geen uitspraak doen over doelstellingen van het beleid. Deze zijn immers vaak politiek bepaald en de evaluatieonderzoeker begeeft zich op een hellend vlak wanneer hij zijn onderzoek gebruikt om deze doelstellingen inhoudelijk te beoordelen. Wel is het mogelijk om aan te geven of de doelstellingen makkelijker of goedkoper gerealiseerd hadden kunnen worden of dat gegeven de omstandigheden meer mogelijk was geweest. Wanneer daar structureel sprake van is, kan de conclusie getrokken worden dat beleidsdoelstellingen weinig ambitieus zijn en dat zou het resultaat kunnen zijn van preventief gedrag.

In het begin van dit hoofdstuk is gesteld dat beleidsvoerders bestaan uit de politiek-ambtelijke leiding en de uitvoerende staf. Tussen deze groepen – vooral tussen politici en ambtenaren – kunnen verschillende inzichten bestaan over het gebruik en effecten van beleidsevaluaties op gedragingen van beide groepen. Voor de politicus zal een beleidsevaluatie vooral betekenen dat er verantwoording wordt afgelegd over het gevoerde beleid. Daaraan kleven natuurlijk politieke oordelen door de volksvertegenwoordiging. Ambtenaren kunnen zich “verschuilen” achter de minister, provinciebestuurder of wethouder; ondanks dat ze vaak langer en intensiever bij een bepaald dossier zijn betrokken. Aangezien de verantwoordingsfunctie van beleidsevaluatie meer politieke lading heeft, zal de verantwoordelijke politicus meer behoefte hebben aan een positief oordeel. Tenzij hij pas kort betrokken is bij het dossier, want dan heeft hij in ieder geval munitie om het anders aan te pakken. Voor ambtenaren zal het leren voor de organisatie en het beleid meer centraal staan. De meerwaarde van een beleidsevaluatie zal meer op deze onderdelen beoordeeld worden. Het effect van een evaluatie op het beleid zal in dit geval diffuser zijn; en wellicht ook minder snel leiden tot preventief gedrag. In dit verband kan gesteld worden dat een beleidsevaluatie vooral leidt tot “Aufklärung” of “Enlightment”: de betrokken partijen krijgen conceptueel beter inzicht in de werking en effecten van het beleid zonder dat hun gedrag direct anders wordt.

## 5. Tot besluit

Beleidsvaluaties vervullen een nuttige en noodzakelijke functie in het beleidsproces. Zij geven vorm aan politieke verantwoording en kunnen bijdragen aan verbetering van de kwaliteit van het openbaar bestuur en van overheidsbeleid. Ondanks

deze meerwaarde is het niet verkeerd om stil te staan bij mogelijke ongewenste effecten van beleidsevaluaties. Eén van deze ongewenste effecten is preventief gedrag van beleidsvoerders. Beleidsdoelstellingen en inzet van middelen worden aangepast aan de evaluatie (en mogelijk zelfs de daarbij gehanteerde methoden en criteria). Dit preventief gedrag is echter geen vanzelsprekendheid en niet per definitie ongewenst en wordt bovendien bemoeilijkt vanwege allerlei factoren op institutioneel, beleidsmatig en individueel niveau. Voorts is preventief gedrag –voor zover present, intens en schadelijk– niet louter afhankelijk van een aangekondigde beleidsevaluatie. Dat is een grove overschatting van het effect van evaluaties op het beleid en op bestuurders. Het is ook geen automatisme dat een aangekondigde evaluatie resulteert in ongewenst preventief gedrag. Een positief effect kan zijn dat beleidsvoerders door deze kennis vooraf beter nadenken over wat het beleid moet bereiken en de bakens verzetten om later een goede evaluatie mogelijk te maken. Het lastige punt bij onderzoek naar preventief gedrag is echter dat dit gedrag uitermate lastig meetbaar is. Sociaal-wetenschappelijk onderzoek kent de uitdaging dat respondenten gewenst gedrag kunnen gaan vertonen (men weet dat men onderzocht wordt) dan wel sociaal-wenselijke antwoorden geven in vragenlijsten. Het is zeer goed denkbaar dat deze fenomenen zich versterkt voordoen bij onderzoek naar preventief gedrag. Het is moeilijk hier een kant en klare oplossing voor aan te reiken. In de vorige paragraaf is daar een handreiking voor gedaan: een eenduidige interpretatie van ongewenst gedrag, een experimentele opzet, inlevingsvermogen van de onderzoeker, etc. Aan deze voorwaarden is in praktijkonderzoek niet altijd te voldoen. Vooral de opzet van veldexperimenten in beleidsprocessen is moeilijk. De onderzoeker zal daarom moeten putten uiteenlopende bronnen, ofwel triangulatie, een methode die niet ongewoon is in beleidsevaluaties.

Ondanks meetproblemen bij onderzoek naar preventief gedrag zijn de beleidssociologie en de beleidspraktijk gediend bij een kritische reflectie op nut, noodzaak en effecten van beleidsevaluaties en verwant daaraan preventief gedrag van beleidsmakers. De kwaliteit van evaluaties zijn immers niet gegarandeerd en zal alleen maar meer te lijden hebben onder een verkeerde beeldvorming over beleidsevaluatie, bijvoorbeeld dat evaluaties zich laten leiden door politiek gunstige doelen of onvoldoende gebruikt worden. Op dat terrein is er dus nog voldoende werk te verzetten, bijvoorbeeld door meer afstand te creëren tussen beleidsvoerders en onderzoekers, maar ook door kwaliteitseisen te stellen aan evaluatieonderzoek en bewustwording stimuleren over mogelijk preventief gedrag bij evaluaties. Verbetering van beleidsevaluaties is een doel en beleidsevaluaties zelf zijn slechts het middel om ons openbaar bestuur transparant, effectief en democratisch gelegitimeerd te maken en te houden.



## Bibliografie

Abma, T.A. (1996). *Responsief evalueren. Discoursen, controversen en allianties in het postmoderne*. Delft: Eburon.

B&A Groep (2007). *Leren afrekenen. Een voorlopige evaluatie van lokale en provinciale rekenkamer(commissie)s*. Den Haag: B&A Groep.

Bongers, F. (2008). Waar komen deze effecten vandaan? In: *Openbaar Bestuur*, jrg. 18 (in voorbereiding).

Bongers, F. & den Hertog, P. (2007). Vernieuwing van methoden voor beleidsevaluatie. In: *Bestuurswetenschappen*, 61 (3), 29-44.

Bressers, J. T. (1998). De evaluatie van beleid. In: Hoogerwerf, A. & Herweijer, M. *Overheidsbeleid: een inleiding in de beleidswetenschap*. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink, 119- 140.

Bressers, J. Th. A. & A. Hoogerwerf (1991). *Beleidsvaluatie*. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.

Dialogic (2006). *Beleidsvaluatie Kenniswijk*. Utrecht: Dialogic.

Dunn, N.D. (1994). *Public policy analysis. An introduction*. Englewood Cliffs: Prentice Hall.

Guba, E.G. & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park: Sage Publications.

Hoessel, P. van & Mevissen, J. (2005). Beleidsonderzoek onder vuur. In: Hoesel, P. van, Leeuw, F.L. & Mevissen, J. *Beleidsonderzoek in Nederland*. Assen: Van Gorcum, 145-161.

Hoogerwerf, A. & Herweijer, M. (1998). *Overheidsbeleid: een inleiding in de beleidswetenschap*. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.

Huppe, P. (2007). *Overheidsbeleid als politiek. Over de grondslagen van beleid*. Assen: Van Gorcum.

Jong, A.H. de & Kabel, D.L. (2005). De rol van evaluaties in het begrotingsproces. In: *Justitiële Verkenningen*, 31 (5), 100-109.

Morgan, G. (1986). *Images of Organization*. Newbury Park: Sage Publications.

Pestman, P. & Van Wijk, A. (2008). Mogelijkheden beleidsdoorlichting beperkt benut. In: *Openbaar Bestuur*, 18 (3), 27-29.

Pröpper, I.M. (1998). Terugkoppeling: de vertaling van evaluatie naar beleid. In: Hoogerwerf, A. & Herweijer, M. *Overheidsbeleid: een inleiding in de beleidswetenschap*. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink, 141-162.

Trusted, J. (1987). *Inquiry and understanding. An introduction to explanation in physical and human sciences*. Londen: MacMillan Education.

Verlet, D. & Devos, C. (2008). *Beleidsevaluatie: doel of middel?* Toelichting workshop 8. Nijmegen, Politicologenetmaal 2008.

# **Hoofdstuk 4: Risico's op perverse effecten bij beleidsevaluatie? Een vergelijking vanuit mogelijke evaluatiemotieven tegen de achtergrond van de Vlaamse hervormingscontext**

Bart De Peuter, Valérie Pattyn & Marleen Brans  
Instituut voor de Overheid (KULeuven)

<b>1. Inleiding .....</b>	<b>78</b>
<b>2. Evaluatiemotieven .....</b>	<b>78</b>
<b>3. Evaluatiemotieven en risico's op meetproblemen .....</b>	<b>80</b>
<b>4. Evaluatiemotieven bij de Vlaamse overheid .....</b>	<b>81</b>
4.1. Evaluatiemotieven in een hervormingscontext .....	81
4.2. Drijfveren .....	83
4.3. Casestudy: Vlaanderen .....	84
<b>5. Risico's verkleinen .....</b>	<b>91</b>
<b>Bibliografie .....</b>	<b>104</b>

## 1. Inleiding

In de bestuurskundige literatuur wordt enerzijds herhaaldelijk gewezen op onwenselijke neveneffecten die verbonden zijn aan de praktijk van prestatiemeting (Bouckaert, 1999; Van Dooren, 2006). Hierbij past meteen een terminologische uitklaring. Onder de term prestatiemeting verstaan we in deze bijdrage de praktijk van monitoring ten behoeve van beleid en beheer via zogenaamde input-, proces-, output- en effectindicatoren. Monitoring of prestatiemeting levert beschrijvende informatie op, terwijl evaluatie als ambitie kan hebben het achterhalen van causale verbanden of verklaringen.

In tegenstelling tot wat sommige meettechnici zouden verwachten, is het introduceren van een monitoring- of meetsysteem niet per definitie neutraal. De lijst van mogelijke ongewenste effecten die kunnen ontstaan tengevolge van de introductie van een meetsysteem is bijzonder lang. Verlet & Devos geven in hun hoofdstuk een overzicht van deze mogelijke disfuncties. Aangezien monitoring een belangrijke informatiebron voor beleidsevaluatie kan zijn, is deze problematiek ook relevant voor de praktijk van beleidsevaluatie (De Peuter e.a., 2007).

Anderzijds vermeldt de literatuur een brede waaier aan motieven om beleid te evalueren. Uiteenlopende gebruikswijzen van beleidsevaluaties worden onderkend, waaronder verantwoording en beleidsleren maar ook symbolisch gebruik (bv. Balthasar, 2006; Chelimsky, 1997).

Deze bijdrage beoogt deze twee analyses te koppelen door te onderzoeken of en in welke mate de verschillende motieven om beleid te evalueren meer of minder risico's op perverse effecten inhouden, met name in de vorm van distorties tijdens de processen van het meten zelf en/of bij het vormen van waardeoordelen over het beleid als evaluandum. Deze vergelijkende risicoanalyse wordt gebaseerd op de aandachtspunten die de relevante literatuur aanreikt.

De bijdrage brengt eerst de motieven in kaart die aan de basis kunnen liggen van een evaluatie. Een volgende stap is het formuleren van een aantal assumpties omtrent evaluatiemotieven en mogelijke meetproblemen. Vervolgens verkennen we de praktijk in welke mate deze veronderstellingen gegrond zijn. We benaderen de praktijk indirect, namelijk door een screening van de relevante informatie in beleids- en beheersdocumenten, meer bepaald verwijzingen naar vooropgestelde beleidsevaluaties. Vanuit welke evaluatiemotieven worden deze evaluatieintenties geformuleerd en welke motieven prevaleren? En kunnen we assumpties inzake de risico's op perverse effecten valideren?

Ten slotte beogen de auteurs pistes aan te geven om de risico's op perverse effecten te verkleinen. Enerzijds is de doelstelling om theoretische reflecties te transponeren naar adviezen ter verbetering, gegeven het bestaande profiel van evaluatie zoals het uit de motievenanalyse naar voor komt. Anderzijds wordt een lange termijn scenario verkend van een verder evoluerende evaluatiecultuur waarin de oorzaken van perverse effecten kunnen ingedijkt worden. De auteurs schetsen ook pistes voor aanvullend en verdiepend onderzoek.

## 2. Evaluatiemotieven

Beleidsmakers worden geconfronteerd met interne en externe druk om beslissingen te nemen die onderbouwd zijn. In dit streven naar “evidence based” beleid kan beleidsevaluatie een belangrijke bijdrage leveren en dit vanuit verschillende perspectieven: een beter beleidsmanagement, de versterking van de democratische principes en het bereiken van een consensus tussen verschillende stakeholders betrokken bij het publiek beleid.

Daaronder kan een hele waaier van motieven en principes geplaatst worden van waaruit belang wordt gehecht aan beleidsevaluatie. Wanneer we motieven voor beleidsevaluatie zoals aangehaald in de literatuur en praktijk overlopen, kunnen deze teruggebracht worden tot volgend kwartet (De Peuter e.a., 2007):

- Ondersteuning van beleidsplanning
- Verbeteren van de beleidsuitvoering en bijdragen tot institutionele ontwikkeling van de betrokken organisaties
- Verantwoording afleggen
- Beleidsleren

Het motief om beleid te evalueren is een belangrijk vertrekpunt. Zoals Chelimsky (1997) immers stelt: *‘The purpose of an evaluation, conditions the use that can be expected of it’*. Met andere woorden: het motief achter een evaluatie bepaalt mee op welke wijze de resultaten en conclusies van deze evaluatie gebruikt zullen kunnen worden. Afhankelijk van het motief zal het evaluatieproces op een verschillende wijze gestalte krijgen, zullen andere vragen centraal staan en zal de focus van conclusies en eventuele aanbevelingen verschillen. De hoger vermelde motieven kunnen worden gekoppeld aan bepaalde maar niet alle mogelijk gebruikswijzen. Balthasar (2006) onderscheidt een viertal vormen van gebruik van evaluaties:

- Instrumenteel gebruik: evaluaties verschaffen een impuls die gebruikt wordt om het beleidsproces te informeren op een aantoonbare wijze. Conclusies van de evaluatie worden rechtstreeks toegepast. Deze gebruikswijze is minder waarschijnlijk wanneer het voorwerp van evaluatie sterk gepolitiseerd is.
- Conceptueel gebruik: evaluatieresultaten beïnvloeden de attitudes en overtuigingen van beleidsmakers met betrekking tot het voorwerp van de evaluatie. Dit is meestal het geval wanneer het om een externe evaluatie gaat.
- Procesgerelateerd gebruik: het uitvoeren van de evaluatie op zich brengt voordelen mee voor de betrokken stakeholders, bijvoorbeeld het creëren van een gemeenschappelijk begrip, de ontwikkeling van een netwerk, of leerprocessen.
- Symbolisch gebruik: de evaluatie wordt gehanteerd om bijkomende legitimatie van een beleidsinitiatief te verschaffen. Verbetering van beleid door middel van de evaluatie wordt daarbij niet beoogd. Indien de resultaten van evaluaties niet legitimerend zijn, worden ze niet openbaar gemaakt.

De hoger vermelde evaluatiemotieven kunnen voornamelijk aan de eerste drie gebruikswijzen gekoppeld worden. Bij instrumenteel gebruik kan men spreken van direct of daadwerkelijk gebruik. Bij conceptueel en procesgerelateerd gebruik is “invloed” wellicht de meest geëigende beschrijving van de doorwerking van een evaluatie. Symbolisch gebruik bestaat ook en mag niet worden genegeerd. Inherent aan symbolisch gebruik is dat er vooraf de intentie bestond om de resultaten en conclusies niet te gebruiken, tenzij voor onderbouwing van vooraf ingenomen posities. Het uitvoeren van de evaluatie is in dit geval het doel, niet het middel. Het fundamentele verschil met de andere gebruikswijzen is dat de evaluatie in se géén meerwaarde verschaft voor het beleid.

### 3. Evaluatiemotieven en risico's op meetproblemen

In deze paragraaf leiden we een aantal assumpties af uit de literatuurstudie over evaluatiemotieven en gerelateerde perverse effecten bij het meten. Uitgangspunt is dat betrokken actoren interpretaties zullen maken over het nut en het doel van de metingen. Zeker wanneer metingen dienen als input in het kader van beleidsevaluaties, kan het meten zelf leiden tot een aantal problemen. Bepaalde disfuncties zullen zich vermoedelijk sneller of sterker stellen naargelang het motief voor het meten en evalueren. Onze basisveronderstelling is dat het verantwoordingsmotief, gezien haar vaak negatievere connotatie vanuit het perspectief van de betrokkenen, de meeste meetproblemen met zich mee zal brengen.

Zo komen we tot een eerste set van assumpties.

Assumptie 1: Wanneer verantwoording het evaluatiemotief is, zal de geëvalueerde bij prestatie meting ten behoeve van zulke evaluaties eerder opteren voor outputindicatoren dan voor effectindicatoren.

De redenering is dat een grotere beheersbaarheid van output mogelijk is in vergelijking tot effecten, waarop omgevingsfactoren een relatief grotere invloed kunnen hebben, en dat men liever wenst afgerekend te worden op datgene wat men kan controleren.

Assumptie 2: Wanneer verantwoording het evaluatiemotief is bestaat het risico op een pervers effect, namelijk dat beleidsdoelstellingen bewust niet worden geformuleerd volgens het SMART-principe, dat staat voor specifiek, meetbaar, afgesproken, realistisch en tijdsgebonden. In het bijzonder wat de specificiteit, de meetbaarheid en de tijdsgebondenheid betreft, zijn de grootste problemen te verwachten.

Assumptie 3: Beleidsevaluaties geïnitieerd door de verantwoordelijken voor het beleid zelf, gebeuren eerder vanuit de motieven van het ondersteunen van de beleidsplanning, het verbeteren van beleidsimplementatie en beleidsleren, dan wel vanuit het motief van verantwoording. Verantwoording gebeurt eerder vanuit een externe druk.

Assumptie 4: De aandacht voor de verschillende evaluatiemotieven verschilt

naargelang het beleidsdomein. Die vaststelling werd al in meerdere studies onderschreven (zie bv. Furubo e.a., 2002). Omdat deze aandacht vanuit verschillende drijfveren kan ontstaan, verwachten we ook dat de mate van aandacht voor elk van de evaluatiemotieven zal verschillen. In sommige sectoren kan beleidsleren bijvoorbeeld meer benadrukt worden, terwijl in andere sectoren ondersteuning van de planning belangrijker wordt geacht.

Een bijkomende assumptie ten slotte, heeft betrekking op de Vlaamse praktijk van beleidsevaluatie.

Assumptie 5: Binnen de Vlaamse overheid is het risico op perverse effecten bij meten en evalueren groter geworden sinds de hervorming Beter Bestuurlijk Beleid (BBB). Deze hervorming draagt duidelijke karakteristieken van het New Public Management-paradigma waar verantwoording een belangrijke grondtoon vormt.

Omdat in deze hervorming evaluatie sterker verankerd werd in de beleidsopdracht van de overheid en in de beheerscyclus, zal de aandacht voor evaluatie groeien en de praktijk ervan versterken. Daarmee stijgt ook het *risico* op perverse effecten. Met deze assumptie stellen we echter niet dat we een stijging verwachten in de praktijk, omdat men ook dient rekening te houden met een groeiende ervaring om risico's op te vangen en met mechanismen die worden ontwikkeld om perverse effecten in te dijken.

Wanneer we kijken naar de aard van perverse effecten, verwachten we meer bepaald problemen die betrekking hebben op vertekeningen van de output en een verschuiving naar het tijdspectief verbonden aan de verantwoording.

## 4. Evaluatiemotieven bij de Vlaamse overheid

### 4.1. Evaluatiemotieven in een hervormingscontext

De trend naar een sterkere aandacht in het discours en de praktijk voor beleidsevaluatie is te onderkennen via meerdere kanalen of indicatoren: referenties in beleidsnota's aan evaluaties van maatregelen, projecten of programma's, het aantal opdrachten voor externe evaluaties in het bulletin der aanbestedingen, het aantal evaluatierapporten dat openbaar wordt gemaakt op websites, het aantal evaluatiefuncties binnen de administratie en afdelingen of cellen die beleidsevaluatie onder hun takenpakket vermelden, modules rond beleidsevaluatie in onderwijscurricula en initiatieven naar postacademische vorming rond het thema.

Ook in het wetgevende kader rond de organisatie van de overheid vinden we meer en meer verwijzingen naar evaluatie. Beleidsevaluatie wordt in toenemende mate verankerd in de regelgeving en gekoppeld aan het wetgevende proces zelf.

Voor Vlaanderen is de reguleringsimpactanalyse (RIA) daarvan een schoolvoorbeeld. De RIA wordt uitgevoerd voor elk voorontwerp van decreet en ontwerpbesluit met

een regulerend effect op de burger, het bedrijfsleven en non-profit organisaties. Het omvat naast een probleemformulering, een omschrijving van de beleidsdoelstelling en een overzicht van de opties om deze laatste te bereiken. De opties omvatten de nuloptie (geen beleidsinitiatief), traditionele regulering en alternatieven. Bij deze laatste onderscheidt men zelfregulering, sociaalcommunicatieve, economische, privaatrechtelijke of institutionele instrumenten, en collectieve maatregelen. Het evaluatieve aspect van een RIA ligt in de analyse van de voor- en nadelen van elke optie (gevolgen voor de doelgroep, directe en indirecte effecten en eventuele herverdelingseffecten) en meer bepaald in de onderlinge afweging van de opties op basis van de informatie over verwachte effecten en in functie van de beleidsdoelstelling. RIA is dus opgevat als een vorm van ex ante evaluatie. Het vormt ook een kader om andere, thematische evaluatiecriteria of “toetsen” te integreren, bv. kindeffectrapportage, de weerslag op de lokale besturen (Belfort-principe), armoedebestrijding et cetera. Ook op federaal vlak werden recent dergelijke “testen” ontworpen, bv. de duurzame ontwikkeling effectbeoordeling (DOEB-test) en genderevaluatie.

In het kader van de algemene hervorming van de Vlaamse Overheid, Beter Bestuurlijk Beleid (BBB), werd beleidsevaluatie algemeen verankerd in de beleidscyclus. Het kaderdecreet belast de departementen met beleidsondersteunende taken. De Memorie van Toelichting (art.4) vermeldt daarbij de evaluatie van de beleidsuitvoering, in het bijzonder het nagaan van de effectiviteit van de ingezette instrumenten, en de relaties tussen output en effecten. De agentschappen dienen met het oog op de beleidsevaluatie relevante beleids- en beheersinformatie als input te genereren. Het uitvoeringsbesluit bij het kaderdecreet omschrijft de taakstelling van de departementen nader op dit vlak. Ze dienen de minister te ondersteunen:

*‘1° bij het uitwerken van diens beleid: daartoe heeft het departement een beleidsvoorbereidende en beleidsevaluerende opdracht  
2° bij de aansturing en opvolging van de beleidsuitvoering, die in de regel toevertrouwd wordt aan agentschappen.’*

Op het vlak van beleidsvoorbereiding en -evaluatie ten behoeve van de minister hebben departementen onder meer als taak in een permanente monitoring en omgevingsanalyse van het beleidsdomein te voorzien, met inbegrip van het beheer van databanken en de verzameling en analyse van statistische informatie. Ten tweede dienen zij de beleidsuitvoering (ingezette instrumenten, effecten, et cetera) op macroniveau te evalueren met het oog op eventuele bijsturing van het beleid of bijsturing van de aansturing van agentschappen. Ook de aansturing van beleidsgericht wetenschappelijk onderzoek en de aanwending van de resultaten behoort tot dit takenpakket. Tevens staan de departementen in voor de ontwikkeling van een beleidsvoorbereidend instrumentarium (modellen, scenarioanalyse, benchmarking, et cetera). Ten slotte moeten zij ook voorstellen uitwerken voor de beleidsbepaling en de toetsing van beleidsintenties.

Het comptabiliteitsdecreet (art.6 §3) vermeldt de beleidseffectenrapportage: een rapportering waarin de realisaties van de beleidsopties gedurende het vorige begrotingsjaar geëvalueerd worden aan de hand van indicatoren en kengetallen



die gekoppeld zijn aan de strategische doelstellingen (maatschappelijke effecten) en operationele doelstellingen (beleidsprestaties of output) en die vermeld staan in de beleidsnota's. De beleidseffectenrapportage dient te gebeuren op het niveau van het beleidsdomein.

Voor de lokale besturen vinden we, gelijktijdig met een versterkte aandacht voor strategische planning, ook referenties aan beleidsevaluatie terug in het Gemeentebestuur en het Provinciedecreet. Zo hebben de gemeentesecretaris en provinciegriffier onder meer verantwoordelijkheden op het vlak van de evaluatie van het beleid, en moet het managementteam de coördinatie van de diensten ondersteunen bij de beleidsevaluatie.

Via wettelijke verankering kan evaluatie een systematisch karakter krijgen en geïnstitutionaliseerd worden. Daarnaast speelt ook mee dat Vlaanderen een gewoonte heeft om veel beleid in regelgeving om te zetten. De verankering in kaderwetgeving maakt ook dat beleidsevaluatie gelijktijdig voor een brede waaier van beleidssectoren een aandachtspunt wordt.

Expliciete verwijzingen naar motieven voor de verankering van beleidsevaluatie vinden we echter zelden terug in de beleidsdocumenten en wetgeving die de hervormingen binnen de diverse bestuurslagen introduceren. Dat neemt niet weg dat in concrete beleidsdomeinen ad hoc beslissingen en meer structurele initiatieven worden genomen om te evalueren.

## **4.2. Drijfveren**

Op het vlak van de introductie van een evaluatiepraktijk binnen de overheid zijn België en haar verschillende bestuurslagen te situeren achteraan in het peloton van Europese landen, die eind jaren '90 een evaluatiecultuur ontwikkelden (De Peuter e.a., 2007). Een beperkte kopgroep van landen deed dat al eind jaren '60, begin jaren '70. De verklaring voor het feit dat de verspreiding van een evaluatiepraxis in Europa relatief traag verliep en in twee golfbewegingen, is terug te voeren op een verschil tussen interne druk (pionierlanden) en externe druk (peloton) (Furubo e.a., 2002).

Bij de voorlopers werd het evaluatiedenken gepromoot door een mix van factoren: een relatief groter geloof in de maakbaarheid in de samenleving, de sociaalwetenschappelijke achtergrond van het gros van het ambtelijke apparaat, de intensere contacten met de VS (vanwaar het evaluatiedenken overwaaid naar Europa), en een grotere rol van de overheid (afgelezen aan het comparatieve uitgavenpeil).

Bij de grote groep van landen die enkele decennia later volgden, waaronder België, speelde vooral een externe druk om beleidsevaluatie te introduceren binnen de overheid. Een eerste impuls kwam er met de verspreiding van het ideeëngoed rond New Public Management. In het bijzonder het verantwoordingsmotief om te evalueren werd daarbij naar voor geschoven. Evaluatie werd benaderd vanuit een managementperspectief of sterk gekoppeld aan het organisatieniveau.

Een tweede belangrijke impuls kwam vanuit de Europese Unie. Tal van Europese beleidsprogramma's en initiatieven vereisen een actieve participatie van verschillende bestuursniveaus in de lidstaten. De Europese Unie heeft een sterke evaluatiefunctie en -cultuur ontwikkeld binnen de Commissie, die verder uitstraalt op de binnenlandse overheden. De Europese Commissie, vaak in de rol van beleidsinitiator en geldschietster, is dan ook zeer sterk geïnteresseerd in de goede besteding van de middelen en de mate van doelbereiking. Dit mechanisme heeft ervoor gezorgd dat de lidstaten moeten meestappen in het evaluatiegebeuren. Hoewel voor Europa het federale niveau het aanspreekpunt blijft, vergt de bevoegdheidsverdeling de betrokkenheid van de andere bestuurslagen. In Vlaanderen hebben dan ook heel wat overheden en administraties praktijkervaring met evaluaties voor Europa, onder meer in de sectoren milieu en natuur, plattelandsontwikkeling, landbouw en visserij, werkgelegenheid, regionale ontwikkeling en onderwijs.

Een derde impuls kwam er via de talrijke hervormingsprocessen die recent werden doorgevoerd in de verschillende bestuurslagen in België. Varone & Jacob (2003) wezen op het belang van bijkomende inspanningen op het vlak van organisatorische inbedding en ontwikkeling van evaluatie. Het zijn niet alleen voorwaarden om op een kwaliteitsvolle en verantwoorde wijze evaluaties te kunnen uitvoeren. Ook de politieke en publieke zichtbaarheid van evaluatieresultaten en eraan gekoppelde beleidsverbeteringen vergroot op die manier.

Een vierde drijfveer vormt een evoluerende attitude bij beleidsmakers en een groeiende overtuiging dat evaluatie een meerwaarde kan betekenen voor beleid. Men ziet daarin ook een opportuniteit om de al geleverde inspanningen op het vlak van monitoringsystemen te valoriseren. De keuze voor een zogenaamd "evidence-based" of geïnformeerd beleid neemt een belangrijke plaats in vanuit de twee krachtlijnen waarlangs hervormingen binnen de overheid zich aftekenen: professionalisering en vermaatschappelijking (Brans & Vancoppenolle, 2005). Evaluaties vormen een belangrijke bron om beslissingen van beleidsmakers over toekomstig, lopend of uitgevoerd beleid te informeren, te onderbouwen en te beoordelen. Ze kunnen ook perspectieven van verschillende stakeholders meenemen en de zoektocht naar een consensus ondersteunen (De Peuter, 2008).

### **4.3. Casestudy: Vlaanderen**

Welke evaluatiemotieven prevaleren in Vlaanderen, gegeven de hervormingscontext en de drijfveren naar een meer geïnformeerd beleid? En welke risico's op perverse effecten zijn bijgevolg –op basis van de literatuur hierover– in te calculeren?

De betrachting is een kwalitatieve inschatting te maken van risico's op perverse effecten van beleidsevaluatie. We ambiëren dus geen concluderende uitspraken over het bestaan van perverse effecten in de Vlaamse praktijk van beleidsevaluatie. De "ingang" voor onze analyse zijn de evaluatiemotieven. Om zicht te krijgen hierop hebben we een documentenanalyse gevoerd, gekoppeld aan het recente hervormingstraject dat de Vlaamse overheid doorlopen heeft. Meer bepaald hebben we gekeken naar beleidsnota's en beheersovereenkomsten. De beleidsnota's zijn de meest recente op Vlaams niveau (2005-2009). Voor de beheersovereenkomsten

hebben we ons beperkt tot zeven organisaties waarvan de beheersovereenkomst ook voorwerp uitmaakte van een analyse door het Rekenhof in 2001. Op deze wijze kunnen we de evolutie voor en na BBB bekijken. Deze werkwijze heeft uiteraard zijn beperkingen vermits het gaat om intenties van evaluaties en een discours in verband met evaluatiemotieven. In beleidsnota's kan men bovendien redelijkerwijs relatief weinig verwijzingen naar verantwoording verwachten. Een alternatieve empirische screening, bv. via evaluatierapporten, stelt echter evenzeer beperkingen inzake het achterhalen van evaluatiemotieven en garandeert geen volledig beeld. Verder kunnen beleidsnota's ook gezien worden als een instrument van verantwoording voor de besteding van publieke middelen naar de burger. In onze conclusie komen we terug op alternatieve onderzoekspistes.

#### 4.3.1. *Evaluatiemotieven in Vlaamse beleidsnota's*

In totaal werden een totaal van 844 citaten<sup>38</sup> uit dertig beleidsnota's van de Vlaamse Regering (2004-2009) gescreend op evaluatiemotieven. In 254 van deze citaten wordt melding gemaakt van een evaluatiemotief. Tabel 4.1 geeft een overzicht van het aantal geïdentificeerde citaten over evaluatie per beleidsnota, opgesplitst naar de vier eerder aangehaalde evaluatiemotieven: het ondersteunen van beleidsplanning (PL), het verbeteren van de beleidsuitvoering (VB), het afleggen van verantwoording (VA) en beleidsleren (BL).

Ten eerste valt op dat in sommige beleidsnota's evaluatie prominenter aanwezig is dan in andere sectoren. We onderscheiden drie groepen: nota's waarin veel verwezen wordt naar evaluatiemotieven (meer dan 20 citaten), nota's waar weinig wordt verwezen naar evaluatiemotieven (minder dan 10 citaten) en een middengroep (tussen 10 en 20 citaten).

In de koplopers vinden we beleidsdomeinen met ofwel een langere traditie van evaluatiepraktijken (onderwijs), met een typische planningscultuur (openbare werken, ruimtelijke ordening), en met een focus die veelal terugkomt bij evaluatieonderzoek, zoals bv. efficiëntie en zuinigheid (financiën en begroting).

In de middengroep vinden we domeinen terug waarvan we oorspronkelijk nog meer aandacht voor evaluatie verwachtten, en waar *evidence-based policy-making* ook internationaal in de lift zit, nl. leefmilieu; welzijn, volksgezondheid en gezin; wetenschap, economie en innovatie.

De groep met het minste verwijzingen naar evaluatie en haar motieven vinden we opvallend genoeg bij de beleidsnota's waar evaluatievragen in principe zeer pertinent zijn omdat de effecten van het beleid op bepaalde doelgroepen (inburgering, jeugd) of streken (stedenbeleid, Brussel) voorwerp zouden kunnen uitmaken van evaluatievragen. Er zijn ook enkele nota's bij waarbij het typisch gaat over domeinoverschrijdende beleidsvoorwerpen, waar transversale evaluatie een

---

<sup>38</sup> Deze 844 citaten zijn in de studie van Conings e.a. (2005) verrat onder categorie 2: Beleidsplanning en –opvolging.

rol zou kunnen spelen. Vooral het gebrek aan verwijzingen naar evaluatiemotieven voor duurzame ontwikkeling en in iets mindere mate gelijke kansen springt hierbij in het oog.

Vervolgens zijn we nagegaan welk type evaluatiemotief prevaleert, en of hierbij ook de voorkeuren voor motieven domeinspecifiek zijn. Beleidsleren komt het vaakst terug als evaluatiemotief (104), gevolgd door ondersteuning van de planning (65) en vervolgens door verbetering (61). Zoals verwacht is verantwoording het motief waar het minst naar wordt verwezen (24). Wat dit laatste betreft zijn er evenwel twee uitzonderingen (met de hoogste scores op verantwoording). Van het totaal van 24 verwijzingen komen er 10 uit “onderwijs” en 3 uit openbare werken. Dat “onderwijs” verantwoording als evaluatiemotief hoog in het vaandel draagt, kan ons inziens te maken hebben met de sterk ingebedde inspectiepraktijken in de onderwijssector. Voor “openbare werken” kunnen we hier wellicht ook wel verwachten dat de beroepsmogelijkheden van burgers een aandacht op verantwoording hebben gevestigd. Feit blijft dat in 19 nota’s geen enkele verwijzing naar verantwoording is terug te vinden. Beleidsleren komt dus vaak terug als motief. En bij die nota’s waar in ieder geval evaluatie sterker aanwezig is, vinden we bij beleidsleren veel verwijzingen terug naar het gebruik van proeftuinen en wetenschappelijk onderzoek (“onderwijs”, en “werk”, steunpunten inclusief (zeker bij Financiën en Begroting). Bij openbare werken vinden we een zeer sterke nadruk op het gebruik van evaluaties voor planning en bij (sociale) economie aandacht voor verbetering via voor deze domeinen voor de handliggende vragen rond efficiëntie.

Wanneer we terugkoppelen naar de assumpties geformuleerd in punt 3, stellen we vast dat zowel de derde als de vierde assumptie gevalideerd worden door de data. In beleidsnota’s wordt verantwoording als evaluatiemotief het minst vermeld in vergelijking met de andere drie geïdentificeerde motieven van beleidsplanning, verbeteren van de uitvoering en beleidsleren. Bovendien verschilt de onderlinge verdeling van evaluatiemotieven naargelang het beleidsdomein.

**Tabel 4.1: Evaluatiemotieven vermeld in beleidsnota's van de Vlaamse Regering (2004-2009)**

beleidsdomein	PL	VB	VA	BL	TOT
BESTUURSZAKEN	0	2	0	1	4
BINNENLANDS BESTUUR	1	2	0	1	4
BRUSSEL	0	0	0	0	0
BUITENLANDS BELEID	0	0	1	0	1
CULTUUR	1	1	0	1	3
DUURZAME ONTWIKKELING	0	0	0	0	0
ECONOMIE WETENSCHAP EN INNOVATIE	2	6	1	2	11
ENERGIE	2	3	0	5	10
FINANCIEN EN BEGROTING	6	2	1	11	20
GELIJKE KANSEN	1	0	0	1	2
INBURGERING	0	2	0	0	2
JEUGD	0	4	0	0	4
LANDBOUW	3	1	0	3	7
MEDIA	1	1	0	5	7
MOBILITEIT	4	1	0	2	7
ONDERWIJS	5	7	10	23	45
OPENBARE WERKEN	18	3	3	8	32
OVERHEIDSCOMMUNICATIE	0	0	0	0	0
PLANNING EN STATISTIEK	0	1	0	5	6
PPS	0	1	1	0	2
RUIMTELIJKE ORDENING	8	6	2	4	20
SOCIALE ECONOMIE	2	3	1	2	8
SPORT	0	4	1	0	5
STEDENBELEID	0	0	0	0	0
TOURISME	1	0	0	2	3
VLAAMSE RAND	0	1	0	2	3
WELZIJN VOLKSGEZONDHEID EN GEZIN	4	2	1	3	10
WERK	2	1	0	17	20
WONEN	0	0	1	1	2
LEEFMILIEU	3	7	1	5	16
TOTAAL	65	61	24	104	254

#### **4.3.2. Meten en evalueren in het kader van beheersovereenkomsten**

Hoewel beheersovereenkomsten al in 1992 als managementtool werden geïntroduceerd binnen de Vlaamse Overheid, kent het gebruik van dit instrument vernieuwde aandacht, sinds het Regeerakkoord van 13 juli 1999 en de daaropvolgende beslissing van de Vlaamse Regering van 19 januari 2001 om de figuur van “beheersovereenkomst” veralgemeend in te voeren als sturingsmechanisme voor alle verzelfstandigde entiteiten, met vastlegging van een aantal verplicht op te nemen bepalingen. Deze beslissing kaderde al in de aanloop naar de hervormingen BBB (cf. infra) (Vlaamse Regering, 22-3-2002).

Mede naar aanleiding van deze veralgemeende invoering nam het Rekenhof het initiatief om een selectie van beheersovereenkomsten aan een kritische analyse te onderwerpen (werkingsjaar 1999- verslag van 3 juli 2001) en aanbevelingen te formuleren voor de verfijning van het instrument. Meer bepaald wilde men onderzoeken in hoeverre de voorwaarden werden vervuld voor een optimaal gebruik van de beheersovereenkomsten als sturingsinstrument voor de (toenmalige) Vlaamse Openbare Instellingen (VOI's). De analyse beperkte zich tot de destijds afgesloten zeven beheersovereenkomsten in verschillende beleidsdomeinen (Vlaamse Vervoermaatschappij (VVM), Vlaamse Dienst voor Arbeidsbemiddeling (VDAB), Vlaamse Radio en –Televisieomroep (VRT), Vlaamse Instelling voor Technologisch Onderzoek (VITO), Vlaamse Opera (VLOpera), Dienst Scheepvaart (DS), Vlaams Fonds voor de Letteren (VFL)). Op basis van de studie concludeerde het Rekenhof dat het noodzakelijk was belangrijke lacunes aan te pakken, opdat beheersovereenkomsten op een doelmatige manier kunnen worden gebruikt als sturings- en opvolgingsinstrument. De voor ons onderzoek meest relevante vaststellingen zijn de volgende:

- Onvoldoende toetsbaar geformuleerde doelstellingen in de beheersovereenkomsten en bijhorende planningsdocumenten. Hoewel op het niveau van de VOI's over het algemeen een coherent geheel van doelstellingen wordt geformuleerd, blijft de graad van concretisering ervan vrij beperkt. De doelstellingen zijn vaak onvoldoende geoperationaliseerd door middel van adequate indicatoren. Waar indicatoren worden vermeld, zijn deze dikwijls onaangepast aan de doelstelling die zij moeten meten; en sommige indicatoren bevatten geen normstelling.
- Er worden verschillende soorten indicatoren aangewend: proces, input, output en effect.
- Een ontoereikende organisatorische inbedding van de beheersovereenkomsten in de VOI's die de strategische doelstellingen nog niet volledig hebben doorvertaald in interne doelstellingen, indicatoren en normen voor de verschillende organisatiedoelstellingen. Dit bemoeilijkt de formele toewijzing van verantwoordelijkheden voor de realisatie van de doelstellingen.
- De opvolging en evaluatie van de uitvoering van de beheersovereenkomsten verloopt gebrekkig en leidt niet tot een gestructureerde interne en externe rapportering. Geen enkele beheersovereenkomst legt een globale evaluatie bij het einde van de beheersovereenkomst op. Wel wordt voorzien in een jaarlijkse evaluatie van de uitvoering
- De meeste instellingen beschikken over meetsystemen. De betrouwbaarheid van de prestatie-informatie wordt echter niet systematisch overal gecontroleerd. Ook werken niet alle instellingen met boordtabellen of periodieke uitvoeringsrapporten die de realisatie van de doelstellingen en verplichtingen meegeven.

- Hoewel de meeste beheersovereenkomsten sancties bevatten voor het niet naleven van verplichtingen (meestal in de vorm van financiële (dis)incentives), zijn deze niet verplicht van aard. Beloningen zijn echter wel verplicht.
- De verplichtingen van de VOI's neergeschreven in de beheersovereenkomsten betreffen vaak inspanningsverbintenissen, in plaats van resultaatsverbintenissen.

De vaststellingen gelden in eerste instantie voor de bestudeerde beheersovereenkomsten, maar gelden mutatis mutandis ook voor beheersovereenkomsten met andere autonome entiteiten.

Bij gebrek aan een uniforme regelgeving met betrekking tot de inhoud en vormgeving van beheersovereenkomsten, vertoonden de zelfstandige entiteiten evenwel niet allemaal dezelfde karakteristieken. Tabel 4.2 geeft in de eerste kolom een overzicht van de belangrijkste conclusies van het Rekenhof per VOI voor de voor ons meest relevante aspecten (Rekenhof, 2001).

De opmerkingen geformuleerd door het Rekenhof bleven niet zonder gevolg. Ze fungeerden mede als input voor de verdere reflectie over de optimalisering van dit sturingsinstrument.

De Visienota Beheersovereenkomsten legde hiertoe de conceptuele basis. De specifieke modaliteiten, inhoud, definitie, vorm, opvolging en evaluatie werden juridisch gefixeerd in het Kaderdecreet Beter Bestuurlijk Beleid<sup>39</sup> (Vlaamse Regering, 22-7-2005a). In haar Memorie van Toelichting definieert ze beheersovereenkomsten als een *'wederzijdse overeenkomst die handelt over de bijzondere regels en voorwaarden waaronder het agentschap zijn taken vervult. Het is een sturings- en opvolgingsinstrument, in het bijzonder gericht op een doelmatige uitvoering of dienstverlening, opvolging en evaluatie'* (Vlaams Parlement, 18-7-2003).

Voortaan worden alle Intern Verzelfstandigde Agentschappen (IVA's) en Extern Verzelfstandigde Agentschappen (EVA's) geacht een beheersovereenkomst af te sluiten met hun bevoegde Minister.<sup>40</sup> Teneinde een zekere mate van uniformiteit en transparantie te garanderen tussen de diverse beheersovereenkomsten, en rekening houdend met de tot nog toe beperkte expertise van Vlaanderen inzake beheersovereenkomsten, werd ervoor geopteerd een "Handleiding" ter beschikking te stellen aan de betrokken partijen met de uitwerking van een sjabloon en generieke elementen voor elke beheersovereenkomst. Eerder dan de specifieke modaliteiten

<sup>39</sup> Met het Kaderdecreet Beter Bestuurlijk Beleid werd eveneens de nieuwe organisatiestructuur van BBB vastgelegd. Waar VOI's voorheen relatief losstonden van de organisatiestructuur van het Ministerie van de Vlaamse Gemeenschap, sluiten de verzelfstandigde entiteiten nu nauwer aan bij hun beleidsdomein. Een functionele specialisatie werd doorgevoerd. De beleidsondersteunende taken worden door de Departementen uitgevoerd; agentschappen verzorgen de beleidsuitvoering. Verder werd besloten tot differentiatie tussen Intern Verzelfstandigde Agentschappen (IVA's), Extern Verzelfstandigde Agentschappen (EVA's) en Privaatrechterlijk vormgegeven Extern Verzelfstandigde Agentschappen (EVA's).

<sup>40</sup> In tegenstelling tot de Intern Verzelfstandigde Agentschappen en de Extern Verzelfstandigde Agentschappen, worden de Privaatrechterlijk vormgegeven Extern Verzelfstandigde Agentschappen niet aangestuurd door een beheersovereenkomst, dan wel door een samenwerkingsovereenkomst.

van de beheersovereenkomst vast te leggen bij besluit van de Vlaamse Regering, biedt het werken met een handleiding de flexibiliteit en mogelijkheid tot creatieve inbreng van ministers en de administratie en laat deze ruimte voor de diversiteit en eigenheid van de diverse beleidsdomeinen (Vlaamse Regering, 22-7-2005a). Toch kan deze handleiding helemaal niet als vrijblijvend worden beschouwd. De sjablonen dienen in principe strikt te worden gevolgd (Vlaamse Regering, 22-7-2005b).

De in het Kaderdecreet betreffende artikelen over de beheersovereenkomst, evenals de verplicht te incorporeren onderdelen van het sjabloon komen op het eerste zicht tegemoet aan de destijds vastgestelde tekortkomingen door het Rekenhof. Wat rapportering betreft, verplicht het Kaderdecreet voortaan elk agentschap te voorzien in:

- *'Een jaarlijks rapport vóór 31 maart van het volgende jaar, betreffende de uitvoering van de beheersovereenkomst gedurende het afgelopen kalenderjaar;*
- *Een eindrapport betreffende de uitvoering van de beheersovereenkomst sinds de inwerkingtreding ervan, zes maanden vóór het aflopen van de duurtijd van de beheersovereenkomst.'* (art. 9 §1)

Bovendien wordt nu ook een jaarlijkse evaluatie verwacht van de overeenkomst en de implementatie ervan door de Vlaamse Regering (art. 9, §3). De resultaten van de evaluatie zullen daarenboven als belangrijke input dienen voor de persoonlijke evaluatie van de leidend ambtenaar (Vlaamse Regering, 22-7-2005b, p.8).

Het Kaderdecreet en het sjabloon vermelden daarnaast expliciet de noodzaak van rapportage op basis van beleidsrelevante indicatoren en kengetallen: *'Strategische en operationele organisatiedoelstellingen, de te leveren prestaties en de te verrichten processen kunnen en moeten worden geconcretiseerd, tot op een niveau dat haalbaar en relevant is, in indicatoren en te bereiken normen (streefwaarden). Indicatoren en normen kunnen zowel betrekking hebben op kwantiteits- als kwaliteitselementen'* (Vlaamse Regering, 22-7-2005b, 6).

Hoewel het gebruik van indicatoren verplicht zal zijn, doch met bepaalde vrijheidsgraden, is het interessant te onderzoeken hoe agentschappen en hun bevoegde ministers omgaan met deze toegenomen druk tot inclusie van indicatoren.<sup>41</sup> De eerder bestudeerde set van zeven agentschappen biedt in deze zin een ideale (vergelijkings)basis ter toetsing van onze assumpties.

Tabel 4.2 bevat de vergelijking van de bevindingen van de huidige beheersovereenkomsten van de zeven agentschappen met de bevindingen van de bovenvermelde Rekenhofstudie. In het licht van de assumptietoetsing zijn vooral de

---

<sup>41</sup> Weliswaar geldt er momenteel een overgangssituatie, waarbij enkel de nieuw opgerichte agentschappen al de bepalingen in de handleiding worden geacht te volgen. Instellingen die een lopende beheersovereenkomst hebben hoeven geen nieuwe beheersovereenkomst op te maken. Dit neemt echter niet weg dat het klimaat betreffende inclusie van indicatoren in beheersovereenkomsten al langere tijd een kentering kent (supra), en dito op die manier eveneens normaliter in rekening moet zijn gebracht in de lopende beheersovereenkomsten met bestaande agentschappen.



secties “doelstellingen van de VOI”; “verplichtingen, prestaties, indicatoren, normen en timing”; “rapporteringsverplichtingen”, “evaluatie” en “sancties en beloningen” relevant.

Wat leert een vergelijking van het profiel van de beheersovereenkomsten van de 7 agentschappen voor en na BBB? De betrokken agentschappen hebben een verschillend pad afgelegd op het vlak van het expliciteren en het meetbaar maken van de doelstellingen. Sommige hadden reeds een traditie daarin en hebben deze lijn doorgetrokken in de recente beheersovereenkomst. Best practices in de bestudeerde groep waren en blijven de VRT en de VDAB. Andere agentschappen hebben een inhaalbeweging gemaakt, namelijk VITO, DS en de VVM. Bij enkele agentschappen valt slechts met moeite een vooruitgang terzake te onderkennen, namelijk bij Vlopera en het VFL. De vraag rijst in welke mate de aard van het beleidsdomein (Cultuur) hier een verklarende factor is. Hetzelfde geldt voor de attitude van beleidsmakers in deze sector ten aanzien van meetbare doelstellingen. De problematiek blijft echter niet beperkt tot het technisch ontwikkelen van indicatoren, maar omvat ook de uitwerking van strategische en daarbij aansluitend operationele doelstellingen.

Wanneer we terugkoppelen naar de assumpties geformuleerd in punt 4, stellen we vast dat onze eerste assumptie toch wel wordt bevestigd. Beheersovereenkomsten en hun evaluatie passen onder het verantwoordingsmotief. Wanneer indicatoren worden geformuleerd, blijkt het merendeel inderdaad betrekking te hebben op de output van agentschappen, en niet zozeer op effecten. De VDAB vormt daarop een uitzonderling met een evenwichtige verdeling tussen output- en effectindicatoren. Wanneer we echter de bredere indicatorenset (incl. de zogenaamde monitoringsindicatoren) nemen is er ook bij dit agentschap een overwicht van outputindicatoren. Assumptie twee, namelijk dat met verantwoording als evaluatiemotief, het risico groter is dat doelstellingen niet SMART worden geformuleerd, kan niet worden bevestigd op basis van de informatie in de beheersovereenkomsten. We zien namelijk een ongelijke evolutie in het SMART definiëren van doelstellingen over de agentschappen heen. Bovendien verwijst het regelgevend kader voor de hervorming BBB op de noodzaak van het gebruik van beleidsrelevante indicatoren en kengetallen. Uit de screening van de beheersovereenkomsten en de vergelijking met de bevindingen van het Rekenhof is af te leiden dat op dit punt wel wat vooruitgang is geboekt, zij het niet in dezelfde mate voor elk agentschap. Hieruit concluderen dat in bepaalde agentschappen een strategie hebben ontwikkeld om doelstellingen zo weinig mogelijk meetbaar te maken, is echter een brug te ver.

Tot slot is er de vijfde assumptie dat het risico op perverse effecten met verantwoording als evaluatiemotief groter is geworden na BBB. Feit is dat de beheersovereenkomsten groter gewicht hebben gekregen via de hervorming. Een vergelijking van de beheersovereenkomsten die het Rekenhof bestudeerde met de meest recente versies, leert overigens dat de rubriek “evaluatie” meer gewicht heeft gekregen in de beheersdocumenten. De nadruk op evaluatie vanuit het verantwoordingsmotief is dus gestegen.

In theorie is dus ook het risico op perverse effecten gegroeid. Anderzijds kunnen op basis van de empirische gegevens die werden verzameld daarover geen eenduidige uitspraken worden gedaan. We zien namelijk tegenstrijdige evoluties in de rubriek “sancties en beloningen”. Slechts één beheersovereenkomst (DS) voorziet nog steeds expliciet in beloningen (d.m.v. verhoogde dotatie). Bij Vlopera en VITO is dergelijke verwijzing weggevallen sedert de studie van het Rekenhof. De evolutie in de passages over sanctionering is interessant. Bij de VDAB en Vlopera werd en wordt expliciet verwezen naar een vermindering van de dotatie bij een ongunstige evaluatie. Bij het VFL spreekt men van remediëring op financieel en/of management vlak. Bij de VRT en de VVM was dergelijke expliciete link aanwezig ten tijde van de Rekenhofstudie, maar niet meer in de meest recente beheersovereenkomst. Men spreekt in tegendeel van procedures voor heronderhandeling en overleg en maatregelen met het oog op bijsturing of eventuele sancties. Bij het VITO houdt men het op het bepalen van de voorwaarde voor een nieuwe BO mede op basis van de resultaten op de performantiecriteriën. Bij de VRT en de VDAB is bovendien een expliciete link met de evaluatie van de leidend ambtenaar verdwenen in de meest recente versie van de beheersovereenkomst.

De vraag rijst dan ook hoe deze ongelijke evolutie te verklaren is. Het opstellen van een beheersovereenkomst gebeurt in een krachtenveld tussen minister, departement en agentschap. De uitkomst van onderhandelingen over de beheersovereenkomst zal dan ook verschillen per agentschap naargelang de verdeling van de krachten en de professionele verhoudingen tussen de minister en beleidsverantwoordelijken in departement en agentschap.

## 5. Risico's verkleinen

De empirische analyse van beleidsnota's en beheersovereenkomsten levert ons een beeld op van de te verwachten mogelijke risico's in deze post-BBB periode van bestuur. Onze assumpties kunnen evenwel niet allemaal worden bevestigd. Uit de screening van beleidsnota's kunnen we besluiten dat diverse evaluatiemotieven aandacht krijgen (assumptie 3). De NPM oriëntatie van de hervormingen vindt bijgevolg geen eenduidige reflectie in een overwicht van verantwoordingsgerichte motieven in beleidsdocumenten. Deze algemene vaststelling dient echter te worden genuanceerd indien we de evaluatiemotieven bekijken per sector. De cijfers onderschrijven het bestaan van variatie naargelang beleidsdomein, in de lijn met internationale bevindingen (assumptie 4). Waar het verantwoordingsmotief primeert, zal, zoals verondersteld, eerder geopteerd worden voor outputindicatoren, dan wel effectindicatoren (assumptie 1).

Om in de verantwoordingscontext het risico in te perken, kan men inzetten op verschillende pistes, die elders al hun succes hebben bewezen. Ten eerste kunnen mechanismen van interne controle en audit bijdragen tot het indijken van perverse effecten met name op het vlak van de problematiek van betrouwbaarheid van meetgegevens en de validiteit van indicatoren die voor beleidsevaluatie worden ontwikkeld. Ten tweede dient te worden gewaakt over een adequate SMART

formulering van doelstellingen, om eventueel pervers gedrag te vermijden. De empirische screening liet in dit verband niet toe assumptie 2 eenduidig te toetsen. Ten derde, hieraan gelinkt, is eigenaarschap bij de ontwikkeling van indicatoren in het kader van beleidsevaluatie noodzakelijk. Men kan dit element nog uitbreiden door te stellen dat ook indicatoren SMART geformuleerd moeten worden; ze moeten duidelijk zijn, praktisch meetbaar, afgesproken, relevant in het kader van het evaluatiemotief en een tijds kader meekrijgen. Ten slotte is eigenaarschap ook een cruciaal issue doorheen het evaluatieproces; beleidsverantwoordelijken moeten betrokken worden, bijvoorbeeld in een stuurgroep die de evaluatie opvolgt, zodat keuzes doorheen dat proces gezamenlijk kunnen worden genomen. Op deze wijze wordt ook het gebruik of de invloed van evaluatie in het kader van beleidsontwikkeling en –beoordeling bevorderd.

Op langere termijn komt het erop aan om de hefboomen te bewegen die de evaluatiecultuur kunnen versterken. Een rijpe evaluatiecultuur ontstaat door een evenwichtige aandacht voor de verschillende evaluatiemotieven, maar ook door een versterkte evaluatiecapaciteit. Netwerking –we denken aan de rol van het recent opgerichte Vlaams Evaluatieplatform– en een verdere verankering van beleidsevaluatie in de beleidscyclus kunnen daartoe bijdragen. De relatie tussen evaluatiecapaciteit en –cultuur is echter complexer dan dat. Om een capaciteit op te bouwen zijn ook culturele elementen noodzakelijk, zoals politiek engagement, de overtuiging bij beleidsmakers van de meerwaarde van evaluatie voor het beleid en –vanzelfsprekend- ook een concrete vraag naar beleidsevaluatie.<sup>42</sup>

Verder onderzoek kan helpen een nog effectievere en efficiëntere strategie voor risicobeperking uit te bouwen. Teneinde aangepaste maatregelen per beleidssector te nemen, is het noodzakelijk een beter zicht te krijgen over de precieze dynamieken, ten grondslag aan de contingente en variabele aanpak van beleidsevaluatie in beleidssectoren, die onze empirische screening lijkt te suggereren. In dit verband is het eveneens primordiaal op te volgen hoe de verhouding van evaluatiemotieven zich zal ontwikkelen in de toekomst. Gezien het relatief recente karakter van de BBB hervormingen, is het tot nog toe onvoldoende duidelijk welke tendensen en mogelijke neveneffecten bij meten en evalueren zich op lange termijn kunnen ontwikkelen (assumptie 5). Naast een analyse van het discours, is het hierbij uiteraard van belang, oog te hebben voor de effectieve ontwikkeling van mogelijke risico's in de eigenlijke evaluatiepraktijk.

---

<sup>42</sup> Voor meer over dit thema, zie: De Peuter B. & Pattyn V. (2008), Evaluation capacity: enabler or exponent of evaluation culture? Paper for the Symposium "Policy and programme evaluation in Europe: cultures and prospects." Strasbourg, July 3 & 4, 2008.

**Tabel 4.2: Vergelijkende inhoudsanalyse van 7 beheersovereenkomsten (Rekenhofstudie en eigen studie)**

<p><i>Doelstellingen van de VOI (na BBB: agentischap)</i></p>	<p>VRT (Rekenhof)</p> <p>Samenvatting van de openbare omroepopdracht omschreven in het maxidecreet (art. 1). Beheersovereenkomst bevat de strategische doelstellingen en de bijhorende performantiemaatstaven.</p>	<p>VRT (BO 2007-2011)</p> <p>De openbare omroepopdracht bestaat uit de inhoudelijke opdracht (het aanbod van de VRT) die vastgelegd is in art. 6, §2, van de gecoördineerde mediadecreten en omschreven is in hoofdstuk 2 van de voorliggende beheersovereenkomst. (art.1) Beheersovereenkomst bevat de strategische doelstellingen en de bijhorende performantiemaatstaven.</p>
<p><i>Verplichtingen, prestaties, indicatoren, normen en timing</i></p>	<p>De beheersovereenkomst bepaalt voor elke strategische doelstelling verschillende performantiemaatstaven. De meeste zijn meetbaar gemaakt door indicatoren en normen (zowel kwantitatieve en kwalitatieve, output en effect indicatoren.) - Voor de meeste indicatoren is een timing voorzien (meestal te behalen tegen einde van overeenkomst). - Daarbuiten bevat de beheersovereenkomst nog een aantal andere verplichtingen: *enkele projecten voor radio en televisie uit te voeren binnen looptijd eerste overeenkomst (art. 2,§ 1 &amp; art. 3, §1); *duidelijke profilering radio- en televisienetten; *Participatie in Vlaamse onafhankelijke audio-visuele productie (art. 3, §1); *Besparing op personeelskosten (art. 6); *Hernegotiatie pensioencontract (art. 7);</p>	<p>Strategische doelstellingen performantiemaatstaven hebben naast de inhoudelijke opdracht ook betrekking op (hoofdstuk 4-7): engagements m.b.t. technologie transmissie van de inhoudelijke opdracht personeelsbeleid financiële afspraken tussen VRT en Vlaams Gemeenschap Verder is sprake van “toegevoegde opdrachten”: onderzoek en innovatie in de media, conservering en ontsluiting van het archief door digitalisering en renovatie van het omroepgebouw. (art. 1 en hoofdstuk 3) De VRT moet jaarlijks vóór 1 juni een rapport voorleggen dat voor elk van de performanciecriteria aangeeft in hoeverre de vooropgestelde doelstellingen zijn bereikt (art. 42). De VRT moet jaarlijks vóór 1 juni een nota voorleggen dat voor elk van de performanciecriteria aangeeft in hoeverre de vooropgestelde doelstellingen zijn bereikt (art. 42).</p>
<p><i>Rapporteringsverplichtingen</i></p>	<p>De VRT moet jaarlijks vóór 1 juni een rapport voorleggen dat voor elk van de performanciecriteria aangeeft in hoeverre de vooropgestelde doelstellingen zijn bereikt (art. 13). -De VRT moet jaarlijks vóór 1 juni een rapport uitbrengen over de mate waarin de performanciecriteria zijn bereikt (art. 13). - Er wordt nergens bepaald wie de evaluatie moet uitvoeren. - Geen tussenkomst van Regeringscommissaris voorzien, maar in de praktijk maakt hij wel een evaluatie voor de voogdijminister</p>	<p>De gemeenschapsafgevaardigde stelt eveneens een evaluatierapport op dat hij voorlegt aan de Vlaamse regering en waarvan een kopie aan de VRT wordt bezorgd. (art. 42)</p>
<p><i>Evaluatie</i></p>		

		<p>De Vlaamse Regulator voor de Media rapport jaarlijks aan de Vlaamse Regering m.b.t. inhoudelijk toezicht op de naleving door de VRT van de BO, en met kopie aan de VRT. (art. 42)</p> <p>De Vlaamse Regering legt de nota van de VRT samen met het evaluatierapport van de gemeenschapsafvaardigde en het rapport van de VRM voor aan het Vlaams Parlement vóór 30 september. De gedelegeerd bestuurder van de VRT licht de jaarlijkse rapportering toe in de bevoegde parlementaire commissie. (art. 42)</p> <p>Het Rekenhof kan de getrouwheid van de gegevens en van de meetsystemen die worden gebruikt in de rapportering controleren. De VRT verbindt er zich toe hieraan de nodige medewerking te verlenen. (art. 42)</p> <p>In 2011 wordt de rapporteringsnota van de VRT over het voorgaande jaar uitgebreid met een globale evaluatie over de periode 2007-2011. (art. 42)</p> <p>Rekening houdend met de voortdurende belangrijke evoluties in de mediemarkt en in de technologie zal deze beheersovereenkomst na 3 jaar geëvalueerd worden, m.b.t. evoluerende medialandschap en de rol van de VRT daarin.. Het evaluatiemoment maakt het mogelijk om de openbare omroepdracht van de VRT eventueel bij te sturen en eventueel uit te breiden in functie van de evoluties van de mediemarkt. (art. 3 ) Ingeval van uitbreiding van de opdracht wordt gezamenlijk een addendum bij de BO besloten en de financiering vastgesteld. (art. 17)</p>
<p><i>Sancties en beloningen</i></p>	<ul style="list-style-type: none"> <li>-Toekennen van de jaarlijkse verhoging van 4% van de dotatie indien de vooropgestelde begrotingsnorm is gehaald (BO, art. 8).</li> <li>- Link tussen doelstellingen en verplichtingen in beheersovereenkomst en evaluatie van de gedelegeerd bestuurder gelegd in het arbeidscontract.</li> </ul>	<p>Heronderhandeling kan aangevraagd worden zo de omstandigheden dermate ingrijpend gewijzigd zijn dat de openbare omroepdracht en de financiering ervan niet langer adequaat zijn, zonder dat de basisfinanciering enveloppe kan verminderd worden.</p> <p>In het geval dat één van de partijen manifest zijn engagement niet zou nakomen en beide partijen niet tot een vergelijk komen, wordt een arbitragecommissie opgericht om in functie van de specifieke omstandigheden de gepast maatregelen te bepalen. (art. 43).</p>

<i>Type van indicatoren</i>	---	42 indicatoren, 10 input, 25 output en 7 effect 25 kwantitatief, 17 kwalitatief geformuleerd ( <i>eigen berekening</i> )
Doelstellingen van de VOI ( <i>na BBB: agentschap</i> )	VDAB (Rekenhof) -Verwijzing naar oprichtingsdecreten en andere regelgeving in verband met taken VDAB (art. 1). -Oprichting van de VDAB nader omschreven in het ondernemingsplan. -De beheersovereenkomst bevat geen strategische doelstellingen maar wel verbintenissen (art. 5).	VDAB (BO 2007-2011) Verwijzing naar oprichtingsdecreten en andere regelgeving in verband met taken VDAB (art. 4). De vertaling van engagementen in deze BO in concrete, kwantitatieve objectieven in overleg en opgenomen in addendum. Verdere vertaling van de BO in het meerjarenondernemingsplan met zelfde looptijd en verrijft in een jaartaactieplan. (art. 7)
Verplichtingen, prestaties, indicatoren, normen en timing	-Beheersovereenkomst bevat 6 verbintenissen. De meeste zijn toetsbaar gemaakt door middel van kwantitatieve parameters (hoofdzakelijk outputindicatoren) en genormeerd (optimalisatie%). Timing voorzien. (art. 5 en bijlage 1) - Kwalitatieve parameters moeten worden ontwikkeld (art. 5) - Andere verbintenissen (niet-trajectgerelateerde activiteiten, optimaal beheer van werkingsmiddelen, ...) zijn niet genormeerd, enkele zijn moeilijk meetbaar.	BO bevat engagementen. Ze worden geoperationaliseerd in concrete parameters. Daarbij wordt onderscheid gemaakt tussen: - doelstellingsparameters: parameters waarvoor een norm wordt afgesproken en waarop de VDAB kan worden afgekend - monitoringsindicatoren: parameters waarvoor geen normen worden vooropgesteld maar die wel van belang zijn in de opvolging van de BO. Aanpassing in definities en berekening mogelijk, in overleg.
Rapporteringsverplichtingen	- De gemeenschapscommissarissen moeten jaarlijks vóór 31 maart rapporteren aan voogdijminister en minster van financiën en begroting over controle op de uitvoering van de beheersovereenkomst. Hun verslag wordt betrokken in de jaarlijkse evaluatie. (art. 8) - Geen andere rapporteringsverplichtingen.	Zie evaluatie.
Evaluatie	- Jaarlijks vóór 30 april evalueren beide contractanten de beheersovereenkomst en de resultaten ervan, rekening houdend met het evaluatieverslag van de gemeenschapscommissarissen (art. 7)	Jaarlijkse evaluatie via: (art. 22) - rapportering van VDAB met stand van zaken en vooruitzichten mbt engagementen adhv opgenomen parameters, ook ivm financiën en personeel - beleidsmatige toetsingsrapport van de administratie - werkgelegenheid, obv gegevens door VDAB aangeleverd. - auditrapport van het Rekenhof, indien dit rapport betrekking heeft op de werking van de VDAB

Sancties en beloningen	<ul style="list-style-type: none"> <li>- Dotatievermindering voor volgend werkjaar indien prestaties 5% lager dan norm. (art. 7)</li> <li>- Jaarlijkse evaluatie van leidend ambtenaar. Geen expliciete link tussen evaluatie en doelstellingen van beheersovereenkomst.</li> </ul>	<p>Indien uit jaarlijkse evaluatie blijkt dat de VDAB zijn verbintenissen niet is nagekomen, kan de regering de dotatie voor volgende jaren bijsturen. Er wordt in zo'n geval rekening gehouden met vaste en variabele kosten en met effect op dienstverlening.</p>
Type van indicatoren	---	<p>28 doelstellingsparameters, 3 input, 12 output en 12 effect</p> <p>25 kwantitatief, 3 kwalitatief geformuleerd</p> <p>Monitoringsindicatoren: overwegend output (eigen berekening)</p>
<i>Doelstellingen van de VOI (na BBB: agentschap)</i>	<p>Vlopera (Rekenhof)</p> <ul style="list-style-type: none"> <li>- Verwijzing naar opdrachtingschrijving in het oprichtingsdecreet (art. 1&amp;3)</li> <li>- Beheersovereenkomst bevat twee strategische doelstellingen en bijhorende performantie-indicatoren (art.4)</li> <li>- De opdracht van de Vlopera wordt nader omschreven in een ondernemingsplan, waarvan de concrete invulling ieder jaar moet worden meegedeeld aan de Vlaamse Regering (art.3)</li> </ul>	<p>Vlopera (BO 2006-2010)</p> <ul style="list-style-type: none"> <li>- Verwijzing naar opdrachtingschrijving in het Kunstendecreet (art.5)</li> <li>- Geen systematische oplijsting van strategische en operationele doelstellingen in BO. Indirect kunnen 8 strategische doelstellingen en 30 operationele worden geïdentificeerd doorheen de tekst.</li> <li>- De doelstellingen worden nader toegelicht in een meerjarig beleidsplan. Deze planning vormt een onderdeel uit van de beheersovereenkomst (art. 6)</li> <li>- Jaarlijks dient een geactualiseerd beleidsplan te worden ingediend (art. 6).</li> </ul>
<i>Verplichtingen, prestaties, indicatoren, normen en timing</i>	<ul style="list-style-type: none"> <li>- De beheersovereenkomst bevat een beperkt aantal indicatoren voor de twee strategische doelstellingen (art.4). De indicatoren zijn vaag geformuleerd en niet genormeerd. Het zijn kwalitatieve en effectindicatoren.</li> <li>- Het ondernemingsplan bevat geen duidelijke indicatoren.</li> <li>- Daarnaast bevat de overeenkomst nog een aantal concrete verbintenissen (art.5): <ul style="list-style-type: none"> <li>* minimumaantal operaproducties, eigen producties en voorstellingen per seizoen; minimumaantal concerten per seizoen;</li> <li>* Spreiding voorstellingen tussen beide plateaus;</li> <li>* contingent betrekkingen waarmee de instelling kan werken;</li> <li>* Minimumpercentage eigen inkomsten 12,5% punt 4.6).</li> <li>* minimumaantal concerten per seizoen;</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>- Operationele doelstellingen en dito indicatoren zijn overwegend kwalitatief (n=24), en zijn veelal niet SMART gedefinieerd. BO bevat slechts 6 kwantitatief geformuleerde doelstellingen.</li> <li>- De beheersovereenkomst bevat een beperkt aantal expliciete indicatoren. Deze zijn echter vaak vaag geformuleerd en meestal niet genormeerd.</li> <li>- Indicatoren zijn overwegend outputindicatoren.</li> </ul>



	<ul style="list-style-type: none"> <li>* Spreiding voorstellingen tussen beide plateaus;</li> <li>* contingent betrekkingen waarmee de instelling kan werken;</li> <li>* Minimumpercentage eigen inkomsten 12,5%</li> </ul>	
<i>Rapporteringsverplichtingen</i>	<ul style="list-style-type: none"> <li>- Vlopera moet een performantierapport voorleggen vóór 30 april (art. 4) en een rapport vóór 30 mei (art. 7).</li> <li>- Beide gemeenschaps-commissarissen moeten vóór 30 mei een evaluatierapport voorleggen (art. 7).</li> </ul>	<ul style="list-style-type: none"> <li>- Vlopera moet een jaarverslag voorleggen van elk werkingsjaar vóór 31 mei (art.7)</li> </ul>
<i>Evaluatie</i>	<ul style="list-style-type: none"> <li>- Jaarlijks zal vóór 30 mei een evaluatie plaatsvinden over de uitvoering van de beheersovereenkomst (art. 7).</li> <li>- Daartoe moet de Vlopera vóór 30 mei een rapport indienen, waaraan de evaluatierapporten van beide gemeenschapscommissarissen zijn toegevoegd.</li> <li>- Er is niet bepaald wie de eindevaluatie uitvoert.</li> </ul>	<ul style="list-style-type: none"> <li>- De administratie zal jaarlijks een evaluatie uitvoeren van de zakelijke werking van Vlopera. Indien ze het nodig acht, kan eveneens een evaluatie van de artistieke werking worden uitgevoerd (art. 9 §1). Halverwege de looptijd van de BO is er een tussentijdse evaluatie van de uitvoering van het meerjarig beleidsplan (art. 9 §2).</li> <li>- De evaluatie gebeurt aan de hand van de (weliswaar moeilijk te identificeren en vaak afwezige) beoordelingscriteria gestipuleerd in de BO en aan de hand van criteria uit het Kunstendecreet. Een normstelling ontbreekt veelal.</li> <li>- Wanneer de organisatie niet voldoet aan de afspraken in de beheersovereenkomst, kan worden overgegaan tot een financiële remediering en/of een remediering van het management (art.7 §6).</li> <li>- Geen vermelding van performantieboni, noch vermelding van eventuele jaarlijkse evaluatie van leiding van Vlopera.</li> </ul>
<i>Sancties en beloningen</i>	<ul style="list-style-type: none"> <li>- Mogelijke toekenning van performantiebonus indien voldaan is aan de performantie-indicatoren (art. 6, §4.)</li> <li>- herziening dotatie indien niet is voldaan aan de verplichtingen van de beheersovereenkomst. (art. 6, § 7).</li> <li>- Geen jaarlijkse evaluatie van leiding van Vlopera</li> </ul>	
<i>Type van indicatoren</i>	--	<ul style="list-style-type: none"> <li>Beperkt aantal expliciete indicatoren, vaak vaag geformuleerd en meestal niet genormeerd. Overwegend outputindicatoren.</li> </ul>
<i>Doelstellingen van de VOI (na BBB: agentschap)</i>	<p>VITO (Rekenhof)</p> <ul style="list-style-type: none"> <li>- Het "protocol" bepaalt performantie-indicatoren voor het meten van kwaliteit, relevantie, efficiëntie en flexibiliteit van het VITO-onderzoek te meten (art.4)</li> <li>- Het "technisch dossier" omvat de strategische doelstellingen en krachtlijnen van het wetenschappelijk onderzoek, inclusief bijhorende resultaatsverintenissen.</li> <li>- Er moet ook jaarlijks een gedetailleerde doelstellingsnota worden opgemaakt</li> </ul>	<p>VITO (BO 2008-2012)</p> <ul style="list-style-type: none"> <li>Rekening houdende met de suggesties van een externe doorlichting van het VITO in 2005 worden in het addendum van de beheersovereenkomsten drie strategische doelstellingen gedefinieerd. De operationele vertaling gebeurt aan de hand van "key performance indicators" (art. 5, addendum bij BO)</li> <li>Daarnaast moet het VITO ook referentietaken uitvoeren tby de Vlaamse Overheid (technische metingen). (Art. 2 Addendum onder</li> </ul>



<p><i>Verplichtingen, prestaties, indicatoren, normen en timing</i></p>	<ul style="list-style-type: none"> <li>- Het technisch dossier omvat voor elke onderzoeksafdeling concrete resultaatverbintenissen. Deze zijn niet altijd toetsbaar gemaakt.</li> <li>- Het protocol bepaalt 8 prestatie-indicatoren. Er is geen duidelijke link met de resultaatverbintenissen in het technisch dossier. Er is één globale en relatieve norm: prestatie evolueert gunstig als minstens 5 van de 8 indicatoren gunstig evolueren.</li> <li>- Art. 5, §4 bepaalt dat in overleg met de voogdijminister specifieke streefcijfers inzake prestatie-indicatoren worden vooropgesteld</li> <li>- De beheersovereenkomst bevat een beperkt aantal andere verplichtingen (ter beschikking stellen van doctoraatsmandaten - art.9, ...)</li> </ul>	<p>algemene bepalingen/s</p> <ul style="list-style-type: none"> <li>- In de BO wordt VITO eraan gehouden de doelstellingen na te streven</li> <li>- Voor twee van de drie strategische doelstellingen zijn in het addendum van de beheersovereenkomst output- en effectindicatoren opgenomen, drie voor de eerste doelstelling, twee voor de tweede doelstelling.</li> <li>- deze indicatoren zijn uitgewerkt met timing en normering.</li> <li>- de indicatoren zijn kwantitatief van aard</li> <li>- voor de derde operationele doelstelling voorziet men nog een set van indicatoren per referentietask op te nemen in een addendum</li> </ul>
<p><i>Rapporteringsverplichting en</i></p>	<ul style="list-style-type: none"> <li>- Te voegen bij begrotingsontwerp voor 31 mei: activiteitenverslag en samenwerkingsverslag vorig jaar, evaluatieverslag prestatie-indicatoren (art.6, §1).</li> </ul>	<ul style="list-style-type: none"> <li>- Jaarlijks tegen 30 april bezorgt VITO een 'Lentebegroting' (art. 19)</li> <li>- Jaarlijks op 1 juni bezorgt het VITO een jaarrekening en jaarverslag, waarin de 'mate waarin aan operationele doelstellingen werd tegemoetgekomen' wordt omschreven (art. 20)</li> <li>- Daarnaast is de instelling ook gehouden aan het actief (voor brede publiek art. 22) en passief ter beschikking stellen van informatie (voor minister en IvF, art. 21)</li> </ul>
<p><i>Evaluatie</i></p>	<ul style="list-style-type: none"> <li>- Jaarlijks evaluatieverslag m.b.t. de prestatie-indicatoren opgesteld door VITO, in samenspraak met minister, en mede gebaseerd op evaluatie door wetenschappelijke adviesraden (art.7).</li> </ul>	<ul style="list-style-type: none"> <li>- Jaarlijks voert het Departement een doorlichting uit op basis van een rapport van de onderzoeksinstelling mbt de resultaten op de prestatie-indicatoren, waaruit moet blijken hoe deze laatste evolueren (addendum 1, art. 6)</li> <li>- VITO zal in 2012 worden geëvalueerd (BO, art. 24).</li> <li>- Voor de aansturing van evaluatie zal een stuurgroep worden aangesteld door de Minister(BO, art. 24).</li> <li>- De BO (art. 24) vermeldt dat de evaluatie zal rekening houden met de jaarverslagen, een bibliometrische analyse, een oordeel van een internationaal panel van experts, een benchmarking, en een analyse van de impact van VITO op de Vlaamse economie</li> <li>- in addendum 1 (art. 19) worden voorwaarden gesteld aan het samenstellen van het expertenpanel</li> <li>- in addendum 1 (art. 19), wordt vooropgesteld dat het VITO een zelfevaluatie zal maken, met inbegrip van een analyse van prestaties, ter voorlegging aan het expertenteam.</li> </ul>

<i>Sancities en beloningen</i>	<ul style="list-style-type: none"> <li>- Als performantie van de VITO in het jaar t gunstig evolueert (= wanneer 5 van de 8 indicatoren gunstig evolueren), dan bonus bij opmaak begroting jaar t+2 (art.5, §4).</li> <li>- Geen jaarlijkse evaluatie van leiding van VITO.</li> </ul>	De evaluatie zal mede de voorwaarden bepalen van een nieuwe beheersovereenkomst met inachtneming van de algemene resultaten en de resultaten op de performantiecriteria (art. 24)
<i>Indicatoren</i>		5 kwantitatieve indicatoren waarvan 4 output, 1 effect

<i>Doelstellingen van de VOI (na BBB: agentschap)</i>	<p>DS (Rekenhof)</p> <ul style="list-style-type: none"> <li>- Strategische-doelstellingen, kernopdrachten en performantiemaatstaven voor de VOI expliciet opgesomd in beheersovereenkomst (punt 3.1 en 3.2).</li> <li>- Performantiemaatstaven (= operationele doelstellingen) achteraf te concretiseren via actieplannen (tegen eind 1999)</li> </ul>	<p>DS (BO 2007-2010)</p> <p>De BO vermeldt 6 strategische doelstellingen (art.9). Zij maakt hierbij onder bijlage 1 een koppeling met de doelstellingen in de beleidsnota van de Minister.</p> <ul style="list-style-type: none"> <li>- De strategische ze doelstellingen worden in detail geoperationaliseerd via strategische projecten (met kwalitatieve omschrijving) en met 22 prestatie-indicatoren voor de recurrente prestaties.(art. 10). Per prestatie-indicator is er in bijlage 3 een toelichting.</li> </ul>
<i>Verplichtingen, prestaties, indicatoren, normen en timing</i>	<ul style="list-style-type: none"> <li>- Voor alle doelstellingen zijn één of meerdere indicatoren vermeld (output- en effectindicatoren).</li> <li>- Voor alle indicatoren is norm en timing vastgelegd.</li> <li>- Niet alle indicatoren zijn reeds uitgewerkt in beheersovereenkomst</li> <li>- Andere verplichtingen zijn: <ul style="list-style-type: none"> <li>* De DS commercialiseert de watergebonden gronden. (punt 4.2)</li> <li>* De DS treedt op als bouwheer voor investeringswerken en stelt jaarlijks investeringsprogramma voor (punt 4.3 en 4.4). Deelname aan beleid en overleg (punt 4.6).</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>- Naast het nastreven van de strategische doelstellingen engageert de DS. zich in de beheersovereenkomst voor de volgende horizontale doelstellingen: samenwerking, kwaliteitsbeleid, correct werkgeverschap, diversiteit, duurzaam en ethisch aankoopbeleid (art. 11)</li> <li>- Voor de strategische doelstellingen zijn zowel in de tekst van BO zelf als in de bijlage de prestaties en indicatoren opgesteld via het SMART principe. Ze zijn alle begeleid van timing en normering. Deze 22 indicatoren zijn op 3 na kwantitatief, betreffend overwegend output, en slechts in 4 gevallen proces, in 1 geval input.</li> <li>- Voor de horizontale doelstellingen worden geen indicatoren vermeld.</li> <li>In art. 11 engageert de DS zich ook tot effectmeting voor een drietal doelstellingen alsook tot het aanbrengen van gegevens (inputindicatoren) die het departement moeten toelaten over beleids effecten te rapporteren (begroting, personeelsbezetting, diversiteitskengetallen). Ze verbindt er zich ook in het algemeen</li> </ul>

		<p>toe alle outputgegevens aan het departement te verstrekken die nuttig kunnen zijn voor het meten, opvolgen, en evalueren van beleidseffecten.</p> <p>Noch in de BO en bijlagen zijn er effectindicatoren gedefinieerd, en genormeerd. (bijlage 4), maar alleen vaag omschreven.</p>
<i>Rapporteringverplichtingen</i>	<ul style="list-style-type: none"> <li>- Jaarlijks tegen 15 maart aan Vlaamse regering, na certificering door bedrijfsrevisor (punt 5): <ul style="list-style-type: none"> <li>* Jaarrekening</li> <li>* Toestand reservefonds</li> <li>* Staat cijfermatige performantiemaatstaven</li> </ul> </li> <li>- Wijze van rapportering niet nader bepaald.</li> </ul>	<ul style="list-style-type: none"> <li>- Art. 23 maakt melding van de volgende rapporteringsmodaliteiten <ul style="list-style-type: none"> <li>- Een uit te bouwen MIS systeem zal rapporteringen over prestaties genereren voor de minister.</li> </ul> </li> <li>- Jaarrapport (feb.) en eindrapport hebben als voorwerp: de performantiematen, (niet)-realisatie van doelstellingen, horizontale projecten, interveniërende factoren, uitvoering van de begroting.</li> </ul>
<i>Evaluatie</i>	<ul style="list-style-type: none"> <li>- Controleverslag van regeringscommissarissen over uitvoering van beheersovereenkomst (jaarlijks tegen 1 april aan bevoegde minister) (punt 5).</li> <li>- Evaluatie niet nader geregeld.</li> </ul>	<ul style="list-style-type: none"> <li>- Een jaarlijkse evaluatie in overlegmoment. Wanneer verplichtingen niet zijn nagekomen wordt overlegd over bijsturingen.</li> <li>- In 2009 eindevaluatie op basis van realisaties van strategische organisatiedoelstellingen in waarvoor een evaluatieorgaan kan worden aangesteld (art. 24)</li> </ul>
<i>Sancities en beloningen</i>	<ul style="list-style-type: none"> <li>- Aanpassing bonusdotatie in functie van realisatie van beoordelingscriteria vanaf jaar 2000 (punt 6).</li> <li>- Jaarlijkse evaluatie van leidend ambtenaar. Geen expliciete link tussen evaluatie en doelstellingen van beheersovereenkomst</li> </ul>	<ul style="list-style-type: none"> <li>- Art. 20 bepaalt dat 4 (van de 22) indicatoren gebruikt worden als beoordelingscriteria bij de dotaties. Bij elk van de vier maatstaven waarvan de norm wordt gehaald, zal de dotatie met 0,5 percent worden verhoogd. Het betreft drie outputindicatoren en 1 procesindicator.</li> </ul>
<i>Indicatoren</i>		<ul style="list-style-type: none"> <li>- 22 indicatoren zijn op 3 na kwantitatief, overwegend output, en slechts in 4 gevallen proces, in 1 geval input.</li> <li>- Daarnaast is er ook impliciet melding van verschillende andere indicatoren maar deze zijn niet gedefinieerd noch genormeerd, maar moeten effectmeting door het Departement ondersteunen.</li> </ul>

<p><i>Doelstellingen van de VOI (na BBB: agentschap)</i></p>	<p>Vlaamse Vervoermaatschappij De Lijn (Rekenhof)</p> <ul style="list-style-type: none"> <li>- De VVM voert haar beleid uitgaand van haar strategisch ondernemingsplan (art. 5);</li> <li>- De doelstellingen en opties uit het ondernemingsplan worden vermeld in de beheersovereenkomst (art. 5)</li> <li>- Strategisch ondernemingsplan moet worden geconcretiseerd in (interne) operationele plannen en jaarlijkse actieplannen (art. 5), dit is tot op heden slechts gedeeltelijk gebeurd.</li> </ul>	<p>VVM (BO 2003-2009)</p> <ul style="list-style-type: none"> <li>- De VVM voert haar beleid uitgaande van haar strategisch ondernemingsplan, wat integraal deel uitmaakt van de BO (art. 7).</li> <li>- Het ondernemingsplan concretiseert de bepalingen, vastgelegd in de beheersovereenkomst (art. 7).</li> </ul>
<p><i>Verplichtingen, prestaties, indicatoren, normen en timing</i></p>	<ul style="list-style-type: none"> <li>- De beheersovereenkomst legt een aantal verplichtingen en uit te voeren prestaties op aan de VVM. Er is geen duidelijke relatie met strategische doelstellingen in beheersovereenkomst.</li> <li>- De overeenkomst omvat een aantal kwantitatieve indicatoren met normwaarden: <ul style="list-style-type: none"> <li>- Minimaal aantal kilometers (art.7, §5);</li> <li>- Gebruik van aanbod (vervoersonvangsten - art.7, §6);</li> <li>- Evenwicht in exploitatierkening (art.7, §2);</li> <li>- Levensduur rollend materieel (art.14)</li> </ul> </li> </ul> <p>De normen moeten in principe jaarlijks worden gehaald.</p> <ul style="list-style-type: none"> <li>- Overeenkomst omvat ook aantal andere verplichtingen: <ul style="list-style-type: none"> <li>* Lastenboek uitwerken en mededinging organiseren voor contracten exploitanten (art.8);</li> <li>* Samenwerking met andere vervoermaatschappijen (art.13);</li> <li>* Uitwerken van kwaliteitsmeetsysteem (art. 7, §4);</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>- De beheersovereenkomst vermeldt onder '<i>verplichtingen van de maatschappij</i>' 8 strategische doelstellingen, en 22 operationele doelstellingen. De expliciete relatie tussen de strategische en operationele doelstellingen wordt geëxpliciteerd in annex van de BO (annex 10).</li> <li>- Voor elk van de operationele doelstellingen is een meetinstrument beschikbaar. De meerderheid van de doelstellingen bevat eveneens een meetbare norm. (annex 10)</li> <li>- De meerderheid van indicatoren zijn outputindicatoren.</li> </ul>
<p><i>Rapporteringsverplichtingen</i></p>	<ul style="list-style-type: none"> <li>- Hoofdstuk 3 (verplichtingen van de VVM) bepaalt wanneer de VVM over de uitvoering van de verplichting moet rapporteren. Over het algemeen is aangegeven wanneer en aan wie moet worden gerapporteerd, niet op welke wijze dit moet gebeuren.</li> <li>- De jaarlijkse evaluatie van de overeenkomst zal gebeuren op basis van voormelde rapportering (art.21, §1).</li> </ul>	<ul style="list-style-type: none"> <li>- De maatschappij rapporteert over de uitvoering van het afgelopen jaar. De rapportering omvat de jaarrekening en het jaarverslag met het evaluatierapport van de beheersovereenkomst.</li> <li>- Het laatste jaar wordt er parallel met de jaarlijkse evaluatie ook een evaluatie van de volledige looptijd van de BO uitgevoerd.</li> <li>- Voor eind september van het laatste jaar evalueert de Vlaamse Regering de uitvoering van de BO en de overeenkomst zelf op basis van de rapportering (art. 26)</li> </ul> <p>Zie rapportering</p>
<p><i>Evaluatie</i></p>	<ul style="list-style-type: none"> <li>- De uitvoering over het voorbije jaar wordt jaarlijks vóór 15 juni geëvalueerd, dit op basis van de rapportering over de uitvoering van de verplichtingen (art.21, §1);</li> <li>- Er wordt nergens bepaald wie de evaluatie moet uitvoeren.</li> <li>- Geen tussenkomst regeringscommissaris voorzien.</li> </ul>	

<i>Sancties en beloningen</i>	<ul style="list-style-type: none"> <li>- Lijst van maatregelen indien één van de partijen verplichtingen niet nakomt, zowel voor VVM als voor Vlaamse regering (art. 22): <ul style="list-style-type: none"> <li>*vermindering dotatie</li> <li>*verminderen aanbod</li> <li>*...</li> </ul> </li> <li>- Link tussen doelstellingen en verplichtingen in BO en evaluatie van de leidend ambtenaren.</li> </ul>	<ul style="list-style-type: none"> <li>- Wanneer de in de overeenkomst voorziene verplichtingen door een partij niet worden nageleefd, plegen de partijen overleg over de te nemen bijsturingmaatregelen. Indien deze binnen de zes maanden geen resultaat hebben opgeleverd, komen beide partijen overeen “welke extra maatregelen of welke sanctie uitgevoerd moeten worden”. De overeenkomst maakt geen specificatie van de mogelijke sancties (art. 27).</li> </ul>
<i>Type van indicatoren</i>	--	28 indicatoren: input (3); proces (4); output (16); effect (5)
<i>Doelstellingen van de VOI (na BBB: agentschap)</i>	<p>Vlaams Fonds voor de Letteren (VFL) (Rekenhof)</p> <ul style="list-style-type: none"> <li>- De beheersovereenkomst verwijst naar de opdracht van het VFL opgenomen in het oprichtingsdecreet, maar bevat geen strategische doelstellingen.</li> <li>- De doelstellingen worden vastgelegd in het meerjarenplan (art. 1 §2).</li> <li>- Het meerjarenplan moet worden geconcretiseerd in jaarlijkse actieplannen (decreet, art. 1, 18°)</li> </ul>	<p>VFL (BO 2006-2010)</p> <ul style="list-style-type: none"> <li>- De beheersovereenkomst verwijst naar de opdracht van het VFL opgenomen in de beheersovereenkomst, wat als strategische doelstelling kan worden beschouwd. (art. 5)</li> <li>- Een 6-tal operationele doelstellingen worden vermeld. Deze zijn echter niet allemaal SMART geformuleerd (art. 6)</li> <li>- De doelstellingen worden vastgelegd in het meerjarenplan (art. 13 §1)</li> <li>- Het meerjarenplan moet worden geconcretiseerd in jaarlijkse actieplannen (art. 13 §2).</li> </ul>
<i>Verplichtingen, prestaties, indicatoren, normen en timing</i>	<ul style="list-style-type: none"> <li>- De beheersovereenkomst bevat geen meetbare doelstellingen (indicatoren en normen).</li> <li>- De overeenkomst bevat wel een beperkt aantal andere concrete verplichtingen: <ul style="list-style-type: none"> <li>* bestaande verbintenissen tussen de Vlaamse Gemeenschap en literaire organisaties met betrekking tot de subsidiëring overnemen. (art. 3)</li> <li>* deelnemen aan initiatieven genomen door de Nederlandse en Vlaamse ministers van Cultuur en door de Nederlandse Taalunie met het oog op de integratie van het Vlaamse en het Nederlandse letterenbeleid (art. 4)</li> <li>* Het totaal van de werkingskosten, personeelskosten en vergoedingen mag ten hoogste 12% van de totale dotatie bedragen (art. 6 §2);</li> <li>* Op eigen initiatief of op verzoek alle nuttige en nodige informatie meedelen aan de bevoegde administratie (art. 21)</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>- De vermelde doelstellingen zijn niet SMART gedefinieerd.</li> <li>- De aangegeven indicatoren (art. 17), sluiten niet onmiddellijk aan bij de doelstellingen (art. 6). Zij betreffen eerder inspanningsverbintenissen dan resultaatverbintenissen. Er is slechts één kwantitatief geformuleerde procesindicator (doorlooptijd van aanvraag beslissing). Normstelling ontbreekt veelal.</li> <li>- BO maakt wel melding van resultaatverbintenissen die dienen te worden aangegaan teneinde ‘culturele diversiteit’ na te streven. Deze worden gefixeerd in het jaarlijks actieplan (art. 9).</li> <li>- De BO vermeldt wel een beperkt aantal kwalitatief geformuleerde concrete verplichtingen</li> </ul>

<i>Rapporterings- verplichtingen</i>	- De beheersovereenkomst voorziet geen rapportering over de uitvoering van de beheersovereenkomst en het meerjarenplan.	- VFL moet een jaarverslag opstellen van elk werkingsjaar (art. 14) vóór 31 mei. -De BO voorziet geen rapportering van het gehele meerjarenplan.
<i>Evaluatie</i>	- De beheersovereenkomst voorziet geen evaluatie van de overeenkomst. De jaarrekening en het jaarverslag moeten door de gemeenschapscommissaris o.m. getoetst worden aan de beheersovereenkomst (BO, art. 10, §2).	- De administratie voert jaarlijks een evaluatie uit van de zakelijke werking van het VFL (art.16). - De inhoudelijke werking wordt jaarlijks geëvalueerd door het Vlaams Parlement op basis van het jaarverslag (art. 17).
<i>Sancties en beloningen</i>	Geen beloningen of sancties voorzien.	Indien het VFL niet voldoet aan de gemaakte afspraken in de beheersovereenkomst, kan worden overgegaan tot het opleggen van een financiële (geheel of gedeeltelijk staken van de dotatie) en/of remediering van het management (remediteren van de leidinggevende functies) (art. 14 §6)
<i>Type van indicatoren</i>	--	9 expliciete indicatoren: 8 kwalitatief – 1 kwantitatief

## Bibliografie

Balthasar, A. (2006). The effect of the institutional design on the utilization of evaluation: Evidenced using qualitative comparative analysis (QCA). In: *The International Journal of Theory, Research and Practice*, 12 (3), 353-371.

Bouckaert, G. & Auwers, T. (1999). *Prestaties meten in de overheid* (Overheidsmanagement nr. 5). Brugge: Die Keure.

Brans, M. & Vancoppenolle, D. (2005). Policy-making reforms and civil service: an exploration of agendas and consequences. In: Painter, M. & Pierre, J. (Eds.) *Challenges in State Policy Capacity: Global Trends and Comparative Perspectives*. Basingstoke: Palgrave MacMillan, 164-184.

Chelimsky, E. & Shadish, W.R. (eds.) (1997). *Evaluation for the 21<sup>st</sup> Century – A Handbook*. Thousand Oaks: Sage.

Conings, V., Sterck, M., Van Dooren, W. & Bouckaert, G. (2005). *Beleidsnota's versus Beter Bestuurlijk Beleid. Een toetsing van de beleidsnota's aan het Comptabiliteitsdecreet en het Kaderdecreet Bestuurlijk Beleid*. Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

De Peuter, B. (2008). Beleidsevaluatie: vraag en aanbod afstemmen en versterken via netwerking. De ambitie van het Vlaams Evaluatieplatform. In: *Vlaams Tijdschrift voor Overheidsmanagement*, 13(1), 36-44.

De Peuter, B., De Smedt, J. & Bouckaert, G. (2007). *Handleiding beleidsevaluatie – Deel 1: evaluatiedesign en –management*. Leuven, Steunpunt Bestuurlijke Organisatie Vlaanderen.

De Scheepvaart (2007). *Beheersovereenkomst 1 januari 2007 – 30 juni 2010 tussen de Vlaamse Regering en nv De Scheepvaart*.

Furubo, J.-E., Rist, R.C. & Sandahl, R. (eds.) (2002). *International Atlas of Evaluation*. New Jersey: Transaction Publishers.

Rekenhof (2001). *Auditrapport over het onderzoek naar de beheersovereenkomsten met Vlaamse openbare instellingen*.

Van Dooren, W. (2006). *Performance measurement in the Flemish public sector: A supply and demand approach*. K.U. Leuven.

Varone, F. & Jacob, S. (2003). *Evaluer l'action publique. Etat des lieux et perspectives en Belgique*. Gent: Academia Press.



VFL (2006). *Beheersovereenkomst 2006-2010 tussen de Vlaamse Regering en het Vlaams Fonds voor de Letteren.*

VITO (2008). *Beheersovereenkomst 2008-2012 met de Vlaamse Instelling voor Technologisch Onderzoek (VITO).*

Vlaams Parlement (18-07-2003). *Decreet Kaderdecreet Beter Bestuurlijk Beleid.*

Vlaams Parlement (2003). *Decreet Comptabiliteitsdecreet.*

Vlaamse Dienst voor Arbeidsbemiddeling (2005). *Beheersovereenkomst tussen de Vlaamse Regering en de Vlaamse Dienst voor Arbeidsbemiddeling en Beroepsopleiding voor de periode 2005-2009.*

Vlaamse Opera (2006). *Beheersovereenkomst 2006-2010 tussen de Vlaamse Gemeenschap en de Vlaamse Opera.*

Vlaamse Regering (22-03-2002). *Visienota Beheersovereenkomsten.*

Vlaamse Regering (22-07-2005a) *Nota Beheersovereenkomsten in de Vlaamse overheid: Concept, implementatie en model.*

Vlaamse Regering (22-07-2005b). Handleiding bij de opmaak (en de opvolging) van een beheersovereenkomst, in: *Nota Beheersovereenkomsten in de Vlaamse Overheid: Concept, implementatie en model, bijlage.*

Vlaamse Vervoersmaatschappij (2003). *Beheersovereenkomst voor 2003-2009 tussen de Vlaamse regering en de Vlaamse Vervoermaatschappij De Lijn.*

VRT (2006). *Beheersovereenkomst 2007 - 2011 tussen de VRT en de Vlaamse Gemeenschap.*



# **Hoofdstuk 5: De plaats en betekenis van beleidsevaluaties in de Vlaamse overheid –de domeinen Binnenlands Bestuur en Stedenbeleid verkend**

Ellen Wayenberg  
Hogeschool Gent en Universiteit Gent

<b>1. Inleiding .....</b>	<b>107</b>
<b>2. Evalueren van het Vlaamse beleid inzake binnenlands bestuur en steden: de puntjes op de i .....</b>	<b>107</b>
2.1. Het kernstuk van beleidsevaluatie .....	108
2.2. Het voorwerp van beleidsevaluatie .....	108
2.3. Evalueren van “eigen” of andermans beleid .....	109
2.4. Evalueren: zelf doen of laten doen .....	110
2.5. Evalueren: samen doen of toch alleen .....	110
2.6. Evalueren niet enkel als sluitstuk van beleidsvoering .....	112
2.7. Kwaliteitsvol evalueren .....	113
<b>3. Evalueren: ook om te leren? .....</b>	<b>114</b>
<b>4. Besluit .....</b>	<b>118</b>
<b>Bibliografie .....</b>	<b>120</b>

## 1. Inleiding

Evalueren is “in”. Voor velen lijkt het haast een levensstijl geworden om alles en iedereen te evalueren, te beoordelen naar effect, impact, (meer-)waarde. Ook het overheidsbeleid is niet aan die trend ontsnapt. Net zoals ingeval van andere beleidsontwikkelingen fungeerde het Verenigd Koninkrijk ook in verband met de evaluatie van overheidsbeleid als Europese voortrekker. De aandacht voor beleidsevaluatie bereikte er een hoogtepunt in 1997 met het aantreden van New Labour. Die partij wilde af van beleidsvoering op basis van ideologie en werken vanuit het adagium ‘*what really matters, is what works*’ (Davies e.a., 2000). Vandaag gaan er echter steeds meer stemmen op dat beleidsevaluatie in het Verenigd Koninkrijk over haar hoogtepunt heen is want zoals verwoord door Healy (2002, 97-98) ‘... *there is a strong requirement to [always] assess whether it works or not [and] as a result, evaluation is sometimes an obligation whereas a true desire to learn and improve is not present anymore.*’

Met dit veeleer pessimistische oordeel ten opzichte van het Verenigd Koninkrijk in het achterhoofd willen wij in deze bijdrage verkennen hoe het vandaag staat met beleidsevaluatie in Vlaanderen en in het bijzonder in de Vlaamse overheid. Geeft die overheid aandacht aan het (laten) evalueren van haar beleid? En hoe gebruikt ze de resultaten uit beleidsevaluaties dan? Om er echt uit te leren ter verbetering van haar beleid? Het antwoord op die vragen hebben we in deze verkennende bijdrage beperkt tot twee domeinen van Vlaamse beleidsvoering. Het gaat om Binnenlands Bestuur en Stedenbeleid die beiden ressorteren onder de bevoegdheid van dezelfde Vlaamse minister en inzake (interbestuurlijke) beleidsvoering al menig jaar vlot meevolueren met de gangbare ontwikkelingen en trends op Europees vlak (Wayenberg, 2004, 2006). Dit laatste maakte beide domeinen dan ook erg interessant voor onze verkenning van de plaats en betekenis van beleidsevaluaties binnen de Vlaamse overheid.

Aan die verkenning geven we vorm in twee stappen. Eerst focussen we op het voorkomen van beleidsevaluaties in zowel Binnenlands Bestuur als Stedenbeleid. En daarna zoomen we in op het gebruik dat binnen beide domeinen wordt gemaakt van de resultaten uit evaluaties. Beide stappen zullen worden onderbouwd door een combinatie van verwijzingen naar de literatuur met bevindingen uit een eigen analyse van Vlaamse beleidsnota’s en –brieven en met citaten van bevroegde politieke en ambtelijke beleidsverantwoordelijken op Vlaams niveau.<sup>43</sup> Deze laatsten werden tijdens hun bevraging aangemaand om in eigen boezem te kijken in verband met de plaats en betekenis die zij aan evaluatie van “hun” beleid toekennen. En dit leverde op momenten een erg kritische zelfevaluatie op zoals zal blijken uit onderhavige tekst.

---

<sup>43</sup> Graag willen wij de bevroegde politieke en ambtelijke beleidsverantwoordelijken danken voor hun bereidheid tot medewerking en hun openhartigheid tijdens ons gesprek.

## 2. Evalueren van het Vlaamse beleid inzake binnenlands bestuur en steden: de puntjes op de i

Om de puntjes op de i te zetten omtrent de beleidsevaluaties in de domeinen Binnenlands Bestuur en Stedenbeleid staan we stil bij een zevental aspecten van het evalueren van beleid. Achtereenvolgend zal het gaan om: het kernstuk van een beleidsevaluatie, het voorwerp van zo'n evaluatie, het eigenaarschap van het te evalueren beleid, de uitvoerder(s) van een beleidsevaluatie, de betrokkene(n) erbij, de plaats van evaluatie in de beleidscyclus en de erbij in acht te nemen kwaliteitsstandaarden.

### 2.1. Het kernstuk van beleidsevaluatie

Of de Vlaamse overheid aandacht geeft aan de evaluatie van haar beleid wordt vanuit het kabinet van minister Keulen, bevoegd in de huidige Vlaamse regering voor Binnenlands Bestuur en Stedenbeleid, beantwoord met een volmondig *'ja! Want evalueren is heel belangrijk om het verdere beleid vorm te geven.'* Waarom dat zo is, sloot aan bij de essentie van evaluatie zoals die tot uiting komt in de volgende definitie van Scriven (1980): *'evaluation is the systematic and objective determination of the worth or merit of an object.'* Bij een beleidsevaluatie wordt dus de (meer-)waarde of verdienste van het al gevoerde beleid bepaald, wat essentiële informatie kan opleveren voor de besluitvorming over het nog verder te voeren beleid.

### 2.2. Het voorwerp van beleidsevaluatie

Het precieze object of voorwerp van een beleidsevaluatie kan driedelig zijn. Volgens Bressers (2003, 179) kan het gaan om de inhoud van beleid, het proces van beleidsvoering of de effecten van beleid. Voorbeelden van elk type zijn terug te vinden in de beleidsnota Binnenlands Bestuur waarin minister Keulen zijn beleidslijnen voor de huidige legislatuurperiode 2004-2009 heeft uitgetekend (Keulen, 2004a).

Zo kondigt de minister –die binnen de Vlaamse regering ook bevoegd is voor Wonen– in die nota aan dat hij de lokale fiscaliteit wil inzetten als een instrument van beleidssturing naar een betere samenhang ingeval beleid gelijktijdig op Vlaams en lokaal niveau wordt gevoerd. Dit is het geval inzake het vrijmaken van bouwkavels. De Vlaamse overheid voert hieromtrent een beleid maar ook de gemeenten doen dit door jaarlijks een belasting te heffen op niet-bebouwde percelen of gronden gelegen in een niet-vervallen verkaveling en in gebieden bestemd voor wonen en industrie. Zij ontlenen die belastingsbevoegdheid aan de Vlaamse regelgeving op de ruimtelijke ordening die de modaliteiten van die gemeentebelasting bepaalt evenals de (Vlaamse) vrijstellingen die verplicht over te nemen zijn op het lokale niveau. Minister Keulen wil die regelgeving en de erin verankerde beleidskeuzes evalueren, wat duidelijk een voorbeeld is van een aangekondigde evaluatie van de inhoud van beleid (Keulen, 2004a, 28).

Een duidelijke aankondiging van een evaluatie van (een deel van) het beleidsproces is in diezelfde nota terug te vinden in verband met het decreet op de intergemeentelijke samenwerking. Dat decreet dateert van 6 juli 2001 en de minister heeft niet de bedoeling om *'zo kort na het van kracht worden van een toch volledig vernieuwd concept [van intergemeentelijke samenwerking] de principes ervan opnieuw in vraag te stellen. Wel kunnen, in functie van de evaluatie, desgewenst partiële aanpassingen gebeuren om de toepassing van de aangenomen principes te verbeteren.'* (Keulen, 2004a, 23). Dit kondigt dan ook een evaluatie aan van de naleving van die principes oftewel van de uitvoering van het door dit decreet gevoerde beleid.

Effecten van beleid zijn doorgaans het meest "verwachte" voorwerp van een beleidsevaluatie. Men wil immers weten of al gevoerd beleid echt "gewerkt" heeft, of het de impact heeft teweeggebracht die men ervan had verwacht en of het onbedoelde, al dan niet positief te beoordelen neveneffecten heeft gegenereerd. Die intentie drukt minister Keulen in zijn beleidsnota Binnenlands Bestuur meermaals uit waaronder in verband met de in de jaren '70 doorgevoerde fusie-operatie op lokaal vlak. Die fusies zijn nu meer dan 30 jaar oud. Een grondige wetenschappelijke evaluatie over hun effect op de lokale bestuurskracht heeft echter nog nooit plaatsgevonden en lijkt de minister dan ook *'meer dan gewenst'* (Keulen, 2004a, 32).

### 2.3. Evalueren van "eigen" of andermans beleid

De drie bovenstaande voorbeelden geven meteen ook aan dat het bij een aankondiging van een beleidsevaluatie van overheidswege niet altijd gaat om het evalueren van "eigen" beleid. Zo kondigt minister Keulen met die evaluatie van Vlaamse regelgeving inzake ruimtelijke ordening een beleidsevaluatie aan op het bevoegdheidsdomein van minister Van Mechelen, een collega-minister in de huidige Vlaamse regering. Het te evalueren decreet op de intergemeentelijke samenwerking is tot stand gekomen in 2001 en dus tijdens een voorbije Vlaamse regeerperiode (1999-2003).<sup>44</sup> Aldus kondigt minister Keulen een evaluatie aan van beleid van een vorige Vlaamse regering waarvan hij toen geen deel uitmaakte. De grote fusie-operatie uit de jaren '70 ten slotte is een voorbeeld van nationaal beleid dat de minister wil evalueren. Pas in 1981 trad immers de eerste Vlaamse regering aan die slechts vanaf 1993 de bevoegdheid verwierf tot organisatie van de intergemeentelijke samenwerking in het Vlaamse Gewest. Evalueren van vroeger beleid is op zich niet vreemd want zoals verwoord in de ministeriële beleidsnota Stedenbeleid *'hebben beleidsvoerders vaak heel concreet te maken met de gevolgen van een bepaald soort beleid uit het verleden zoals dat is neergeslagen in maatschappelijk gedrag, beleidseffecten en administratieve praktijken.'* (Keulen, 2004b, 4).

Toch zijn er ook voorbeelden van (aangekondigde) evaluaties van beleid waarvan het eigenaarschap in belangrijke mate aan minister Keulen toekomt. Die voorbeelden laten zich het makkelijkst traceren in het domein Binnenlands Bestuur waar ze vooral de nieuwe regelgeving betreffen die in dat domein tot stand is gebracht. Immers, in

---

<sup>44</sup> Voor een overzicht van de opeenvolgende Vlaamse regeringen en hun regeertermijn, zie: [http://www.vlaanderen.be/servlet/Satellite?c=Solution\\_C&cid=1181701910138&context=1141721623065---1191211215889-1191211215089--1181701910138&p=1186804409590&pagename=Infolijn%2FView](http://www.vlaanderen.be/servlet/Satellite?c=Solution_C&cid=1181701910138&context=1141721623065---1191211215889-1191211215089--1181701910138&p=1186804409590&pagename=Infolijn%2FView)

de woorden van een kabinetsmedewerker van minister Keulen ‘zal die 5 jaar [van de huidige Vlaamse legislatuur] in het domein Binnenlands Bestuur zeker voor een groot stuk besteed zijn aan het maken van nieuwe regelgeving ten opzichte van de lokale besturen. Als je aan steden en gemeenten zou vragen wat er het meeste opvalt aan de legislatuur van Keulen, dan zal het dit zijn: dat hele pakket aan nieuwe regelgeving.’ Tot dat pakket behoren het kiesdecreet ten aanzien van de lokale en provinciale verkiezingen, het gemeentedecreet, het provinciedecreet, diverse decreten tot wijziging van de OCMW-wet en daarbij horende uitvoeringsbesluiten, omzendbrieven en al gemaakte decretale aanpassingen.<sup>45</sup>

In die regelgeving zijn op initiatief van minister Keulen diverse maatregelen verankerd om de participatie van vrouwen aan de lokale beleidsvoering te verhogen. Zo verplicht het kiesdecreet tot een pariteit tussen mannelijke en vrouwelijke kandidaten op de gemeentelijke kieslijsten<sup>46</sup> en schrijft het gemeentedecreet voor dat beide geslachten aanwezig moeten zijn in het College van Burgemeester en Schepenen (CBS).<sup>47</sup> De minister wil inzicht in het effect van al de genomen maatregelen op de samenstelling van lokale raden en colleges en heeft terzake dan ook een evaluatieonderzoek uitbesteed (Keulen, 2007a, 38).

Zo’n “effectenonderzoek” is eveneens uitbesteed ingeval van andere innovaties die het gemeentedecreet heeft geïntroduceerd op het vlak van zowel de politieke als administratieve werking van de gemeente.<sup>48</sup> Tot die innovaties behoren onder andere de gecreëerde mogelijkheden tot delegatie van de gemeenteraad aan het CBS<sup>49</sup> en van het CBS aan de gemeentesecretaris,<sup>50</sup> de invoering van de afsprakennota tussen het CBS en het op te richten managementteam<sup>51</sup> en de vereiste tot opstelling van een strategisch meerjarenplan bestaande uit een strategische en financiële nota en op maat van de gehele gemeente.<sup>52</sup>

#### 2.4. Evalueren: zelf doen of laten doen

De vaststelling dat een evaluatie “eigen” beleid betreft, wil nog niet zeggen dat die evaluatie ook zelf wordt uitgevoerd. Dit brengt ons bij het onderscheid tussen een externe en een interne evaluatie.

De twee bovenstaande voorbeelden zijn illustratief voor een externe evaluatie waarbij de evaluator niet behoort tot de organisatie waarvan het beleidsinitiatief wordt geëvalueerd (De Peuter e.a., 2007, 39-40). Zo is die externe evaluator het Steunpunt Bestuurlijke Organisatie Vlaanderen (SBOV) ingeval van de maatregelen ter innovatie van de politieke en administratieve werking van de gemeente.

<sup>45</sup> Voor een overzicht van de tot stand gebrachte regelgeving in het domein Binnenlands Bestuur, zie Keulen, 2007a, 7-12.

<sup>46</sup> Art. 9 § 3 van het decreet van het Vlaams Gewest van 10 februari 2006 tot wijziging van de Gemeentekieswet, gecoördineerd op 4 augustus 1932, de wet van 19 oktober 1921 tot regeling van de provincieraadsverkiezingen, de wet tot organisatie van het geautomatiseerd stemmen van 11 april 1994 en het decreet van 7 mei 2004 houdende regeling van de controle van de verkiezingsuitgaven en de herkomst van de geldmiddelen voor de verkiezing van het Vlaams parlement, (B.S. 10 maart 2006).

<sup>47</sup> Art. 44 § 4 van het gemeentedecreet van het Vlaams Gewest van 15 juli 2005, (B.S. 31 augustus 2005).

<sup>48</sup> <http://soc.kuleuven.be/sbov/ned/onderzoek/B200707.htm>

<sup>49</sup> Art. 43 van het gemeentedecreet van het Vlaams Gewest van 15 juli 2005, (B.S. 31 augustus 2005).

<sup>50</sup> Art. 58 van het gemeentedecreet van het Vlaams Gewest van 15 juli 2005, (B.S. 31 augustus 2005).

<sup>51</sup> Art. 87 van het gemeentedecreet van het Vlaams Gewest van 15 juli 2005, (B.S. 31 augustus 2005).

<sup>52</sup> Art. 146 van het gemeentedecreet van het Vlaams Gewest van 15 juli 2005, (B.S. 31 augustus 2005).

Aankondigingen van interne evaluaties –waarbij de evaluator dus wel behoort tot de organisatie waarvan het beleidsinitiatief wordt geëvalueerd– zijn er ook zoals ingeval van de communicatie-inspanningen ter creatie van een maatschappelijk en politiek draagvlak voor het stedelijk beleid (Keulen, 2004b, 16). Één van die inspanningen is bijvoorbeeld de website [www.thuisindestad.be](http://www.thuisindestad.be) die op initiatief van minister Keulen ‘permanent wordt bijgehouden naar bezoekersaantal en geoptimaliseerd’ (Keulen, 2007b, 26).

“Intern” is in de huidige “governance”-context van beleidsvoering wel een ruim invulbaar begrip want het kan betekenen dat meerdere (overheids-)organisaties tegelijk als evaluator optreden van gezamenlijke beleidsinitiatieven. Dit leidt ons naar een belangrijk verschil tussen de twee geanalyseerde beleidsdomeinen van Binnenlands Bestuur en Stedenbeleid.

### 2.5. Evalueren: samen doen of toch alleen

De beide bestudeerde domeinen staan centraal in de Vlaamse interbestuurlijke verhoudingen met steden en gemeenten.<sup>53</sup> Uiteraard worden er ook vanuit andere Vlaamse beleidsdomeinen relaties onderhouden met d(i)e lokale besturen, zij het vooral ter realisatie van doelstellingen geconcentreerd rond een bepaald inhoudelijk domein van beleidsvoering zoals “milieu” of “welzijn”. De doelstellingen vanuit Stedenbeleid en nog meer vanuit Binnenlands Bestuur betreffen echter primair de lokale besturen zelf als een belangrijk onderdeel van het institutionele landschap.<sup>54</sup> En ondanks het feit dat beiden ressorteren onder de bevoegdheid van dezelfde Vlaamse minister Keulen verschillen ze toch in hun relatie met die lokale besturen ter realisatie van die doelstellingen, een verschil dat ook doorwerkt op de in die domeinen heersende evaluatiecontext en –praktijk.

Een uitstapje naar Rosenau’s conceptualisering van “government” en “governance” helpt om dit verschil te duiden. Volgens Rosenau (1992) verwijst “government” naar ‘activities that are backed up by formal authority’ terwijl “governance” refereert aan ‘activities backed up by shared goals.’ In het licht van die conceptualisering speelt het domein Binnenlands Bestuur zich in vergelijking met Stedenbeleid veeleer af binnen een traditionele “government”-context van interbestuurlijke verhoudingen. De er ontsponnen activiteiten berustten dan ook nog in sterke mate op de formele autoriteit die de centrale (Vlaamse) overheid heeft over een subcentraal (lokaal) bestuur en zich onder andere uit in de afwezigheid van platforms voor systematisch tweerichtingsverkeer bij raakpunten tussen beiden. Immers, de voorbereiding van nieuwe Vlaamse regelgeving gebeurde in belangrijke mate op Vlaams niveau en ook het centrale toezicht op de lokale naleving van (die) regelgeving is geen voorwerp van systematisch tweerichtingsverkeer. Precies dat toezicht is de meest voorkomende vorm van interne evaluatie binnen het domein Binnenlands Bestuur waarbij dus enkel de Vlaamse overheid als evaluator optreedt, in lijn met de er heersende “government”-context van interbestuurlijke verhoudingen.

<sup>53</sup> Met ‘steden’ wordt hier bedoeld op de 13 Vlaamse centrumsteden: Aalst, Antwerpen, Brugge, Genk, Gent, Hasselt, Kortrijk, Leuven, Mechelen, Oostende, Roeselare, Sint-Niklaas en Turnhout. Daarnaast is er soms sprake van “provinciesteden” maar die worden hier onder de noemer ‘gemeenten’ gevat.

<sup>54</sup> Uiteraard betreffen de doelstellingen in het domein Binnenlands Bestuur ook andere institutionele actoren op subcentraal niveau zoals de provincies, OCMW’s en de kerkfabrieken.

Weliswaar lijken er in dit domein ook al tekenen dat die context evolueert naar een “post-government” en dus meer in de richting van een “governance”-fase. Zo zouden de lokale besturen steeds meer betrokken (geworden) zijn in de voorbereidingsfase van nieuwe regelgeving die hen aanbelangt, zij het dan wel op een indirecte manier. Volgens de bevroegde medewerker van het kabinet Keulen is immers *‘de Vlaamse Vereniging van Steden en Gemeenten (VVSG) onze belangrijkste gesprekspartner en daar hebben wij ook fors in geïnvesteerd. In het regeerakkoord was uitdrukkelijk ingeschreven dat bij de uitwerking van nieuwe regelgeving de vertegenwoordigers van steden en gemeenten zouden betrokken worden. En dat is ook echt gebeurd, tot op heden. Voor wij [vanuit het kabinet Keulen] naar een InterKabinettenWerkgroep (IKW) gaan, zitten wij met de VVSG aan tafel en vergaderen wij over de concepten.’*

Maar ook ingeval van andere van de ontplooidde activiteiten in het domein Binnenlands Bestuur is er sprake van dat indirecte interbestuurlijke tweerichtingsverkeer. Denk bijvoorbeeld aan het fiscaal pact dat minister Keulen samen met zijn collega-minister Van Mechelen, (ook) bevoegd voor Financiën en Begroting, wil afsluiten met vertegenwoordigers van de lokale besturen en dit met het oog op een meer bedrijfsvriendelijke lokale fiscaliteit, mogelijkheden voor de lokale besturen om die fiscaliteit meer te differentiëren en de afschaffing van een aantal kleinere gemeentelijke belastingen die vooral voor bedrijven veel administratieve rompslomp veroorzaken (Keulen, 2004a, 10).

Dat pact illustreert overigens dat tweerichtingsverkeer in het domein Binnenlands Bestuur niet alleen voorkomt tijdens de voorbereiding van een activiteit maar ook in de fase van besluitvorming eromtrent want een “pact” tussen partijen is naar aard op twee richtingen georiënteerd. Ook de uitvoering van activiteiten in dit domein verloopt steeds meer bilateraal. Het is zelfs één van de strategische doelstellingen van minister Keulen om de administratie Binnenlands Bestuur te heroriënteren naar meer ondersteuning en coaching ten aanzien van de lokale besturen, wat niet in het minst aangewezen is voor een vlotte implementatie van de talrijke nieuwe regelgeving die hen aanbelangt (Keulen, 2004a, 35). Tijdens de evaluatiefase van de activiteiten in dit domein overheerst evenwel éénrichtingsverkeer ingeval die evaluatie zich tenminste intern voltrekt en bekeken wordt vanuit een één-op-één relatie tussen de Vlaamse overheid en een lokaal bestuur.

Precies op dit laatste punt lijkt het domein Stedenbeleid te verschillen van Binnenlands Bestuur want volgens de bevroegde kabinetsmedewerker over dit domein gebeurt er in het Vlaamse Stedenbeleid *‘heel veel in partnerschap met de lokale besturen.’* En dit niet alleen tijdens de fasen van voorbereiding, besluitvorming en uitvoering van activiteiten maar ook tijdens de evaluatie ervan. Op te merken valt dat dit in het Stedenbeleid niet alleen zo lijkt op het meso-niveau van de afzonderlijke beleidsinitiatieven zoals het Stedenfonds en de strategische stadscontracten maar ook op het meta-niveau dat het hele domein overspant. Immers, in het eerste werkjaar van de huidige legislatuur vond een beleidsevaluatie plaats van de globale beleidslijnen die waren uitgetekend tijdens de vorige legislatuur en dit, zoals aangekondigd, vanuit de Vlaamse overheid in partnerschap met de twee Vlaamse grootsteden, de 11 centrumsteden en de Vlaamse GemeenschapsCommissie (VGC) (Keulen, 2006, 5). Het voorkomen van zo’n meta-evaluatie, waarvoor intussen ook een platform ter systematische organisatie is ontstaan, laat dan ook extra toe te stellen dat het Stedenbeleid de “post-government” fase voorbij lijkt op het vlak van



de interbestuurlijke verhoudingen. Dit domein blijkt veeleer te opereren volgens een “governance”-context, wat inhoudt dat Vlaamse overheid en lokale besturen als gezamenlijke evaluatoren optreden bij beleid dat zij voeren vanuit gedeelde doelstellingen.

Het vastgestelde verschil tussen beide domeinen verklaarden de bevroegde medewerkers van het kabinet Keulen vooral door te wijzen op het aantal lokale besturen dat vanuit “hun” domein wordt gevisieerd. Het domein Binnenlands Bestuur ontplooit activiteiten die vrijwel steeds gericht zijn op alle lokale besturen in het Vlaamse Gewest. Dat zijn er 308, ruim 250 meer dan het aantal dat in het Stedenbeleid als Vlaamse partner aantreedt en kan aantreden *‘want dit lager aantal maakt een één-op-één relatie bij [interbestuurlijke] beleidsvoering precies ook mogelijk.’*

## **2.6. Evalueren niet enkel als sluitstuk van beleidsvoering**

Bovenstaand voorbeeld van “meta-evaluatie” is ook exemplarisch voor het feit dat een evaluatie niet altijd te situeren valt op het eind van de beleidscyclus zoals in de standaard-voorstelling van die cyclus wordt gesuggereerd. Het gaat dan om een “ex post”-evaluatie die plaatsvindt nadat de uitvoering van het betreffende beleid achter de rug is. Evalueren van beleid kan echter ook “ex ante”, voor de besluitvorming rond dat beleid plaatsvindt, en tussentijds, tijdens de uitvoering ervan (De Peuter e.a., 2007, 36). Het voorbeeld van de “meta-evaluatie” illustreert deze beide tijdstippen want die evaluatie vond niet eenmalig plaats in de periode van scherpstelling door de minister van zijn beleidslijnen voor de komende legislatuur maar ook op geregelde tijdstippen tijdens die legislatuur en dus tijdens de uitvoering van het door de minister uitgetekende beleid.

## **2.7. Kwaliteitsvol evalueren**

Of een evaluatie nu de beleidsinhoud, het proces of de effecten van beleidsvoering tot voorwerp heeft, eigen of andermans beleid betreft, intern of extern gebeurt en al dan niet door meerdere evaluatoren en als sluitstuk van de beleidsvoering; een evaluatie hoort kwaliteitsvol te zijn. Waaruit die kwaliteit dan precies moet blijken, is een vraag met vele antwoorden want er zijn talrijke opsommingen van kwaliteitsvereisten of -standaarden op de evaluatiemarkt (De Peuter e.a. 2007, 188-196). Vermits we hier geen ruime kwaliteitstoets van de beleidsevaluaties in de domeinen Binnenlands Bestuur en Stedenbeleid beogen, staan we enkel bij een tweetal van die vereisten stil.

Een eerste vereiste waaraan een beleidsevaluatie idealiter voldoet, volgt direct uit Scrivens definitie ervan. Evalueren betekent immers het systematisch bepalen van de meerwaarde of verdienste van het evaluandum. Precies die systematiek is niet altijd aanwezig ingeval men er de tekst van de ministeriële beleidsnota's op naleest. In de nota Binnenlands Bestuur bijvoorbeeld wordt melding gemaakt van de resultaten van een interne evaluatie van de organisatie van de financiële controle in de gemeente volgens de (toen nog geldende) gemeentewet. Volgens die resultaten voldoet die organisatie niet langer, zij het dat dit oordeel niet op een systematische evaluatie berust maar op *‘meerdere feiten van de jongste jaren’* (Keulen, 2004a, 22).



Die feiten worden bovendien niet nader geëxpliciteerd evenmin als het criterium op grond waarvan de organisatie van de gemeentelijke financiële controle volgens diezelfde feiten een onvoldoende verdient. Transparantie is nochtans een andere, in ruime kring onderkende standaard voor een kwaliteitsvolle evaluatie. Uiteraard is een beleidsnota of -brief niet de aangewezen plaats om de opzet, het verloop en de resultaten van een evaluatieonderzoek uitvoerig uit de doeken te doen. Maar met een eenvoudige bronverwijzing kan echter al aan die vereiste tot transparantie worden tegemoet gekomen, op voorwaarde dan natuurlijk dat het betreffende onderzoek een formeel karakter heeft en niet op veeleer informele wijze is gevoerd. En precies die voorwaarde is ook niet altijd vervuld zoals ingeval van de door de minister afgelegde ronde langs de Vlaamse steden ter evaluatie van (onderdelen van) het Vlaamse Stedenbeleid. De globale resultaten van die evaluatie worden via de Beleidsbrief wel gedeeld met het parlement en de burger maar een duidelijke explicitering van de redenen op grond waarvan die stedenronde volgens de minister *'de duidelijke meerwaarde van het Stedenfonds heeft aangetoond als hefboom van maatschappelijke en interne veranderingen en het aanvoelen heeft bevestigd dat de bestuurlijke kwaliteit in de steden de voorbije jaren gevoelig is toegenomen'* blijft daarbij achterwege (Keulen, 2007b, 2).

Tot zover onze eerste verkenning van beide Vlaamse bevoegdheidsdomeinen inzake beleids-evaluatie. Het is duidelijk dat er van overheidswege aandacht is voor evaluatie en dat variatie daarbij niet wordt geschuwd. De aangekondigde evaluaties verschillen immers naar voorwerp, soort en aantal evaluatoren, eigenaarschap, tijdstip en kwaliteit. Niet elk aangekondigd beleidsinitiatief wordt standaard geëvalueerd maar van evaluatie is dus toch frequent sprake in beide ministeriële beleidsnota's, zij het soms onder een andere noemer zoals die van "toetsing" of "onderzoek". Louter kwantitatief bekeken wordt in de nota Binnenlands Bestuur wel meer melding gemaakt van beleidsevaluatie dan in die met betrekking tot Stedenbeleid, wat meteen ook verklaart waarom het merendeel van de gegeven voorbeelden uit het eerste domein afkomstig is.

Het aankondigen van een evaluatie is natuurlijk één zaak; het effectief doorvoeren ervan en vervolgens gebruiken van de evaluatieresultaten nog een totaal andere. Van de 10 hierboven aangehaalde voorbeelden van beleidsevaluaties is vandaag bekend dat de helft extern gebeurt of is gebeurd zoals aangeduid in tabel 5.1. Maar hoe staat het eigenlijk met het gebruik van de resultaten uit extern gevoerd evaluatieonderzoek? Deze vraag staat centraal in het tweede deel van dit hoofdstuk.

**Tabel 5.1: Voorwerp van aangekondigde beleidsevaluaties in de domeinen Binnenlands Bestuur en Stedenbeleid (en van extern uitgevoerd evaluatieonderzoek)**

Domein Binnenlands Bestuur	Domein Stedenbeleid
<p>Regelgeving op de ruimtelijke ordening</p> <p><i>Decreet op de intergemeentelijke samenwerking</i></p> <p><i>Doorgevoerde fusie-operatie tijdens jaren '70</i></p> <p><i>Maatregelen ter verhoging van vrouwelijke participatie aan lokale politiek</i></p> <p><i>Innovaties uit gemeentedecreet op stuk van de politieke en administratieve werking van de gemeente</i></p> <p>Bepalingen uit de gemeentewet ter organisatie van financiële controle in de gemeente</p>	<p>Communicatie-inspanningen over het Vlaamse stedenbeleid</p> <p>Vlaams Stedenfonds</p> <p>Strategische stadscontracten</p> <p><i>De tijdens de vorige legislatuur uitgetekende beleidslijnen voor het Vlaamse Stedenbeleid (meta-evaluatie)</i></p>

(Bron: Keulen, 2004a; 2004b; 2007a; 2007b)

### 3. Evalueren: ook om te leren?

Evalueren is leren, claimen sommigen (Bressers, 2003, 192). Maar is dit wel zo? En vooral bij het evalueren van beleid? Leren beleidsvoerders wel uit evaluaties van beleid die ze vaak ook zelf besteld hebben? Deze vragen zijn niet nieuw. Ze zijn al eerder gesteld en beantwoord, zowel in negatieve als positieve zin. Een negatief antwoord volgt er wanneer het bestellen van een (evaluatie-)studie van overheidswege wordt gebruikt als een vorm van “non decision-making” (Weimer & Vining, 1999). Als beleidsvoerders een studie bestellen omdat ze geen beslissing willen nemen, lijkt het weinig waarschijnlijk dat ze snel geneigd zullen zijn om resultaten uit die studie via de nodige beslissingen beleidsmatig te verankeren. Positief is er ook geantwoord op bovenstaande vragen. Dit heeft geleid tot het scherpstellen van wat onder “an influential evaluation” of een invloedrijke evaluatie(-studie) moet worden begrepen. Volgens Henry (2003, 515) is er hiervan sprake *‘whenever there is an indication that the evaluation provided a change between the pre- and the post-evaluation states [knowing that] these states could be individual attitudes, behaviors of organization members who serve as change agents, or government policies.’*

In het licht van deze duiding is het interessant na te gaan hoe invloedrijk de vijf extern bestelde evaluatiestudies uit tabel 5.1 zijn of waren. Bevatten de door ons geraadpleegde beleidsdocumenten een indicatie dat deze studies (zullen) instaan of –stonden voor een verandering tussen het pre- en post-evaluatiestadium van het betreffende beleid? Zo'n indicatie is steeds aanwezig, zij het variërend naar intensiteit van de (mogelijke) verandering waarop wordt gealludeerd. Ingeval van de aangekondigde evaluatie van het decreet op de intergemeentelijke samenwerking en van de maatregelen ter verhoging van de vrouwelijke participatie aan de lokale politiek is die indicatie het minst uitgesproken. Er valt immers telkens te lezen dat evaluatieresultaten *'dienstig kunnen zijn voor eventuele beleidsbijstellingen'* in de toekomst (Keulen, 2004a, 23; 2007a, 38). De gebruikte toon wordt wat meer "utilitair" van aard ingeval van de bestelde studie inzake de doorgevoerde fusie-operatie tijdens de jaren '70. In dat verband valt immers te lezen dat die evaluatiestudie *'zeker gewenst is [want] het ligt in de bedoeling er lering uit te trekken voor het toekomstige beleid'* (Keulen, 2004a, 33). Hetzelfde gaat ook op voor de bij het SBOV bestelde studie ter evaluatie van innovaties van het gemeentedecreet. Over de SBOV –studies wordt immers in het algemeen gezegd dat *'de resultaten uit de onderzoeksprojecten een belangrijke bijdrage zullen leveren aan het toekomstige beleid inzake binnenlands bestuur.'* (Keulen, 2007a, 37-38). Zo'n bijdrage is intussen al geleverd inzake het stedenbeleid door het rapport van de Visitatiecommissie Stedenbeleid die dit domein in 2006 evalueerde. In zijn recentste beleidsbrief Stedenbeleid geeft de minister een overzicht van de gemaakte aanbevelingen door de Visitatiecommissie, inclusief een weergave hoe op die aanbevelingen is gereageerd. En die reactie komt ingeval van meerdere aanbevelingen neer op een directe overname ervan in het gevoerde beleid (Keulen, 2007b, 9-11).

Kortom, in 2 van de 5 gevallen van bestelde evaluatiestudie drukt de minister zijn intentie uit tot benutting van de onderzoeksresultaten en in één geval heeft hij die al benut. Dit betekent dat in het merendeel van de gevallen een evaluatiestudie is besteld met het oog op een verdere beleidsevolutie die meer "evidence-based" van aard is. Dit beantwoordt aan het klassieke doel van een beleidsevaluatie. Dit is een erg nobel doel, wat in theorie ressorteert onder het zogenaamde "knowledge-driven model" ter conceptualisering van de relatie tussen onderzoek en beleidspraktijk. Volgens dat model wordt die praktijk immers voortgestuwd door de resultaten van objectief en methodologisch verantwoord onderzoek naar maatschappelijke noden en problemen die al dan niet het voorwerp uitmaakten van reeds (uit-)gevoerd beleid.

**Tabel 5.2: Modellen van “research utilization”**

Model	Duiding
Knowledge-driven model	derives from the natural sciences. The fact that knowledge exists sets up pressures for its development and use
Problem-solving model	involves the direct application of the results of a specific study to a pending decision
Interactive model	researchers are just one set of participants among many. The use of research is only one part of a complicated process that also uses experience, political insight, pressure, social technologies and judgement
Political model	research as political ammunition; using research to support a predetermined position
Tactical model	research as a delaying tactic in order to avoid responsibility for unpopular policy outcomes
Enlightenment model	the indirect influence of research rather than the direct impact of particular findings in the policy process. Thus the concepts and theoretical perspectives that social science research engenders pervade the policy-making process.

(Bron: Weiss, 1979)

Naast het “knowledge-driven” model worden er doorgaans nog diverse andere modellen onderscheiden met betrekking tot de relatie onderzoek – beleidspraktijk. Zo onderscheidde Carol Weiss, als grondlegster van deze modellen, er nog een vijftal andere die zij duidde overeenkomstig tabel 5.2 (Weiss, 1979). Anderen na Weiss hebben vaak het “political” en het “tactical” model verenigd tot één enkel “political-tactical model”. Zo ook Young, Ashby, Boaz en Grayson (2002, 216) die dit model kenmerkten *‘as seeing policy as the outcome of a political process. The research agenda is then politically driven and studies are commissioned and/or used to support the position adopted by the government of the day, the relevant minister or perhaps the civil servants most closely concerned.’* Beide modellen laten zich verenigen omdat zij de machtsrelatie tussen onderzoekers en beleidsvoerders op dezelfde wijze conceptualiseren: zowel ingeval van het “political” als van het “tactical model” ligt het machtsoverwicht duidelijk bij de beleidsvoerders.

Precies het tegenovergestelde doet zich voor ingeval van de eerste twee modellen die Weiss en, in haar navolging, ook Young e.a. onderscheidden: het “knowledge-driven” en het “problem-solving model”. In beide gevallen berust het machtsoverwicht immers bij de onderzoekers, zij het misschien nog meer ingeval van het “knowledge-driven” dan van het “problem-solving model”. Immers, in het laatste geval brengen de beleidsvoerders nog de problemen aan waarrond onderzoek wordt gevoerd terwijl van deze beleidsinput ook geen sprake meer is ingeval van het “knowledge-driven

model". Het "interaction model" geeft dan weer een totaal andere machtsrelatie aan tussen onderzoekers en beleidsvoerders die past binnen een typische "governance"-context want volgens dit model beïnvloeden beide partijen mekaar *'with the agenda for both shaped within policy communities which contain both practitioners and researchers'* (Young e. a., 2002, 215). Het "enlightenment model" ten slotte is een buitenbeentje ten opzichte van de rest want dat model veronderstelt dat onderzoek een indirecte invloed heeft op beleid: *'research is standing at a little distance from the hothouse of immediate policy concerns. Its benefits are indirect because it tackles the context within which decisions need to be taken rather than the decision problem itself.'* (Young e.a., 2002, 216).

Samengevat lijken er dus drie (soorten) modellen volgens dewelke er van overheidswege direct gebruik kan worden gemaakt van onderzoek bij de beleidsvoering. Naar toenemende mate van machtsuitoefening door de overheid gaat het overeenkomstig figuur 5.1 om het "knowledge-driven model", het "interaction model" en het "political-tactical model". Beantwoordt het gebruik door de Vlaamse overheid van de resultaten van extern bestelde evaluatiestudies nu aan één van deze drie modellen?

**Figuur 5.1: Modellen inzake het gebruik van onderzoek van overheidswege**



<i>Machtsoverzicht bij</i>	Onderzoek(-ers)	Onderzoek(-ers) & beleid(-svoerders)	Beleid(-svoerders)
<i>Model</i>	Knowledge-driven model	Interaction model	Political-tactical model

Gevraagd naar hun gebruik van onderzoeksresultaten waren de bevroegde medewerkers van het kabinet Keulen het eens over het belang van het "political-tactical model" om dat gebruik te vatten want zoals in de woorden van één van hen: *'in onze soort job worden evidence-based zaken zoals evaluatiestudies e.d. gebruikt als hulpmiddel als het past maar even makkelijk aan de kant geschoven als het niet goed uitkomt. Neem twee dossiers die even belangrijk zijn. In één dossier kan het zijn dat wij enorm zullen zwaaien met cijfers, met resultaten van wetenschappelijk onderzoek enz. En in een ander dossier kan het zijn dat dit op basis van pure politieke voorkeuren tot stand komt en dat op geen enkele manier, hoewel daar misschien gegevens over zijn of wat dan ook, gegevens echt een rol spelen. Dus 'evidence' wordt gebruikt op een eerder selectieve manier om uw dossier kracht bij te zetten. ... Soms worden proactief studies besteld of laten we gegevens opzoeken waarvan we weten dat ze ons dossier zullen ondersteunen. Maar even goed hebben we dus hier al beslissingen genomen met zware financiële consequenties waar geen simulaties, cijfers, evaluaties of wat dan ook aan ten grondslag lagen.'* Het kabinet Keulen is volgens beide medewerkers omwille van dat politiek-tactisch omspringen met evaluatiestudies *'niet anders dan andere kabinetten'*. Bijgevolg was het ook weinig verwonderlijk één van hen te horen opmerken dat *'cijfers, al het*

*mogelijke onderzoeksmateriaal weinig waarde [hebben] in politieke besluitvorming genre interkabinettenwerkgroepen (IKW's). Ik heb dat gemerkt ten tijde van het gemeentedecreet – er waren dan draagvlakstudies. Dat soort evidence wordt daar een beetje weggelachen, het maakt daar zeer weinig indruk als je zou zeggen van: er is helemaal geen draagvlak voor die of die maatregel. IKW's hebben echt niet de cultuur van het werken met cijfers of onderzoeken als een echt goed argument aanvaard wordt. Zeker het eerste jaar dat ik hier werkte, ben ik erg geschrokken van de manier waarop besluitvorming in dat soort van fora werkt. Het gaat om politieke keuzes en compromissen sluiten. Maar argumenten rond onderzoeken, niet nee. Het gaat van: we gaan voor die keuze en niet anders en werkt dit niet in Nederland, wel tant pis.'*

Met die verwijzing naar het maken van politieke keuzes en het daarbij moeten sluiten van compromissen is meteen een belangrijke factor bekend waarom beleidsvoering dus weinig "evidence-based" gebeurt zoals op basis van resultaten uit bestelde evaluatie-studies. Die factor labelt Leicester (1999) zelfs als een vijand van "evidence-based policy-making (EBP)" want als "politics" domineert, primeert *'the art of the possible rather than what is rational or what might work best.'*

In de EBP-literatuur zijn intussen nog diverse andere van die zogenaamde vijanden bekend. De twee meest onderkende, naast 'politics', zijn 'time' en 'the civil service culture' (Nutley e.a. 2000, 36). Die laatste vijand uit zich volgens Leicester (1999) *'in a strong distrust [among civil servants] of information generated outside the system'*. Die factor wordt niet als cruciaal beschouwd voor het gebrek aan gebruik van evaluatieresultaten in de domeinen Binnenlands Bestuur en Stedenbeleid, noch vanuit het kabinet noch vanuit de bevoegde administratie. Volgens een ambtenaar uit die administratie zou die cultuur, indien al aanwezig, aldus ook moeilijk kunnen doorwerken in de beleidsvoering want *'ondanks het hele Beter Bestuurlijk Beleid-project hebben de kabinetten de hele beleidsvoorbereiding terug naar zich toe gezogen. Ze zijn groot, nooit zo groot geweest als vandaag. En de administratie is terug een uitvoerende administratie geworden.'* Het kabinet op zijn beurt spreekt die realiteit niet tegen, zij het dat men er de administratie in het algemeen *'liever proactiever [zou zien]. De administratie levert wel een hele grote insteek voor de beleidsbrieven [in de domeinen Binnenlands Bestuur en Stedenbeleid]. Maar over het algemeen is de houding ervan reactief. Zij zullen iets bevragen op het moment dat je de opdracht geeft. Maar proactief van: OK, er komt een bepaalde maatregel aan en we gaan nu eens kijken naar het draagvlak ervoor of naar studies erover of naar het buitenland, nee dat gebeurt niet. Die cultuur is er niet.'*

Wat er volgens de kabinetairs ook niet (altijd) is, is de nodige tijd om meer "evidence-based" beleid te voeren: *'zelf kan je niet bij al die studies over je beleid betrokken zijn of die allemaal opvolgen. Maar je kan wel plannen: inschatten wat je van plan bent en proactief daar onderzoek op afstemmen. ... Als je weet dat je een grootschalige evaluatie gaat doen van het decreet op de intergemeentelijke samenwerking en je weet dat je dit gaat doen in 2008, dan is het toch gigantisch interessant om in 2006 of 2007 te zeggen: we gaan een universiteit nu eens een jaar laten onderzoeken of er aanpassingen moeten gebeuren. Op basis waarvan ga je anders dat decreet wijzigen? Want basically, als beleidsmaker heb je altijd heel onverwachte dingen, dat is een deel van onze job, maar de grote dossiers zijn echt al vanaf het begin*

*gepland met een jaartal erop enz. Dus je kan eigenlijk echt wel plannen en proactief zaken laten onderzoeken [op hun (meer-)waarde]. Ik vind dat we die cultuur nog veel sterker moeten gaan ontwikkelen.’*

#### **4. Besluit**

Dit hoofdstuk omvat een eerste verkenning van de plaats en betekenis van beleidsevaluaties in de Vlaamse overheid. We beperkten die verkenning tot de domeinen Binnenlands Bestuur en Stedenbeleid die op vlak van beleidsvoering doorgaans vlot mee-evolueren met gangbare Europese trends en ontwikkelingen. Is dit ook het geval in verband met die trend tot (meer) beleidsevaluatie die zich doorheen Europa sinds de eeuwwisseling manifesteerde? Om dit te achterhalen werd een eigen analyse van beleidsdocumenten in beide domeinen aangevuld met een bevraging van politieke en ambtelijke beleidsverantwoordelijken. Daaruit bleek dat beleidsevaluatie vandaag duidelijk haar plaats heeft verworven in beide domeinen want de aandacht voor en wil tot het (laten) evalueren van beleidsmaatregelen en –initiatieven is er onmiskenbaar. En terecht volgens de bevroagde beleidsverantwoordelijken die in evalueren duidelijk een weg erkenden tot het verbeteren van “hun” beleid. Zij erkenden echter ook dat die weg vandaag nog onvoldoende wordt bewandeld want resultaten uit evaluatiestudies werken nog te weinig door in de beleidsvoering. Daarmee sloegen ze zichzelf als beleids-verantwoordelijken op de borst en getuigden ze niet terug te schrikken van enige kritische zelfevaluatie. En precies zo’n attitude is onontbeerlijk om ervoor te zorgen dat beleidsevaluaties in de toekomst niet alleen een plaats hebben in de Vlaamse beleidsvoering maar ook voldoende betekenis.



## Bibliografie

Bressers, J. T. (2003). Evaluatie van beleid. In: Hoogerwerf, A. & Herweijer, M. (eds.). *Overheidsbeleid. Een inleiding in de beleidswetenschap*. Alphen aan den Rijn: Kluwer, 173-196.

Davies, H., Nutley, S. & Smith, P. (2000). Chapter one: Introducing evidence-based policy and practice in public services. In: Davies, H., Nutley, S. & Smith, P. (eds.). *What works? Evidence-based policy and practice in public services*. Bristol: The Policy Press, 1-11.

De Peuter, B., De Smedt, J. & Bouckaert, G. (2007). *Handleiding beleidsevaluatie. Deel 1: evaluatiedesign en management*. Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

Healy, A. (2002). Commentary: Evidence-based policy – The latest form of inertia and control? In: *Planning Theory and Practice*, 3 (1), 97-98.

Henry, G.T. (2003). Influential evaluations. In: *The American Journal of Evaluation*, 24 (4), 515-524.

Keulen, M. (2004a). *Beleidsnota 2004-2009 – Binnenlands Bestuur*. Brussel, Kabinet minister Keulen.

Keulen, M. (2004b). *Beleidsnota 2004-2009 – Stedenbeleid*. Brussel, Kabinet minister Keulen.

Keulen, M. (2005). *Beleidsbrief Binnenlands Bestuur – Beleidsprioriteiten 2006-2007*. Brussel, Kabinet minister Keulen.

Keulen, M. (2007a). *Beleidsbrief Binnenlands Bestuur – Beleidsprioriteiten 2007-2008*. Brussel, Kabinet minister Keulen.

Keulen, M. (2007b). *Beleidsbrief Stedenbeleid – Beleidsprioriteiten 2007-2008*. Brussel, Kabinet minister Keulen.

Leicester, G. (1999). The seven enemies of evidence-based policy. In: *Public Money and Management*, 19 (1), 5-7.

Nutley, S. & Webb, J. (2000). Chapter two: Evidence and the policy process. In: Davies, H., Nutley, S. & Smith, P. (eds.). *What works? Evidence-based policy and practice in public services*. Bristol: The Policy Press, 13-42.

Rosenau, J. (1992). Governance, order and change in world politics. In: Rosenau,



J. & Czempiel, E.O. (eds.). *Governance without government: order and change in world politics*. Cambridge: CUP, 1-30.

Scriven, M. (1980). *The logic of evaluation*. Inverness: CA Edgepress.

Wayenberg, E. (2004). *Vernieuwingen in de Vlaamse centrale-lokale verhoudingen: op weg naar partnerschap? Een kwalitatieve studie van de totstandkoming en uitvoering van het sociaal impulsbeleid*. Onuitgegeven doctoraal proefschrift, Departement Politieke Wetenschappen, KULeuven.

Wayenberg, E. (2006). The Flemish Government's Responsibility for Local Government Modernization. In: *Public Management Review*, 8 (1), 47-65.

Weimer, D.L. & Vining, A.R. (1999). *Policy analysis: Concepts and practice*. New Jersey: Prentice Hall.

Young, K., Ashby, D., Boaz, A. & Grayson, L. (2002). Social science and the evidence-based policy movement. In: *Social Policy & Society*, 1 (3), 215-224.

## **Hoofdstuk 6: Evaluatie van antiradicaliseringsprojecten. Hoe meet je of radicalisering wordt tegengegaan?**

Amy-Jane Gielen, A.G. Advies en Universiteit van Amsterdam  
Gerd Junne, Universiteit van Amsterdam

<b>1. Introductie</b> .....	<b>123</b>
<b>2. Projecten om radicalisering tegen te gaan</b> .....	<b>124</b>
2.1. Het actieplan ‘Wij Amsterdammers’ .....	124
2.2. Antiradicalisering nader uitgewerkt in Stadsdeel Slotervaart .....	125
<b>3. Traditionele evaluatiemethoden in vogelvlucht</b> .....	<b>127</b>
3.1. Meten is weten .....	127
3.2. Pragmatische en beschrijvende evaluatie .....	129
3.3. Interpreteren .....	129
3.4. Onderhandelen .....	129
3.5. Realistische evaluatie .....	130
<b>4. Evaluatiemethoden van internationale onderzoeksprogramma’s</b> .....	<b>131</b>
4.1. Het “Local Capacities for Peace Project” (LCPP) .....	131
4.2. Peace and Conflict Assessment (PCA) .....	137
4.3. Outcome mapping .....	139
<b>5. Aanzet tot een evaluatie</b> .....	<b>141</b>
<b>6. Het trainingsproject ‘Identiteit en weerbaarheid’</b> .....	<b>144</b>
<b>7. Conclusie</b> .....	<b>149</b>
<b>Bibliografie</b> .....	<b>150</b>

## 1. Introductie

Na de moord op Theo van Gogh kwam in Amsterdam het actieplan ‘*Wij Amsterdammers*’ tot stand. Het plan was opgesteld om radicalisering onder Amsterdamse moslimjongeren te voorkomen. Centrale begrippen uit dit plan waren het bevorderen van sociale cohesie, tegengaan van polarisatie en participatie. Initiatieven om het interetnisch en interreligieus contact te vergroten, worden met behulp van het actieplan gestimuleerd. Niet alleen in Amsterdam, maar overal in Nederland zijn de laatste jaren op vooral lokaal niveau veel initiatieven en maatregelen genomen om de contacten tussen allochtonen en autochtonen te bevorderen. Een veelgestelde vraag is wat nou het effect is van dit soort initiatieven. Het Sociaal Cultureel Planbureau (SCP) deed hier in 2007 onderzoek naar door meer dan 400 (gedocumenteerde) initiatieven te bekijken. Het onderzoek werd echter belemmerd door een gebrek aan (deugdelijke) evaluaties.

In het licht van toenemende verantwoordingsvragen en een streven van beleidsmakers naar “evidence-based policy” dat een (indruk van) hoge “return on public investment” moet garanderen, is het markant dat initiatieven op het gebied van antiradicalisering niet worden geëvalueerd. Ministeries, gemeentes en politie tonen echter in toenemende mate belangstelling voor de effecten van de in de afgelopen jaren uitgevoerde projecten. Na een eerste golf van pilotprojecten is men nu benieuwd, wat uit deze projecten te leren valt om in de toekomst beter onderbouwd beleid te ontwikkelen. De ervaring met de evaluatie van projecten op dit terrein is echter beperkt. Immers, hoe meet je het effect van activiteiten met vaag geformuleerde doelstellingen als “het bevorderen van de sociale cohesie” en het “wegnemen van de voedingsbodem voor radicalisering”?

Is een evaluatie van dit soort activiteiten dan werkelijk niet mogelijk? Deze bijdrage betoogt op basis van literatuur over de evaluatie van peacebuilding projecten, dat een effectmeting van dit soort initiatieven weliswaar moeilijk, maar toch degelijk mogelijk is.

In dit hoofdstuk willen wij laten zien dat de bestaande methoden en technieken onvoldoende toereikend zijn om antiradicaliseringsbeleid te evalueren. Met behulp van methodieken, die voor de evaluatie van “vredesopbouw”-projecten elders ontwikkeld zijn, willen wij een conceptueel kader schetsen voor de evaluatie van lokale projecten, die in Nederland in het algemeen en Amsterdam in het bijzondere zijn opgezet. Concreet gaat het daarbij om de evaluatie van een reeks projecten in het stadsdeel Slotervaart die onderdeel vormen van ‘*Actieplan Slotervaart: het tegengaan van radicalisering*’ (2007), want Slotervaart vervult binnen Amsterdam een soort voortrekkersrol en is een experimenteelveld. Mochten projecten daar slagen, dan zullen zij ook elders navolging vinden.<sup>55</sup>

In deze bijdrage zal eerst worden ingegaan op de maatregelen die in Slotervaart

---

<sup>55</sup> Het stadsdeel heeft voor het uitvoeren van dergelijke projecten een startbedrag van € 30.000 ontvangen van de gemeente Amsterdam. Tevens heeft het ministerie van Binnenlandse Zaken € 400.000 ter beschikking gesteld.

genomen zijn. In de volgende paragraaf wordt een overzicht gegeven van bestaande evaluatiemethoden, die ontoereikend blijken voor de evaluatie van antiradicaliseringsprojecten. Aansluitend worden verschillende evaluatiemethoden voorgesteld, die voor de evaluatie van “vredesprojecten” elders zijn ontwikkeld. Afsluitend wordt bekeken, of deze methoden hulp kunnen bieden bij de evaluatie van de projecten die in het “Actieplan Slotervaart” zijn opgenomen. Daarbij zal een conceptueel kader geschetst worden, om zowel het antiradicaliseringsbeleid als geheel te evalueren, alsmede de afzonderlijke deelactiviteiten.

## 2. Projecten om radicalisering tegen te gaan

### 2.1. Het actieplan ‘Wij Amsterdammers’

De eerste gemeente die na de moord op Theo van Gogh onmiddellijk actie ondernam, was Amsterdam. Twee weken na de moord op Theo van Gogh stelde het College van Burgemeester en Wethouders in Amsterdam de analyse en de uitgangspunten van het actieplan ‘Wij Amsterdammers’ (2004) vast. Het probleem ligt volgens de gemeente vooral bij de radicale islam, die streeft naar ingrijpende veranderingen in de verhouding tussen overheid en burgers (een andere staatsinrichtingen) maar ook naar ingrijpende veranderingen in de verhouding tussen burgers onderling in de samenleving. De (dreigende) polarisatie en radicalisering moest doorbroken worden. Zo ontstond het actieplan dat zich richtte op de mogelijkheden die de gemeente heeft om bij te dragen aan het bestrijden van terreur, tegengaan van radicalisering en het voorkomen van polarisatie.

Maar wat is eigenlijk het “tegengaan van radicalisering” en het “voorkomen van polarisatie”? Het antwoord op deze vraag ligt volgens het actieplan bij de “wij-groep”. Om dreigende polarisatie tussen moslims en niet-moslims te voorkomen moet de wij-groep zo groot mogelijk worden. De wij-groep zijn ‘*alle Amsterdammers die geweld afwijzen en de basisregels van onze samenleving accepteren*’ (2004, 6). In het actieplan wordt benadrukt dat gematigde moslims uiteraard ook bij de “wij-groep” horen. ‘*Niet loslaten, maar vasthouden*’, dat is de kern van de Amsterdamse ‘*hard-zachte aanpak*’ (2004, 16). Het tegengaan van uitsluiting betekent insluiting bevorderen. Het gemeentebestuur werkte met het Actieplan ‘Wij Amsterdammers’ op vele fronten aan het bestrijden van terreur en het tegengaan van radicalisering. Tegelijkertijd wilde de gemeente polarisatie voorkomen en positieve krachten mobiliseren. De acties van ‘Wij Amsterdammers’ moesten zo worden uitgedragen dat enerzijds de harde aanpak en de grenzen duidelijk zouden worden, terwijl anderzijds groepen en individuen zouden worden gestimuleerd om met elkaar in gesprek te gaan en acties te initiëren die het “wij-gevoel” in de stad versterken.

Omdat de doelstellingen van het actieplan niet gemakkelijk te realiseren waren, werd het stedelijk integratiebeleid onder de loep genomen, wat resulteerde in de nota ‘Wij Amsterdammers II: Investeren in mensen en grenzen’. In vergelijking met de eerste nota is een discourseverschuiving waar te nemen (van der Welle, 2007, 9). Waar het eerste actieplan vooral een schrikreactie was op de moord op Theo van Gogh en met name werd gesproken over de gevolgen van de terroristische aanslag en het radicaliseringspotentieel, wordt bij het tweede actieplan de aandacht verschoven naar het investeren in sociaal kapitaal en sociaal vertrouwen. Vooral het

stimuleren van een lokale Amsterdamse identiteit is noodzakelijk. (Jonge) moslims worden in deze nota niet meer expliciet genoemd, er wordt nu gesproken van jonge Amsterdammers. Een groot deel van de activiteiten is volledig gericht op jongeren, zoals een zomerschool voor jonge moslimstudenten of pas afgestudeerden om hun weerbaarheid te vergroten, coaching voor pas afgestudeerde hoogopgeleide jongeren, gesprekken tussen jonge moslims en hun ouders, een transparant deurbelid in het Amsterdamse uitgaansleven, stimulering van de activiteiten van de organisatie Moslim Jongeren Amsterdam en het ondersteunen van de radiozender FunX.

Wij Amsterdammers I en II maken onderdeel uit van '*Amsterdam tegen radicalisering*' (2007). In deze beleidsnota worden drie doelgroepen onderscheiden (zie Figuur 6.1):

1) Individuen die radicaliseren of die een radicaal gedachtegoed actief verspreiden. Deze groep dient te deradicaliseren met behulp van een curatieve aanpak. De essentie van de aanpak is het weer positief verbinden van deze personen aan onze samenleving.

2) Dan zijn er de "zoekenden" die weerbaar moeten worden gemaakt. Deze groep is enigszins gevoelig voor radicalisering en kan ten goede of ten kwade beïnvloed worden. Deze groep (moslims en niet-moslims) worstelt met hun identiteit en kan moeite hebben met hun plaats in de samenleving of juist met de plaats van anderen daarin. Op hen is de specifieke preventieve aanpak gericht waardoor zij weerbaarder moeten worden tegen negatieve invloeden en radicale gedachten.

3) Een derde doelgroep bestaat uit al degenen die geraakt worden door negatieve ontwikkelingen die hun binding met de samenleving aan kunnen tasten. Voor deze groep moet de voedingsbodem voor onvrede verkleind worden. De algemene preventieve aanpak is op deze groep van toepassing en richt zich in principe op alle Amsterdammers en benadrukt wederzijdse tolerantie en acceptatie.

## 2.2. Antiradicalisering nader uitgewerkt in Stadsdeel Slotervaart

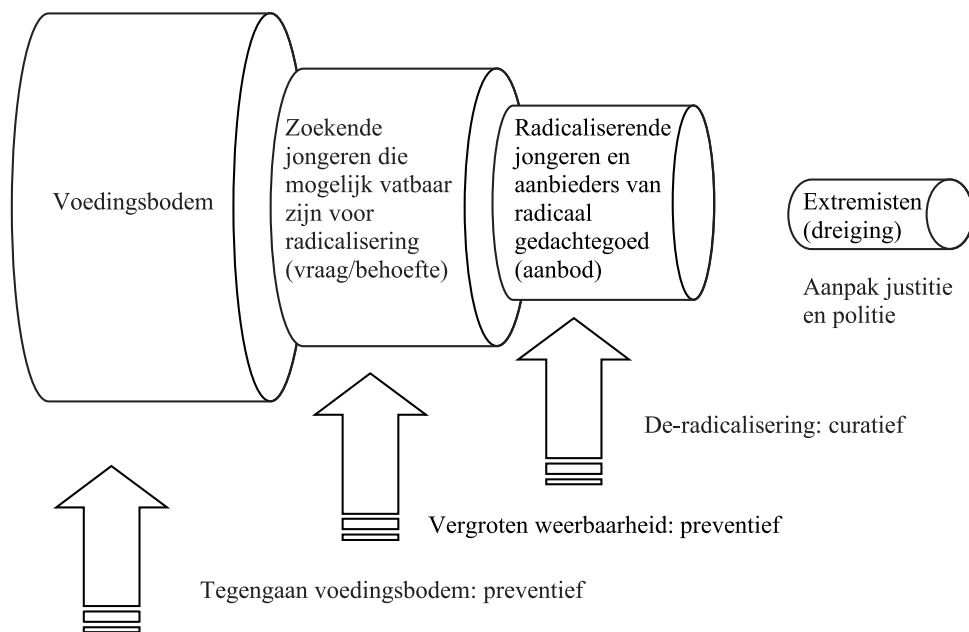
Omdat enkele leden van de Hofstadgroep uit Slotervaart<sup>56</sup> komen en het IMES-rapport van Sloodman en Tillie (2006) liet zien dat er in Slotervaart jongeren wonen die vatbaar en toegankelijk zijn voor radicalisering, besloot stadsdeelvoorzitter Ahmed Marcouch zelf aan de slag te gaan met een actieplan om radicalisering te voorkomen. Er werd een radicaliseringdeskundige aangesteld en het '*Actieplan Slotervaart: het tegengaan van radicalisering*' (2007) werd geboren. Dit actieplan werd in februari 2007 gepresenteerd en legt expliciet het probleem van radicalisering bij het identiteitsvraagstuk:

*'Radicalisering is een complex proces, waarin een persoon gefaseerd een radicale geesteshouding ontwikkelt en zich een nieuwe identiteit of zelfbeeld 'aanmeet'. Men gaat de wereld om zich heen met andere ogen zien en begint de eigen relatie tot die wereld opnieuw te definiëren. (Inter)persoonlijke factoren, maatschappelijke en nationale en internationale politieke ontwikkeling bepalen hoe dit proces in een individueel geval verloopt.'* (2007, 4).

---

<sup>56</sup> De Slotervaart is een stadsdeel in Amsterdam.

**Figuur 6.1: Het antiradicaliseringsbeleid van de Gemeente Amsterdam**  
(Gemeente Amsterdam, 2007, 29)



Het dagelijks bestuur van Stadsdeel Slotervaart stelt dat het voorkomen en tegengaan van radicalisering primair een bestuurlijke verantwoordelijkheid is. Waar radicalisering zich manifesteert in de bereidheid om geweld aan te wenden, is het tegengaan van radicalisering een verantwoordelijkheid van de politie en het Openbaar Ministerie. Voor de rest legt Stadsdeel Slotervaart primair de verantwoordelijkheid bij haarzelf als het gaat om het tegengaan van radicalisering van Slotervaartse jongeren. Preventie is daarbij het sleutelwoord: de aanpak van groepen die vatbaar en ontvankelijk zijn voor radicaliseringprocessen. De preventieve aanpak is erop gericht (moslim)jongeren bewust te maken van radicaliseringprocessen en de wijze waarop ze daartegen weerbaar gemaakt kunnen worden. Daarnaast geven ze aan hoe de ouders van deze (moslim)jongeren ondersteund kunnen worden in hun taak als opvoeder, welke partijen betrokken moeten worden in deze aanpak en waar de accenten komen te liggen (2007, 7-8).

Het stadsdeel realiseert zich dat Slotervaart radicalisering niet in haar eentje kan voorkomen. Het tegengaan van radicalisering vraagt om een gezamenlijke aanpak met andere overheden: een samenwerking met het project gemeentelijke informatiehuishouding radicalisering en het platform Amsterdam Samen, alsmede andere (lokale) partners zoals ouders, jongeren, scholen, moskeeën, imams en (migranten-) en jongeren zelforganisaties (2007, 9). Het actieplan moet gezien

worden als pilotproject, waarbij wordt geëxperimenteerd met nieuwe methodieken. Opgedane (positieve) ervaringen zullen worden overgebracht naar andere stadsdelen. In het actieplan worden zeven beleidsmaatregelen aangekondigd:

- Het bevorderen van de bewustwording onder de (moslim)jongeren en hun sociale omgeving van de gevaren van radicaliseringprocessen.
- Het faciliteren van het aanbod dat gericht is op het ontwikkelen van de competenties die de weerbaarheid van jongeren en hun ouders vergroten.
- Het ondersteunen van ouders in hun opvoedingstaak.
- Het ondersteunen van moskeeën en imams in het omgaan met (vatbaar) radicaliserende jongeren en in het tegengaan van radicalisering.
- Het sluitend maken van de radicaliseringsmeldingssystemen.
- Het faciliteren van de samenwerking met het onderwijs en de jeugdvoorzieningen.
- Het bevorderen van de interactie tussen de verschillende bevolkings- en levensbeschouwelijke groepen.

Om te beoordelen of de ervaring met het actieplan in het geheel en de afzonderlijke deelprojecten positief waren, is een gedegen evaluatie nodig. Maar tot nu toe is weinig ervaring opgedaan met de evaluatie van dergelijke programma's. Hoe meet men of radicalisering is voorkomen? In de volgende paragraaf zullen we laten zien dat traditionele evaluatiemethoden niet voldoende zijn om te meten of radicalisering is tegengegaan. Daarom kijken wij in de daaropvolgende paragraaf naar methoden die buiten Nederland met projecten in crisisgebieden zijn opgedaan, die daar een bijdrage aan conflicttransformatie en aan duurzame vrede tussen verschillende bevolkingsgroepen trachten te leveren.

### 3. Traditionele evaluatiemethoden in vogelvlucht

De vraag hoe beleidseffecten te meten domineert de discussie over beleidsevaluaties. In de literatuur worden vaak verschillende evaluaties naast elkaar gepresenteerd als onverenigbare polen (zie bijv. Guba & Lincoln, 1989; Pawson & Tilley 1997). Zo wordt de klassieke rationale evaluatie tegenover de post-moderne constructivistische evaluatie geplaatst. Een keuze maken voor een specifieke evaluatiemethode, betekent partij kiezen in de oorlog van paradigma's. Guba & Lincoln spreken van verschillende generaties van beleidsevaluatie, namelijk: "meten is weten" (eerste generatie), "beschrijven" (tweede generatie), "beoordelen c.q. interpreteren" (derde generatie) en tot slot hun eigen vierde generatie evaluatiemethode, "onderhandelen". Pawson & Tilley maken een soortgelijke indeling: de experimentele evaluatie, de pragmatische, de constructivistische en tot slot hun eigen "realistische evaluatie".

Ieder paradigma c.q. generatie heeft zo z'n voor- en nadelen die wij in deze paragraaf kort zullen bespreken, waarbij we zullen ingaan op de bruikbaarheid van deze methoden om antiradicaliseringsbeleid te evalueren.

### 3.1. Meten is weten

Deze klassieke methode van het kwantitatief georiënteerde (quasi) experiment verdeelt de doelgroep van het beleid (in principe aselect) over een experimentele en controlegroep, waarbij de eerste wel en de tweede niet aan het beleid wordt blootgesteld. Het beleidseffect wordt uitgedrukt in het verschil tussen de na verloop van tijd gemeten effectvariabelen in beiden groepen. Met deze methode kan in beginsel getoetst worden of er een causale relatie bestaat tussen beleid en verandering. Omdat de experimentele groep en de controle groep voorafgaand aan het beleid/interventie vrijwel hetzelfde zijn, kan iedere geconstateerde verandering worden toegeschreven aan het beleid.

Deze methode is om verschillende redenen buitengewoon problematisch om het effect van antiradicaliseringsbeleid te meten. Ten eerste wordt bij deze methode de context helemaal niet meegenomen. In tegenstelling tot de medische wetenschap waarin deze methode veelvuldig wordt gebruikt, vindt antiradicaliseringsbeleid niet binnen een geïsoleerde klinische omgeving plaats, maar binnen een sociale context die zeer sterk van invloed is op het proces van (de)radicalisering. Ten tweede is het beleid gericht op een doelgroep die zeer wantrouwend is naar de overheid en de wetenschap toe, met als gevolg dat het al moeilijk genoeg is om de relevante doelgroep te bereiken, laat staan zowel een experimentele én een controle groep te onderwerpen aan onderzoek voorafgaand en na het beleid.

Een manier om deze twee nadelen te omzeilen is om niet te werken met een controlegroep en met behulp van een grootschalige survey voorafgaand aan het beleid en na de implementatie van het beleid te meten in hoeverre er verandering is opgetreden in de mate van radicalisering onder de bevolking, met andere worden een quasi-experimenteel onderzoek. Het Instituut voor Migratie en Etnische Studies (IMES) heeft in 2006 door een ex-ante survey de mate van vatbaarheid van radicalisering onder Amsterdammers in kaart gebracht. Het onderzoek concludeerde dat 2 procent van de Amsterdammers vatbaar was voor radicalisering. Door na enkele jaren nogmaals zo een onderzoek uit te voeren, zou gekeken kunnen worden of er verandering is gekomen in dat percentage. Helaas is met deze (quasi)-experimentele methode geen causaliteit te bewijzen. Als er verandering is opgetreden in dat percentage dan is er geen causaal verband aangetoond tussen het antiradicaliseringsbeleid en de opgetreden verandering in vatbaarheid voor radicalisering. Zoals reeds gezegd, is radicalisering zeer context gebonden, zowel op micro- (persoonlijk en interpersoonlijk) als macroniveau (nationaal en internationaal). Zo verzuchtte de voormalige Amsterdamse wethouder Ahmed Aboutaleb ooit eens dat je in Amsterdam-West erg je best kon doen om radicalisering te voorkomen, maar als Israël Libanon binnenvalt, dan slaat de vlam in de pan en zijn alle lokale maatregelen in de stad weer voor niets geweest. Met andere woorden: het beleid zou op zich succesvol kunnen zijn geweest en de vatbaarheid voor radicalisering hebben kunnen verminderen, maar het survey onderzoek zou in een dergelijke situatie toch een toenemende vatbaarheid kunnen constateren, omdat de externe invloeden radicale tendensen versterkt hebben.



Tot slot is er nog de vraag, wanneer er van effect gesproken kan worden. Wanneer kan men een beleidsinterventie om radicalisering tegen te gaan succesvol noemen? Antiradicalisering is bijvoorbeeld niet hetzelfde als criminaliteitspreventie. Om te kijken of criminaliteitspreventie succesvol is, zou bijvoorbeeld gekeken kunnen worden naar de criminaliteitscijfers voor en na de implementatie van het beleid. Zulke cijfers bestaan echter niet, omdat tot op heden geen soortgelijke vragenlijst is ontwikkeld die meet in hoeverre er sprake is van radicalisering.<sup>57</sup> Wat zijn de indicatoren voor (anti)radicalisering? In tegenstelling tot bij criminaliteit kan er niet worden gekeken naar het aantal aangiften, omdat radicalisering an sich geen strafbaar feit is.

### **3.2. Pragmatische en beschrijvende evaluatie**

De vele kritiek op bovenstaande evaluatiemethode, leidde tot een nieuwe vorm van evalueren die bovenal gekenmerkt wordt door pragmatisme, waarvoor de basis is gelegd door Patton (1978) met de “utilization-focused evaluation”. Een experimentele evaluatiemethode is “theory-driven”, terwijl bij de “utilization-focused” evaluatie de nadruk vooral ligt op de bruikbaarheid. Deze evaluatievorm vertoont veel overeenkomsten met een programma-evaluatie, maar met een belangrijk verschil: het gaat om het beoogde gebruik door de beoogde gebruiker (Patton, 1997, 20). De evaluatie is primair gericht op de besluitvormers die informatie willen hebben en de bestaande onzekerheid willen verminderen, om op basis van die informatie eventuele wijzigingen in het beleid aan te brengen.

De tweede-generatie-methode van “beschrijven” is eigenlijk ook een vorm van pragmatisme. Het gaat hier vooral om het aanduiden van de kenmerken van een beleid of programma. Ook een beschrijvende vergelijking van doelen en resultaten/effecten valt hier onder, zoals hoeveel mensen met het beleid zijn bereikt. Denk in geval van een onderwijsprogramma aan een kwalitatieve beschrijving van ondernomen activiteiten, docenten, studenten en interacties in de klas, eventueel aangevuld met beschrijvende statistieken over doorstroom, uitval et cetera. Deze methode kan nuttig zijn om in kaart te brengen hoe men bij antiradicaliseringsactiviteiten te werk gaat en wat de reikwijdte hiervan is, maar verschaft weinig tot geen inzicht in de “outcome” van het beleid.

### **3.3. Interpreteren**

Bij deze vorm van evalueren gaat het om het uitspreken van een finaal oordeel over de kwaliteit. De evaluator fungeert dan als een soort rechter, waarbij de onderzoeksresultaten worden geïnterpreteerd door bijvoorbeeld een reflectie op de vergelijking van doelen en resultaten/effecten (zijn de doelen gehaald? Waardoor wel/niet?). Bijvoorbeeld door een vergelijking van de feitelijke effecten van het onderwijsprogramma met de beoogde doeleinden of door de kwaliteit van de doelen

---

<sup>57</sup> De vragenlijst van het IMES komt in de buurt, door de vatbaarheid voor radicalisering te meten.

te wegen aan de hand van achterliggende wensen (zijn de juiste doelen gesteld?). Deze methode is wat Guba & Lincoln de derde generatie evaluatiemethode noemen. Alle drie de generaties (meten, beschrijven en beoordelen) hebben volgens Guba & Lincoln hun langste tijd gehad, omdat een “management-bias” bestaat: doeleinden en intenties van beleidsmakers worden overgenomen als ijkpunten voor de beoordeling. Er wordt consensus verondersteld terwijl dat in een plurale samenleving vaak niet het geval is, en daardoor ontstaat een eenzijdig vertekend beeld.

### 3.4. Onderhandelen

Doordat er echter een verandering is opgetreden in de locus en focus van politiek en beleid, wordt beleid niet langer gemaakt door de klassiek moderne instituties en heeft er een democratisering in het beleidsproces plaatsgevonden, die wordt gekenmerkt door de shift van government naar governance. Deze democratisering heeft ook zijn doorwerking gehad in beleidsevaluaties en is o.a. uitgewerkt door Guba & Lincoln (1989) die spreken van een “vierde-generatie evaluatie methode” die is gestoeld op het constructivistische gedachtegoed.

Volgens het constructivistische gedachtegoed wordt de “werkelijkheid” door onze waarnemingen geconstrueerd en bestaat er niet zoiets als een fundamentele realiteit. Het gaat in de evaluatie om het leren kennen van de verschillende werkelijkheden en het creëren van een nieuwe werkelijkheid die onderschreven wordt door het grootste deel van de actoren en leidt tot actie. Uitgangspunt bij deze methode is dan dat er niet één te onderzoeken waarheid is, maar dat betrokkenen (“stakeholders”) bij het beleid eigen ervaringen en ideeën hebben die allemaal even “waar” kunnen zijn. Alle betrokkenen hebben een eigen reconstructie, opvattingen, beelden en verhalen. In de vierde generatie evaluatie wordt een gezamenlijke reconstructie gemaakt, in een soort “hermeneutische dialoog”.

De merite van deze methode is dat de vele stakeholders bij het antiradicaliseringsbeleid (lokale overheid, scholen/docenten, jongerenwerkers, moskee et cetera) allemaal hun zegje kunnen doen over de effectiviteit van het beleid. Tevens kan daardoor niet alleen de output, maar ook de outcome van het beleid gemeten worden. Immers, de docent kan mede beoordelen of de geleerde competenties en vaardigheden tijdens socratische gesprekken ook op school worden toegepast door jongeren. Daarnaast kunnen gezamenlijk met al deze practitioners indicatoren geformuleerd worden waaruit blijkt of het antiradicaliseringsbeleid effectief is.

De nadelen van deze methode worden eigenlijk al door Guba & Lincoln zelf verwoord:

*“Evaluation data derived from constructivists inquiry have neither special status nor legitimation; they represent simply another construction to be taken into account in the move toward consensus” (1989, 45)*

*“Phenomena can be understood only within the context in which they are studied;*

*findings from one context cannot be generalized to another; neither problems nor their solutions can be generalized from one setting to another”(1989, 45).*

Kortom, waar met (quasi-)experimentele methoden helemaal geen rekening wordt gehouden met de context, stelt de vierde generatie evaluatiemethode dat de uitkomsten van de evaluatie alleen geldig zijn voor die specifieke context. In de praktijk zou dit betekenen dat (positieve) resultaten van het Slotervaartse antiradicaliseringsbeleid niet overdraagbaar zijn naar andere stadsdelen en gemeenten. Dat is erg ongelukkig gezien het stadsdeel van het ministerie van Binnenlandse zaken juist subsidie heeft gekregen om een experimentele methode op te zetten die vervolgens in het hele land gebruikt zou kunnen worden.

### **3.5. Realistische evaluatie**

In reactie op deze benadering, die nauwelijks ruimte laat om uit evaluaties te leren, hebben Pawson & Tilley een nieuw model ontwikkeld waarin niet langer de vraag centraal staat “wat werkt?” maar de vraag “wat werkt in welke context voor wie onder welke omstandigheden?”. Deze “realistische benadering” heeft een grotere aandacht voor de beleidscontext, het realiteitsgehalte van het beleid, de werkzame elementen ervan en de beleidsuitvoering, zonder overigens de vraag naar de effecten uit de weg te gaan. Sterker, de genoemde aandachtsgebieden van de realistische evaluatie dragen bij aan de vaststelling en het verklaren van de effecten. Bij deze onderzoekstraditie worden interventies dus beschouwd als (micro)theorieën over wat maakt dat een interventie, onder bepaalde condities, de gewenste effecten heeft. De kernbegrippen van een realistische evaluatie zijn mechanisme, context en outcome. Kennis over context en mechanismen zijn immers onontbeerlijk, bijvoorbeeld om succesvol gebleken interventies te vertalen naar andere situaties, maar ook om te begrijpen wat mis ging als een interventie geen of averechtse effecten heeft. Een interventie is alleen causaal als de outcome wordt getriggerd door een mechanisme in de context. Zoals reeds naar voren is gekomen is het bij anti-radicaliseringsbeleid nu net de outcome die zo moeilijk te definiëren is. Met welke indicatoren kan gemeten worden of een persoon bijvoorbeeld weerbaar is geworden tegen radicale gedachten en personen?

Kortom, dit overzicht heeft laten zien dat causaliteit, validiteit, effectiviteit, pluraliteit, generaliseerbaarheid, outcome, context en mechanisme, belangrijke begrippen zijn in een evaluatie. Echter, geen van de beschreven methodes biedt (volledige) uitkomst om een antiradicaliseringsbeleid te evalueren. Daarom zullen we in het volgende deel van dit hoofdstuk een conceptueel kader aanrijken, die voor de evaluatie van “vredesopbouw” projecten elders ontwikkeld is, maar ook, een inspiratiebron zou kunnen vormen voor de evaluatie van Amsterdams anti-radicaliseringsbeleid.

## 4. Evaluatiemethoden van internationale onderzoeksprogramma's

In deze paragraaf zullen drie internationale onderzoeksprogramma's worden voorgesteld, die een schat van ervaringen met evaluaties in crisisgebieden rond om de wereld bij elkaar hebben gebracht. Daarbij gaat het om

- a) het "Local Capacities for Peace" project,
- b) de "Peace and Conflict Assessment"-benadering, en
- c) "Outcome mapping" als een methode voor de evaluatie van "peacebuilding" projecten.

### 4.1. Het "Local Capacities for Peace Project" (LCPP)

Het LCPP werd in 1994 gestart door het initiatief "Collaborative for Development Action", een samenwerkingsverband tussen VN-organisaties, nationale ontwikkelingsorganisaties en meer dan 50 NGO's. Het doel was meer inzicht te krijgen in de gevolgen van ontwikkelingshulp in conflictgebieden. Centraal stond de vraag hoe hulp ingezet kan worden om conflictpartijen dichter bij elkaar te brengen.<sup>58</sup> De in dit project opgeworpen vragen hebben duidelijk raakvlakken met projecten die radicalisering tegen moeten gaan. In beide gevallen gaat het om projecten die de afstand tussen verschillende bevolkingsgroepen moeten verminderen, vijandbeelden weg moeten nemen en de bereidheid om geweld te gebruiken in moeten dammen. De resultaten van het project werden door Mary B. Anderson, directeur van het project, in het boek 'Do No Harm' (1999) samengevat.

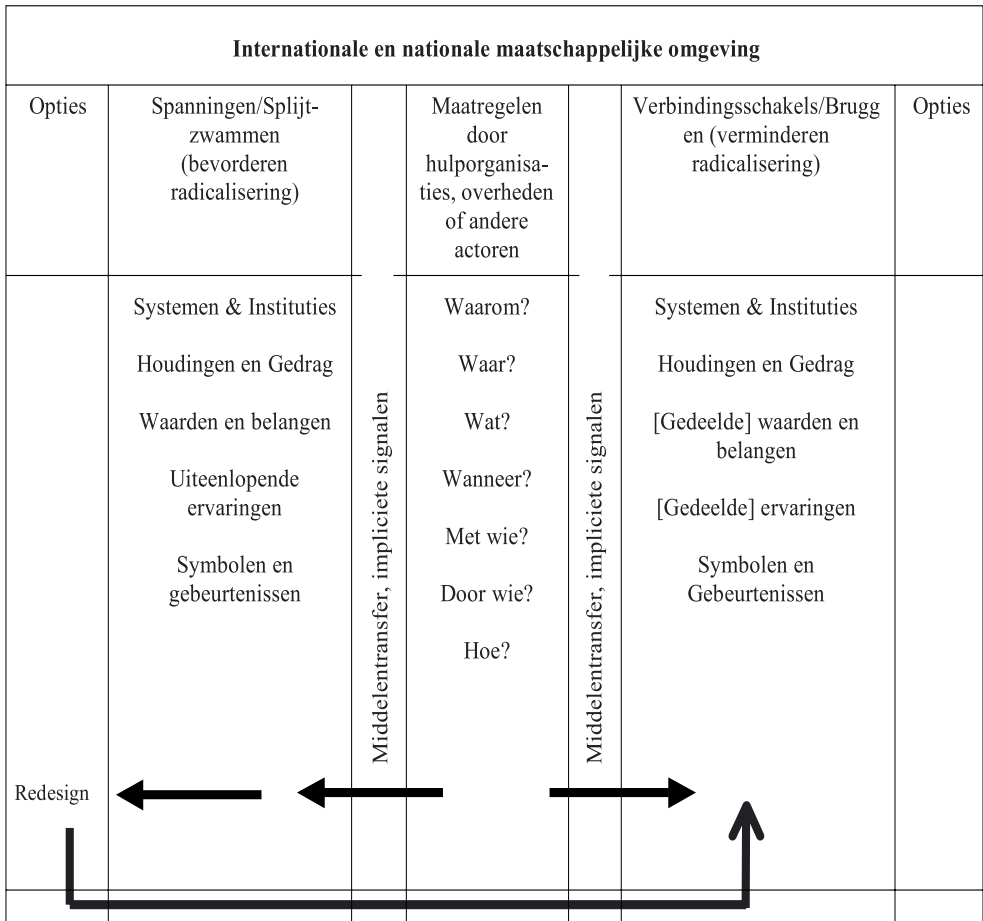
Het LCPP onderzocht hoe interventies (in het geval van LCPP ontwikkelingshulp) ertoe bijdragen om ofwel "dividers" (splitszwammen) of "connectors" (bruggen) in een maatschappij te versterken of te verzwakken. Deze "dividers" en "connectors" kunnen zijn: systemen en instituties, houdingen en gedragingen, waarden en belangen, of symbolen en gebeurtenissen. Indien interventies onbewust ertoe bijdragen, de spanningen op te voeren, dan zouden zij meteen bijgesteld moeten worden.

Mary Anderson (1999, 74) geeft deze terugkoppeling in het volgende – hier vertaalde en wat vereenvoudigde - schema weer:

---

<sup>58</sup> De vraagstelling luidde: 'How may aid be provided in conflict settings in ways that, rather than feeding into and exacerbating the conflict, help local people disengage from the violence that surrounds them and begin to develop alternative systems for addressing the problems that underlie the conflict?' Zie [www.donoharm.info](http://www.donoharm.info).

**Figuur 6.2: Spleitzwammen vs. verbindingsschakels: een kader voor conflicttransformatie**



Figuur 6.3 geeft illustratieve voorbeelden van “connectors” en “dividers” zoals door Mary Anderson e.a. in de context van internationale hulpprojecten beschreven.

**Figuur 6.3: Spleitzwammen vs. verbindingsschakels in internationale conflictgebieden**

(Anderson, M.B. 1999, 24-33)

Spleitzwammen (“dividers”)	Verbindingsschakels (“connectors”)
<p><i>Systemen en instituties:</i> Legers, bendes, wapenproductie, propagandamedia, discriminerende wetgeving, stelselmatige uitsluiting</p> <p><i>Houdingen en gedrag:</i> Geweld, dreigingen, brutaliteit, repressie, wetteloosheid, wantrouwen, verdachtmaking</p> <p><i>Divergerende waarden en belangen:</i> Religieus extremisme, sektarische oriëntaties, conflicterende ideologieën, vooroordelen, hebzucht</p> <p><i>Uiteenlopende ervaringen:</i> Concurrerende historische mythes, ervaringen van onrechtvaardigheid, dominantie, vernedering, repressie, eenzijdige beschrijvingen in schoolboeken</p> <p><i>Symbolen en gebeurtenissen :</i> Beladene symbolen, vieringen en herdenkingen (bijv. Oranjemars in Noord-Ierland; bezoek aan Yasukuni tempel in Japan)</p>	<p><i>Systemen en instituties:</i> markten, bus- en spoorverbindingen, elektriciteitsnetten, telefoonlijnen, , professionele associaties</p> <p><i>Houdingen en gedrag:</i> Tolerantie, acceptatie, buurthulp, vriendschappen, gemengde huwelijken, professionalisme</p> <p><i>Gedeelde waarden en belangen:</i> Liefde voor kinderen, voor muziek &amp; kunst, belang bij werking van gezondheidsdiensten</p> <p><i>Gedeelde ervaringen:</i> Ervaringen van vrouwen, van moeders, van gewonde strijders, ervaringen met hulpverleners, met internationale organisaties</p> <p><i>Symbolen en gebeurtenissen</i> Nationaal historisch erfgoed, monumenten, feestdagen, vredesacties</p>

Iedere interventie zou voortdurend bekeken moeten worden welke signalen ervan uitgaan en welke spleitzwammen of verbindingsschakels door de ter beschikking gestelde middelen en door de keuze voor bepaalde projecten, partners of locaties versterkt of verzwakt worden.

Je kunt dit analyseschema ook “vertalen” naar antiradicaliseringsprogramma’s. Het resultaat is dan een schema zoals in figuur 6.4 weergegeven.

**Figuur 6.4: Splijtzwammen vs. verbindingsschakels: een kader voor conflictransformatie**

Impact van maatregelen in het kader van het antiradicaliseringsbeleid						
Opties	Spanningen/Splijtzwammen (bevorderen radicalisering)		Maatregelen door maatschappelijke organisaties, overheden of andere actoren		Verbindingsschakels/Bruggen (verminderen radicalisering)	Opties
	<p><i>Systemen &amp; Instituties:</i> Immigratiebeleid, splitsing tussen “zwarte” en “blanke” scholen; radicale moskeen, media</p> <p><i>Houdingen en Gedrag:</i> Crimineel gedrag, agressie, hangjongeren, overlast, spijbelen, vooroordelen, discriminatie, haatpreken, hand weigeren, Islam bashing,</p> <p><i>Divergerende waarden en belangen:</i> Exclusief gebruik van ruimte, subsidieclaims, genderverhoudingen, voorkeursbehandeling, kijk op homosexualiteit, dierenliefde, concurrentie</p> <p><i>Uiteenlopende ervaringen:</i> Achterstelling, vernedering, relatieve deprivatie, bedreiging, achtervolging door politie, gevangenis, reclassering, overlast</p> <p><i>Symbolen en gebeurtenissen:</i> Moorden, veroordelingen, martelaarschap, oorlogen, invasies, koloniale geschiedenis, schotels, hoofddoekjes, cartoons, Wilders film</p>	Middelentransfer, impliciete signalen	Interventiologica  Keuze van - doelgroep - locatie - maatregel - partners - actoren - middelen - manier van interventie	Middelentransfer, impliciete signalen	<p><i>Systemen &amp; Instituties:</i> Markten, gezondheidszorg, woningbouwcoöperaties, sportcentra, stagebemiddeling, hulp bij inburgering, antidiscriminatie-wetgeving</p> <p><i>Houdingen en Gedrag:</i> Tolerantie, buurthulp, acceptatie, professionalisme, waardering van diversiteit</p> <p><i>Gedeelde waarden en belangen:</i> Gezondheid, educatie, stabiliteit, orde, rechtvaardigheid, welvaart, levenskwaliteit</p> <p><i>Gedeelde ervaringen:</i> ontmoetingen, dialogen, gesprekken, civil society activiteiten, gedeelde successen (bijv. in sport), Stadspelen, gedeeld leed (bijv. ontslagen)</p> <p><i>Symbolen en Gebeurtenissen:</i> buurtfeesten, culturele manifestaties, suikerfeest, tentoonstellingen</p>	
	Redesign					

Het LCPP project heeft met name duidelijk gemaakt, dat veel goed bedoelde maatregelen (in het kader van LCPP met name humanitaire hulpverlening) uiteindelijk negatieve uitwerkingen kunnen hebben in conflictgebieden. Dezelfde waarschuwing is misschien ook voor maatregelen in het kader van het antiradicaliseringsbeleid op zijn plaats. Amartya Sen (2006) heeft bijvoorbeeld gewaarschuwd dat projecten, die bedoeld zijn om de dialoog tussen verschillende bevolkingsgroepen te verbeteren, soms averechtse gevolgen hebben. Al door het uitnodigen van verschillende personen als representanten van afzonderlijke groepen wordt de groepsidentiteit versterkt in plaats van voorop te stellen dat iedere mens een complexe identiteit bezit, die voortdurend in ontwikkeling is en zeer uiteenlopende aspecten bevat (zoals gender, leeftijd, afkomst, gezinssituatie, buurt, opleiding, beroep, taal, dialect, religie, huidskleur, levensgeschiedenis, seksuele geaardheid, eetgewoontes, vrijetijdsbesteding, etc.). Ook veel andere maatregelen tegen radicalisering kunnen tegenstrijdige effecten hebben. Zij worden vaak als discriminerend ondervonden door de doelgroepen, waardoor de voedingsbodem voor radicalisering juist kan toenemen.

Een vervolg op het LCPP project was het project “Reflecting on Peace Practice” (RPP), eveneens gedragen door de Collaborative Learning Projects (CDA) at Cambridge, Mass. Ook dit project heeft tot een schat van ervaringen uit tal van conflictgebieden geleid. In het kader van het RPP project is het volgende schema ontworpen, dat een bijdrage kan leveren aan het ontwerpen en de analyse van interventiestrategieën.

Het schema helpt in beeld te brengen waar de verschillende programma-activiteiten aanzetten, hoe zij onder elkaar samenhangen, welke veranderingen zij bewerkstelligen en in hoeverre de output van de projecten daadwerkelijk het gewenste maatschappelijke resultaat dichterbij brengt.



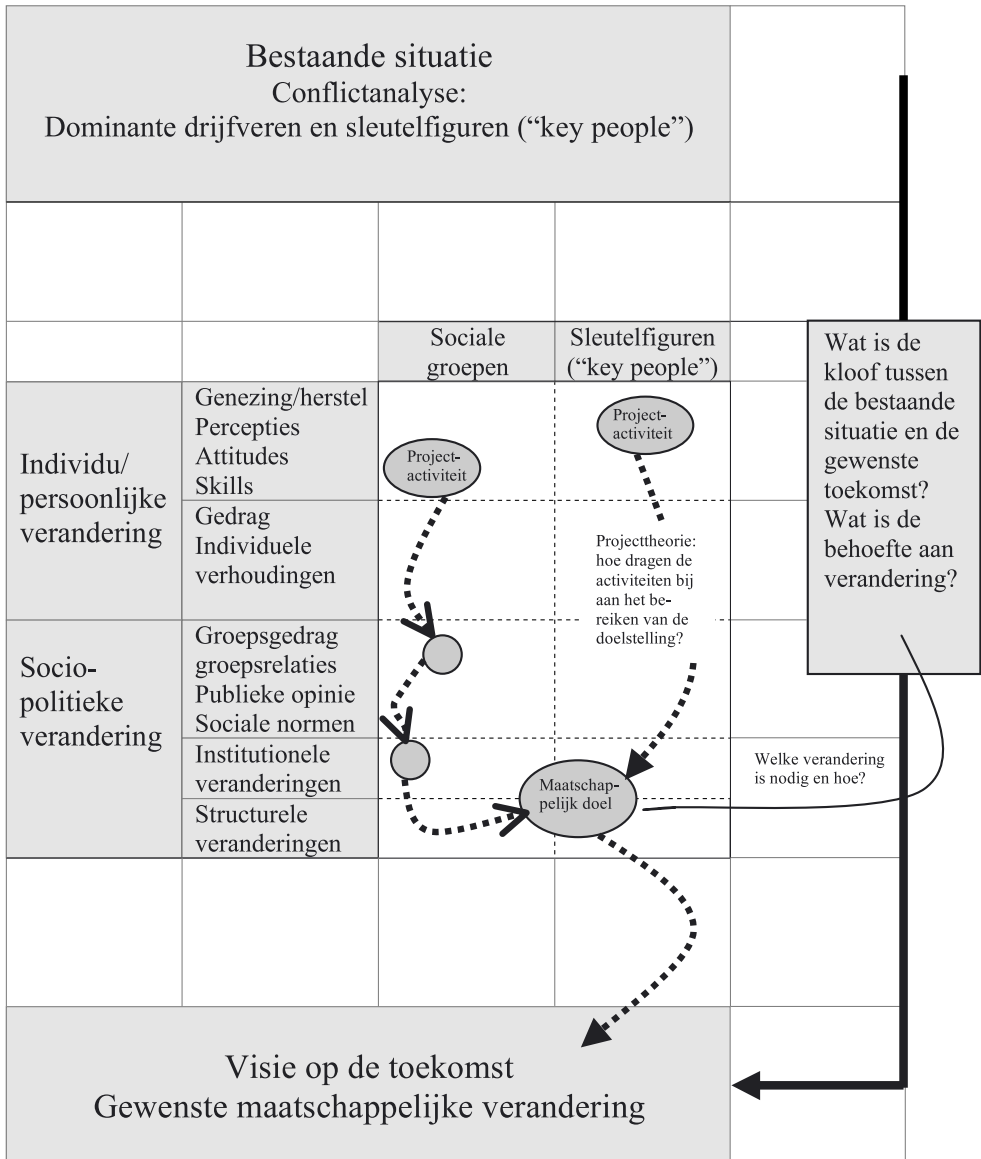
**Figuur 6.5: Analyse kader voor ontwerp en evaluatie van interventiestrategieën**

Veranderingstheorie: Hoe verhoudt het onmiddellijke projectdoel zich tot de nagestreefde maatschappelijke verandering?

(Reflecting on Peace Practice project,

[http://www.cdainc.com/cdawww/pdf/manual/rpp\\_training\\_matrix20plus\\_200803181\\_Pdf.doc](http://www.cdainc.com/cdawww/pdf/manual/rpp_training_matrix20plus_200803181_Pdf.doc),

het schema is vertaald en licht aangepast)



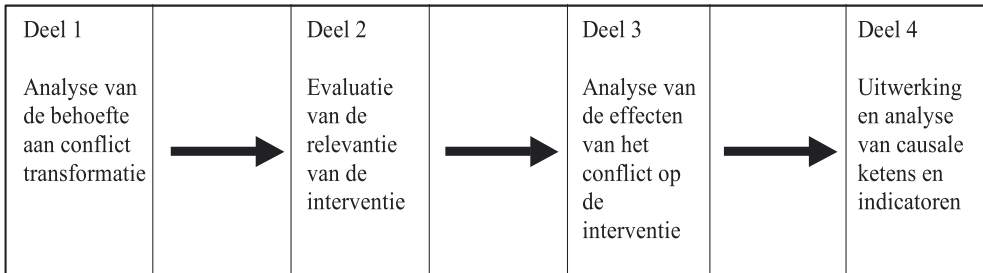
## 4.2. Peace and Conflict Assessment (PCA)

Deze benadering is in 1996 onder de label “Conflict Impact Assessment Systems” (CIAS) begonnen om een kader te bieden voor de evaluatie van politieke interventies op macroniveau in conflictsituaties en vredesprocessen (Reychler, 1999b) De aanpak werd voortdurend verder ontwikkeld en sterk door de ervaringen met andere evaluatiemethoden verrijkt, waaronder de inzichten uit de “Do no harm”-benadering van Anderson. Het Berghof Handbook for Conflict Transformation<sup>59</sup> bevat talloze artikelen die een bijdrage leveren aan de discussie over Peace and Conflict Impact Assessment (PCIA).<sup>60</sup>

De verzamelde inzichten hebben Paffenholz en Rechler in het boek ‘*Aid for Peace. A Guide to Planning and Evaluation for Conflict Zones*’ (2007) samengevat. In “Aid for Peace” geven de auteurs een complete handleiding voor het plannen van ontwikkelingsprojecten in conflictgebieden en voor de evaluatie van dergelijke projecten. Ook hier bestaan veel raakvlakken met antiradicaliseringsprojecten.

Het Peace and Conflict Assessment framework bestaat uit vier delen (Paffenholz & Reychler 2007, 17):

**Figuur 6.6: Vier delen van het Peace and Conflict Assessment framework**



Over het eerste deel kunnen we hier kort zijn. We vooronderstellen dat de behoefte aan antiradicaliseringsprojecten bestaat. Onderzoek vanuit verschillende invalshoeken heeft laten zien dat er in Nederland de polarisatie tussen verschillende bevolkingsgroepen is toegenomen. Een indicator is dat de belangstelling voor het salafistische gedachtegoed is gegroeid.<sup>61</sup>

Deel 2 is helemaal niet vanzelfsprekend. Bij de relevantie van veel projecten die gestart zijn om radicalisering tegen te gaan kunnen grote vraagtekens gezet worden. Vaak zijn het algemene projecten die een bijdrage leveren aan sociale cohesie en

<sup>59</sup> <http://www.berghof-handbook.net/>

<sup>60</sup> Peace and Conflict Assessment (PCA) heeft zich uit Peace and Conflict Impact Assessment (PCIA) ontwikkeld, zie Paffenholz & Reychler 2007, 84 en 145-146.

<sup>61</sup> Volgens de AIVD (oktober 2007) is het aantal jongeren dat salafistische lessen volgt in twee jaar tijd minstens verdubbeld: zo'n twintig- tot dertigduizend jongeren zijn volgens de dienst vatbaar voor het salafistische gedachtegoed. Volkskrant, 17 mei 2008, 37.

daarmee –hoopt men– indirect ook op de een of andere manier aan de vermindering van radicale tendensen. In veel gevallen is dit zeer discutabel. Om beter zicht te krijgen op de relevantie van projecten moeten vragen worden gesteld zoals:

- Wat is de doelgroep van het project?
- Wie zijn de uitvoerende partners?
- Welk inzicht bestaat bij de partners in de dynamiek van radicalisering en deradicalisering?
- Werden soortgelijke projecten al eerder uitgevoerd?
- Wat weten wij over de resultaten van deze projecten?

Aandacht voor de relevantie van de interventie moet veilig stellen dat een project ook daadwerkelijk aansluit bij de analyse van het probleem (Paffenholz & Reyhler 2007, 21). Er moet duidelijk zijn dat de beoogde effecten, als zij gehaald worden, daadwerkelijk tot de nagestreefde verandering van de situatie leiden. En er moet duidelijkheid zijn, dat de beoogde effecten ook daadwerkelijk met de ingezette middelen bereikt kunnen worden. Hiervoor zijn bijvoorbeeld de volgende vragen te beantwoorden:

- Is de doelgroep daadwerkelijk bereikbaar?
- Is de doelgroep bereid tot medewerking aan het project?
- Wordt er op een zodanige schaal medewerking verleend dat het project ook zoden aan de dijk zet?
- Zullen de effecten duurzaam zijn of gaat het slechts om een vluchtige verandering?
- Welke follow-up moet er straks aan het project worden gegeven?
- Zijn hiervoor de benodigde financiële middelen beschikbaar?
- Is er voldoende expertise in huis om het project uit te voeren en tot een goed einde te brengen?

Een andere belangrijke vraag is of het project onderdeel uitmaakt van een omvattende aanpak of het slechts om een op zich zelf staand project gaat. De kans dat geïsoleerde projecten een noemenswaardige bijdrage leveren aan een omvattende doelstelling zoals het tegengaan van radicalisering is zeer beperkt.

Bij sommige van de in Slotervaart uitgevoerde projecten zou men de relevantie in twijfel kunnen trekken (zie hieronder). Waar de relevantie met recht betwijfeld mag worden, kan het om projecten gaan die primair een andere doelstelling hebben: Soms gaat het er niet zo zeer om een bepaald effect te bereiken, maar om te laten zien dat er überhaupt iets ondernomen wordt, d.w.z. om een “symbolic use of politics”.

Deel 3 is slechts voor risicorijk onderzoek in conflictgebieden interessant, waar de onderzoekers en hun partners gevaar zouden kunnen lopen. Je kunt je weliswaar situaties voorstellen, waar dit ook geldt voor onderzoek naar radicalisering en deradicalisering, maar op deze randgevallen zullen we hier niet ingaan.

Deel 4 gaat ervan uit dat een goede evaluatie vraagt naar de interventielogica die aan een project ten grondslag ligt. Is er goed over nagedacht hoe de causale ketens er uit zien die tot het ongewenste verschijnsel (radicalisering) leiden? Is duidelijk, op welk punt van deze ketens de interventie moet inwerken? Alleen als je dit weet, dan

kun je naar indicatoren kijken die de aannames over de causale ketens helpen te toetsen en die kunnen laten zijn of er een verandering in de keten opgetreden is na de interventie. Of deze verandering dan ook aan de interventie te danken is, dat is nog een andere vraag.

### 4.3. Outcome mapping

Het zonet besproken *Peace and Conflict Assessment* is vooral gericht op de bijdrage van projecten aan het aanwakkeren of verminderen van conflicten, de bijdrage aan “vrede”. Maar “vrede” (net als de toename of afname van radicalisering) hangt van heel veel factoren af. Het *attributieprobleem*, waarmee vrijwel alle evaluaties op dit gebied te kampen hebben, is levensgroot. Ieder project kan maar een klein aantal factoren beïnvloeden, die een bijdrage leveren aan “vrede” of aan “radicalisering”. We leven in een complexe samenleving, waarin talloze processen van invloed zijn op het gedachtegoed en het gedrag van mensen. Meestal is het een interactie tussen verschillende invloeden, die een bepaald resultaat teweeg brengt.

Om op een gepaste manier met deze problematiek om te gaan, volgt “outcome mapping” een bescheidenere benadering. In plaats van de invloed van bepaalde interventies op “conflict transformatie” of “vermindering van radicalisering” in het algemeen te proberen te meten, beperkt deze methode zich op de analyse van gedragsveranderingen, acties, activiteiten van personen en organisaties, die direct met de interventie in aanraking zijn gekomen. Hoe heeft de doelgroep gereageerd? Gedraagt zij zich nu anders? Is het gedrag van omstanders veranderd (vriendengroepen, medescholieren, burens...)? Zien de contactpersonen (leraren, politieagenten, jeugdwerkers) een verandering? Op deze manier wordt in kaart gebracht, welke veranderingen in de directe omgeving hebben plaatsgevonden, die nog het meest met het project zelf in verband kan worden gebracht. Speculaties over verdergaande maatschappelijke repercussies worden vermeden, omdat hierover geen goed onderbouwde uitspraken kunnen worden gedaan.

Outcome mapping werd met name door de Canadese Organisatie IDRC (International Development Research Centre)<sup>62</sup> ontwikkeld en er is haast een hele industrie erom heen ontstaan met trainingspakketten, een virtual learning community en een groot aantal artikelen.<sup>63</sup> In Nederland geeft bijvoorbeeld MDF Training & Consultancy in Ede cursussen in Outcome Mapping.<sup>64</sup>

Als we de outcome mapping methode toepassen op antiradicaliseringsprojecten, dan zou aan de volgende indicatoren gedacht kunnen worden (gebaseerd op kenmerken van radicalisering, zie Gielen 2008 voor een overzicht).

---

<sup>62</sup> Zie de uitgebreide website over outcome mapping: [http://www.idrc.ca/en/ev-26586-201-1-DO\\_TOPIC.html](http://www.idrc.ca/en/ev-26586-201-1-DO_TOPIC.html).

<sup>63</sup> Zie de website van de Outcome Mapping Learning Community: [www.outcomemapping.ca](http://www.outcomemapping.ca).

<sup>64</sup> Zie de website [http://www.mdf.nl/index.php/page/85/?mod\[MDFCourseCalendarModule\]\[item\]=95](http://www.mdf.nl/index.php/page/85/?mod[MDFCourseCalendarModule][item]=95).

## Figuur 6.7: Indicatoren voor de toetsing van antiradicaliseringsprojecten

### *Indicatoren op het individuele niveau:*

- toename van contacten met andersdenkenden,
- toenemende belangstelling voor andere aspecten van het leven (familie, werk, school, sport, kunst, buurt, bioscoop, ... )
- minder strikte kleding- en gedragsregels
- hoger tolerantieniveau tegenover andersdenkenden (af te lezen aan bijv. bezoek van discussiebijeenkomsten, bezoek aan “gemengde”(m/f) bijeenkomsten, hand geven )
- vaker op bezoek bij ouders/familieleden
- kennisname van verschillende media
- minder schoolverzuim
- bereidheid tot contacten met “de politiek”

### *Indicatoren op het niveau van de groep:*

- grotere openheid van de groep (voor bezoekers, vrienden, leden met andere opvattingen)
- uitnodiging van sprekers met andere meningen
- minder hiërarchie en meer diversiteit in de groep
- minder strikt verwachtingspatroon t.a.v. het handelen van groepsleden
- breder programma van de groep, nieuwe discussiethema's
- minder neiging zich af te zetten tegen andere groepen
- wegtrekken radicale sleutelfiguren (of minder aanhang radicale sleutelfiguren)
- nieuwe rolmodellen
- uitnodigingen door andere groepen

### *Indicatoren op het niveau van de omgeving:*

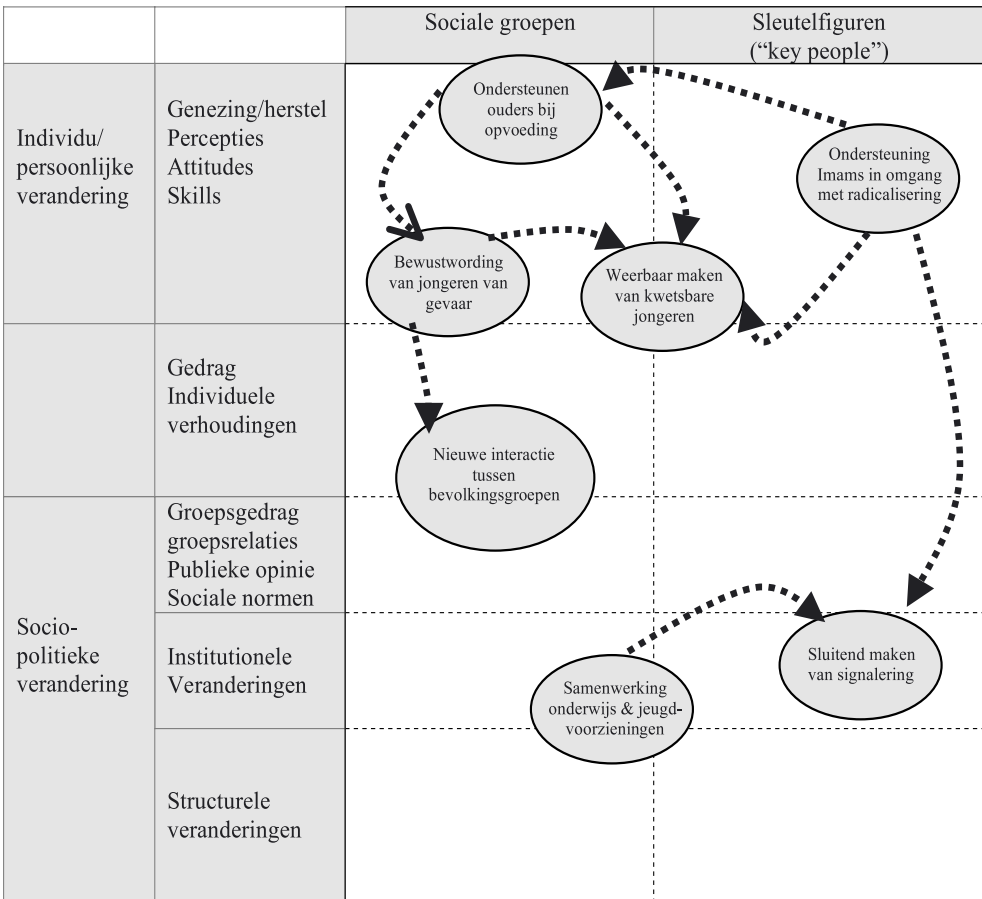
- positieve waarneming door jeugdwerkers, leraren, politieagenten, imams
- meer contact met de burens
- actievere participatie in de samenleving
- minder opvallend op het werk, beter contact met collega's
- minder discriminatie
- positievere berichtgeving in de media

## 5. Aanzet tot een evaluatie

Als we nu de in paragraaf II genoemde projecten in Amsterdam Slotervaart om radicalisering tegen te gaan willen evalueren, hebben we dan iets aan de zonet beschreven evaluatiemethoden?

In ieder geval is het schema, dat het *Reflecting on Peace Practices* (RPP) Project aanreikt, van nut om de verschillende deelprojecten van het actieplan in Slotervaart in kaart te brengen en hun onderlinge verhouding duidelijk te maken.

**Figuur 6.8: De projecten in Slotervaart in kaart gebracht**



Een dergelijke "mapping exercise" helpt om bij de evaluatie een aantal kritische vragen te stellen, die verband houden op de onderlinge samenhang tussen de projecten. In dit geval kunnen dit bijvoorbeeld vragen zijn zoals:

- Heeft de ondersteuning aan de imams positieve gevolgen gehad voor de

ondersteuning van de ouders, - hebben de imams op basis van het project de ouders goed advies kunnen geven?

- Hebben de imams acties ondernomen om de weerbaarheid van kwetsbare jongeren te verhogen?

- Hebben ouders hun kinderen (en hun vrienden) attent gemaakt op de gevaren van radicalisering?

- Wat hebben de ouders op basis van het project gedaan om de weerbaarheid van kwetsbare jongeren te verhogen?

- Heeft de toegenomen bewustwording van jongeren t.a.v. de gevaren van radicalisering bijgedragen tot nieuwe interactie tussen verschillende bevolkingsgroepen?

- Hebben de imams een bijdrage geleverd aan de signalering van radicaal gedrag?

- Heeft de samenwerking tussen onderwijs & jeugdvoorzieningen bijgedragen aan een beter signaleringssysteem?

Naast de mogelijke interactie tussen de verschillende deelprojecten helpt een dergelijke kaart ook om mogelijke leemtes in het actieplan te constateren en ideeën aan te reiken voor nieuwe projecten.

In het schema wordt bijvoorbeeld duidelijk dat het stadsdeel ook zou kunnen denken aan institutionele en structurele veranderingen. Sommige veranderingen zijn op het niveau van het staddeel denkbaar (bv. in scholen, woonsituatie), andere structurele maatregelen (bijv. t.a.v. de arbeidsmarktsituatie) kunnen waarschijnlijk slechts op het niveau van de gemeente Amsterdam of zelfs een hoger niveau worden gerealiseerd.<sup>65</sup> Er zijn ook geen deelprojecten die individuele sleutelfiguren aanspreken op hun gedrag, maar dat is waarschijnlijk overgelaten aan politie en justitie. Misschien zouden ook nog andere institutionele veranderingen denkbaar zijn die helpen om radicalisering tegen te gaan.

Uit dezelfde koker als het RPP project komt het LCPP project, waarvan Anderson's boek *'Do No Harm'* het meest belangrijke resultaat was. Het met name voor humanitaire hulporganisaties bestemde boek maakte hen erop attent dat de goed bedoelde humanitaire hulpverlening ook ertoe bij kan dragen conflicten te rekken en daarmee uiteindelijk de humanitaire situatie te verergeren in plaats van te verbeteren.

Meer in het algemeen leert dit project om bij evaluaties niet alleen aandacht te besteden aan de vraag of de verwachte positieve resultaten ook daadwerkelijk optreden, maar om ook te letten op onbedoelde neveneffecten. Daarom is het goed om stil te blijven staan bij mogelijke "dividers" (splitszwammen), die eventueel door de projecten versterkt zouden kunnen worden.

In het geval van de projecten in Slotervaart zou bijvoorbeeld gevraagd kunnen worden:

---

<sup>65</sup> Een wél op betere kansen op de arbeidsmarkt gericht lokaal project is Campus Nieuw West, dat ervoor zorgt dat jongeren kennis maken met het bedrijfsleven, zie <http://www.campusnieuwwest.nl/>.

- Zouden de imams zelf radicaler worden in hun opvattingen als de Nederlandse overheid (in dit geval het stadsdeel) zou proberen om hen een rol te geven in het tegengaan van radicalisering?
- Zouden jongeren zich nog meer afzonderen van hun ouders als de ouders proberen hen met andere opvoedingspraktijken te bejegenen?
- Zouden jongeren door projecten, die ze op de gevaren van radicalisering moeten wijzen, misschien juist nieuwsgierig worden om meer radicaal gedachtegoed ter kennis te nemen?
- Zouden jongeren door de poging om weerbaarder te worden en zelfbewust eigen beslissingen te kunnen nemen misschien “weloverwogen” kunnen besluiten om zich bij een radicale groepering aan te sluiten?
- Zou een sluitend signaleringssysteem ertoe bijdragen dat jongeren zich nog meer gewantrouwd en achtervolgd voelen?
- Zou de samenwerking tussen onderwijs en jeugdvoorziening ten laste gaan van andere samenwerkingsverbanden? Er is immers tijd gemoeid met een goede samenwerking, en daardoor blijven misschien andere taken liggen.

Het “Do No Harm”-perspectief leert dus ook heel bewust te kijken naar mogelijke negatieve effecten van projecten, die vol goede bedoelingen zijn opgezet. Juist door de goede bedoelingen komt het bij de projectuitvoerders vaak helemaal niet op dat hun projecten ook andere uitwerkingen zouden kunnen hebben (en wordt informatie die in deze richting wijst, niet graag ter kennis genomen).

Van de “Peace and Conflict Assessment”-benadering vallen weer andere dingen te leren. Hier wordt met name naar de relevantie van de projecten gevraagd. Daarbij gaat het om tenminste vier deelvragen: Wordt de juiste *doelgroep* bereikt? Is de interventie *doelmatig*, d.w.z. zou zij ook de beoogde effecten kunnen bereiken? Wordt niet steeds weer *hetzelfde* gedaan? En staan de projecten zodanig in een *context* dat ook daadwerkelijk verreikende consequenties verwacht kunnen worden?

In het geval van Slotervaart zouden deze vragen er als volgt uit kunnen zien:

### *Doelgroep*

Zijn de jongeren die in het kader van het antiradicaliseringsbeleid benaderd worden, een relevante doelgroep? Worden de juiste groepen ouders benaderd? Wordt ondersteuning geboden aan imams die hier behoefte aan hebben en er ook iets mee doen? Op vrijwel alle gebieden kennen we het probleem van “preken voor de eigen parochie”. Er worden vooral mensen benaderd die toch al de goede intenties hebben, vrij goed geïnformeerd zijn, over de juiste skills beschikken en contact hebben met de relevante instituties. Maar dat zijn waarschijnlijk niet die mensen die met name bereikt zouden moeten worden.

### *Doelmatigheid*

In het hieronder beschreven trainingsproject is het primaire doel om orthodoxe moslima’s aan zelfstandige oordeelsvorming te helpen (en hen daardoor minder volgzaam te maken voor leiders met radicale ideeën). Maar hoe hangt de zelfstandige oordeelsvorming samen met radicalisering? Zullen mensen die onafhankelijk nadenken per se tot andere conclusies komen? Zij emanciperen immers tegelijk



ook t.a.v. de invloeden van imams, ouders, leraren, politici en andere autoriteiten die radicalisering tegen willen gaan. Vaak heeft een trainingsbureau al een bepaald programma in huis en dit programma wordt slechts licht aangepast voor de nieuwe doelgroep, zonder voldoende te behartigen of dit daadwerkelijk een bijdrage zou kunnen leveren aan het nagestreefde veranderingsproces.

### *Herhaling van projecten*

Vaak wordt niet voldoende stil gestaan bij de resultaten die eerdere projecten hebben gehad (of juist niet hebben gehad). Als eerdere projecten geen succes waren, waarom wordt dan van de lopende projecten een succes verwacht? Wat is deze keer anders? Als eerdere projecten wél een succes waren, waarom moet er dan nog een schepje bovenop worden gelegd? Waren de positieve gevolgen slechts tijdelijk? Is er een nieuwe groep aan wie hetzelfde project aangeboden moet worden? Werden ouders eerder al in hun opvoedingspraktijken ondersteund? Heeft dit blijvende effecten gehad? Wat waren de problemen in de samenwerking met de moskeen? Is uit deze projecten geleerd? Waarom is de samenwerking tussen onderwijs en jeugdvoorzieningen niet eerder tot stand gekomen? Zijn er structurele obstakels voor de samenwerking die tot nu toe over het hoofd werden gezien?

### *Context*

Alleenstaande projecten hebben meestal weinig uitzicht op succes. Zij moeten ingebed zijn in een heel programma, waarmee tegelijk veel factoren die ongewenst gedrag (kunnen) veroorzaken, aangepakt kunnen worden. Een overzicht zoals in figuur 6.8 kan helpen bij het bepalen, of er een omvattend programma bestaat, waarvan de afzonderlijke projecten deel uitmaken. Dat is in Slotervaart zeker het geval, ook al zou het programma nog omvattender kunnen zijn. Maar dit is in ieder geval een poging om projecten deel uit te laten maken van een groter geheel zodat de kans op succes hoger is.

## **6. Het trainingsproject 'Identiteit en weerbaarheid'**

Nu wij het conceptuele kader voor het evalueren van een volledig antiradicaliseringsbeleid hebben geschetst, zullen wij in deze paragraaf de lezer instrumenten aanreiken om de deelactiviteiten te evalueren, middels de bespreking van een casus. Gielen (2008) heeft diverse deeleactiviteiten van het antiradicaliseringsprogramma in Stadsdeel Slotervaart geëvalueerd. Voor de concrete toepassing hebben wij gekozen voor het project 'Identiteit en weerbaarheid'. In deze paragraaf zullen wij bespreken hoe het geschetste bovenstaande conceptuele kader is gebruikt bij de evaluatie van deze deelactiviteit.

In feite is het zaak om dezelfde vragen te blijven stellen zoals in de vorige paragraaf zijn geformuleerd, maar dan specifiek toegepast op de deelactiviteit. De methode van Peace and Conflict Assessment leert ons dat een aantal deelvragen relevant zijn, om te beginnen: wordt de juiste doelgroep wel bereikt?

### *De doelgroep*

De doelgroep voor de training '*Identiteit en weerbaarheid*' was een groep zeer praktiserende moslima's (tweede en derde generatie) in de leeftijdscategorie van 17-33 jaar. De meeste meiden leefden geïsoleerd: ze hadden alleen contact met hun "zusters" (andere zeer praktiserende moslimmeiden). Met autochtonen hadden deze meiden over het algemeen nauwelijks of geen contact. Ze leefden niet in het hier en nu, maar slechts voor het hiernamaals. Er was weinig tot geen interactie met mannen, collega's of medescholieren/studenten of mensen uit de buurt. Soms vielen ouders zelfs buiten de boot. De meiden spraken namelijk met *dédain* over hun ouders die "slechts" een culturele religieuze beleving hebben en niet de "zuivere islam" praktiseren. De meiden waren weinig thuis en trokken van lezing naar lezing. Ook waren ze zeer actief op internet (religieuze fora). De oudere meiden hadden hun baan opgegeven en houden zich slechts met de islam bezig. Zij gaan in een zwarte *hijaab* of *ghimaar* gekleed en sommige gaven mannen geen hand en keken ze niet aan. Muziek luisteren was in hun ogen niet toegestaan. Competenties zoals vragen stellen, zelfstandig antwoorden, reflecteren, kritisch zijn, ontbraken volledig. Enkele meiden volgden ook een salafistische opleiding waarbij ze een contract hebben moeten ondertekenen dat ze niet zelfstandig mogen nadenken. Tijdens die opleiding werden ze "opgevoed" om heden ten dage te leven zoals de Profeet en zijn metgezellen 1400 jaar geleden leefden. Er werd geleerd "daden van ongeloof te haten" en de voorkeur te geven aan islamitische vrienden.<sup>66</sup>

Kortom, op basis van deze nulmeting (op basis van interviews met de meiden en trainsters), kan worden vastgesteld dat deze meiden tot de doelgroep van het antiradicaliseringsbeleid behoorden. Hoewel zij zelf (nog) niet geradicaliseerd waren, waren zij vatbaar voor radicale personen en gedachten.

Daarom is besloten deze meiden een training aan te bieden, om ze een eigen identiteit te laten ontwikkelen (in plaats van een identiteit van buitenaf opgelegd te krijgen) en ze weerbaar te maken tegen invloeden van buitenaf (radicale predikers, dominante ouders etc). In dat licht is de derde deelvraag van Peace and Conflict Assessment relevant: wordt niet steeds hetzelfde gedaan? Om deze vraag te beantwoorden, zal de inhoud van de training nu nader worden toegelicht.

### *De training*

Voorafgaand aan een inhoudelijke beschrijving van de training, is het belangrijk te vermelden dat er zeer weinig activiteiten voor tweede generatie moslima's worden georganiseerd. Er zijn zeer veel activiteiten voor migrantenvrouwen van de eerste generatie (taalverwerving, empowerment etc.) en daarnaast wordt er ook veel georganiseerd voor jongens van de tweede generatie. De meiden van de tweede generatie zijn tot op heden het achtergebleven kindje gebleven. Simpelweg omdat men van deze meiden geen last heeft: ze veroorzaken geen overlast, geen

---

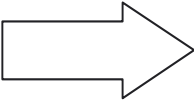

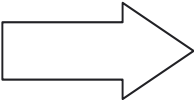
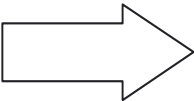
<sup>66</sup> Dit is gebleken uit lesmateriaal wat is ingezien en interviewgesprekken met de meiden zelf.

criminaliteit, ze beheersen de Nederlandse taal uitstekend en de meerderheid maakt de opleiding af. Kortom, dat er überhaupt een speciale activiteit wordt georganiseerd voor tweede generatie moslima's, is al vrij uniek te noemen.

De training die de meiden kregen aangeboden, bestond uit drie onderdelen. Er werd gestart met de training Turning Point. In circa 16 bijeenkomsten en met behulp van individuele coachingsgesprekken was het doel om bij de meiden een proces van bewustwording en reflectie te stimuleren, een positief zelfbeeld op te bouwen en zelfvertrouwen te verwerven, het vermogen te versterken om eigen gevoelens en gedachten onder woorden te brengen, het verwerven van inzicht in het eigen doen en laten in relatie tot dat van de ander te bevorderen en het probleemoplossend vermogen te verhogen. Deze cursus is – met succes – gegeven aan migrantenvrouwen van de eerste generatie. Tot op heden is een dergelijke cursus nog nooit aan vrouwen van de tweede generatie aangeboden. Vervolgens kregen de meiden een cursus “Morele oordeelsvorming” aangeboden om te leren niet alleen religieus, maar ook moreel te oordelen. Om het zelfstandige en morele denken te begeleiden, hebben ze een stappenplan aangereikt gekregen als hulpmiddel bij het analyseren van morele beslissingen en leerden ze het stappenplan toepassen op zelf ingebrachte praktijkvoorbeelden. De trainers bieden deze cursus normaal gesproken aan ambtenaren. De cursus is nog nooit aan deze doelgroep aangeboden. Gezien deze cursus aan orthodoxe moslimmeiden werd gegeven, is de cursus op enkele punten aangepast. Er is een les gegeven over moderniteit en hoe de Nederlandse westerse samenleving een handelingsperspectief heeft ontwikkeld –vrij van religie– in termen van vrijheid, rijkdom en geluk. Tevens is op de behoefte van de meiden ingespeeld om vragen met betrekking tot religie in te bedden in de training, de zogenaamde religieuze vorming. Tot slot hebben de meiden ook een training conflicthantering gevolgd. De doelen van deze training waren om de meiden inzicht te laten krijgen in competenties die behoren bij morele opvoeding en religieuze opvoeding, zicht krijgen op conflict, geweld en agressie, omgaan met elkaar in situaties van conflict en stress, emoties, kracht en interventies in een conflictsituatie en alternatieven voor geweld, bijdragen aan een vredescultuur.

Nu de inhoud van de training helder is, kan men teruggaan naar de tweede vraag van de Peace and Conflict Assessment: is de interventie doelmatig? Met andere woorden is de interventie effectief? Om deze vraag te behandelen, biedt de derde geschetste methode een uitkomst: Outcome mapping. De hypothesen over de output en outcome van de training zijn verwoord in figuur 6.9.

**Figuur 6.9: Hypotheses t.a.v. het effect**

	<b>Doelen van de training (micro-niveau)</b>	<b>Strategische doelen van de training c.q maatschappelijk effect (macro-niveau)</b>	
<b>Turning point</b> 	<ul style="list-style-type: none"> <li>○ Bewustzijn van het dragen van een persoonlijke geschiedenis.</li> <li>○ Acceptatie van eigen identiteit.</li> <li>○ In staat zijn om zich vrij te maken van knellende banden.</li> <li>○ Zelfbewust kiezen</li> <li>○ Communicatie.</li> </ul>	<ul style="list-style-type: none"> <li>○ Zichtbaar worden als een individu.</li> <li>○ Gemotiveerd participeren in de samenleving, waarbij het zelfopgeworpen isolement wordt doorbroken.</li> <li>○ Belangstelling voor de ander en het andere: luisteren naar de ander en in gesprek gaan met de ander.</li> </ul>	 <div data-bbox="892 820 1082 897" style="border: 1px solid black; padding: 5px; display: inline-block;"> <b>Het voorkomen van radicalisering</b> </div>
<b>Morele oordeelsvorming</b> 	<ul style="list-style-type: none"> <li>○ Het bewustzijn van de criteria waaraan je morele vragen toetst om tot een moreel oordeel te komen.</li> <li>○ Onafhankelijk vrij van (religieus) gezag een mening durven en kunnen formuleren.</li> </ul>	<ul style="list-style-type: none"> <li>○ Isolement verder doorbreken.</li> <li>○ Open voor, en deelname aan dialoog.</li> </ul>	
<b>Conflicthantering &amp; Weerbaarheid</b> 	<ul style="list-style-type: none"> <li>○ Leren om met eigen angst om te gaan.</li> <li>○ Het ontwikkelen van lef en moed om jezelf te zijn.</li> <li>○ Leren omgaan met conflict.</li> </ul>	<ul style="list-style-type: none"> <li>○ Meer zelfstandigheid, weerbaarheid en grenzen durven stellen t.a.v. omgeving.</li> <li>○ Conflicten niet uit de weg gaan.</li> </ul>	

Bij een dergelijk project is het overduidelijk dat de gevolgen van het project niet door een opinieonderzoek van meningen in het stadsdeel getoetst kunnen worden, maar dat de gevolgen alleen in de directe omgeving van de meiden zelf onderzocht kunnen worden.

Volgens de “Outcome Mapping”-methode zouden dan bijv. de volgende vragen kunnen worden gesteld:

- Hebben de deelnemers andere interesses en activiteiten ontwikkeld?
- Zijn zij nieuwsgierig geworden naar de inhoud van andere opvattingen; hebben zij ook lezingen bezocht van sprekers die andere opvattingen aanhangen?
- Nemen zij deel aan andere culturele of politieke bijeenkomsten?
- Gaan de deelnemers vaker om met andere mensen?
- Hoe heeft het contact met andere familieleden zich ontwikkeld?
- Spreken zij open over hun ontmoetingen en opvattingen?
- Laten zij ander gedrag zien tegenover leraren, medescholieren, werkgevers, collega's?
- Worden zij iets losser t.a.v. hun kledings- en gedragscodes?

Daarnaast zijn uit de Outcome Mapping aanpak ook suggesties af te leiden *aan wie* de evaluatievragen gesteld zouden kunnen worden. Volgens deze methode zouden dit personen moeten zijn die zo dicht mogelijk bij de doelgroep staan. Nemen we het trainingsproject als uitgangspunt, dan gaat het bijvoorbeeld om:

- de deelnemers aan de training,
- de trainers,
- de ouders van de deelnemers en andere directe familieleden (broers, zusters, neven)
- vrienden uit de kring van de deelnemers,
- docenten van lezingen die de deelnemers hebben bijgewoond,
- andere deelnemers aan dezelfde lezingen,
- leraren op de scholen van de deelnemers,
- leerplichtambtenaren,
- werkgevers en collega's van de deelnemers,
- burens van de deelnemers,
- winkeliers, bibliotheekpersoneel.

Omdat deze opdracht werd uitgevoerd in opdracht van Stadsdeel Slotervaart en de Gemeente Amsterdam, was er slechts een beperkt budget voor de evaluatie beschikbaar. Het was derhalve onmogelijk om ook iedereen in de omgeving van deze meiden te interviewen. Dat vereiste een creatieve oplossing. Voor de evaluatie zijn zowel de meiden als de trainsters geïnterviewd. De meiden werd gevraagd te reflecteren op de verandering die ze zelf en hun omgeving hadden ervaren sinds de training. Daarnaast is de trainers gevraagd te reflecteren op het proces dat de meiden van begin tot het einde van de deelname aan de training hadden doorlopen. Als aanvullende methode is gebruik gemaakt van participerende observatie om zelf te kunnen toetsen hoe de meiden gedurende de training veranderden. Vanwege de lengte van deze bijdrage is het niet mogelijk in te gaan op het daadwerkelijke effect van de training, maar met behulp van bovenstaande vragen en methoden is gebleken dat de training zowel op het individuele alsmede op het groepsniveau en contextuele niveau, effect heeft gehad.

Er is in de lijn van het LCPP-project tevens ook gekeken naar de onbedoelde neveneffecten. De meiden waren voor aanvang van de training zeer timide en zeer

gevoelig voor gezag. Zij pasten zich aan, aan wat cultureel en religieus als juist en wenselijk werd geacht. In het bijzonder door de training Turning Point –waarin deze meiden hun eigen identiteit leerden te bespreken– was het effect dat deze timide dames zich ontpopten tot mondige vrouwen. Niet iedereen in hun omgeving was daarvan even gecharmeerd, in het bijzonder de ouders van deze meiden. Sommige ouders dreigden de meiden niet meer naar de training te laten gaan, omdat ze daar “feministisch” werden gemaakt. Het onbedoelde effect van de training was dat er thuis soms heftige conflictsituaties ontstonden.

Tot slot wordt daarmee automatisch de vierde vraag van het Peace and Conflict Assessment en het schema van Reflecting on Peace Practices relevant: de context, binnen welk programma is deze activiteit ingebed? Nu was deze activiteit wel ingebed binnen een breder programma, maar binnen dat programma wordt niet dezelfde doelgroep bereikt. Idealitair hadden de ouders van deze meiden de opvoedingsondersteuningscursus moeten volgen van het actieprogramma. De leraren van deze meiden hadden een cursus moeten volgen om mogelijke radicalisering te signaleren. Op dit punt valt dus nog veel winst te behalen, want alleen op deze wijze kan een maximaal effect worden bereikt met de deelactiviteiten.

## 7. Conclusie

In dit hoofdstuk is bekeken in hoeverre evaluatiemethoden die voor “peace building” projecten in conflictgebieden in het buitenland werden ontwikkeld, ideeën kunnen aanreiken voor de evaluatie van projecten in Nederland die op het tegengaan van radicalisering gericht zijn.

Er zijn drie benaderingen gepresenteerd (het Local Capacities for Peace Project (en het vervolgproject Reflecting on Peace Practice), Peace and Conflict Assessment and Outcome mapping). Van alle drie benaderingen bleken aspecten bruikbaar voor de planning van evaluaties in Nederland, concreet in Amsterdam Slotervaart waar recentelijk met de uitvoering van het *Actieplan Slotervaart: het tegengaan van radicalisering* werd begonnen.

Allereerst reiken deze benaderingen een kader aan dat het vergemakkelijkt afzonderlijke deelprojecten in kaart te brengen, hun onderlinge verhouding te schetsen en attent te maken op lacunes in een programmatische aanpak. Verder brengen zij in herinnering dat het niet alleen noodzakelijk is om te toetsen in hoeverre een project ook daadwerkelijk de verwachte positieve resultaten oplevert, maar dat ook expliciet gevraagd moet worden of er ook onverwachte negatieve neveneffecten optreden. Ten slotte vallen er een aantal vragen uit af te leiden, die ingaan op de relevantie van de te evalueren projecten. Daarbij gaat het om vragen naar de juiste doelgroepen, de doelmatigheid van de projecten, de lessen die uit eerdere projecten werden getrokken en de inbedding van afzonderlijke projecten in een omvattend programma.

Deze ideeën zijn nog vrij algemeen en liggen ook aan veel evaluaties ten grondslag, die nooit naar de genoemde methoden hebben gekeken. Dat neemt niet weg dat de

drie benaderingen een bron van inspiratie kunnen vormen voor concrete vragen die je aan projecten in Nederland zou willen stellen en dat zij helpen bij het zoeken naar indicatoren, waarmee de resultaten van projecten op het gebied van antiradicaliseringsbeleid getoetst kunnen worden.

Wij hebben in dit hoofdstuk slechts naar de publicaties gekeken, waarin een algemene beschrijving van de gehanteerde evaluatiemethoden werd gepresenteerd. Een gedetailleerde analyse van het grote aantal evaluaties, die op basis van deze methoden hebben plaatsgevonden, zou ongetwijfeld een nog aanzienlijk grotere rijkdom aan ideeën en indicatoren opleveren.

## Bibliografie

AIVD (2006). *De gewelddadige jihad in Nederland. Actuele trends in de islamitisch-terroristische dreiging*. Den Haag: Algemene Inlichtingen- en Veiligheidsdienst. <https://www.aivd.nl/actueel-publicaties/aivd-publicaties>, geraadpleegd op 23 april 2007.

Anderson, M.B. (1999). *Do No Harm. How Aid Can Support Peace – or War*. Boulder/London: Lynne Rienner.

Barakat, Sultan e.a. (2006). *Evaluatierapport Thematisch Medefinancieringsprogramma (TMF), Deelstudie (A) Vredesopbouw*. Amsterdam: Triple L.

Buijs, F.J., Demant, F. & Hamdy, A. (2006a). *Strijders van eigen bodem. Radicale en Democratische moslims in Nederland*. Amsterdam: Amsterdam University Press.

COT Instituut voor Veiligheids- en Crisismanagement (2007). *Ondersteuning bij de lokale aanpak van polarisatie en radicalisering*. De gemeentelijke behoefte(n) verkend, door M. Zannoni & Van der Varst, L., 29 oktober.

Gemeente Amsterdam (2004). *Actieplan Wij Amsterdammers*.

Gemeente Amsterdam (2006). *Wij Amsterdammers II. Investeren in mensen en grenzen*. Een vervolg op Wij Amsterdammers.

Gemeente Amsterdam (2007). *Amsterdam tegen radicalisering*, Redactieteam: PAS, IHH en COT (Zannoni, M.). <http://www.hetccv.nl/binaries/ccv/dossiers/samenleven-en-wonen/radicalisering/amsterdam-tegen-radicalisering-1-.pdf>, geraadpleegd op 15 november 2007.

Gielen, A.J. (2008). *Radicalisering en Identiteit. Radicale rechtse en moslimjongeren vergeleken*. Amsterdam: Aksant.

Glaser, S. & Pfeiffer, T. red. (2007). *Erlebniswelt Rechtsextremismus. Menschenverachtung mit Unterhaltungswert. Hintergründe, Methoden, Praxis der Prävention*, Schwalbach/Ts: Wochenschau Verlag.

Guba, E.G. & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park: Sage Publications.

Paffenholz, T. & Reyhler, L. (2007). *Aid for Peace. A Guide to Planning and Evaluation for Conflict Zones*. Baden-Baden: Nomos.

Patton, M.Q. (1997). *Utilization-Focused Evaluation. The Third Century Text*. Beverly Hills, CA: Sage.



Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.

Sen, A. (2006). *Identity and Violence. The Illusion of Destiny*. London: Allen Lane.

Slotman, M. & Tillie, J. (2006). *Processen van Radicalisering. Waarom sommige Amsterdamse moslims radicaal worden*. Amsterdam: Instituut voor Migratie- en Etnische Studies, Universiteit van Amsterdam. <http://www.imes.uva.nl/staff/documents/061010DefinitieveversierapportPOMAmsterdam.pdf>, geraadpleegd op 18 maart 2007.

Stadsdeel Slotervaart (2007). *Actieplan Slotervaart: het tegengaan van radicalisering*. Amsterdam: Stadsdeel Slotervaart. <http://www.slotervaart.amsterdam.nl/contents/pages/00007436/actieplanslotervaarttegengaanvanradicalisering.pdf>, geraadpleegd op 2 maart 2007.

Veldhuis, T. & Bakker, E. (2007). Causale factoren van radicalisering en hun onderlinge samenhang. In: *Vrede en Veiligheid*, 36 (4), 447-469.

Welle, I. van der (2007). Amsterdamse identiteitspolitiek en (probleem)jongeren. In: *Agora* 23 (1), 8-11.

Withuis, J. (2007). Sekte en Sekse. Moslima-terrorisme geanalyseerd en vergeleken. In: *Radicaliserende vrouwen. Nederlandse communistische vrouwen (1945-1970) en Nederlandse Islamitische vrouwen (1989-heden), een comparatieve analyse*. Den Haag: Bureau Strategische Kennisontwikkeling (SKO). [http://www.hetccv.nl/dossiers/Samenleven\\_en\\_wonen/Radicalisering/rijksbeleid/Landelijk---Radicaliserende-vrouwen.html](http://www.hetccv.nl/dossiers/Samenleven_en_wonen/Radicalisering/rijksbeleid/Landelijk---Radicaliserende-vrouwen.html), geraadpleegd op 21 juli 2007.

# Hoofdstuk 7: Doelmatige publieke diensten in een democratische samenleving: een geïntegreerde benadering en praktische evaluatie-instrumenten

Walter Van Dongen

Studiedienst van de Vlaamse Regering (SVR)

<b>Inleiding</b> .....	<b>153</b>
<b>1. Dienstverlening in de dagelijkse leefsituatie</b> .....	<b>154</b>
1.1. De leefsituatie en levensloop van mannen en vrouwen .....	154
1.2. Tijdsverdeling binnen het marktsysteem .....	155
<b>2. Vraag en aanbod van publieke diensten in het marktsysteem</b> .....	<b>157</b>
2.1. Actoren en aspecten van de marktwerking .....	157
2.2. Positie van de aanbieders en de ontvangers .....	158
2.3. Mate van gelijke beschikbaarheid en collectieve financiering .....	161
<b>3. Doelmatigheid van publieke diensten</b> .....	<b>162</b>
3.1. Betekenis .....	162
3.2. Meten van effectiviteit en efficiëntie .....	164
3.3. Ontwikkeling van evaluatie-instrumenten voor publieke diensten .....	164
<b>4. Toepassing: evolutie van de dagopvang voor jonge kinderen in het Vlaamse Gewest</b> .....	<b>167</b>
4.1. Beschikbaarheid .....	167
4.2. Participatie .....	168
4.3. Omvang van het gebruik .....	169
4.4. Bezetting .....	170
4.5. Productiviteit .....	170
4.6. Kostenefficiëntie .....	171
<b>5. Besluit</b> .....	<b>174</b>
<b>Bibliografie</b> .....	<b>175</b>

## Inleiding

In elke maatschappij worden allerlei soorten diensten aangeboden en/of gefinancierd door de diverse overheden, die verweven zijn in het omvattende marktsysteem. In dit artikel gaan we uit van een geïntegreerde benadering van de diverse soorten maatschappelijke diensten die worden geproduceerd en aangeboden, en als onderdeel daarvan de diensten van de overheden en sterk gesubsidieerde organisaties, hier kort “publieke diensten” genoemd. Overheden en gesubsidieerde organisaties worden beschouwd als volwaardige, productieve actoren in het marktsysteem die waardevolle publieke diensten voortbrengen en aanbieden aan diverse andere actoren, uitgaande van de wisselwerking tussen de vraagzijde en aanbodzijde. Bijgevolg is de discussie over de doelmatigheid van de publieke diensten even belangrijk als voor de andere actoren in de maatschappij. De centrale vraag daarbij is in welke mate de diverse diensten doelmatig of efficiënt worden geproduceerd en aangeboden “op de markt”, waarbij de noden van de diverse “vragers” naar deze diensten voldoende worden beantwoord en rekening wordt gehouden met de specifieke waarden en doelstellingen van de publieke diensten. De publieke diensten moeten op eenzelfde positieve maar kritische en realistische manier worden bekeken en geëvalueerd als de niet- of minder publieke diensten. De doelmatigheid van publieke diensten verhogen is mogelijk indien voldoende rekening wordt gehouden met hun intrinsieke mogelijkheden en indien constructief wordt gezocht naar instrumenten om de doelmatigheid in kaart te brengen en te verbeteren.

In dit artikel willen we een aanzet geven om de effectiviteit en efficiëntie van publieke diensten te bekijken vanuit een microperspectief, met het oog op een correcte evaluatie en verbetering van de dagelijkse productieprocessen in de organisaties zelf, en als basis voor een evaluatie op mesoniveau (bijv. sectoraal) en op macroniveau (bijv. Vlaamse Gemeenschap, België, Benelux, Europa, ...).

In deze bijdrage presenteren we een algemene benadering voor de diverse soorten dienstverlening in de maatschappij. Voor mee toelichting bij de achtergrond en ontwikkeling van deze benadering verwijzen we naar Van Dongen (2006a, 2008a, 2008b). Centraal daarbij staat de bepaling van een aantal basisindicatoren met betrekking tot de doelmatigheid van de dienstverlening in een democratische samenleving. In een eerste sectie wordt een brede omschrijving gegeven van het concept dienstverlening binnen een algemene benadering van de tijdsverdeling van mannen en vrouwen binnen de diverse leefvormen (gezinnen, families, buurten, verenigingen, bedrijven, overheidsdiensten enz.). Vervolgens wordt een overzicht gegeven van de relevante aspecten van de arbeidsverdeling of vraag-aanbodssituatie. Daarna geven we aan wat wordt verstaan onder het begrip efficiëntie of doelmatigheid. Op basis daarvan kan een aantal basisindicatoren worden uitgewerkt voor de doelmatigheid van de diverse diensten: beschikbaarheid, participatie, benutting, bezetting, productiviteit, kosten en baten en gelijke verdeling. Hierbij wordt gepleit voor de ontwikkeling van praktische evaluatie-instrumenten voor de registratie, analyse en rapportering van de indicatoren op micro-, meso- en macroniveau, als een (minimale) kwantitatieve basis voor een brede beleidsevaluatie. Ten slotte wordt een concrete toepassing toegelicht met betrekking tot de evolutie van de dagopvang van kinderen jonger dan drie jaar in het Vlaamse Gewest in de periode 1990-2002.

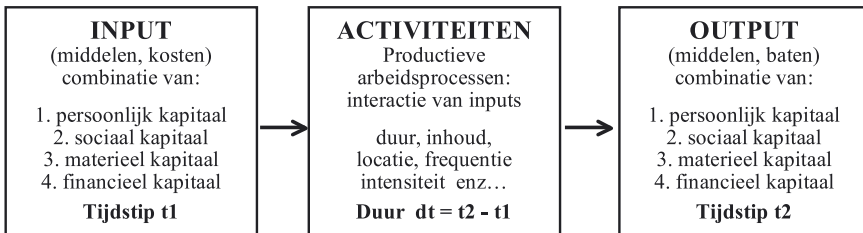
# 1. Dienstverlening in de dagelijkse leefsituatie

## 1.1. De leefsituatie en levensloop van mannen en vrouwen

De *leefsituatie* van mannen en vrouwen wordt in onze benadering omschreven als de dagelijkse verdeling van de diverse activiteiten (de tijdsverdeling) en van de resultaten ervan. De *levensloop* is de omvattende term voor de ontwikkeling of verandering van die leefsituatie of tijdsverdeling in de tijd. Al de activiteiten van individuen worden uitgevoerd binnen de diverse soorten leefvormen of organisaties van de maatschappij: gezinnen, families, buurten, verenigingen, clubs, organisaties, bedrijven, overheidsinstanties enz. (Van Dongen, 2008a, 2008b). De *activiteiten* van de actoren vormen het “dynamische draagvlak” van hun leefsituatie en levensloop. *Alle* activiteiten van *alle* subjecten (individuen, gezinnen, bedrijven, organisaties enz.) worden beschouwd als productieve arbeidsprocessen die worden gereguleerd via het algemene mechanisme van menselijke interactie of ruil, of van vraag en aanbod. Elke activiteit levert een specifieke output op die zelf een input is voor andere activiteiten (figuur 7.1). De combinatie van basisactiviteiten moet de gewenste samenstelling van persoonlijke, sociale, materiële en financiële middelen of opbrengsten opleveren, in alle fasen van de levensloop. Alle activiteiten zijn dus principieel even belangrijk in de dagelijkse tijdsverdeling. De specifieke inhoud, vorm, duur et cetera bepaalt de concrete functie, impact en waarde ervan. Het is van belang dat alle activiteiten zo doelmatig mogelijk verlopen opdat ze de best mogelijke output realiseren.

De activiteiten worden inhoudelijk gezien ingedeeld in enkele basiscategorieën: betaalde arbeid of beroepsarbeid, sociale arbeid, externe opvoeding of vorming, gezinsarbeid, persoonlijke verzorging en vrije tijd. De betekenis of waarde van de activiteiten wordt bepaald door de specifieke combinatie van middelen die ze opleveren. *Beroepsarbeid* biedt de meeste monetaire middelen voor de aankoop van inpulementen van de andere activiteiten, maar ook de persoonlijke, sociale en materiële opbrengsten zijn belangrijk. *Gezinsarbeid* levert directe diensten *binnen* het gezin die indirect een output leveren voor andere activiteiten. De arbeid wordt binnen het gezin zelf “vergoed”. *Sociale arbeid* betreft activiteiten ten behoeve van personen of organisaties buiten het gezin, waarbij de persoon geen of maar weinig monetaire opbrengsten nastreeft, maar vooral persoonlijke en sociale opbrengsten. Een analoge redenering geldt voor externe opvoeding, vrije tijd en persoonlijke verzorging.

**Figuur 7.1: Een algemeen model voor menselijke activiteiten**



## 1.2. Tijdsverdeling binnen het marktsysteem

De tijdsverdeling van mannen en vrouwen vindt altijd plaats binnen de context van of in interactie met de *maatschappelijke leefvormen*. Hiermee komen we tot de *interne en externe tijds- of arbeidsverdeling*: de wijze waarop de activiteiten en de resultaten ervan onder de maatschappelijke actoren worden verdeeld, in tijd en ruimte: het eigen gezin of andere gezinnen (familie, buurt, vrienden enz.), verenigingen, bedrijven en overheidsinstanties. Interne en externe arbeidsverdeling worden analytisch onderscheiden maar zijn permanent met elkaar verbonden. De *interne arbeidsverdeling* verwijst naar de verdeling van de activiteiten onder de diverse leden van een bepaalde actor of leefvorm (gezin, bedrijf, vereniging enz.) en noemen we de arbeidsorganisatie. Belangrijke aspecten daarbij zijn de wijze van besluitvorming binnen de leefvormen, de kwaliteit van de arbeidsprocessen, ruilprocessen en waardevorming. De *externe maatschappelijke arbeidsverdeling* verwijst naar de verdeling van activiteiten of de interacties tussen de diverse actoren of subjecten en noemen we de *marktwerking*, als het complement van de arbeidsorganisatie. In onze benadering betekent het begrip “markt” in de brede, neutrale zin “locus van de maatschappelijke ruil” van goederen en diensten. Marktactiviteiten brengen goederen en diensten voort die met andere externe subjecten worden geruild (markttransacties), waarbij de marktwaarde tot stand komt. De marktprijs is dan het monetaire gedeelte van de marktwaarde van het geruilde object. Dit impliceert dat alle (min of meer) autonome subjecten een bepaalde vorm van marktwerking hebben, dus ook de diverse overheidsinstanties en gesubsidieerde organisaties. Zo kan de marktwerking variëren van maximaal tot minimaal commercieel en van maximaal tot minimaal professioneel.

Het *marktsysteem* is dan het *complexe maatschappelijke netwerk van actoren* die als juridisch autonome instanties goederen en diensten uitwisselen, waarbij de diverse soorten marktwerking in een bepaalde verhouding voorkomen. De concrete vorm van een marktsysteem kan worden beschreven aan de hand van een aantal graduele dimensies: de keuze van kredietvormen, de mate van productsegmentatie, de openheid van de deelmarkten, de concurrentiedruk, de verhouding tussen de basiswaarden vrijheid, gelijkheid, solidariteit en efficiëntie, het aandeel van de overheid, het aandeel van sterk commerciële organisaties et cetera.

In elk marktsysteem (als maatschappelijk ruilsysteem) neemt de *overheid of het collectieve systeem* een specifieke positie in. Het (nationale) collectieve systeem (met al zijn deelsystemen) levert aan de “particuliere” subjecten diverse collectieve diensten. De financiering ervan geschiedt via een bepaalde combinatie van algemene fiscale middelen en directe betaling door de particuliere actoren. Het collectieve systeem, met aan de ontvangtzijde het fiscaal systeem (met inbegrip van de sociale zekerheid) en aan de uitgavenzijde de aangeboden publieke diensten, is een normaal en noodzakelijk onderdeel van elk marktsysteem dat in grote mate de werking van de andere actoren bepaalt. Het antwoord op de vraag welk aandeel de overheid moet hebben in het marktsysteem is afhankelijk van de visie op de maatschappij en van de vraag in welke mate de overheid doelmatig de collectieve diensten kan aanbieden aan de meer particuliere actoren, zonder afbreuk te doen aan de andere basiswaarden.

Op basis van deze aspecten kan men diverse marktsystemen onderscheiden, zowel theoretisch en feitelijk. Theoretisch gezien werd binnen de economische theorie vaak gebruik gemaakt van het extreme model van de perfecte mededinging, waarin alle actoren worden beschouwd als gelijke individuele spelers op een maximaal open

markt (Gravelle & Rees, 1987; Kreps, 1990). Organisaties en collectieve systemen worden in dit model als het ware weggedacht en gereduceerd tot individuele actoren. Dit extreme atomistische en individualistische model kan worden geplaatst tegenover een extreem collectivistisch model, waarbij de staat als centrale actor van de markt wordt gezien en de rol of impact van de particuliere actoren tot een minimum wordt herleid.

De feitelijke marktsystemen verschillen evenwel grondig van deze extreme theoretische modellen. Tijdens de negentiende eeuw was het *liberale vrijemarktsysteem* dominant in West-Europa en de Verenigde Staten. In een "vrij marktsysteem" (met diverse varianten) staat de waarde vrijheid bovenaan en zijn de andere waarden ondergeschikt. Dit betekent dat men bovenal het vrije initiatief van de actoren laat primeren en pas daarna aandacht heeft voor andere waarden. Bovendien krijgt de waarde "efficiëntie" een enge economische betekenis. Het is gekenmerkt door de *laissez-faire* staat, die weinig aandacht heeft voor de dagelijkse leefsituatie van mannen en vrouwen. In die periode waren de arbeidsmarkten weinig gereguleerd en de beroeps participatie was zeer hoog bij volwassen mannen en vrouwen en bij kinderen. Tegelijk waren de maatschappelijke voorzieningen voor sociale bescherming zeer beperkt. Voor gezinnen kwam dit systeem neer op een overlevingsmodel waarbij alle gezinsleden actief moesten bijdragen om het gezin voldoende overlevingskansen te bieden (Van Dongen, 1993, 2008a).

Tijdens de twintigste eeuw was in West-Europa het *sociaal gecorrigeerde vrijemarktsysteem of de traditionele verzorgingsstaat* dominant. Het systeem toonde een grotere bezorgdheid voor de dagelijkse leefsituatie van mannen en vrouwen binnen gezinnen, bedrijven et cetera. Het legde daarbij een sterke klemtoon op de regulatie van de arbeidsmarkten en op de bescherming van kinderen en vrouwen (gehuwde moeders). Dit komt neer op een lage beroeps participatie van (gehuwde) vrouwen en kinderen. Ter compensatie werden de maatschappelijke voorzieningen voor sociale bescherming sterk uitgebreid. Voor de gezinnen werd het kostwinnersmodel dominant, met gehuwde mannen in de rol van kostwinner en gehuwde vrouwen vooral in de rol van huisvrouw. Voor de bedrijven vertaalde dit model zich in de fordistische arbeidsorganisatie.

De voorbije decennia werd de traditionele welvaartsstaat gecontesteerd en zoekt men vanuit diverse hoeken naar een *meer rechtvaardig en doelmatig systeem*. Dit mogelijke nieuwe toekomstmodel kan men omschrijven als een *democratisch marktsysteem*. In een "democratisch marktsysteem" is de waarde vrijheid niet langer dominant, maar zijn de andere basiswaarden gelijkheid, solidariteit en efficiëntie evenwaardig. Iedere speler (individu, gezin, vereniging, bedrijven, overheidsinstantie, et cetera) moet voldoende vrijheid hebben om initiatieven te nemen en zijn bestaan of werking te organiseren, maar er moet tegelijkertijd op alle niveaus voldoende gelijkheid en solidariteit bestaan. Bovendien moeten alle spelers voldoende efficiënt zijn in hun activiteiten, in de ruime zin van het woord, om een maatschappelijke meerwaarde te realiseren, als basis voor een positieve ontwikkeling. Deze basiswaarden moeten elkaar begrenzen en ondersteunen. Een marktsysteem is in dat opzicht pas voldoende democratisch als er tegelijkertijd voldoende vrijheid, gelijkheid, solidariteit en efficiëntie bestaat (Van Dongen, 2008a). Een democratische maatschappij veronderstelt democratische overheidsinstanties, gezinnen, verenigingen en bedrijven. Een democratische overheid streeft voor al haar activiteiten en diensten naar een goed evenwicht tussen de basiswaarden vrije keuze, gelijkheid, solidariteit en doelmatigheid. Uiteraard zijn varianten mogelijk bij de concrete invulling ervan.

## 2. Vraag en aanbod van publieke diensten in het marktsysteem

Het debat over de maatschappelijke diensten is een onderdeel van het ruimere debat over de tijdsverdeling: de verdeling van de beroepsarbeid, de gezinsarbeid, de sociale arbeid, de vrije tijd en de persoonlijke verzorging. De verdeling van diensten zit verweven in de dagelijkse tijdsverdeling van alle mannen en vrouwen, binnen en buiten hun gezin. Of anders uitgedrukt, de verdeling van de diverse diensten is een onderdeel van het omvattende marktsysteem waarin personen en organisaties diverse soorten “diensten” en “goederen” aanbieden en waarin alle mannen en vrouwen diverse soorten “diensten” en “goederen” zoeken die ze nodig hebben.

De diverse soorten dienstverlening (activiteiten) kunnen worden uitgevoerd in de vorm van (volwaardig) betaalde beroepsarbeid, niet of weinig betaalde sociale arbeid (voor derden of verenigingen) of als onbetaalde gezinsarbeid. In principe kunnen alle soorten diensten op een professionele manier worden aangeboden. Feitelijk gezien worden diverse diensten aangeboden in de vorm van *beroepsarbeid*: bejaardenzorg, kinderopvang, gehandicaptenzorg, gezondheidszorg, onderwijs, sportbegeleiding enz. De dienstverlener (individueel en/of organisatie) ontvangt voor deze diensten een professionele vergoeding. Een aantal diensten wordt als *gezinsarbeid* uitgevoerd, waarbij de output (grotendeels) bedoeld is voor de gezinsleden: verzorging en opvoeding van kinderen, verzorging en begeleiding van andere inwonende personen, huishoudelijke taken en karweien in en om de woning. De diensten voor gezinsleden zijn “onbetaald”, in de zin dat er geen directe financiële ruil is tussen de partners op het ogenblik van de dienstverlening. Wel is er een algemene financiële ruil tussen de partners via de specifieke verdeling van de beroeps- en gezinsarbeid, het beroepsinkomen en de andere inkomensvormen, bijvoorbeeld vervangingsinkomens en fiscale tegemoetkomingen. Ten slotte worden diverse diensten als *sociale arbeid* gerealiseerd voor personen buiten het eigen gezin en verenigingen. Ze worden financieel niet vergoed (vrijwillige arbeid) of beperkt. Voorbeelden zijn het sociale werk voor familieleden, burens en vrienden en in/voor verenigingen (sport, sociale en artistieke activiteiten, milieu, jeugd, et cetera).

### 2.1. Actoren en aspecten van de marktwerking

De keuze om de diensten uit te voeren als beroepsarbeid, gezinsarbeid of sociale arbeid wordt bepaald door de wisselwerking tussen de vraagzijde en de aanbodzijde, d.w.z. tussen de intentie en oriëntatie van de aanbieders en de noden en wensen van de vragers/ontvangers. Om de verdeling van de “menselijke en materiële dienstverlening” weer te geven, kunnen modellen van de marktwerking worden ontwikkeld, waarin de diverse actoren en aspecten en de onderlinge wisselwerking worden opgenomen. Het uitgangspunt is dat elke persoon de drie soorten diensten (beroepsarbeid, sociale arbeid of gezinsarbeid) in zeker mate aanbiedt en deze tegelijk in zekere mate nodig heeft en gebruikt. Zowel aan de vraagzijde als aan de aanbodzijde worden de volgende aspecten onderscheiden:

- actoren: individuen, gezinnen, groepen, verenigingen, bedrijven, overheidsinstanties,
- vorm: beroepsarbeid, sociale arbeid of gezinsarbeid,
- inhoud: gezondheid (lichamelijk, psychisch, sociaal), onderwijs of opvoeding, huisvesting, mobiliteit, sport, kunsten, ontspanning, enz.,
- kwantiteit van de dienstverlening (aantal eenheden)
- kwaliteit van de dienstverlening (op basis van een aantal inhoudelijke



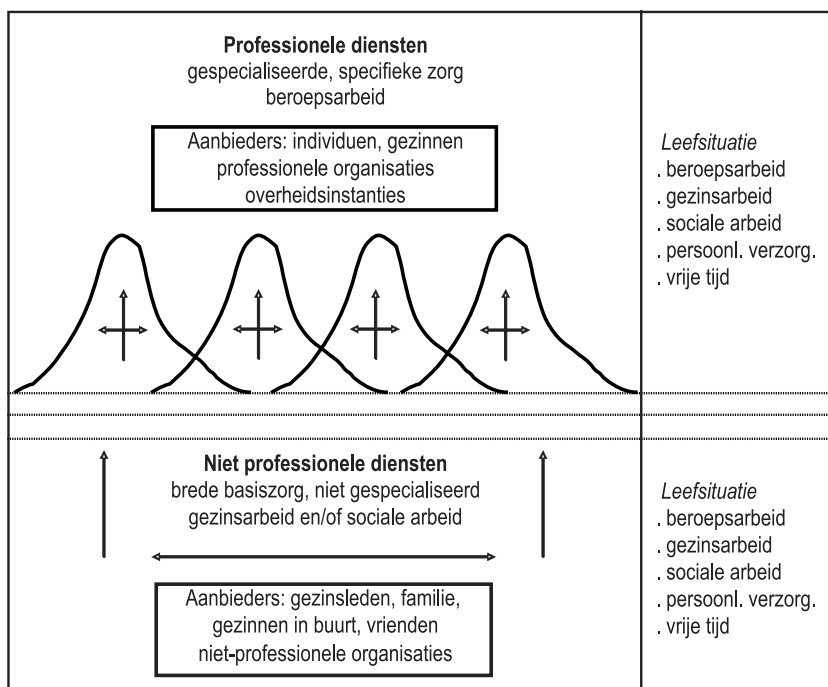
kenmerken).

Voor de wisselwerking tussen vraag en aanbod worden de volgende aspecten vermeld:

- kwantiteit of omvang van de diensten per dag, week, maand,
- graduele kwalificatie en specialisatie van de vraag en het aanbod,
- productiviteit: aantal outputeenheden van een bepaalde kwaliteit per VTE of per totale inpuiteenheid,
- normatieve keuze met betrekking tot gelijke beschikbaarheid, vrije keuze en solidariteit,
- soort financiering: gradatie van persoonlijke betaling, via graduele retributies, tot volledig collectieve financiering.

Een cruciale vraag in de brede evaluatie is of de dienstverlening doelmatig wordt aangeboden, waarbij de noden van de mensen met de gepaste dienstverlening worden beantwoord, uitgaande van een bepaalde verhouding van de basiswaarden (vrije keuze, gelijkheid en solidariteit) die men nastreeft en de doelstellingen die daaruit voortvloeien. Figuur 7.2 geeft ter illustratie een eenvoudig model weer voor de verdeling van de dienstverlening onder de mogelijke aanbieders.

**Figuur 7.2: Verdeling van de dienstverlening onder de diverse aanbieders**



## 2.2. Positie van de aanbieders en de ontvangers

Op basis van figuur 7.2 kan men diverse combinaties van dienstverlening uittekenen, uitgaande van de positie van de aanbieders en de ontvangers.



### 2.2.1. De situatie van de aanbieder

#### a) Aanbod als beroepsarbeid

De persoon heeft gekozen voor een beroep waarin dienstverlening centraal staat, uitgaande van een bepaalde opleiding en specialisatie. De persoon streeft daarbij onder meer naar een bepaald aantal uren per week, een bepaald beroepsinkomen, een bepaalde kwaliteit van het werk. Hierbij wordt een *professionele doelmatigheid* verwacht en nagestreefd, gegeven de kenmerken van het arbeidsproces: de aangeboden diensten moeten kunnen worden geleverd tegen een betaalbare “kostprijs” per eenheid (bijvoorbeeld per uur, per dag, enz.), uitgaande van “normale” volwaardige lonen voor de geleverde arbeid. Dit impliceert dat de aangeboden diensten in hoge mate zijn gespecialiseerd (verticale pijlen in bovenste deel van figuur 7.2) en bijgevolg gelimiteerd worden in tijd: met zo weinig mogelijk dure tijd en middelen zo veel mogelijk de gewenste kwaliteit leveren. Wel is het zinvol en wenselijk om voldoende samenwerking en uitwisseling (horizontale pijlen) te realiseren tussen de diverse soorten dienstverlening, ten behoeve van de ontvangers. Horizontale samenwerking kan zowel binnen de organisaties als tussen de organisaties worden tot stand gebracht (Bronselaer, 2003).

#### b) Aanbod als gezinsarbeid

De persoon besteedt naast de beroepsarbeid ook een bepaalde tijd aan diensten voor de eigen gezinsleden (vooral partner en kinderen). Elk gezinslid heeft een zeker engagement en een zekere verplichting ontwikkeld ten aanzien van de andere gezinsleden. De persoon is beschikbaar voor een bepaald aantal uren per week, afhankelijk van de tijd die nodig is voor andere activiteiten. Hij streeft hierbij vooral naar een voldoende “persoonlijke, sociale en materiële opbrengst” binnen de gezinscontext, afhankelijk van de soort taak, maar ook het financiële aspect is belangrijk, namelijk de mogelijkheid om geld te besparen. Voor deze diensten wordt bijgevolg voldoende *sociale doelmatigheid* verwacht en nagestreefd, gegeven de kenmerken van het arbeidsproces. Dit impliceert weinig gespecialiseerde basisdiensten (horizontale pijl in onderste deel van figuur 7.2) die vooral gelimiteerd worden door de beschikbare tijd en middelen en door de verwachte opbrengst: met de beschikbare tijd en middelen zo veel mogelijk de gewenste kwaliteit leveren. Toch is het belangrijk dat de persoon over de juiste competenties beschikt om de diverse taken behoorlijk uit te voeren. Naarmate men meer ervaring heeft met deze diensten, nemen de competenties toe en wordt de kloof met de professionele diensten kleiner (verticale pijlen in onderste deel van figuur 7.2).

#### c) Aanbod als vrijwillige of sociale arbeid

De persoon kan naast zijn beroepsarbeid en gezinsarbeid nog bepaalde diensten aanbieden aan personen buiten het eigen gezin, al dan niet in het kader van een organisatie. De persoon stelt zich beschikbaar voor een bepaald aantal uren per week, afhankelijk van de tijd besteed aan andere activiteiten. Hij streeft hierbij vooral naar een “persoonlijke, sociale en materiële opbrengst” en eventueel wordt een beperkte financiële vergoeding betaald, afhankelijk van de soort taak, de vereiste opleiding, de frequentie en duur enz. (bijvoorbeeld schoonmaken bij een familielid, een jeugdtrainer of bestuurder in een sportclub of een vrijwillige brandweerman). De aanbieder moet voldoende persoonlijke en sociale baten genieten om de activiteit

“(bijna) gratis” te blijven doen. Zodra men voor een taak heeft gekozen, heeft men een zekere verplichting ten aanzien van de ontvanger en de organisatie. Voor deze dienstverlening wordt eveneens voldoende “sociale doelmatigheid” verwacht en nagestreefd, gegeven de kenmerken van het arbeidsproces, zodat de aangeboden diensten kunnen worden geleverd tegen een betaalbare “kostprijs” per eenheid (bijvoorbeeld per uur, per dag, enz.). Bijgevolg zijn de aangeboden diensten minder gespecialiseerd dan professionele arbeid en relatief meer gelimiteerd door de beschikbare tijd en de verwachte persoonlijke en sociale baten dan door de financiële kosten.

#### *d) Het totale aanbod van alle aanbieders*

Elke aanbieder—elke (volwassen) persoon—realiseert een bepaalde combinatie van de drie soorten dienstverlening, uitgaande van de concrete invulling van de tijdsverdeling binnen de context van het eigen gezin en andere maatschappelijke organisaties. De feitelijke combinaties variëren sterk en worden tijdens de levensloop permanent aangepast (deels bewust en deels onbewust) op basis van de relatieve meerwaarde van de activiteiten voor de betrokken persoon en de geboden mogelijkheden in de omgeving. De complexiteit en dynamiek van vraag en aanbod impliceren dat er in diverse gradaties *substitutiemogelijkheden* zijn voor de diensten: beroepsarbeid vervangen door gezinsarbeid of sociale arbeid, gezinsarbeid door beroepsarbeid of sociale arbeid en sociale arbeid door beroepsarbeid of gezinsarbeid. Essentieel hierbij is dat de betekenis van de drie soorten arbeid binnen de tijdsverdeling van elke aanbieder wordt gerespecteerd, gekoppeld aan het evenwicht tussen de basiswaarden. Zo kan een dienst als beroepsarbeid niet zonder gevolgen vervangen worden door een dienst als gezinsarbeid of als sociale arbeid, en omgekeerd. Het is niet doelmatig mensen (meer) diensten als vrijwillige arbeid of gezinsarbeid te laten presteren, wanneer ze vooral diensten als beroepsarbeid willen aanbieden. Maar ook het omgekeerde is ondoelmatig. Belangrijk hierbij is het verschil in de directe kostprijs van een uur beroepsarbeid versus een uur gezinsarbeid of sociale arbeid en het ongelijke aandeel van de persoonlijke en sociale opbrengst. Bij de vervanging van beroepsarbeid door niet-betalde gezinsarbeid of sociale arbeid moet men rekening houden met de verschuiving van de persoonlijke, sociale, materiële en financiële inkomsten voor de betrokkenen en voor de overheid. Wanneer bijvoorbeeld iemand zijn beroepsarbeid opgeeft om voltijds een langdurig ziek gezinslid of familielid te verzorgen, verliest hij zijn inkomen, maar ook de persoonlijke en sociale opbrengsten van de betaalde baan. Tegelijk verliest de overheid de fiscale en sociale bijdragen. Bovendien moet worden gekeken naar de situatie van de zorgontvanger: welke diensten heeft die nodig, wat is de gewenste frequentie en duur, wie kan welke diensten het meest doelmatig aanbieden?

### **2.2.2. De situatie van de ontvanger**

De situatie van de *ontvangers* moet op een analoge manier worden voorgesteld, uitgaande van de betekenis van de dienstverlening in hun tijdsverdeling. Een “dienst ontvangen” is immers vervat in een bepaalde activiteit en impliceert dus ook een specifieke verhouding van persoonlijke, sociale, materiële en financiële opbrengsten.

- *Inhoud* van de diensten: gezondheid (lichamelijk, psychisch, sociaal), onderwijs

- of opvoeding, huisvesting, mobiliteit, sport, kunsten, ontspanning, et cetera.
- *Afhankelijkheid of behoefte*, uitgedrukt in termen van de omvang en kwaliteit van de gevraagde dienstverlening. Hierbij kan men rekening houden met factoren als leeftijd, geslacht, onderwijsniveau, nationaliteit, beroepssituatie en inkomen. Essentieel hierbij is of de afhankelijkheid verbonden is met gedeeltelijke of volledige beroepsinactiviteit.

### 2.2.3. De ruil tussen aanbieders en ontvangers

Uitgaande van deze factoren kan/moet men bepalen welke combinatie van beroepsarbeid, gezinsarbeid en sociale arbeid doelmatig is voor het aanbod van de gewenste dienstverlening. De *totale vraag van elke ontvanger* (elke persoon) is dus altijd een bepaalde combinatie van de drie soorten dienstverlening, uitgaande van de concrete invulling van de tijdsverdeling binnen de context van het eigen gezin en andere maatschappelijke leefvormen. Deze gevraagde combinatie is altijd aan verandering onderhevig.

Elke soort dienstverlening impliceert een maatschappelijk ruilproces tussen de aanbieder en de ontvanger, waar telkens een bepaalde combinatie van “persoonlijke, sociale, materiële en financiële” elementen worden uitgewisseld, uitgaande van hun initiële noden en mogelijkheden. De *ontvanger* wil de juiste “persoonlijke, sociale en materiële” elementen die zijn noden beantwoorden. Hiervoor wil hij een bepaalde “vergoeding” geven, in de vorm van een specifieke combinatie van “persoonlijke, sociale, materiële en financiële” middelen. De *aanbieder* wil de juiste “persoonlijke, sociale en materiële” diensten leveren die een antwoord zijn op de noden van de ontvanger. In ruil daarvoor wil hij zelf de juiste “persoonlijke, sociale, materiële en financiële” middelen als “vergoeding” ontvangen die zijn eigen noden beantwoorden. In termen van personen overlappen ontvangers en aanbieders elkaar in grote mate: alle ontvangers zijn tegelijkertijd in zekere mate aanbieders, en omgekeerd. Een marktwerking is doelmatig in de mate dat de totale vraag naar dienstverlening van alle ontvangers maximaal wordt beantwoord door het totale aanbod aan dienstverlening van alle aanbieders. Tegelijkertijd moet de totale gewenste opbrengst van alle aanbieders voldoende worden beantwoord door de totale “vergoeding” van alle ontvangers.

### 2.3. Mate van gelijke beschikbaarheid en collectieve financiering

Een cruciaal aspect van alle soorten maatschappelijke dienstverlening is de wijze van financiering, meer specifiek de mate van financiering via het collectieve systeem. Zoals gezegd, gaan we uit van een gradueel onderscheid tussen maximaal private goederen/diensten aan de ene kant en maximaal collectieve goederen/diensten aan de andere kant, waarbij vooral de diverse tussenvormen relevant zijn. Het is niet eenvoudig om te bepalen in welke mate een dienst collectief is of zou moeten zijn. Dit is deels een empirische en deels een normatieve vraag. We kunnen hiervoor drie formele criteria toepassen (Van Dongen, 2008a; Vandenbroucke, 2002). Goederen en diensten zijn meer collectief

- 1) naarmate de beschikbaarheid voor de doelgroep in grotere mate gelijk is, m.a.w. naarmate het gebruik door bepaalde actoren het gebruik voor andere leden van de doelgroep minder beperkt, of naarmate minder mensen van de doelgroep

(kunnen) worden uitgesloten;

- 2) naarmate ze minder deelbaar zijn vanuit het oogpunt van het gebruik, d.w.z. ze zijn voor de gebruiker maar zinvol als een complex geheel;
- 3) naarmate het maximale gebruik voor de doelgroep begrensd is zodat overconsumptie wordt geminimaliseerd.

De toepassing van deze criteria maakt het mogelijk een gradueel onderscheid te maken tussen maximaal collectieve diensten aan de ene zijde en maximaal particuliere diensten aan de andere zijde. Maximaal collectieve diensten kunnen dan volledig via algemene collectieve inkomsten worden gefinancierd en “gratis” aangeboden aan de doelgroep op het ogenblik van de aanschaf of het gebruik. Voor minder collectieve diensten is het doelmatiger een directe prijs aan te rekenen aan de gebruiker om de vraag en het aanbod op elkaar af te stemmen. Naarmate de nood aan gelijk gebruik binnen de doelgroep groter is dan het gevaar voor overconsumptie, kan men de prijzen gradueel gezien laag houden. We laten hier in het midden in welke mate de publieke financiële middelen meer naar de ontvanger dan naar de aanbieder moeten gaan.

Deze benadering kan men toepassen op de diverse soorten maatschappelijke diensten. Een eerste opdracht is voor de diverse soorten diensten empirisch na te gaan in welke mate ze als collectieve dienst worden aangeboden en in welke mate dit overeenstemt met de vermelde criteria. Uitgaande van de empirische studie kan de normatieve vraag worden gesteld: in welke mate moeten deze diensten als meer of minder “collectieve” diensten worden beschouwd en gefinancierd?

### 3. Doelmatigheid van publieke diensten

#### 3.1. Betekenis

Overheden en gesubsidieerde organisaties worden beschouwd als volwaardige, productieve actoren in het marktsysteem die waardevolle diensten voortbrengen en aanbieden aan diverse andere actoren, uitgaande van de wisselwerking tussen de vraagzijde en aanbodzijde. De publieke sector bestaat uit diverse deelsystemen die bepaalde diensten/voorzieningen “produceren” en “aanbieden” of “verkopen” aan diverse maatschappelijke actoren (individuen, gezinnen, verenigingen, bedrijven, andere overheidsinstanties). In ruil daarvoor ontvangen zij bepaalde financiële middelen of vergoedingen. De uitdaging is de juiste publieke diensten op een voldoende doelmatige en correcte manier te produceren en aan te rekenen aan de diverse (particuliere) actoren.

Voor de publieke diensten wordt vaak het MAPE-model gebruikt, dat een onderscheid maakt tussen de middelen (M) die worden ingezet tijdens de activiteiten (A) en leiden tot de producten of prestaties (P) met het oog op bepaalde effecten (E) in de maatschappelijke omgeving (Bouckaert & Thijs, 2003; Verlet en Devos in dit boek). Efficiëntie of doelmatigheid is in deze benadering dan de mate waarin met de ingezette middelen (input) een bepaald product (output) wordt bereikt. Efficiëntie of doelmatigheid geeft aan in welke mate het “uiteindelijk” doel (outcome) wordt gerealiseerd door het gebruik van de output.

Hoewel het genoemde MAPE-model vaak als internationale standaard wordt voorgesteld (bijvoorbeeld in "Government at a Glance van de OESO, OECD, 2007), zijn er ook kritische geluiden te horen over dit model (zie bijvoorbeeld in 't Veld, 2003; Van Dongen, 2008b). Meer specifiek heeft de kritiek betrekking op de niet consistente toepassing van de maatstaven op de korte en lange termijn waarin de output en outcome zijn te situeren. In deze alternatieve benadering heeft men eveneens aandacht voor de verwevenheid van de inputs en outputs van diverse actoren. De output van de ene organisatie is immers een input van een andere. Deze alternatieve benadering geeft een andere invulling aan de concepten "effectiviteit" en "efficiëntie". Effectiviteit is in het MAPE-model niet van toepassing op de realisatie van de eigenlijke output van het productieproces van de betreffende organisatie (korte termijn) maar alleen op de effecten van die output op lange termijn voor de maatschappij, d.w.z. voor de activiteiten van andere organisaties waar die output als input wordt ingezet. Efficiëntie wordt alleen toegepast op het eigenlijke productieproces van de organisatie op korte termijn en niet op de processen op langere termijn.

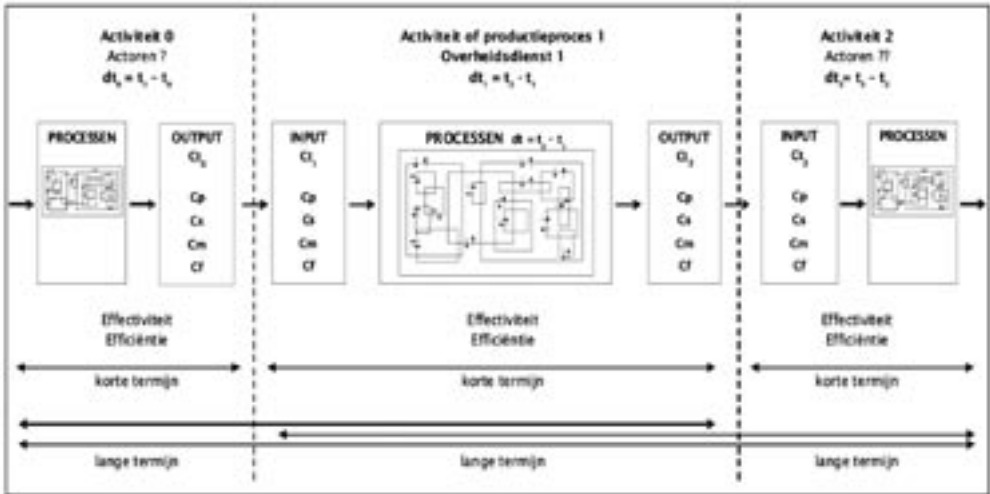
Figuur 7.3 geeft in het verlengde van figuur 7.1 een bredere *benadering* weer (in 't Veld 2002; Van Dongen, 2008b) waarin effectiviteit en efficiëntie zowel van toepassing zijn op de eigenlijke productieprocessen van de betrokken organisaties op korte termijn (activiteit 1 tussen de stippellijnen) als op de effecten van de output voor de maatschappij op langere termijn, rekening houdend met de persoonlijke, sociale, materiële en financiële aspecten van elke activiteit, aan de input- en de outputzijde.

*Effectiviteit of doeltreffendheid* geeft volgens deze benadering in de eerste plaats weer in welke mate de beoogde output van de activiteit is gerealiseerd met de ingezette middelen. *Efficiëntie of doelmatigheid* geeft weer in welke mate de (totale) waarde van de (beoogde) output van een activiteit hoger is dan die van de input, m.a.w. in welke mate een maatschappelijke meerwaarde wordt gerealiseerd. Normaal gezien probeert men processen zo te organiseren dat ze zowel effectief als efficiënt zijn. Bij de evaluatie ervan is effectiviteit op korte termijn (de geplande output realiseren) de eerste voorwaarde en efficiëntie (verhouding tussen de waarde van de input en de output) de volgende.

De activiteiten van een publieke dienst (activiteit 1, korte termijn) vormen een onderdeel van een ruimer maatschappelijk proces op langere termijn (tussen activiteit 0 en activiteit 2 in figuur 7.3) dat gestalte krijgt via de activiteiten van diverse andere maatschappelijke organisaties. In dat opzicht moet men ook voldoende aandacht besteden aan de effectiviteit en efficiëntie op langere termijn, ten aanzien van ruimere maatschappelijke doelstellingen. Deze evaluatie is per definitie moeilijker dan de evaluatie van het eigenlijke productieproces op korte termijn.

Bij de evaluatie van de activiteiten is de samenstelling van de componenten aan de input- en outputzijde afhankelijk van de normatieve keuzes van de actoren, uitgaande van de combinatie van een aantal basiswaarden (vrije keuze, gelijkheid, solidariteit enz.). Bijgevolg is elke maat voor efficiëntie beïnvloed door de normatieve visie van de actoren. Men kan dus een gradueel onderscheid maken tussen een zeer enge efficiëntiebenadering aan de ene kant (bijvoorbeeld alleen geldelijke winst als criterium) en een zeer brede benadering aan de andere kant (alle aspecten worden evenwaardig opgenomen in de evaluatie), met zeer veel tussenposities.

**Figuur 7.3: Breder model voor menselijke activiteiten en efficiëntie/ effectiviteit**



$C_t$  = beschikbaar kapitaal op tijdstip  $t$

$C_p$  = persoonlijk kapitaal;  $C_s$  = sociaal kapitaal;  $C_m$  = materiaal kapitaal;  $C_f$  = financieel kapitaal

### 3.2. Meten van effectiviteit en efficiëntie

Het meten van de effectiviteit en efficiëntie van publieke diensten vereist in de eerste plaats de correcte meting/registratie van de eigen productieprocessen binnen de afgebakende diensten: de elementen aan de input- en outputzijde en de kenmerken van de arbeidsprocessen van activiteit 1 tussen de stippellijnen in figuur 7.3. Deze interne evaluatie kan/moet een onderdeel zijn van een bredere maatschappelijke evaluatie, ex-post (met gegevens over de voorbije periode) en ex-ante (met ramingen voor de volgende periode). De uitdaging is een aantal adequate basisindicatoren samen te stellen van de doelmatigheid van publieke diensten op niveau van de organisatie, uitgaande van de beleidsdoelen die worden vooropgesteld met betrekking tot beschikbaarheid, gebruik, bezettingsgraad, productiviteit, kostenefficiëntie en aandeel van de persoonlijke bijdrage. Op basis daarvan kan men dan betrouwbare indicatoren berekenen op hoger niveau (sectoraal, Vlaamse gemeenschap, ...).

Om de doelmatigheid van de overheid correct te meten, is er dus nood aan adequate registratie van basiskennmerken van de dagelijkse activiteiten en ruilprocessen op microniveau, met aandacht voor de beperkingen en problemen in de publieke sector. De productieprocessen en ruilprocessen van publieke diensten zijn vaak collectiever en complexer en dus minder goed af te bakenen dan die van particuliere actoren. Vaak is de efficiëntiegraad van publieke diensten ruimer dan die van particuliere diensten, omdat men vanuit het sociale beleid de diensten aan een ruimere doelgroep wil/moet aanbieden. Niettemin treden er altijd ruilprocessen op tussen de betrokken



actoren waarbij bepaalde “pakketten” dienstverlening worden geruild tussen relatief autonome actoren. Bijgevolg is er altijd een zinvolle onderverdeling van actoren en activiteiten mogelijk en tevens een bepaalde mate van toewijzing van de kosten en baten. De uitdaging is deze toewijzing zo goed mogelijk te differentiëren, uitgaande van de beschikbare gegevens. Hoewel dit proces niet gemakkelijk is, kan men altijd stappen voorwaarts zetten via doelgerichte verbeteringen in de classificatie, registratie en verwerking van de gegevens, uitgaande van de kenmerken, beperkingen en mogelijkheden van elke organisatie.

Het beleid ten aanzien van de diverse publieke diensten gaat altijd uit van de normatieve vraag in welke mate de diverse soorten diensten (in kwantiteit en kwaliteit) beschikbaar moeten zijn voor de diverse doelgroepen, bijvoorbeeld openbaar bestuur, gezondheidszorg, sociale zorg, mobiliteit, sport, kunst, veiligheid en bescherming, woonomgeving, dagopvang, onderwijs, jeugdvereniging, vakantieopvang enz.? De doelstellingen die hieruit voortvloeien, bepalen grotendeels de hanteerbare en gehanteerde efficiëntiemaat. De doelmatigheid van de diensten kan dus in algemene zin worden gemeten aan de hand van enkele basisindicatoren, uitgaande van de kwalitatieve outputeenheden die worden afgebakend en gerealiseerd, dat kan dienen als een als een kwantitatief raamwerk voor de brede beleidsevaluatie:

- beschikbaarheid: aantal eenheden per persoon van de doelgroep,
- participatie: aandeel gebruikers in de doelgroep,
- omvang gebruik: aantal outputeenheden per gebruiker,
- bezetting: aandeel van gebruikte eenheden ten aanzien van de beschikbare eenheden,
- productiviteit: aantal outputeenheden per VTE, per inpuiteenheid,
- gelijke verdeling: vergoeding per VTE, aandeel van de bijdrage van de gebruikers,
- financiële efficiëntie: totale kosten en baten, kosten en baten per eenheid, per VTE.

### **3.3. Ontwikkeling van evaluatie-instrumenten voor publieke diensten**

De basisgegevens voor de doelmatigheidsindicatoren zitten vervat in het dagelijkse productieproces van de diverse dienstverlenende instanties. Het is dan ook belangrijk dat deze gegevens zo doelmatig mogelijk worden verzameld en verwerkt, zodat ze bruikbaar zijn voor het beheer binnen de afzonderlijke organisaties (microniveau), op sectoraal niveau en op het niveau van de Vlaamse Gemeenschap. Dit impliceert dat men met een minimum aan eenmalig geregistreerde gegevens de basisindicatoren kan bepalen voor alle beleidsniveaus.

Hiervoor zijn drie soorten instrumenten nodig die in het commerciële bedrijfsleven al vrij algemeen worden gebruikt, zij het in diverse vormen en gradaties en met een wisselende doelmatigheid. Uitgaande van de evolutie van de rol en werking van publieke diensten in het marktsysteem, veronderstellen we hier dat deze instrumenten in de publieke diensten minder worden gebruikt en op een minder doelmatige manier. De uitdaging voor de komende jaren is dan een aantal verbeteringen te verwezenlijken via een constructief leer- en verbeterproces. Wel dient men hierbij rekening te houden met de eigenheid van de publieke sector.

### **3.3.1. Doelmatige analytische boekhouding**

Het is belangrijk dat alle instanties die professionele diensten aanbieden een doelmatige analytische boekhouding voeren, die het mogelijk maakt de bedrijfseconomische indicatoren te bepalen. Essentieel is dat een aangepast en voldoende uniform boekhoudkundig stelsel (hoofdrubrieken en subrubrieken van uitgaven en inkomsten) wordt gebruikt binnen en tussen de diverse sectoren van de dienstverlening. Tegelijk moet het stelsel soepel genoeg zijn om de verschillen tussen en binnen de sectoren op te vangen. Zo zou men voor alle organisaties een boekhoudkundig stelsel kunnen uitwerken met uniforme hoofdrubrieken en subrubrieken van eerste orde. De subrubrieken van de tweede en derde orde kan men dan vrijlaten zodat de verschillen in het uitgaven en inkomstenpatroon kunnen worden opgevangen. Voor zover mogelijk kan men in het stelsel ook een onderscheid maken volgens de categorieën van cliënten en medewerkers. Zo kan men een meer gedifferentieerde analyse maken van de gegevens. Essentieel is ook dat er voor alle “bestuurders” een duidelijk overzicht beschikbaar is voor de interne beleidsdiscussie.

### **3.3.2. Instrument voor de registratie van bijkomende aspecten van het productieproces**

Daarnaast moeten een aantal aspecten van het productieproces in kaart worden gebracht die nodig zijn voor de basisindicatoren. Per sector en deelsector (onderwijs, dagopvang, ziekenhuizen, rusthuizen, opvangcentra voor gehandicapten, enz.) moet de differentiatie in de “kwalitatieve outputeenheden” worden bepaald aan de vragende en aanbodzijde. Op die manier krijgt men een meer gedifferentieerd beeld van de diverse outputeenheden. Per sector (en/of deelsector) van de dienstverlening is een voldoende uniform, praktisch registratie-instrument nodig, waarbij op individueel niveau de noodzakelijke gegevens over de ontvangers (cliënten), de aanbieders (medewerkers) en de dienstverlening worden verzameld. Essentieel is dat ook hier voldoende uniformiteit tussen de sectoren mogelijk is voor de hoofdkenmerken en voor vergelijkbare eenheden.

Voor de registratie zijn diverse instrumenten mogelijk, van eenvoudige steekkaarten tot gesofisticeerde digitale registratiesystemen via scanning of via het internet. Essentieel is wel dat de gegevens zo doelmatig mogelijk worden verzameld en opgeslagen in een dataset die bruikbaar is voor analyse en rapportage op diverse niveaus.

### **3.3.3. Instrument voor de berekening van de doelmatigheidsindicatoren**

Op basis van relevante gegevens van de boekhouding en de bijkomende registratie kan men de doelmatigheidsindicatoren berekenen. Het is mogelijk en zinvol om per sector en deelsector een uniform analyse-instrument (computerprogramma) te ontwikkelen voor de integratie en verwerking van de gegevens en voor de berekening van de indicatoren. De resultaten van de analyse worden vervolgens opgenomen in een duidelijk gestandaardiseerd rapport dat voor alle bestuurders beschikbaar is.

Via een geïntegreerd systeem van databeheer kan men per (deel)sector een algemeen bestand maken, als basis voor macroanalyses, met opsplitsing naar



bijvoorbeeld provincie, gemeente, geslacht en leeftijd. Zo kunnen de microgegevens doelmatig worden gebruikt voor het algemene onderzoek en voor het beleid van de Vlaamse regering.

### 3.3.4. Beleidsdiscussie over de wenselijke dienstverlening

De vergelijkende analyse van deze basisindicatoren binnen en tussen diverse sectoren en instanties, vormt een sterke empirische basis voor de *beleidsdiscussie*: in welke richting moeten deze indicatoren wijzigen en in welke mate, binnen het kader van de brede evaluatie van de dienstverlening? In het verlengde daarvan kan worden gezocht naar de instrumenten en middelen om de gewenste wijzigingen te realiseren.

Uitgaande van de feitelijke ontwikkeling van de dienstverlening en de normatieve benadering (bv. “democratische dienstverlening”), moet in de beleidsdiscussie in de eerste plaats de wenselijke *beschikbaarheid* van de diverse vormen van dienstverlening worden bepaald voor de doelgroep. In welke mate moet ieder persoon dezelfde omvang en kwaliteit kunnen genieten met betrekking tot gezondheidszorg, mobiliteit, sport, cultuur et cetera? In welke mate moet ieder kind een zelfde omvang en kwaliteit kunnen gebruiken van bv. dagopvang, basisonderwijs, secundair onderwijs, hoger onderwijs, jeugdverenigingen, vakantieopvang? In welke mate moet elke bejaarde een zelfde omvang en kwaliteit kunnen genieten met betrekking tot specifieke diensten? Vermelden we hier dat *effectieve* beschikbaarheid ook een voldoende fysieke *bereikbaarheid* van de dienstverlening impliceert, uitgedrukt in afstand en tijd. Dit vereist op zijn beurt voldoende informatie over de beschikbare dienstverlening. Vervolgens moet per soort dienstverlening het *gewenste aandeel van beroepsarbeid* worden bepaald, in relatie tot het aandeel van sociale arbeid en gezinsarbeid, uitgaande van de noden, mogelijkheden en wensen van ontvangers en aanbieders. Op basis daarvan kan men per soort dienstverlening de *beleidsdoelstellingen* bepalen voor de vastgelegde “ouputeenheden”. Zo kan de overheid de toekomstige uitgaven doelmatig begroten en elk jaar de feitelijke kosten en baten van de diverse diensten vergelijken met de gebudgetteerde kosten en baten.

### 3.3.5. Casestudies voor de ontwikkeling van evaluatie-instrumenten

De ontwikkeling en toepassing van evaluatie-instrumenten voor publieke diensten heeft als doel de publieke diensten stap voor stap vertrouwd te maken met een realiseerbaar “doelmatigheidsbeleid”, op maat van de (deel)sector en de afzonderlijke organisaties. Deze opdracht vereist veel inzet en volharding bij alle betrokken actoren.

In een eerste fase is het wenselijk enkele prototypes te ontwikkelen voor enkele (deel)sectoren, door middel van een aantal casestudies waarin men uitgaat van de beschikbare gegevens, indicatoren en instrumenten. Men kan hierbij uitgaan van voorbeelden uit het verleden (Bouckaert, 1992, 1993; Bogaert & De Prins, 1996; Probis-Vandorpe-Studies, 2002; Bronselaer, 2003; Van Dongen, 2004). De deelname van organisaties geschiedt op vrijwillige basis. Deelname aan een casestudie moet op zich al voldoende baten opleveren voor het beleid in de organisatie. Later kan dit instrument stap voor stap worden uitgebreid en geoptimaliseerd voor de betrokken sectoren en kan men aangepaste versies ontwikkelen voor andere sectoren. De ontwikkeling/toepassing van een evaluatie-instrument bestaat uit vier onderdelen

die volgens een stappenplan worden uitgevoerd. De *voorbereiding* van het proces moet ertoe leiden dat in de organisatie alles optimaal kan verlopen: contact en informatie, presentatie, doelstellingen, middelen, raming van kosten en baten, planning en overeenkomst. De *diagnose* van de dienstverlening en het huidige efficiëntiebeleid in de organisatie impliceert de verzameling, analyse en rapportage van de nodige gegevens. Het *managementproces* is gericht op een verbetering van het efficiëntiebeleid en de dienstverlening, uitgaande van de resultaten van de diagnose en een interne beleidsevaluatie. Doel is concrete actieplannen uit te werken die leiden tot een win-win situatie voor alle betrokken partijen: gebruikers, aanbieders en de overheid. Ten slotte is een meer algemene *beleidsgerichte analyse* mogelijk ten behoeve van de overheid, uitgaande van de integratie van de diverse casestudies.

## 4. Toepassing: evolutie van de dagopvang voor jonge kinderen in het Vlaamse Gewest

In deze sectie belichten we de ontwikkeling van de professionele dagopvang voor kinderen jonger dan drie jaar in het Vlaamse Gewest in de periode 1990-2002, als onderdeel van de totale opvoeding van jonge kinderen. De uiteenzetting is gebaseerd op een uitvoerige studie (Van Dongen, 2004) waarin op basis van de beschikbare gegevens van Kind en Gezin (1990-2004, Statistische Jaarverslagen) enkele basisindicatoren werden berekend, als empirische basis voor de normatieve vraag: hoe kan en moet de dagopvang in de toekomst verder worden ontwikkeld? We geven toelichting bij de volgende indicatoren:

- beschikbaarheid: aantal volledige plaatsen per 1000 kinderen,
- participatie: aantal gebruikers per 1000 kinderen,
- benutting: aantal volledige verblijfsdagen per gebruiker,
- bezetting: aantal volledige verblijfsdagen per plaats,
- productiviteit: aantal volledige verblijfsdagen per VTE,
- kostenefficiëntie: kosten per volledige plaats en volledige verblijfsdag.

Bij de interpretatie van de empirische resultaten moet men rekening houden met enkele veronderstellingen die nodig zijn wegens het gebrek aan gegevens over bepaalde aspecten. Algemeen gezien gaan we ervan uit dat de kwaliteit van de opvang/opvoeding in het eigen gezin, een opvanggezin of een dagverblijf gelijk is. De kwaliteit kan aan de hand van een aantal meetbare kenmerken worden bepaald. Voor deze toepassing waren de nodige gegevens daarvoor echter niet beschikbaar. Later kan worden nagegaan wat het effect is van een wijziging in deze veronderstelling. Een belangrijk doel van een evaluatie-instrument is dat (in de toekomst) de basisgegevens doelmatig worden geregistreerd, zodat dergelijke veronderstellingen overbodig worden.

### 4.1. Beschikbaarheid

Tabel 7.1 geeft de beschikbaarheid weer via het aantal volledige plaatsen per 1000 kinderen van drie maand tot drie jaar in de gesubsidieerde en niet-gesubsidieerde dagverblijven en opvanggezinnen. De aanduiding “% 90/02” bovenaan de tabel geeft de totale procentuele toe- of afname weer tijdens de periode 1990-2002.

**Tabel 7.1: Aantal volledige plaatsen per 1000 kinderen van drie maand tot drie jaar  
(Vlaams Gewest, 1990-2002)**

Opganginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
<b>Gesubsidieerd</b>								
Dagverblijven	55	45	48	51	58	62	66	70
Gezinnen	89	69	78	90	94	125	127	131
Totaal	76	115	127	141	152	187	193	201
<b>Niet gesubsidieerd</b>								
Dagverblijven	224	19	18	26	33	42	52	61
Gezinnen	21	30	34	38	39	39	37	36
Totaal	99	49	52	64	72	81	89	97
<b>Totaal</b>								
Dagverblijven	105	64	66	77	91	104	118	131
Gezinnen	68	99	112	128	133	163	164	167
Totaal	83	163	178	205	224	268	282	298

Het *aantal gesubsidieerde plaatsen per 1000 kinderen* is toegenomen van 115 in 1990 tot 201 in 2002, een stijging met 76%. Voor 20% van de kinderen was er in 2002 dus een gesubsidieerde “volledige” plaats beschikbaar. Het *totale aantal niet-gesubsidieerde plaatsen per 1000 kinderen* is toegenomen van 49 in 1990 tot 97 in 2002, een stijging met 99%. Voor bijna 10% van de kinderen was er in 2002 dus een niet-gesubsidieerde “volledige” plaats beschikbaar. Het *totale aantal plaatsen per 1000 kinderen* is gestegen van 163 in 1990 tot 298 in 2002 (83%). Voor bijna 30% van de kinderen was er in 2002 dus een volledige plaats beschikbaar. Deze evolutie illustreert het “meersporenbeleid” van de Vlaamse regering tijdens de periode 1990-2002, waarbij vooral werd gekozen voor opvangvormen die minder zwaar wogen op haar eigen budget.

#### 4.2. Participatie

Tabel 7.2 toont de participatie aan de hand van het aantal gebruikers per 1000 kinderen van drie maand tot drie jaar in de gesubsidieerde en niet-gesubsidieerde dagverblijven en opvanggezinnen. Het aantal gebruikers per 1000 kinderen jonger dan drie jaar in de *gesubsidieerde dagopvang* is toegenomen met 83%: van 237 gebruikers in 1990 tot 433 in 2002. De stijging was veel beperkter in de gesubsidieerde dagverblijven (34%), in overeenstemming met de beperkte toename van het aantal plaatsen en het aantal gebruikers. Bij de gesubsidieerde opvanggezinnen is het aantal gebruikers per 1000 kinderen meer dan verdubbeld.

In de *niet-gesubsidieerde dagopvang* is het aantal gebruikers per 1000 kinderen toegenomen met 147%: van 71 gebruikers in 1990 tot 176 in 2002. De stijging bij de dagverblijven (261%) is veel sterker dan bij de opvanggezinnen (56%). Het *totale aantal gebruikers per 1000 kinderen jonger* is toegenomen met 98%: van 308 gebruikers in 1990 tot 609 in 2002. De stijging bij de dagverblijven (85%) is iets beperkter dan in bij opvanggezinnen (109%).

Anno 2002 maakte dus 60% van alle kinderen van drie maand tot drie jaar regelmatig gebruik van de dagopvang. Ruim 71% van de gebruikers doet een beroep op de

gesubsidieerde dagopvang, 29% op de niet-gesubsidieerde dagopvang. Bijna 43% van de gebruikers maakt gebruik van de dagverblijven en 57% van de opvanggezinnen.

**Tabel 7.2: Aantal gebruikers per 1000 kinderen van drie maand tot drie jaar (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
<b>Gesubsidieerd</b>								
Dagverblijven	34	110	111	116	130	134	142	146
Gezinnen	126	127	168	203	240	266	287	287
Totaal	83	237	279	319	370	400	429	433
<b>Niet gesubsidieerd</b>								
Dagverblijven	261	31	30	46	56	75	95	114
Gezinnen	56	40	44	49	53	57	58	62
Totaal	147	71	74	95	109	132	153	176
<b>Totaal</b>								
Dagverblijven	85	141	141	162	186	210	237	260
Gezinnen	109	167	212	252	293	323	345	349
Totaal	98	308	353	414	479	532	582	609

#### 4.3. Omvang van het gebruik

Tabel 7.3 geeft de omvang van het gebruik weer in de *gesubsidieerde dagopvang* via het aantal verblijfsdagen per gebruiker. Voor de niet-gesubsidieerde dagopvang zijn er geen gegevens beschikbaar over het aantal verblijfsdagen. Algemeen is het aantal verblijfsdagen per gebruiker per jaar licht toegenomen: van 67 dagen in 1990 tot 78 dagen in 2002. Bij de dagverblijven lag dit cijfer gemiddeld 9 eenheden hoger in 2002. De tabel toont ook de *gemiddelde benuttingsgraad*, uitgaande van maximaal 220 verblijfsdagen (44 weken x 5 dagen) per gebruiker per jaar. Een wijziging van dit maximum aantal leidt tot een andere benuttingsgraad in alle jaren, maar de verhouding tussen dagverblijven en opvanggezinnen blijft dezelfde. In de dagverblijven is de benuttingsgraad toegenomen van 33% in 1990 tot 39% in 2002, in de opvanggezinnen van 28% in 1990 tot 34% in 2002.

**Tabel 7.3: Volle verblijfsdagen per gebruiker per jaar en de benuttingsgraad in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
Dagverblijven	20	72	81	80	81	84	85	87
benuttingsgraad		33%	37%	36%	37%	38%	39%	39%
Gezinnen	19	62	72	75	72	72	71	74
benuttingsgraad		28%	33%	34%	33%	33%	32%	34%
Totaal	17	67	76	76	75	76	76	78
benuttingsgraad		30%	34%	35%	34%	34%	34%	36%

Tabel 7.4 drukt de omvang van het gebruik in de gesubsidieerde dagopvang uit via het aantal volle verblijfsdagen per gebruiker per week. Deze maat sluit beter aan bij de dagelijkse perceptie van de gezinnen. Algemeen werd door de gebruikers in 2002

1,7 dagen per week effectief gebruik gemaakt van de dagopvang. In de dagverblijven was dit 1,9 dagen per week en in de opvanggezinnen 1,6 dagen per week.

**Tabel 7.4: Aantal volle verblijfsdagen per gebruiker per week in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
Dagverblijven	20	1,6	1,8	1,7	1,8	1,8	1,8	1,9
Gezinnen	19	1,3	1,6	1,6	1,6	1,6	1,5	1,6
Totaal	17	1,5	1,6	1,7	1,6	1,6	1,6	1,7

#### 4.4. Bezetting

Tabel 7.5 geeft voor de *gesubsidieerde dagopvang* het aantal volledige verblijfsdagen per plaats en de bezettingsgraad weer, uitgaande van maximaal 220 verblijfsdagen per jaar. Een wijziging van dit maximum aantal impliceert een andere bezettingsgraad in alle jaren, maar de verhouding tussen dagverblijven en opvanggezinnen blijft gelijk.

In de gesubsidieerde *dagverblijven* was de bezettingsgraad vrijwel constant op een hoog niveau. Met 181 verblijfsdagen per plaats in 2002 was er een bezettingsgraad van 82%. De gesubsidieerde *opvanggezinnen* halen in 2002 met 162 verblijfsdagen een bezettingsgraad van 73%. De bezettingsgraad is sterk toegenomen in de periode 1990-1996, toen men maximaal drie volledige plaatsen mocht aanbieden. De sterke daling in 1997 is te wijten aan de verhoging van het maximale aantal plaatsen per opvanggezin van drie naar vier. De verhoging van de capaciteit in de periode 1996-2002 heeft niet geleid tot een evenredige verhoging van het aantal gebruikers en het aantal verblijfsdagen per opvanggezin.

**Tabel 7.5: Aantal volledige verblijfsdagen per plaats en bezettingsgraad in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
Dagverblijven	3	176	184	179	182	181	184	181
bezettingsgraad		80%	84%	81%	83%	82%	84%	82%
Gezinnen	42	113	155	168	184	153	160	162
bezettingsgraad		52%	70%	77%	83%	70%	73%	73%
Totaal	22	138	166	172	183	162	168	168
bezettingsgraad		63%	75%	78%	83%	74%	76%	77%

#### 4.5. Productiviteit

Tabel 7.6 laat het aantal plaatsen per VTE zien in de gesubsidieerde dagopvang, als een eerste uitdrukking van de productiviteit. Het aantal plaatsen per VTE bedroeg in 2002 6,3 in de *gesubsidieerde dagverblijven*, of een halve plaats minder per VTE dan de officiële norm van 6,8 plaatsen die Kind en Gezin voor gesubsidieerde dagverblijven hanteerde. Het aantal plaatsen per VTE was in de *gesubsidieerde opvanggezinnen* veel lager: ongeveer 3 plaatsen tot in 1996 en vanaf 1997 ongeveer 4 plaatsen, maar deze verhoging heeft niet geleid tot een evenredige stijging van het

aantal verblijfsdagen per VTE (zie tabel 7.7). Een gemiddelde voltijdse baan in een dagverblijf heeft dus een veel hogere productiecapaciteit. Dit heeft een belangrijke invloed op de kostprijs van de twee soorten opvang.

**Tabel 7.6: Aantal plaatsen per VTE in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
Dagverblijven	4	6,0	6,2	5,9	6,1	6,2	5,9	6,3
Gezinnen	-22	5,0	3,5	3,1	2,8	4,3	4,0	3,9
Totaal	-11	5,5	4,4	4,0	3,8	5,2	4,9	4,9

Tabel 7.7 toont het *aantal verblijfsdagen per voltijds equivalente baan* voor de gesubsidieerde dagopvang. Het is de beste indicator van de *productiviteit* van de dagopvang die we kunnen berekenen op basis van de beschikbare gegevens.

**Tabel 7.7: Aantal volledige verblijfsdagen per VTE in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	1990	1992	1994	1996	1998	2000	2002
Dagverblijven	1.058	1.131	1.054	1.117	1.121	1.087	1.133
Gezinnen	562	536	519	518	653	645	626
Totaal	735	690	642	650	773	760	752

Hier valt het grote productiviteitsverschil op tussen dagverblijven en opvanggezinnen. In de dagverblijven worden in 2002 gemiddeld 1.133 volle verblijfsdagen gepresteerd per voltijdse baan, tegenover maar 626 in de opvanggezinnen.

#### 4.6. Kostenefficiëntie

In deze sectie behandelen we de kostprijs per volledige plaats en de kostprijs per volle verblijfsdag in de gesubsidieerde dagopvang. Voor de gesubsidieerde dagverblijven is deze kostprijs berekend op basis van de totale inkomsten, zijnde de som van de subsidies van de Vlaamse Gemeenschap en de ouderbijdragen. Aangezien dagverblijven non-profitorganisaties zijn, veronderstellen we dat de uitgaven gelijk zijn aan de inkomsten en dat bijkomende niet geregistreerde inkomsten (zoals schenkingen en inkomsten uit activiteiten) te verwaarlozen zijn. Voor de gesubsidieerde opvanggezinnen zijn alleen de subsidies en ouderbijdragen gekend, maar niet de gedeerde "fiscale en parafiscale inkomsten" en de gedeerde beroepsinkomsten voor de opvanggezinnen, die verbonden zijn met het beroepsstatuut van de onthaalouders. Aangezien zij geen volwaardig loon ontvangen, worden de "equivalente loonkosten per VTE" geschat als een percentage van de werkelijke totale kosten per VTE in de dagverblijven. Dit percentage kan variëren van 100% tot bijv. 50%. In de analyse werd 70% als equivalentiepercentage gehanteerd.<sup>67</sup> Een

<sup>67</sup> Deze veronderstelde 70% als equivalentiepercentage steunt op de inschatting van de kwalificatieniveau van het personeel tewerkgesteld in de opvanggezinnen (Van Dongen, 2004, 65).

hogere percentage impliceert een hogere kostprijs van de opvang in opvanggezinnen. Voor meer toelichting bij de methodologie verwijzen we naar Van Dongen (2004).

#### 4.6.1. Kostprijs per plaats

Tabel 7.8 geeft de kostprijs per volledige plaats weer in gesubsidieerde dagverblijven en opvanggezinnen. Het zijn nominale prijzen omdat de vergelijking tussen de dagverblijven en opvanggezinnen centraal staat. In de *gesubsidieerde dagverblijven* bedroeg de kostprijs per plaats in 2002 bijna 10.000 euro per jaar, tegenover bijna 4.700 euro in 1990. Een volledige plaats in de *gesubsidieerde opvanggezinnen* kostte in 2002 ruim 11.300 euro, tegenover bijna 4.000 euro in 1990.<sup>68</sup>

**Tabel 7.8: Kostprijs per volledige plaats in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
<b>Gesubsidieerd</b>								
Dagverblijven	113	4.679	5.368	6.152	6.513	6.597	7.698	9.983
Gezinnen	183	3.984	6.690	8.219	9.924	6.704	7.899	11.293
Totaal	154	4.258	6.185	7.468	8.625	6.668	7.830	10.836
<b>Niet gesubsidieerd</b>								
Dagverblijven	177	3.266	3.858	4.848	4.969	5.442	6.442	9.036
Gezinnen	207	2.895	3.997	4.724	5.345	4.635	5.502	8.874
Totaal	195	3.038	3.949	4.775	5.173	5.056	6.055	8.976

#### 4.6.2. Kostprijs per verblijfsdag

Tabel 7.9 toont de kostprijs per volledige verblijfsdag in gesubsidieerde dagverblijven en opvanggezinnen, als de meest accurate indicator voor de kostenefficiëntie van de externe dagopvang.<sup>69</sup> In 2002 bedroeg de totale kostprijs van een voltijdse verblijfsdag in een dagverblijf ruim 55 euro, tegenover bijna 27 euro in 1990. In een opvanggezin bedroeg de totale kostprijs van een voltijdse verblijfsdag in 2002 bijna 70 euro.

<sup>68</sup> Deze gegevens werden afgeleid uit de uitgaven (loonkosten, werkingskosten en investeringen) en de inkomens (subsidies, ouderbijdragen en andere inkomsten) zoals deze uit de jaarrekeningen te puren vallen. Gegevens over het inkomen hebben we enkel op macroniveau. De uitgavenstructuur is op zich niet gekend, maar in deze studie werd er vanuit gegaan dat de uitgaven gelijk zijn aan de inkomsten. Ook voor de gegevens omtrent de opvanggezinnen dienen we noodgedwongen uit te gaan van deze veronderstelling. Een bijkomend probleem hier is dat we niet beschikken over jaarrekeningen en ons enkel kunnen beroepen op de forfaitaire vergoeding per verblijfsdag per kind. Wat de "andere inkomsten" betreft werd bij de opvanggezinnen rekening gehouden met de door de federale overheid gederfde fiscale en parafiscale inkomsten en de voor de opvanggezinnen gederfde beroepsinkomsten. De gederfde fiscale en parafiscale inkomsten werden afgeleid uit de federale subsidies aan de (toenmalige) Vlaamse Gemeenschap die specifiek voorzien zijn voor het statuut van opvangpersonen. De gederfde beroepsinkomsten voor de opvanggezinnen zijn in feite "niet vergoede kosten" van deze activiteit. Inspiratie voor deze werkwijze werd gehaald bij het federaal planbureau in een poging (anno 1997) om de kosten van werkzoekenden in kaart te brengen. Om te komen tot een beeld van de kosten van niet-gesubsidieerde opvang werd een equivalentiemethode gebruikt. Meer specifiek werd verondersteld dat de kosten dezelfde zijn als bij niet-gesubsidieerde dagverblijven, met als enig verschil de financieringswijze, i.c. de hogere ouderbijdrage bij niet-gesubsidieerde opvang (Van Dongen, 2004, 62-66).

<sup>69</sup> Ook hier werd gebruik gemaakt van de equivalentiemethode. We veronderstellen namelijk een gelijke kwaliteit van de dienstverlening in beide soorten opvang. Er werd voor de niet-gesubsidieerde opvang uitgegaan van dezelfde uitgaven per volle verblijfsdag. Daarom worden de gegevens niet afzonderlijk weergegeven in tabel 7.9 (Van Dongen, 2004, 74).



**Tabel 7.9: Kostprijs per volledige verblijfsdag in de gesubsidieerde dagopvang (Vlaams Gewest, 1990-2002)**

Opvanginstantie	% 90/02	1990	1992	1994	1996	1998	2000	2002
Dagverblijven	107	27	29	34	36	36	42	55
Gezinnen	99	35	43	49	54	44	49	70
Totaal	109	31	37	43	47	41	47	64

Dezelfde redenering kan men toepassen op de leefsituatie in kostwinnersgezinnen waar kinderen jonger dan drie jaar thuis worden opgevangen door een van de ouders. In de periode 1990-2002 bedroeg het aandeel van de paargezinnen met één kind jonger dan drie jaar 88% en het aandeel van de paargezinnen met twee kinderen jonger dan drie jaar bijna 12%. Bij de berekening veronderstellen we dat de kwaliteit van de opvang/opvoeding thuis tijdens de normale opvangtijd (8u-17u) gelijk is aan die in een gesubsidieerd opvanggezin. Varianten zijn mogelijk door wijziging van deze veronderstelling. De kosten van de opvanggezinnen worden als basis genomen voor de equivalente kosten van een ouder die voltijds thuis blijft.

Tabel 7.10 laat zien dat de totale kostprijs van een volle opvangdag thuis voor de opvang van één kind in 2002 meer dan drievoud bedroeg van de kostprijs in een gesubsidieerd opvanggezin en bijna het vijfvoud van de kosten in een gesubsidieerd dagverblijf.<sup>70</sup> Bij gezinnen met twee jonge kinderen thuis bedroeg de totale kostprijs van een opvangdag in 2002 bijna het dubbele van de kosten in een gesubsidieerd opvanggezin en bijna het drievoud van de kosten in een gesubsidieerd dagverblijf.

**Tabel 7.10: Kostprijs per verblijfsdag bij de opvang thuis (Vlaams Gewest, 1990-2002)**

Opvanginstantie	1990	1992	1994	1996	1998	2000	2002
Thuis - 1 kind <3 jaar	174	149	151	152	187	199	271
Thuis - 2 kinderen <3 jaar	87	75	75	76	93	100	135
Thuis - 3 kinderen <3 jaar	58	50	50	51	62	66	90
Thuis - 4 kinderen <3 jaar	43	37	38	38	47	50	68

Ongeacht de gemaakte veronderstellingen, raken deze cijfers de kern van de kosten-batenanalyse van de diverse soorten dagopvang van jonge kinderen, verwijzend naar de verschillen in productiviteit. De traditionele opvang door een ouder thuis van één kind, twee of drie kinderen jonger dan drie jaar is voor de maatschappij en de gezinnen veel duurder dan de externe dagopvang in een opvanggezin of een dagverblijf. In de veronderstelling dat de kwaliteit van de opvang/opvoeding in een opvanggezin of een dagverblijf niet lager is dan die in het eigen gezin, is

<sup>70</sup> Bij het berekenen van de totale kosten per voltijds equivalente opvangdag door een ouder thuis, werd er opnieuw gebruik gemaakt van de totale equivalente kosten van de opvanggezinnen als basis voor de totale equivalente kosten van een ouder die voltijds thuis blijft om voor zijn kind(eren) jonger dan drie jaar te zorgen. De veronderstelling hierbij is dat het beroepsprofiel en de professionele verdien capaciteit van opvangpersonen en thuiswerkende ouders (bijna) gelijk is (Van Dongen, 2004, 78).



het maatschappelijk gezien doelmatiger om jonge kinderen gedurende een aantal dagen per week en een aantal uren per dag opvang/opvoeding buiten het gezin aan te bieden, als aanvulling bij de opvang/opvoeding thuis.

Voorts blijkt dat een volle verblijfsdag in een dagverblijf goedkoper is dan in een opvanggezin. Indien de kwaliteit van de opvang/opvoeding in een dagverblijf niet lager wordt ingeschat dan die in een opvanggezin, is het bijgevolg financieel doelmatiger om meer te investeren in goede dagverblijven. In dat perspectief zou men kunnen stellen dat het beleid van de Vlaamse overheid tijdens de voorbije decennia niet zo doelmatig is geweest. Indien men de kwaliteit van de opvang/opvoeding in opvanggezinnen hoger inschat dan die in een dagverblijf, is een andere conclusie mogelijk. Bij de vaststellingen op grond van het voorliggende onderzoek moeten we immers rekening houden met het huidige beleidskader-en keuzes, alsook met de gestelde hypothesen met betrekking tot de kostenomvang, de financiering en de kwaliteit van de dienstverlening.

Deze resultaten betekenen niet automatisch dat de dagopvang in gesubsidieerde dagverblijven voldoende doelmatig is. Het is altijd zinvol te onderzoeken hoe de opvang in dagverblijven doelmatiger en goedkoper kan worden gemaakt, zonder de kwaliteit ervan in het gedrang te brengen.

## 5. Besluit

Deze bijdrage toont dat een geïntegreerde benadering van de publieke diensten in de maatschappij zinvol en wenselijk is. De overheden en gesubsidieerde organisaties worden gezien als volwaardige, productieve actoren in het marktsysteem die waardevolle diensten voortbrengen en aanbieden aan diverse andere actoren, uitgaande van de wisselwerking tussen de vraagzijde en aanbodzijde. De centrale vraag is in welke mate de diverse diensten doelmatig of efficiënt worden geproduceerd en aangeboden “op de markt”, waarbij de noden van de diverse “vragers” naar deze diensten voldoende worden beantwoord.

In het verlengde daarvan presenteerden we een algemene benadering van de diverse soorten dienstverlening in het marktsysteem, met specifieke aandacht voor de verhouding tussen de dienstverlening als gezinsarbeid, vrijwillige arbeid en professionele arbeid. In deze context werden de begrippen effectiviteit en efficiëntie in algemene termen bepaald, met aandacht voor de betekenis van andere basiswaarden zoals vrijheid, gelijkheid en solidariteit en de doelstellingen die daaruit voortvloeien.

Uitgaande van deze algemene benadering, is het meten van de efficiëntie van publieke diensten theoretisch gezien gelijk aan het meten van de efficiëntie van niet-publieke diensten. In de praktijk vereist dit een correcte meting/registratie van een aantal basisgegevens van de productieprocessen binnen de afgebakende diensten, aan de input- en outputzijde. De uitdaging is voor alle publieke diensten de juiste gegevens te registreren en een aantal basisindicatoren te berekenen op organisatieniveau, uitgaande van de beleidsdoelen die worden vooropgesteld met betrekking tot beschikbaarheid, gebruik, bezettingsgraad, productiviteit, kostenefficiëntie en aandeel van de persoonlijke bijdrage. Deze kwantitatieve evaluatie vormt dan een basis voor een bredere maatschappelijke evaluatie, ex-post en ex-ante.

In het verlengde daarvan wordt gepleit voor de ontwikkeling van praktische evaluatie-instrumenten voor de registratie, analyse en rapportering van de basisindicatoren op organisatieniveau, als een kwantitatieve basis voor een brede beleidsevaluatie, op maat van de diverse sectoren en afzonderlijke organisaties. In een eerste fase is het wenselijk om via casestudies in een aantal sectoren enkele prototypes te ontwerpen, mede op basis van de uitwisseling van de aanwezige ervaring en deskundigheid in deze sectoren.

Ten slotte wordt een concrete toepassing toegelicht met betrekking tot de evolutie van de dagopvang van kinderen jonger dan drie jaar in het Vlaamse Gewest in de periode 1990-2002. Op basis van de beschikbare gegevens werden enkele basisindicatoren berekend voor de dagopvang in dagverblijven en opvanggezinnen, als empirische basis voor de normatieve vraag hoe de dagopvang in de toekomst verder kan en moet worden ontwikkeld. We willen hierbij wijzen op de veronderstellingen die nodig zijn wegens het gebrek aan bepaalde gegevens. Dit toont het belang aan van een degelijk evaluatie-instrument waarbij de juiste gegevens worden geregistreerd. Ondanks deze beperkingen bieden de basisindicatoren nuttige informatie over de doelmatigheid van het gevoerde en te voeren beleid.

## Bibliografie

Bouckaert, G. (1992). Productivity analysis in the public sector: the case of fire service. In: *International review of Administrative Sciences*, 58 (2), 175-200.

Bouckaert, G. (1993). Efficiency measurement from a management perspective: a case of the civil registry office in Flanders. In: *International review of Administrative Sciences*, 59 (1), 11-27.

Bouckaert, G. & Thijs, N. (2003). *Kwaliteit in de overheid*. Gent: Academia Press.

Bogaert, G. & De Prins, P. (1996). *Dagprijverschillen in de rustoordsector*. Leuven-Antwerpen: HIVA - RUCA.

Bronselaer, J. (2003). Vergelijking van de middelen en de kosten bij een nursingtehuis en een rust- en verzorgingstehuis. In: Deven, F. & Jacobs, T. (red.) (2006). *Vooruitdenken over zorg in Vlaanderen*. Antwerpen-Apeldoorn: Garant, 179-196.

Gravelle, H. & Rees, R. (1987). *Microeconomics* (7<sup>th</sup> ed.). London-New York: Longman.

in 't Veld, J. (2002). *Analyse van organisatieproblemen*. Groningen: Wolters-Noordhoff.

Kind en Gezin (1990-2004), Statistische Jaarverslagen, Cel Studie en Strategie, Brussel: Kind en Gezin.

Kreps, D.M. (1990). *A course in microeconomic theory*. New York-London: Harvester Wheatsheaf.

OECD (2007). *Issues in Output Measurement for "Government at a Glance"* OECD project on Management in Government: Comparative Country Data. OECD GOV Technical Paper 2 (Second Draft).

Probis-Vandorpe-Studies (2002). *Onderzoek naar de structurele kostprijselementen van openbare rusthuizen in Vlaanderen, Deelrapport 2: vergelijking van de kostenstructuur van openbare en VZW-rusthuizen*, in opdracht van het Vlaamse Infrastructuurfonds voor Persoonsgebonden aangelegenheden en de Vereniging van Vlaamse Steden en Gemeenten, niet gepubliceerd.

Van Dongen, W. (1993). *Nieuwe Krijtlijnen voor Gezin, Markt en maatschappij. Een geïntegreerde benadering*. Leuven-Apeldoorn: Garant.

Van Dongen, W. (2004). *Kinderopvang als basisvoorziening in een democratische samenleving. Scenario's voor de toekomstige ontwikkeling van de dagopvang voor kinderen jonger dan drie jaar*. CBGS Document 2004/4, Brussel: CBGS.

Van Dongen, W. (2006a). Analysemodel en instrumenten voor een doelmatige dienstverlening in een democratische samenleving. In: Deven, F. & Jacobs, T. (red.) (2006). *Vooruitdenken over Zorg in Vlaanderen*. Antwerpen-Apeldoorn: Garant, 151-177.

Van Dongen, W. (2006b). Zorgen op de markt. Een democratisch marktsysteem. In: *Alert*, 32 (4), 12-29.

Van Dongen, W. (2008a). *Naar een democratische arbeidsverdeling? Het Combinatiemodel als basis voor een geïntegreerd beleid in Vlaanderen en Europa*. Brussel: Studiedienst van de Vlaamse Regering (SVR-Studie, 2008/1).

Van Dongen, W. (2008b). *Towards a democratic division of labour in Europe? The Combination Model as a new integrated approach to professional and family life*. Bristol: The Policy Press.

## ***Hoofdstuk 8: Als meten het doel wordt: de vertaling van gender mainstreaming naar de beleidspraktijk***

Petra Meier, Universiteit Antwerpen  
Karen Celis, Hogeschool Gent – Universiteit Gent

<b>1. Inleiding .....</b>	<b>178</b>
<b>2. Gender mainstreaming: een intentioneel en rationeel beleidsproces .....</b>	<b>179</b>
2.1. Intentionaliteit en rationaliteit van het beleidsproces .....	179
2.2. Intentionaliteit en rationaliteit van gender mainstreaming .....	180
2.3. Een gestuurd beleidsproces als instrument voor het realiseren van gelijkheid? .....	183
<b>3. De vertaling van gender mainstreaming naar de Belgische beleidspraktijk .....</b>	<b>185</b>
3.1. De Pekingwet van 1996 (en het Vlaamse Pekingdecreet van 1997) .....	186
3.2. De Vlaamse emancipatie-effect-rapportage van 1997 en 2000 .....	187
3.3. De federale Cel Gender Mainstreaming van 2001 en 2002 .....	188
3.4. De Pekingdecreten en –ordonantie van 2002 en 2006 .....	191
3.5. De Vlaamse Open Coördinatie Methode (2005) .....	193
3.6. De wet gender mainstreaming van 2007 .....	194
<b>4. Verklaringen voor gender mainstreaming als een proces van meten en weten .....</b>	<b>197</b>
<b>5. Gender mainstreaming: een meerwaarde voor het realiseren van gendergelijkheid? .....</b>	<b>200</b>
<b>Bibliografie .....</b>	<b>203</b>

## 1. Inleiding

Sinds het midden van de jaren 1990 staat in het gelijke-kansenbeleid<sup>71</sup> het concept gender mainstreaming centraal. Gender mainstreaming verscheen voor het eerst in internationale teksten volgend op de Derde VN Wereldconferentie in 1985 in Nairobi. Het had in de eerste plaats betrekking op de integratie van vrouwen en vrouwelijke waarden in het ontwikkelingsbeleid. Tien jaar later, in 1995, werd gender mainstreaming opgenomen in het Peking Actieplatform resulterend uit de Vierde VN Wereldvrouwenconferentie in Peking. Dat document engageerde de VN en haar leden om de evolutie inzake gendergelijkheid op te volgen in twaalf actiedomeinen en om de aandacht voor het genderperspectief structureel in te bedden in de verschillende beleidsdomeinen. Sindsdien wordt gender mainstreaming als centrale strategie in het gelijke kansenbeleid geïmplementeerd door nationale beleidsmakers, de Europese Unie, de Raad van Europa, internationale organisaties zoals de VN, UNESCO en de Wereldbank, en NGO's zoals Oxfam (Mazey, 2001; Beveridge & Nott, 2001; Jahan, 1995; Rees, 1998; Hafner-Burton & Pollack, 2002a, 2002b).

Een ruim decennium later laten de resultaten volgens veel critici te wensen over. Bovendien wijst comparatief onderzoek uit dat er, op Zweden na (Daly, 2005; Rubery e.a., 2004), geen sprake is van een volle toepassing van gender mainstreaming in de EU (Mosesdottir & Erlingsdottir, 2005). Verklaringen voor het niet ten volle implementeren van de strategie evenals voor het uitblijven van resultaten worden gezocht in het niet vervuld zijn van politieke en beleidsmatige randvoorwaarden (Rees, 2005; Woodward, 2003); het verzet tegen de intrinsieke doelstelling van gender mainstreaming (Stratigaki, 2005; Oldersma, 2000); de hoge verwachtingen die de definitie van gender mainstreaming oproept (Meier, 2006); en de moeilijkheid om de strategie af te stemmen op de beleidscontext (Behning & Serrano Pascual, 2001; Beveridge, Nott & Stephen, 2000; Pollack & Hafner-Burton, 2000; Lombardo & Meier, 2006).

Deze bijdrage buigt zich ook over de redenen waarom gender mainstreaming als beleidsstrategie er niet in slaagt om de hoge verwachtingen in te lossen. Onze analyse richt zich in eerste instantie op hoe het concept geoperationaliseerd wordt wanneer het in de beleidspraktijk geïmplementeerd wordt. We sluiten zodoende aan bij een aantal buitenlandse studies dat zich boog over de vraag of gender mainstreaming eens geïmplementeerd waarlijk transformatief is en dat concludeerde dat het gevaar van een "technocratisering" van het gender mainstreamingsbeleid reëel is (Beveridge & Nott, 2002; Daly, 2005; Lombardo & Meier, 2006). Gender mainstreaming heeft de neiging om te verworden tot een manier van produceren van specifieke output via het gebruik van bepaalde instrumenten in plaats van een beleidsstrategie gericht op het realiseren van gendergelijkheid als outcome. Gender mainstreaming heeft een uitgesproken intentionele en rationele visie op de beleidscyclus en sluit nauw aan bij "new public management"-praktijken door de nadruk die het legt op het formuleren,

---

<sup>71</sup> In deze bijdrage gebruiken we de term gelijke kansenbeleid voor het beleid ter promotie van de gelijkheid van beide seksen (of van andere maatschappelijke groepen) aangezien die als zodanig ingeburgerd is in België. Tenzij anders aangegeven gaat het om een generieke term.

opvolgen en evalueren van beleidsdoelstellingen. Gender mainstreamingsinstrumenten worden vaak ingevoerd in het kader van administratieve vernieuwingen in het algemeen en in het kader van het vernieuwen, moderniseren of verbeteren van gelijkheidsbeleid in het bijzonder (Daly, 2005). De val die hiermee gepaard gaat, is dat gender mainstreaming eerder een symbool van modernisering wordt dan een streven naar meer gendergelijkheid. Heel wat overheden implementeren gender mainstreaming voorts slechts partieel: enkel gender mainstreamingsinstrumenten, -technieken en -procedures worden ingevoerd en de genderexpertise en -kunde opgebouwd (Verloo, 2000; Daly, 2005). Ze vervolgens inzetten en aanwenden in het kader van ruimere beleidsdoelstelling in het kader van het realiseren van gendergelijkheid is niet per definitie de volgende stap. Gender mainstreaming verwatert zo tot een papieren “tick and bash”-oefening: het aanhechten van een formulier bij beleidsdocumenten waarop het “ja”- of “neen”-vakje wordt aangevinkt naast de vraag “Heeft deze beleidsmaatregel impact op mannen en vrouwen?”. Deze neiging wordt versterkt doordat de operationalisering van een horizontaal of transversaal beleid als gender mainstreaming doorgaans gebeurt in een traditioneel verticaal opgebouwde beleidsstructuur, vaak bevolkt door beleidsactoren die weinig ervaring (en affiniteit) hebben met gelijke kansenbeleid.

In deze bijdrage onderzoeken we of ook in België de implementatie van gender mainstreaming in de beleidspraktijk geleid heeft tot een “doelloos” meten en weten. Zijn het meten van gelijkheid, het in kaart brengen van de acties om gendergelijkheid te realiseren, het monitoren van het beleidsproces en de inherente beleidsevaluatie een doel op zich geworden in plaats van een middel in een proces naar meer gendergelijkheid? In de zoektocht naar een antwoord op deze vraag, analyseert deze bijdrage een aantal sleutelmomenten in de introductie en implementatie van gender mainstreaming in het Belgische beleid sinds het midden van de jaren 1990. Zowel het federale, als het Vlaamse, het Waalse en het Brusselse gelijke-kansenbeleid komen daarbij aan bod. Deze bijdrage richt zich op: i) de mate waarin het monitoren en evalueren van beleid tot een doel verheven wordt; ii) de mate waarin dit inherent is aan gender mainstreaming als concept; iii) de gevolgen hiervan voor het gelijke kansenbeleid. Voorafgaandelijk aan deze analyse situeert de eerste sectie van deze bijdrage gender mainstreaming in een aantal theorieën over het beleidsproces.

## **2. Gender mainstreaming: een intentioneel en rationeel beleidsproces**

### ***2.1. Intentionaliteit en rationaliteit van het beleidsproces***

De beleidswetenschappen hanteren een brede waaier aan visies op het beleidsproces die sterk variëren in de mate waarin ze het beleidsproces en zijn uitkomst rationaliteit en intentionaliteit toedichten. In zijn handboek over beleidsanalyse vertrekt Parsons (1995) van de veronderstelling dat de menselijke rede en inzicht in problemen in staat zijn om beleidsoplossingen aan te dragen die daadwerkelijk de samenleving ten goede veranderen. In deze “klinische benadering” van de “beleidswetenschappen”, zoals Hill en Hupe (2002) Parsons’ visie typeren, worden sociale wetenschappen beschouwd als een remedie voor maatschappelijke ziekten. Beleid wordt gezien als

het resultaat van een relationele inzet van middelen voor het bereiken van specifieke doelen die door de politieke besluitvormers werden bepaald. Het beleidsproces is bijgevolg *'a purposive course of action followed by an actor or set of actors in dealing with a problem or matter of concern'* (Anderson, 1975, 3). De keuze voor een specifiek beleid wordt in deze visie bepaald door een "maximale rationaliteit" waarbij de actoren in het beleidsproces duidelijke doelstellingen voor ogen hebben en, op basis van volledige informatie over het onderwerp, de inzetbare middelen en de effecten van de verschillende opties, beleidskeuzes maken die leiden tot een optimaal resultaat met een zo laag mogelijk kost.

Dit model wordt bekritiseerd omwille van haar onrealistische assumpties. In de realiteit zijn de doelstellingen niet altijd helder, is de beschikbare informatie beperkt en gekleurd, zijn middelen schaars, is het inzicht in de effecten van beleidsopties gelimiteerd en ageren actoren binnen psycho-sociale grenzen (Howlett & Ramesh, 2003). Deze kritiek lag aan de basis van Simon's model van "bounded rationality", dat deze beperkingen erkent maar tevens poneert dat binnen deze grenzen nog steeds een rationele besluitvorming mogelijk is waarbij doelstellingen en acties gekozen worden die het meest geschikt zijn om de op voorhand bepaalde waarden en doelen te realiseren (Parsons, 1995, 278).

Lindbloms' incrementalistische benadering van het beleidsproces verwerpt deze rationaliteit, zelfs in de beperkte vorm. In zijn visie wordt de intentionaliteit in de besluitvorming vervangen door een model dat beleid beschouwt als de uitkomst van permanente onderhandelingen tussen verschillende actoren over ondermeer de probleemdefinitie, de doelen en de kosten. Beleidsmakers verschillen niet enkel van mening over de beleidsuitkomsten en -instrumenten, ze hebben bovendien vaak geen duidelijk beeld van wat ze uiteindelijk willen bereiken en op welke manier. In deze benadering wordt de beleidswetenschap de *'science of muddling through'* (Lindblom, 1959). Verhoogde inzet van middelen om een probleem op te lossen zonder het probleem zelf afdoende te definiëren, wordt beschouwd als de normale wijze waarop beleidsmakers door het beleidsproces "ploeteren".

Nog radicaler in het verwerpen van intentionaliteit en rationaliteit in het beleidsproces, is het "garbage-can"- of prullenmandmodel van March en Olsen (1979). In deze visie is er geen enkel verband tussen middel en doel: de doelstellingen zijn ambigu, de beleidsinstrumenten worden in vraag gesteld en de deelname van beleidsmakers is tijdelijk. Onder deze omstandigheden komt beleid quasi accidenteel tot stand en zijn de oplossingen die voor bepaalde problemen naar voren worden geschoven niet gebaseerd op een keuze voor het meest gepaste middel om het doel te bereiken maar eerder het toevallige resultaat van datgene wat op dat ogenblik aanwezig was in de prullenmand van het beleidproces.

## **2.2. Intentionaliteit en rationaliteit van gender mainstreaming**

Deze (te) korte bespreking van visies op het beleidsproces biedt een aantal interessante aanknopingspunten voor een discussie over gender mainstreaming. Niettegenstaande gender mainstreaming vooral geïntroduceerd en verspreid werd via de VN-conferenties en -netwerken (True & Mintrom, 2001), is de meest frequent



gehanteerde definitie afkomstig van de Raad van Europa: *'Gender mainstreaming is the (re)organization, improvement, development and evaluation of policy processes so that a gender equality perspective is incorporated in all policies at all levels and at all stages, by the actors normally involved in policy-making.'* (Council of Europe, 1998, 15). Deze definitie is tevens de meest omvattende. Het Peking Actieplatform houdt het bij voorbeeld op *'an active and visible policy of mainstreaming a gender perspective, inter alia, in the monitoring and evaluation of all policies and programmes'* (United Nations, 1995, paragraaf 292). Voorts is het opvallend dat heel wat EU-documenten gender mainstreaming naar voren schuiven als een cruciaal onderdeel van het gelijke kansenbeleid, zonder evenwel aan te geven wat gender mainstreaming precies inhoudt. Zo onderlijnt het Verdrag van Amsterdam (1997) de noodzaak om de ongelijkheid van mannen en vrouwen weg te werken en hun gelijkheid te bevorderen (artikel 2 en 3) en sinds het verdrag promoot de EU gender mainstreaming als de meest geschikte beleidsstrategie, maar het definieert deze strategie niet expliciet. De routekaart voor de gelijkheid van vrouwen en mannen voor de periode 2006-2010 (COM, 2006, 92) herneemt de tweeledige aanpak van gender mainstreaming in combinatie met specifieke maatregelen ter bevordering van de gelijkheid van mannen en vrouwen van de vroegere raamstrategie voor de gelijkheid van mannen en vrouwen voor de periode 2001-2005 (COM, 2000, 335). Ook al ontbreekt een eigenlijke definitie van gender mainstreaming in deze en andere formele EU documenten, de invulling die in de teksten aan het concept wordt gegeven, sluit nauw aan bij de definitie zoals de Raad van Europa ze hanteert. Om die reden vertrekken ook wij van deze definitie voor onze analyse.

Gender mainstreaming, zoals gedefinieerd door de Raad van Europa, gaat ervan uit dat beleid en beleidsprocessen in hoge mate intentioneel en rationeel (kunnen) zijn. De intentionaliteit situeert zich vooral op het vlak van de uitkomst van het beleid: gender mainstreaming poneert enerzijds dat genderongelijkheid gegeneerd en in stand gehouden wordt door beleid en, anderzijds, dat het beleid ingezet kan worden om gelijkheid tot stand te brengen. Naast de vaststelling dat de verschillende invullingen van gender mainstreaming het uitgangspunt delen dat gelijkheid door beleid kan en moet bevorderd worden en ze zodoende aan dat beleid een bepaalde intentionaliteit toedichten, moeten we ook concluderen dat er geen consensus bestaat over wat dat doel dan wel is. Er is onduidelijkheid over de concrete invulling van het concept "gelijkheid". De Europese Commissie en het Peking Actieplatform hanteren eerder vrijblijvende definities en houden het op "gelijke kansen" en het aanwezig stellen van een "genderperspectief". De Raad van Europa stelt "gendergelijkheid" centraal, wat *stricto sensu* veel stringenter en uitkomstgerichter is dan gelijke kansen.

De invulling van gelijkheid door het gelijke-kansenbeleid wijzigde doorheen de tijd. Het gelijke kansenbeleid evolueerde van een legalistische aanpak gebaseerd op "gelijke rechten", via "gelijke kansen" (vaak gerealiseerd via positieve acties), naar een meer structurele conceptualisering van effectieve gendergelijkheid die ook de mannelijke norm in vraag stelt. De feministische literatuur situeert gender mainstreaming in die laatste hoek. Het doel van gender mainstreaming is gendergelijkheid; de focus ligt met andere woorden niet op vrouwen, maar op gender, op vrouwen én mannen en op hun onderlinge machtsverhouding (Mazey, 2001). Gender mainstreaming heeft de intentie om de impact van een genderbias op de (re)productie van ongelijkheid

aan te kaarten (Verloof, 2001; Liebert, 2002). Gender mainstreaming vertrekt van het feit dat de levens van mannen en vrouwen verschillen en wil dit verschil erkennen. Gender mainstreaming heeft niet als doel om verschillen tussen mannen en vrouwen uit te wissen, maar net om ze zichtbaar te integreren in de beleidsstructuren en – processen. Dat leidt tot nieuwe standaarden, niet enkel voor vrouwen maar ook voor mannen (Rees, 1998). Ook al stemt niet iedereen in met deze verregerende definitie van gelijkheid als de enige basis voor gender mainstreaming (zie bij voorbeeld Booth & Bennett, 2002) en zit er rek op het concept, er bestaat consensus over het feit dat gender mainstreaming androcentrische standaarden of traditionele rollenpatronen in vraag moet stellen (Squires, 2005). Het uiteindelijke doel van gender mainstreaming is het realiseren van een eerlijkere samenleving waar vrouwen niet moeten voldoen aan de mannelijke norm om gelijk te zijn.

Gender mainstreaming zoals de Raad van Europa het definieert, gaat ook uit van een zekere mate van rationaliteit in beleid en beleidsprocessen. De gender mainstreamingsstrategie veronderstelt dat beleid antwoorden kan bieden op maatschappelijke problemen en deze daadwerkelijk kan (helpen) oplossen. Gender mainstreaming is een strategie die bedacht werd om genderongelijkheid te verhelpen daar waar ander gelijke kansenbeleid faalde. Ook al was dit eerder beleid ontoereikend, gender mainstreaming stelt de grenzen van beleid op zich niet in vraag en evenmin het uitgangspunt dat de samenleving via beleid in een bepaalde richting gestuurd kan worden. Het gaat er veeleer vanuit dat enkel een beter beleid nodig is om bestaande ongelijkheid te temperen of te elimineren en om nieuwe ongelijkheden te voorkomen. De rationaliteit van de gender mainstreamingsstrategie zit vervat in deze connectie tussen middel (beleid) en doel (minder ongelijkheid of meer gelijkheid). Dat uitgangspunt is te situeren in Parsons verlicht model, dat stelt dat menselijke rede en inzicht in staat zijn om problemen te vatten en beleidsoplossingen aan te dragen die daadwerkelijk de samenleving ten goede veranderen. Ook aan het beleidsproces op zich wordt een zekere rationaliteit toegedicht. Het Peking Actieplatform en de EU beleidsdocumenten terzake stellen dat beleidsprocessen gemonitord dienen en kunnen worden en dat het beleid moet en kan geëvalueerd worden. Er wordt dus niet zomaar wat “aangemodderd” of “geploeterd”. Beleid wordt beschouwd als een proces waarbij een bepaald doel voorop gesteld wordt en waarbij de vooruitgang en de eindresultaten opgevolgd en bewaakt kunnen worden. En, zoals de definitie van de Raad van Europa laat uitschijnen, de beleidsprocessen zijn ook zelf aan gerichte sturing onderworpen. Ze worden ten slotte georganiseerd, gereorganiseerd, verbeterd, ontwikkeld en geëvalueerd ten einde gender mainstreaming erin te verankeren.

Deze rationalistische benadering van het beleidsproces is het meest uitgesproken in het expert-bureaucratische model van gender mainstreaming (Beveridge, Nott & Stephen, 2000). In het geval van het expert-bureaucratisch model gebeurt de implementatie van gender mainstreaming in hoofdzaak door interne en externe experts. Het zijn de genderexperten binnen de administraties en/of academische experts die gender mainstreaming sturen (via statistieken, analyses, “gender impact assessments”, gendergevoelig budgetteren, ...). In het geval van het participatief-democratisch model is de focus sterker gericht op het betrekken van individuen en organisaties bij het gender mainstreamingsbeleid. Vrouwenorganisaties worden bijvoorbeeld uitgenodigd om bij te dragen tot een “gender assessment” of hun

mening te kennen te geven op een hoorzitting. Hier worden dus een extra fase van consultatie in het beleidsproces ingelast waarbij men externen betreft wat het proces minder voorspelbaar maakt en de sturing bijgevolg minder gericht. In vele landen prevaleert het expert-bureaucratisch model (Beveridge & Nott, 2002).

Derationele opbouw van het beleidsproces en die gender mainstreaming veronderstelt, wordt ook duidelijk geïllustreerd door het overzicht van de fundamentele veranderingen die gender mainstreaming teweeg zou brengen. Zo stelt Lombardo (2005) dat gender mainstreaming een systematische aanpak impliceert waarbij beleidsinitiatieven op hun gendereffecten doorgelicht worden en waarbij de doelstellingen en middelen waar nodig bijgestuurd worden om de doelstelling (gendergelijkheid) te bereiken. Beleidsdoelen en -objectieven betreffende gendergelijkheid moeten verder ook een relevante inbedding krijgen in betekenisvolle beleidsdomeinen en -initiatieven. Dit veronderstelt al dat beleid geformuleerd wordt in termen van doelstellingen, waarbij duidelijke prioriteiten afgebakend worden. Gender mainstreaming sluit daarmee aan bij de “new public management”-cultuur, bij trends om beleidsprocessen effectief en efficiënt te laten verlopen, waarbij duidelijke doelstellingen concreet opgevolgd worden doorheen de hele beleidscyclus. Volgens Lombardo betekent dit tevens dat de institutionele en organisatorische besluitvormingscultuur waar nodig bijgesteld moet worden.

De rationaliteit die achter het concept van gender mainstreaming schuilt, komt ook tot uiting in het feit dat belangrijke instrumenten betreffende gender mainstreaming zich duidelijk situeren in de logica van “evidence-based” beleid. Een gender mainstreampingsbeleid is gebaseerd op een correcte diagnose van het probleem en prognose van hoe het te remediëren (Meier, 2006). Het omvat een heel aantal analytische instrumenten om het probleem en de potentiële oplossingen in kaart te brengen en te analyseren (Council of Europe, 1998). Het belangrijkste analytisch instrument zijn de seksegesegegreerde statistieken. Naar sekse gesegegreerde statistieken worden aangewend om de huidige situatie van mannen en vrouwen en van genderrelaties in te schatten, om het effect en de effectiviteit van beleid in te schatten en om het bewustzijn inzake genderongelijkheid te vergroten. Ten tweede worden surveys en voorspellingen gehanteerd om belangrijke thema’s te detecteren waarop het beleid moet ingrijpen en om de toekomst te voorspellen (met of zonder ingrijpen vanwege de overheid). Ten derde zijn er de kosten-baten-analyses die zicht geven op de verschillende uitkomsten die beleid kan genereren voor mannen en voor vrouwen omdat hun levenscondities verschillen. Genderonderzoek, een vierde analytisch instrument dat in het kader van gender mainstreaming aangewend wordt, geeft aan wat de belangrijkste problemen zijn inzake genderongelijkheid in een bepaald beleidsdomein en levert inzichten in de oorzaken en de gevolgen ervan. Checklists, richtlijnen en referentiekaders (“benchmarking”, “good practices”) worden vervolgens ingezet door beleidsmakers om gender mainstreaming in praktijk te brengen. Checklists inzake doelen en acties voor gender mainstreaming zijn preciezer dan richtlijnen en referentiekaders. Deze laatste zijn algemener en laten meer ruimte voor eigen inbreng van ambtenaren en politici voor het naar de praktijk vertalen van de doelstelling. Ze hebben als doel een proces van “policy learning” op gang te brengen. Een vijfde instrument is de “gender impact assessment”, in ons taalgebied bekender onder de naam “emancipatie effect rapportage” (cf. infra).

Dat is een screening van de beleidsintentie op gendereffecten zodat ongunstige effecten op genderverhoudingen kunnen weggewerkt worden vooraleer het beleid geïmplementeerd wordt. Het beoogt het structureel inbedden van een “genderbril” of een “genderreflex” (Rees, 2002). Het heeft een ruim toepassingsgebied; het kan onder meer toegepast worden op wetgeving, beleidsplannen en –programma’s, acties en offertes voor onderzoek. Verder kan het zowel op bestaande als op toekomstige wetgeving en beleid uitgevoerd worden. Gender impact assessment hangt nauw samen met een ander instrument: “visioning” (Rees, 2002) dat als doel heeft om de mannelijke norm, die aan veel beleid ten grondslag ligt, te ontmaskeren. Gendergevoelig budgetteren is een specifieke vorm van gender impact assessment. Gendergevoelig budgetteren laat toe na te gaan wat de financiële implicaties van het beleid voor vrouwen en mannen zijn. Het probeert in cijfers te vatten hoe beleid de maatschappelijke verhoudingen tussen mannen en vrouwen beïnvloedt. Daartoe gaat het, analoog aan de logica van een “emancipatie effect rapportage”, op zoek naar de burgers voor wie beleid gemaakt wordt en naar de burgers die er effectief door worden beïnvloed.

### ***2.3. Een gestuurd beleidsproces als instrument voor het realiseren van gelijkheid?***

Bovenstaande paragrafen geven duidelijk aan dat gender mainstreaming geenszins uitgaat van March’ en Olsens prullenmandmodel. Beleid is allesbehalve toeval. Het wordt gekenmerkt door intentionaliteit en rationaliteit. Beleidsprocessen, zoals ze uitdrukking krijgen in het concept gender mainstreaming, dienen een doel, namelijk het realiseren van gendergelijkheid en ze vertrekken daarbij vanuit een duidelijke probleemanalyse inzake genderongelijkheid. Deze erkenning van intentionaliteit en rationaliteit in gender mainstreaming betekent daarom nog niet dat gender mainstreaming Parsons blindelings achterna gaat. Over de mate van rationaliteit en intentionaliteit die aan gender mainstreaming toegeschreven wordt, wordt onder meer genuanceerd door discursieve analyses van het gelijke-kansenbeleid.

Het uitgangspunt van de discursieve analyses van gelijke-kansenbeleid (zie bij voorbeeld in Lombardo, Meier & Verloo, 2009: Bacchi, Bustelo & Verloo; Marx Ferree; Lombardo & Meier; Verloo & van der Vleuten) is dat een beleidsprobleem geen gegeven feit is maar een maatschappelijke constructie. We definiëren expliciet of impliciet bepaalde maatschappelijke aspecten als een probleem en andere niet. We laten ons daarbij leiden door politieke keuzes, maar ook door ons referentiekader (Verloo, 2007). Vooral dit laatste maakt dat in beleid niet enkel intentionele maar ook niet-intentionele elementen vervat zitten. Het gaat om normen en waarden, veronderstellingen, maar ook vooroordelen en stereotypen, die ons denken in een bepaalde richting sturen en tot een (ongewilde) bias leiden waar we ons niet steeds van bewust zijn. Zelfs de doelstellingen en middelen van het gelijke-kansenbeleid zijn soms doordrongen van denkbeelden die ongelijkheid reproduceren. Zo concludeerde Stratigaki (2004) dat het EU beleid dat initieel gericht was op een gelijke verdeling van zorgtaken over mannen en vrouwen geleidelijk aan verschoof naar een neoliberal beleid waarin de toelevering van flexibele vrouwelijke arbeidskracht centraal kwam te staan; het was zodoende niet langer gericht op het realiseren van gendergelijkheid maar op het oplossen van een economisch probleem van

arbeidsmarktschaarste. De combinatie van arbeid en gezin, die vooral door vrouwen werd gerealiseerd en die initieel als een probleem van genderongelijkheid werd beschouwd, werd tot de oplossing geproclameerd (Meier e.a., 2007). Het resultaat was dat de genderongelijkheid in stand werd gehouden en, sterker nog, verankerd werd in het EU arbeidsmarktbeleid. Ook een groot deel van het beleid dat erop gericht is de ongelijke machtsverhoudingen tussen mannen en vrouwen in (politieke) besluitvorming weg te werken, reproduceert de aangeklaagde machtsverhoudingen tussen mannen en vrouwen via zijn discours, de vooropgestelde doelstellingen, de ingezette middelen en de doelgroepen van beleidsmaatregelen (Lombardo & Meier, 2008). Zo zijn het steevast vrouwen die geïmagineerd worden in het beleid: het zijn zij die netwerken moeten uitbouwen, leiderschapskwaliteiten moeten ontwikkelen en zo meer om hun achterstand ten aanzien van mannen in te halen. Uit deze voorbeelden blijkt dat beleidsmakers gevangen zitten in hun referentiekader en respectievelijk in de ideeën dat zorgtaken in eerste instantie een vrouwenaangelegenheid zijn en dat de ideale politicus een mannelijke levensstijl heeft. Deze kritische analyses mogen echter niet gelezen worden als het verwerpen van het geloof in de intentionaliteit van beleid en in beleid als instrument voor het realiseren van gendergelijkheid. Wel wijzen deze kritieken erop dat (gelijke-kansen)beleid naast intentionele ook niet intentionele elementen bevat.

Een tweede nuancering van de rationaliteit van het beleidsproces heeft betrekking op de connecties tussen de diagnose van het probleem, de remedie voor het probleem en de middelen die men ervoor inzet. In het gelijke-kansenbeleid zijn de probleemstelling en de doelstelling vaak niet op elkaar afgestemd, ontbreekt het verband tussen de probleem- en doelstelling enerzijds en de doelgroep van het beleid anderzijds, laat men na concrete middelen om het doel te bereiken te formuleren of worden de middelen ingezet voor een ander doel (Lombardo & Meier in Lombardo, Meier & Verloo, 2009; zie ook Verloo, 2007). Beleid rond huishoudelijk geweld, bij voorbeeld, stelt als doelstelling vaak voorop om het probleem beter te onderzoeken. Het beleid stopt echter vaak met het bestellen van onderzoek. Hoe wenselijk het ook is om beleid te baseren op een gedegen kennis van het betrokken maatschappelijke veld, het resulteert vaak in een soort uitstelgedrag dat het probleem niet verhelpt, zeker niet op korte termijn. In het aangehaalde voorbeeld van genderongelijkheid in de politieke besluitvorming, wordt de ondervertegenwoordiging van vrouwen geproblematiseerd en focussen de maatregelen op vrouwen. Hierdoor blijft een belangrijk deel van het probleem en de oplossingen buiten beeld, met name de androcentrische norm en de oververtegenwoordiging van mannen in de politiek evenals maatregelen die het gedrag van mannen als deel van de oplossing viseren. Dit gebrek aan rationaliteit kan een bewuste keuze zijn, bijvoorbeeld om zodoende een beleid inzake gendergelijkheid te ontwikkelen binnen de grenzen van het politiek haalbare, met andere woorden zonder echt aan de machtsverhoudingen tussen de seksen te raken. De oorzaak van het gebrek aan een logische, rationele koppeling van oorzaak, middelen en doel kan echter ook deels toegeschreven worden aan het eerder vastgestelde feit dat beleid ook niet bedoelde elementen bevat. Wat de oorzaak ook moge zijn, beleid, en misschien gelijke-kansenbeleid bij uitstek, is op een microniveau minder rationeel dan op een macroniveau: op een algemeen niveau gaat het ervan uit dat beleid een instrument is om gelijkheid te realiseren, maar in de concrete uitwerking en implementatie zien we dat een en ander schort

aan de connectie tussen oorzaak, doel en middelen. Een bijkomende verklaring voor deze discrepantie kan bestaan in de aanwezigheid van verschillende en botsende rationaliteiten op de verschillende niveaus. Zo kan het zijn dat de intenties en logica van de initiator van het beleid niet stroken met die van de beleidsuitvoerders. Dat kan de laatsten ertoe aanzetten om bijvoorbeeld een minimale uitvoering te geven aan een grotere ambitie, wat tot de hoger beschreven discrepanties kan leiden.

Ondanks deze nuances bij het gelijke kansenbeleid, houdt de these stand dat gender mainstreaming uitgaat van en zich inschrijft in een logica van rationele en intentionele beleidsprocessen. In de volgende sectie analyseren we hoe Belgische politici en beleidsmensen de gender mainstreamingstrategie operationaliseren. Hierbij focussen we op de vraag of en in welke mate beleidsmonitoring en -evaluatie daarbij in plaats van middel het doel werd. Wat in deze bijdrage geproblematiseerd wordt, is niet dat gender mainstreaming sterk intentioneel en rationeel is op de manier zoals die hierboven beschreven werd. Integendeel, aan de hand van een beschrijvende analyse trachten we in te schatten of ze haar ambities terzake waarmaakt: is het gender mainstreamingsbeleid wel degelijk gericht op het realiseren van gendergelijkheid? Is de rationele aanpak van het beleid gericht op dat doel? Wat dus als problematisch beschouwd wordt, is wanneer de intentionaliteit afgeleid wordt van het doel van het beleid (gendergelijkheid) naar het beleid zelf, meer bepaald naar het voldoen aan de formele vereisten die het rationele beleidsproces vereist.

### **3. De vertaling van gender mainstreaming naar de Belgische beleidspraktijk**

De Belgische overheden hebben sinds de vierde wereldvrouwenconferentie een aantal initiatieven genomen betreffende gender mainstreaming. In deze sectie bespreken we in chronologische volgorde de mijlpalen ervan. Meer punctuele initiatieven inzake gender mainstreaming, zoals het laten screenen op hun gendereffecten van beleidsbrieven, -nota's en wetsontwerpen via een emancipatie-effect-rapportage of gendertrainingen voor beleidsmedewerkers laten we buiten beschouwing.

#### **3.1. De Pekingwet van 1996 (en het Vlaamse Pekingdecreet van 1997)**

De Belgische federale overheid ondernam een eerste stap in het implementeren van gender mainstreaming in 1996, toen het parlement een wet goedkeurde over de toepassing van het VN Peking Actieplatform.<sup>72</sup> De zogenaamde Pekingwet legt een jaarlijkse rapportering op van de regering aan het parlement over de vorderingen die gemaakt worden in de toepassing van de resoluties van het Peking Actieplatform. Gezien deze resoluties zich niet beperken tot een klassiek sectorbeleid gelijke kansen maar heel uiteenlopende beleidsterreinen beslaan, kan het Peking Actieplatform als een eerste aanzet tot gender mainstreamingsbeleid beschouwd worden. De resoluties van het Peking Actieplatform beslaan uiteenlopende beleidsterreinen

---

<sup>72</sup> Wet van 6 maart 1996 strekkende tot controle op de toepassing van de resoluties van de Wereldvrouwenconferentie die van 4 tot 14 september 1995 in Peking heeft plaatsgehad, BS 31-10-1996.



zoals armoedebestrijding, onderwijs, gezondheidszorg, geweld tegen vrouwen, gewapende conflicten, de economie, mensenrechten, media, milieubeleid, zeggenschap en besluitvorming, institutionele mechanismen ter verbetering van de positie van de vrouw en als laatste prioritair interventiedomein de positie en levensomstandigheden van meisjes. Ook nieuw is dat de wet de verschillende bevoegde ministers verplicht om hierover verslag uit te brengen aan de minister van gelijke kansen die deze gegevens vervolgens aan het parlement presenteert. In de jaarlijkse rapportering moeten de ministers toelichten welke vorderingen ze gemaakt hebben op de twaalf actiedomeinen die in het Peking Actieplatform bepaald werden. De andere overheden gingen eveneens over tot het invoeren van dergelijke Pekingrapporten. De Vlaamse overheid legde ze in 1997 op, de Waalse regering en die van de Franse Gemeenschap volgden in 2002 en Brussel in 2006.<sup>73</sup> We zullen de laatste drie later bespreken gezien ze van een iets andere orde zijn. Het Vlaamse decreet behandelen we samen met de federale Pekingwet.

De federale en Vlaamse Pekingrapporten en de eisen, die eraan worden gesteld, leiden niet noodzakelijk tot gender mainstreaming gezien er geen enkele verbintenis is wat de eigenlijke realisaties op het terrein betreft. De verplichtingen beperken zich immers tot een opvolging van de resoluties van het Peking Actieplatform; het voeren van een gender mainstreamingsbeleid, dat gender gelijkheid wil realiseren, is niet vereist. De wet en het Vlaamse decreet verplichten de overheid enkel tot het rapporteren over de vooruitgang die ze boekte, niet om effectief vooruitgang te boeken. De wet bevat enkel technische en organisatorische bepalingen met betrekking tot de rapportering. Er wordt niet verwezen naar gendergelijkheid, in welke vorm dan ook, en welke vorderingen gemaakt moeten worden, worden evenmin geëxpliciteerd. Ook wordt in de praktijk vaak niet vertrokken vanuit de resoluties van het Peking Actieplatform die dan vervolgens omgezet worden in strategische doelstellingen en concrete beleidsmaatregelen. In de verplichte rapportering werd eerder andersom gewerkt waarbij al genomen beleidsinitiatieven gekaderd werden in de resoluties van het Peking Actieplatform. De wet van 1996 en het decreet van 1997 werd met andere woorden vaak aangewend om bestaande beleidsinitiatieven te rechtvaardigen en niet om nieuwe te ontwikkelen. In de praktijk bleef de toepassing van de resoluties van het Peking Actieplatform dan ook vaak beperkt tot een rapportering en zelfs die verliep moeizaam. Het kostte de minister bevoegd voor het gelijke kansenbeleid als de verantwoordelijke voor het samenstellen en presenteren van het rapport en zijn/haar kabinet en administratie veel energie om de deelrapporten van de andere ministers te bekomen. Ten slotte werden de vermelde acties en maatregelen zelden voorzien van de nodige middelen. Al deze vaststelling samen brengen ons tot de conclusie dat de evaluatie van het beleid een finaliteit op zich werd en geen middel om tot een efficiënter beleid te komen dat een reële impact op de maatschappelijke verhoudingen tussen de seksen zou hebben.

---

<sup>73</sup> Vlaamse Gemeenschap: Decreet van 13 mei 1997 houdende de opvolging van de resoluties van de Wereldvrouwenconferentie die van 4 tot 14 september 1995 in Peking heeft plaatsgehad; Waals Gewest: Decreet van 28 november 2002 houdende de opvolging van de resoluties van de Wereldvrouwenconferentie van de Verenigde Naties die in Peking heeft plaatsgehad; Franse Gemeenschap: Decreet van 19 december 2002 houdende de opvolging van de resoluties van de conferentie van de Verenigde Naties over de Vrouwen te Peking; Brussels Hoofdstedelijk Gewest: Ordonnantie van 20 april 2006 betreffende het opstellen, door de Regering, van een jaarlijks evaluatieverslag over haar beleid inzake gelijkheid tussen mannen en vrouwen.

### 3.2. De Vlaamse emancipatie-effect-rapportage van 1997 en 2000

Een van de eerste onderzoeksopdrachten die Gelijke Kansen Vlaanderen, de Vlaamse overheid ter bevordering van gelijke kansen, in 1996 uitbesteedde, was de ontwikkeling van een emancipatie-effect-rapportage (EER). De eerste rechtstreekse verkiezing van de deelstaatparlementen in 1995 en de daarmee gepaard gaande regeringsvorming viel min of meer samen met de vierde Wereldvrouwenconferentie te Peking. De Vlaamse regering richtte toen, in tegenstelling tot de andere deelregeringen en wellicht geholpen door het samensmelten van gewestelijke en gemeenschapsinstellingen, een overheidsadministratie voor het gelijke kansenbeleid in. Sinds 1995 is tevens een van de ministers van de Vlaamse regering belast met de portefeuille van gelijke kansen en heeft daar ook beleidsmedewerkers voor. Mede door de invloed van de Wereldvrouwenconferentie werd gelijke kansen in Vlaanderen van meet af aan als een transversale of horizontale beleidsbevoegdheid gedefinieerd; gender gelijkheid moest via de mainstream van het Vlaamse beleid bevorderd worden (Celis & Meier, 2007a). Ook de EER was een van de innovatieve instrumenten die in Peking werden geïntroduceerd en die in Vlaanderen werden overgenomen.

Zweden en Nederland hadden in Peking de EER op de agenda gezet. EER is een ex ante evaluatie van beleid, waarbij, analoog aan een milieu-effect-rapportage, de effecten van een beleidsvoornemen op de maatschappelijke verhoudingen tussen de seksen ingeschat worden. Via een gerichte ontleding van een beleidsvoornemen en analyse van de doelstellingen, maatregelen, doelgroepen en middelen worden verborgen discriminaties opgespoord. Een EER achterhaalt discriminaties die kunnen stoeien op verkeerde veronderstellingen over de maatschappelijke verhoudingen tussen mannen en vrouwen, maar ook op vooroordelen en stereotypen. Een EER gaat dus deels op zoek naar eventuele niet-intentionele elementen in een beleidsvoornemen. Het gaat erom de eventuele genderbias in te schatten. Aan de hand daarvan kan een beleidsvoornemen, al naar gelang de resultaten van de EER, terug afgevoerd of verder gezet worden, waarbij in het laatste geval de mogelijkheid tot herziening bestaat. Een gerichte toepassing van dat instrument tijdens het proces van beleidsontwikkeling vergemakkelijkt de bevordering van gendergelijkheid op de meest uiteenlopende beleidsterreinen. Nederland ontwikkelde begin jaren 1990 een dergelijk instrument (Verloof & Roggeband, 1994) en de Vlaamse overheid wilde een soortgelijk instrument laten ontwikkelen om het Vlaamse gender mainstreamingsbeleid concreet vorm te geven. Een eerste versie uit 1997 (Meier & Woodward, 1997) werd in 2000 (Meier e.a., 2000) geactualiseerd. Daarnaast werd een lokale variant van de EER ontwikkeld, de LEER (Meier, Meyntjens & Vander Steene, 1999).

In tegenstelling tot de Pekingwet en het Pekingdecreet zit in de EER expliciet de doelstelling van het bevorderen van gendergelijkheid vervat. Beleidsvoornemens worden afgetoetst aan de mate waarin ze tegemoet komen aan de doelstellingen van het gelijke kansenbeleid, of tenminste aan de doelstellingen zoals ze geëxpliciteerd werden in het instrument. In tegenstelling tot de Nederlandse EER tracht de Vlaamse variant tevens gendergelijkheid te verankeren in het beleidsvoornemen. Het Nederlandse instrument belet, in tegenstelling tot de Vlaamse EER, niet om, indien er discriminaties worden vastgesteld, verder te gaan met het beleidsvoornemen.



Beide instrumenten zijn duidelijk gekaderd in een rationele visie op het verloop van de beleidscyclus (het opstellen, screenen, aanpassen, uitvoeren van beleid en het op die wijze bereiken van doelstellingen) en het beleid op zich (de conceptuele zuiverheid van beleid en het op elkaar afstemmen van doel, maatregelen, middelen en doelgroep(en)). Beleidsevaluatie staat ook duidelijk centraal in deze gender main streamingsinstrumenten. Zo gaat met de EER altijd de vraag gepaard naar volgens welke gesegregeerde data, vanuit de optiek dat zonder een gedegen kennis van het beleidsterrein geen adequaat beleid mogelijk is. De beleidsevaluatie trapt hier echter niet in de val van het weten en meten als doel. Het eigenlijke doel, met name meer gendergelijkheid, wordt hier niet uit het oog verloren. Het maken van beleid (met beleidsevaluatie als middel) staat centraal.

Wel is de Vlaamse EER (evenmin als zijn Nederlandse tegenhanger) niet verankerd in de beleidscyclus van de Vlaamse overheid. Er bestaat geen verplichting om beleidsvoornemens door te lichten via een EER. Het instrument is in geen enkele situatie verplicht. In de praktijk geven vooral beleidsverantwoordelijken die meer dan gemiddeld begaan zijn met gender gelijkheid de opdracht om een bepaald beleidsvoornemen te laten doorlichten via een EER en wordt dit instrument vervolgens toegepast op de beleidsteksten die er het minst nood aan hebben. De EER is tot heden voor een groot deel dode letter gebleven. Het instrument wordt als te moeilijk en tijdrovend beschouwd en dat geldt ook nog voor de in 2000 herziene versie. Vooral de kwalitatieve criteria van gender gelijkheid waaraan beleidsvoornemens afgemeten worden, vereisen een hoge genderdeskundigheid en worden dan ook vaak te moeilijk bevonden.

Een sombere tussentijdse conclusie lijkt dan ook dat wanneer een gender main streamingsinstrument daadwerkelijk gendergelijkheid wil en kan realiseren, het niet verplicht en bijgevolg niet toegepast wordt. Dit sluit aan bij een eerdere vaststelling dat in het gelijke kansenbeleid vaak een compromis wordt gezocht tussen wat men op een discursief niveau wil bereiken en de ruimte die de politieke wil creëert om dat daadwerkelijk te bereiken. EER werd ontwikkeld en ter beschikking gesteld door de overheid maar deze verplicht het gebruik ervan niet.

### **3.3. De federale Cel Gender Mainstreaming van 2001 en 2002**

In de jaren 1990 verzette de federale Minister van Gelijke Kansen zich tegen een transversale aanpak van het gelijke kansenbeleid. Daar kwam verandering in met de ministerwissel in 1999. In 2000 groeide op het kabinet van gelijke kansen het idee om in de schoot van de federale regering een breed pilootproject rond gender mainstreaming te starten. Eind 2000 besloot de Ministerraad dat elke Minister een strategische doelstelling met betrekking tot de gelijkheid van mannen en vrouwen diende te formuleren. Voor alle twintig beleidsdomeinen werd minstens één strategische doelstelling bepaald, die vervolgens vertaald werd naar concrete beleidsmaatregelen via een acht-stappenplan. In de eerste stap werden een gemeenschappelijke strategie en referentiekader ontwikkeld. Hierin kaderde het formuleren van de strategische doelstelling. Vervolgens werd de strategische doelstelling ontleed in concrete subdoelstellingen. In de derde stap werden de aan te wenden middelen bepaald. In de vierde en vijfde stap werden een actieplan

ontworpen en aanbevelingen geformuleerd ter uitvoering van het actieplan. Stap zes bestond uit een terugkoppeling naar de hogere hiërarchie, waarin de goedkeuring van het actieplan door de verantwoordelijken, in dit geval de bevoegde minister, centraal stond. Vervolgens werd in de zevende stap het actieplan door de betrokken partijen uitgevoerd en in de achtste en laatste stap werden de genomen maatregelen opgevolgd en hun impact geëvalueerd (Cel Gender Mainstreaming, 2003). Elke minister duidde (eventueel per ministeriële bevoegdheid) een lid van zijn of haar kabinet en administratie aan die instond voor de opvolging van het project en die toezag op de realisatie van de strategische doelstelling. Daarnaast werd voor de gehele periode van het project de Cel Gender Mainstreaming in het leven geroepen die bestond uit academici. Hun taak bestond in het aanreiken van deskundigheid op het vlak van gender mainstreaming en met betrekking tot de verschillende beleidsdomeinen, en het ondersteunen van de beleidsmedewerkers van de kabinetten en van de administraties. De methode in acht stappen was een hulpmiddel die de Cel uitwerkte om het proces van gender mainstreaming te operationaliseren.

Een apart onderdeel van het pilootproject was aan “gender budgeting” of gendergevoelig budgetteren besteed (Holvoet, 2007). Gendergevoelig budgetteren is een vorm van gender mainstreaming dat focust op het budgettaire aspect van het beleid. Het is een analyse vanuit het genderperspectief van alle vormen van overheidsuitgaven en publieke inkomsten en geeft een overzicht van de directe en indirecte gevolgen hiervan voor vrouwen en mannen. Volgens de definitie van de Raad van Europa is gendergevoelig budgetteren *‘an application of gender mainstreaming in the budgetary process. It means a gender-based assessment of budgets, incorporating a gender perspective at all levels of the budgetary process and restructuring revenues and expenditures in order to promote gender equality’* (Council of Europe, 2005, 10). Net zoals gender mainstreaming goed aansluit bij modernisering van het beleidsproces laat gendergevoelig budgetteren zich vooral implementeren in het kader van hervormingen die prestatiegericht begroten beogen (Holvoet, 2007; Morre, 2008). Het pilootproject inzake gendergevoelig budgetteren volgde niet het hierboven beschreven stappenplan, maar bestond uit een trainingsluik, een actieonderzoek (waar met de betrokken ambtenaren een proces van gendergevoelig budgetteren werd doorlopen) en informatieverbreiding.

Het pilootproject had een initiële looptijd van een jaar, die nadien met een jaar verlengd werd aangezien men na het eerste jaar vaststelde dat de strategische doelstellingen verre van bereikt waren. Maar zelfs na het tweede jaar werd een groot aantal doelstellingen uiteindelijk niet gerealiseerd. Het project werd evenwel niet opnieuw verlengd, onder andere omdat in 2003 federale verkiezingen plaatsvonden. De meeste initiatieven stopten op het punt waar acties concreet uitgevoerd moesten worden. Ook van een opvolging van de acties en een evaluatie van hun impact was nauwelijks sprake (Cel Gender Mainstreaming, 2003). Voor de gebrekkige implementatie zijn drie redenen aan te duiden. Een eerste struikelblok vormden de strategische doelstellingen. Deze waren in de meeste gevallen al geformuleerd voordat de Cel Gender Mainstreaming haar werkzaamheden aanpakte. Sommige strategische doelstellingen getuigden eerder van een positieve actie- dan van een gender mainstreaminglogica; andere beperkten zich tot het scheppen van

voorwaarden voor een gender mainstreaming beleid dat in een latere fase zou moeten plaatsvinden. Daarnaast waren de strategische doelstellingen in het merendeel van de gevallen niet aangepast aan het voor het project voorziene tijdsbestek. De bijsturing van de strategische doelstellingen lag echter politiek bijzonder moeilijk. Bovendien was de aanpak van de Cel uitgesproken rationalistisch omdat het stappenplan zich enkel richtte op het beleidsproces en niet op de inhoud ervan. Zo hielp het stappenplan op zich niet om de doelstellingen in de juiste richting te duwen maar werd hiervoor gerekend op de leden van de cel zelf. Een tweede struikelblok voor de implementatie was de taakverdeling tussen en positie van diegenen die in de praktijk verantwoordelijk waren voor het project. Binnen de kabinetten en administratie werd geen extra tijd vrijgemaakt voor het pilootproject waardoor de experts van de eigenlijke Cel als welkome extra werkkracht gezien werden. Deze verwachtingen overstegen echter ruimschoots het tijdsbestek en de bevoegdheden van de experts. Ook was hun kennis van de verschillende beleidsdepartementen soms te beperkt. Daarnaast waren de personen die op de kabinetten en in de administratie belast waren met het project, afhankelijk van de minister en vaak ook van de kabinetschef, waardoor het moeilijk was de nodige stappen te ondernemen, vooral in die gevallen waar er weinig medewerking was van hogerhand. Een derde struikelblok, ten slotte, was het klassieke gebrek aan andere middelen, zoals financiële, om de strategische doelstellingen tot een goed einde te brengen.

De evaluatie van het pilootproject inzake gendergevoelig budgetteren wees op een aantal bijkomende hindernissen voor een succesvolle implementatie van gendergevoelig budgetteren. Zo wijst Holvoet (2007) op een kloof tussen de wereld van de genderexperten en de begrotingsspecialisten. De eersten waren overwegend vrouwelijk en de tweeden overwegend mannelijk en ook hun visie op gelijkheid en de mate waarin budgettaire processen ingezet kunnen worden in een beleid dat gendergelijkheid beoogt, verschilde fundamenteel. Voor de ambtenaren kwam het streven naar gelijkheid in hoofdzaak neer op het vermijden van directe discriminatie van individuele personen. Het systematisch inschakelen van genderanalyses van bij de start van het beleidsproces, i.c. inzake budgetten en begroting, ter preventie van genderongelijkheid werd door de ambtenaren minder aanvaard; voor hen was het voldoende dat ex post geëvalueerd werd of er sprake was van discriminatie. Holvoet onderlijnt verder dat het exclusief situeren van gendergevoelig budgetteren bij de overheid –wat het geval is in het expert-bureaucratisch model van gender mainstreaming– het onwaarschijnlijker maakt dat er een macro-economische analyse gemaakt wordt die kritischer staat ten opzichte van dominante visies op gender die verweven zijn met het algemene economische beleid van een land. Door enkel ambtenaren te betrekken in het proces van gendergevoelig budgetteren, wordt gender mainstreaming ingebed op een meso- of microniveau waardoor er meer nadruk gelegd wordt op gendergevoelig budgetteren in haar managementfunctie, namelijk als een onderdeel van het afleggen van interne verantwoording over het gevoerde beleid. Zo wordt bijvoorbeeld wel nagegaan wat het directe effect van een belastingshervorming op mannen en vrouwen is en kan men nagaan en/of bewijzen dat er al dan niet gediscrimineerd werd. Wat evenwel ongemoeid wordt gelaten is ondermeer de macro-economische effecten ervan op de besluitvorming en de verdeling van betaalde en onbetaalde arbeid over mannen en vrouwen binnen

gezinnen. Zoals ook Budlender (2000) en Sharp & Broomhill (2002) aangeven, heeft een dergelijke diepgaandere analyse die ook macro-economische visies blootlegt, andere actoren die het overheidsbeleid als een buitenstaander kunnen analyseren zoals middenveldorganisaties en onderzoeksinstellingen. Door het operationaliseren van gendergevoelig budgetteren als een intern managementinstrument (door enkel ambtenaren te betrekken in het beleidsproces) sluit het dus de mogelijkheid af om te functioneren als een instrument van transparantie en verantwoording ten aanzien van externe actoren zoals het parlement, het middenveld of supranationale instellingen als de Europese Commissie.

Niettemin is tot op vandaag het pilootproject van de Cel Gender Mainstreaming een van de meest concrete pogingen op het niveau van de federale overheid om de strategie van gender mainstreaming in te passen in de dagelijkse beleidspraktijk. Het project was duidelijk ingebed in een gender mainstreamingsperspectief. Zo werd gewerkt met de definities van de Raad van Europa en van de Europese Commissie en werden alle beleidsdomeinen in beschouwing genomen. Zoals gender mainstreaming was opgevat in het stappenplan was het zeker de betrachting om verder te gaan dan een pro forma formuleren van doelstellingen aangezien ook de middelen voorzien dienden te worden en het proces moest resulteren in een evaluatie van de impact van het beleid. De doelstelling van het beleid is dus wel degelijk meer gendergelijkheid realiseren. Het betrof echter een pilootproject, en geen wettelijke verplichting. Het belangrijkste element dat ons tot de conclusie brengt dat ook in het pilootproject gender mainstreaming de beleidsevaluatie een doel op zich is eerder dan een instrument ter verbetering van het beleid, is dat er geen consequenties verbonden zijn aan de evaluatie. Dat sluit aan bij de vaststelling dat het stappenplan voornamelijk formalistisch was en zich op het beleidsproces richtte eerder dan op de inhoud van de doelstellingen die het trachtte te realiseren. Een expliciete koppeling van de impactmeting van het beleid aan de volgende strategische doelstellingen lijkt ons echter een belangrijke voorwaarde voor beleidsevaluatie die gericht is op beleidsvorming met het realiseren van meer gendergelijkheid als doel. Ook de vaststelling dat de oefening inzake gendergevoelig budgetteren eerder als een intern managementinstrument om discriminatie te vermijden werd beschouwd dan als een instrument om preventief macro-economische genderongelijkheden te voorkomen of weg te werken schraagt de conclusie dat de interesse van de beleidsactoren eerder betrekking heeft op het beleid en de beleidsinstrumenten op zich en minder op beleidsvorming voor het realiseren van een ruimer doel, met name gendergelijkheid.

### **3.4. De Pekingdecreten en –ordonantie van 2002 en 2006**

De decreten van de Waalse overheid en van de Franstalige Gemeenschap van 2002 en de Brusselse ordonnantie van 2006 ter opvolging van het Peking Actieplatform leggen een vorm van Pekingrapportages op die minder vrijblijvend zijn dan die op Vlaams en federaal niveau. Zo moet elke minister van de Waalse regering informatie verstrekken over de strategische doelstellingen van zijn of haar departement betreffende de gelijkheid van vrouwen en mannen, over de daarvoor uitgetrokken financiële middelen, over de uitvoering van de doelstelling en de evaluatie ervan,

evenals over de contactpersoon die op het kabinet of in het departement belast is met deze materie. Ministers van de Franse Gemeenschapsregering en van de Hoofdstedelijke Regering worden tevens expliciet gevraagd de moeilijkheden die ze ondervonden in de verwezenlijking van hun doelstellingen en de vooruitzichten om deze te verhelpen op te nemen in hun rapportering.

Net als in Vlaanderen en op federaal niveau verliep de totstandkoming van de Pekingrapporten moeizaam. Het recente proces van ordonnantie naar het eerste Pekingrapport in het Brussels Hoofdstedelijk Gewest kan dat illustreren (Cortier & Celis, 2008). Reeds begin 2006 kreeg de Cel Gelijke Kansen in het ministerie van het Brussels Hoofdstedelijk Gewest de opdracht een werkmethode voor de realisatie van het Pekingrapport op te maken. Er werd een gids opgesteld, die het ruimer kader rond de internationale akkoorden en de Brusselse verplichtingen schetst, en er werd een praktische werkwijze met stappenplan voorzien voor de toepassing van de theorie op de verschillende gewestelijke bevoegdheden (inclusief concrete voorbeelden ter inspiratie). Een officiële vraag vertrok vervolgens naar de ministers en staatssecretarissen om in het eigen kabinet een persoon aan te stellen die de opmaak van dit evaluatieverslag voor de bevoegdheden van zijn/haar minister zou opvolgen. Op 4 december 2006 werd een interkabinettenwerkgroep ingericht bestaande uit de contactpersonen van de kabinetten. Op deze vergadering werd de methodologie toegelicht, de structuur voor de bijdragen verduidelijkt en een eerste deadline voor het indienen van de bijdragen vooropgesteld. Niet alle ministers/staatssecretarissen waren echter vertegenwoordigd op deze informatievergadering en de eerste deadline voor de opmaak werd door bijna geen enkele minister gehaald. Het tijdig verkrijgen van de bijdragen van alle regeringsleden bleek geen sinecure en uiteindelijk waren er een drietal deadlines, een telefoonronde vanuit de Cel Gelijke Kansen, een aantal officiële brieven vanuit het kabinet en een tweede interkabinettenwerkgroep (waar de al ingediende teksten becommentarieerd werden), nodig om van elk regeringslid een bijdrage te verkrijgen. De laatste bijdrage kwam de dag voor de overhandiging van het rapport aan de Regering bij de verantwoordelijke uit de Cel Gelijke Kansen toe.

De kwaliteit van de rapportages van de ministers is variabel. Meestal werd de bijdrage niet in beide landstalen voorzien en voldoet deze niet aan de structuur zoals die werd vooropgesteld door de Cel Gelijke Kansen (context – opgestarte acties – bereikte resultaten – toekomstige initiatieven; één tekst per beleidsdomein) noch aan de vereisten die de ordonnantie zelf oplegt, namelijk het vermelden van de uitgetrokken financiële middelen, de evaluatie van de doelstellingen, de naam van de contactpersoon op het kabinet of in de administratie belast met die aangelegenheden en het formuleren van de gerezen problemen en mogelijke oplossingen. Verder blijkt dat het merendeel van de bijdragen geen aparte tekst per bevoegdheid opnam, niet vanuit een beginanalyse vertrekt op basis waarvan een probleemstelling ontwikkeld wordt met hieraan gekoppelde beleidsacties en al eerder uitgewerkte algemene beleidsmaatregelen die in het Peking Actieplatform kaderen. Heel wat van de vermelde acties beperken zich tot dataverzameling en het bestellen van studies. Aan het aanpakken van de structurele kant van de ongelijkheid komen de meeste opgelijste initiatieven bijgevolg niet toe. Verder wordt er hoofdzakelijk op vrouwen

gefocus (en niet op genderverhoudingen) en beperkt men gelijkheid tot gelijke kansen en anti-discriminatie. Bovendien werden de acties die in het Pekingrapport vermeld werden niet altijd naar de administratie toe gecommuniceerd; hieruit blijkt weinig tot geen wil om deze acties ook daadwerkelijk te realiseren.

Deze korte schets van de totstandkoming en kwaliteit van het eerste Brussels Pekingrapport maakt duidelijk dat de leden van de Brusselse regering het beleidsinstrument en de beleidsevaluatie en –monitoring die erin vervat zit niet inzetten in een beleidsvorming gericht op de realisatie van gendergelijkheid. Een minimalistisch voldoen aan de verplichtingen opgelegd in de ordonnantie van 2006 bleek de drijfveer achter de bijdragen tot het Brusselse Pekingrapport van de meeste Brusselse ministers.

### **3.5. De Vlaamse Open Coördinatie Methode (2005)**

In 2005 werd een belangrijke impuls gegeven aan het Vlaamse gender mainstreamingsbeleid door de invoering van de open coördinatiemethode (OCM) (Celis & Cortier, 2008). Na een evaluatie van het tot dan toe gevoerde mainstreamingsbeleid werd beslist om een meer structurele verankering voor het voeren van een horizontaal beleid op te zetten dat duurzamer, meer continue en effectiever zou zijn. Daartoe werd in 2006 ook een interdepartementale Commissie Gelijke Kansen opgericht die tweemaandelijks samenkomt (Mertens & Steegmans, 2006). Ze heeft haar ontstaan te danken aan het Vlaams Ministerieel Comité voor Gelijke Kansen bestaande uit alle ministers uit de Vlaamse Regering dat zich bereid verklaarde om een gender mainstreamingsbeleid te voeren op basis van de OCM. De commissie wordt samengesteld uit ambtelijke vertegenwoordigers uit elk Vlaams beleidsdomein en stelt de jaarlijkse actieplannen ter realisatie van de afgesproken doelstellingen op en evalueert het proces (Celis & Meier, 2007b).

De OCM werd al eerder succesvol toegepast op Europees niveau in sociaal-economische beleidsmateries en wordt beschouwd als een werkwijze die “best practices” helpt verspreiden. Het bedeeft de minister van gelijke kansen, een coördinerende rol toe waarbij het vooral haar taak is om haar collega’s in andere beleidsdomeinen te stimuleren en te ondersteunen om initiatieven te nemen die gelijke kansen bevorderen. Concreet houdt OCM in dat de leden van de Vlaamse regering, gemeenschappelijke doelstellingen afspreken, waarna de verschillende betrokkenen vrij bepalen hoe ze die doelstellingen zullen nastreven en realiseren. Op regelmatige basis gebeurt monitoring, evaluatie en peer review zodat de verschillende partners van elkaar kunnen leren en elkaar kunnen stimuleren. Deze manier van werken garandeert een grote vrijheid om vanuit ieders eigenheid bijdragen te leveren binnen een gemeenschappelijk afgesproken en duidelijk kader. Kenmerkend is ook dat deze methode gezien wordt als een leerproces waarbij het uitwisselen van goede praktijken centraal staat. De betrokken beleidsactoren zijn de leden van de Vlaamse regering en de ambtenaren maar het is echter ook de bedoeling dat het middenveld en de academische wereld geconsulteerd wordt (gezien “open” voor de participatie van tal van actoren staat) (Mertens & Steegmans, 2006).



Gevoed door informatie vanuit het gelijke kansenmiddenveld en het Steunpunt Gelijke Kansenbeleid, een wetenschappelijk steunpunt dat beleidsondersteunend onderzoek uitvoert, én de expliciete vraag naar de kabinetten om hun “eigen achterban” te consulteren, vonden bilaterale gesprekken plaats tussen vertegenwoordigers van het Vlaamse gelijke kansenbeleid enerzijds en medewerkers van alle Vlaamse kabinetten anderzijds ter voorbereiding van het formuleren van de doelstellingen. In 2006 werden de doelstellingen goedgekeurd door het Vlaams Ministerieel Comité Gelijke Kansen en deze werden door de Commissie Gelijke Kansen vervolgens in actieplannen gegoten (Celis & Meier, 2007b; Mertens & Steegmans, 2006). De doelstellingen en actieplannen staan volledig in het teken van het dichten van de loon- en loopbaankloof (Steegmans, 2007). In 2007 zou een evaluatie van de ondernomen acties volgen evenals het opstellen van nieuwe actieplannen voor 2008. De evaluatie werd evenwel uitgesteld omdat de creatie van indicatoren meer tijd in beslag nam dan aanvankelijk werd ingeschat. Wel werd een onderwijsmonitor ontwikkeld die dieper ingaat op de indicatoren die voor gender in het onderwijs geformuleerd kunnen worden (Steegmans, 2007).

Het OCM-proces loopt dus niet altijd even vlot; de ministers waren aanvankelijk moeilijk te overhalen om er werk van te maken. Ter verhoging van het politieke draagvlak voor de OCM werd de procedure verankerd in het decreet voor het Vlaamse gelijkekansen- en gelijkebehandelingsbeleid (goedgekeurd door het Vlaams Parlement in juli 2008). Het decreet legt op dat de Vlaamse regering negen maanden na haar aantreden voor elk bevoegdheidsdomein strategische en operationele doelstellingen voor de gehele legislatuur vastlegt. Vervolgens heeft ze zes maanden de tijd om deze doelstellingen te vertalen naar een actieplan dat een contextanalyse, concrete acties, een timing, indicatoren, budgetten en instrumenten omvat. Deze actieplannen dienen minimaal tweejaarlijks herzien te worden.

Het hanteren van de OCM als gender mainstreamingsinstrument wordt als een aanzienlijke vooruitgang ten aanzien van de Pekingrapporten beschouwd.<sup>74</sup> Het vormend en sensibiliserend effect dat van de OCM uitgaat, maakt dat stapgewijze de beleidsmakers komen tot meer performantie beleidsacties in hun beleidsdomein ter realisatie van gendergelijkheid. De beleidsevaluatie op basis van de verslagen van de actieplannen die in het proces aanwezig is, wordt in eerste instantie opgevat als een moment om te leren en vervolgens efficiëntere doelstellingen en actieplannen te formuleren. De format van de fiches voor de opmaak van de verslagen helpen hierbij aangezien de bereikte resultaten naast de intenties van het actieplan opgelijst moeten worden. Eventuele discrepanties tussen ambitie en bereikte resultaten moeten bovendien ook toegelicht worden. Het hanteren van genderindicatoren zou in de toekomst de evaluaties ook minder vrijblijvend maken. Verder maken de tweemaandelijks samenkomsten van de Commissie Gelijke Kansen het moeilijk om zich achter formaliteiten te verbergen. De schriftelijke output (doelstellingen, actieplannen en evaluaties) zijn niet de enige momenten waarop men engagement verwacht van de beleidsactoren. Het is echter eigen aan een “soft tool” als de OCM dat er geen sancties voorzien zijn of geen consequenties worden verbonden aan negatieve resultaten (bijvoorbeeld onder de vorm van het stopzetten van een bepaald

---

<sup>74</sup> Gesprek met Agna Smisdom, OCM-verantwoordelijke van de Cel Gelijke Kansen van het Ministerie van de Vlaamse Gemeenschap, op 8 juli 2008.

beleid). Ook is het welslagen van de beleidsacties afhankelijk van de middelen die de bevoegde ministers binnen hun eigen bevoegdheid willen vrijmaken voor het realiseren ervan. Veel hangt dus af van de wil en de gedrevenheid van de betrokken beleidsactoren, zowel van de trekkers van de OCM (de gelijke-kansenambtenaren en de minister voor Gelijke Kansen) en van de beleidsactoren in de respectievelijke beleidsdomeinen. In de huidige toepassing van de OCM is de beleidsevaluatie en -monitoring wel degelijk ingeschakeld in de beleidsvorming voor het bereiken van meer gendergelijkheid; er zijn evenwel weinig garanties dat dat in de toekomst ook zal blijven.

### **3.6. De wet gender mainstreaming van 2007**

De laatste jaren werd herhaaldelijk geopperd dat de verslaggeving over de opvolging van de resoluties van het Peking Actieplatform niet volstaat en dat er nood is aan een performanter gender mainstreamingsbeleid onder de vorm van een EER en/of gendergevoelig budgetteren. De federale regering volgde daarbij steeds de lijn van een herziening van de wet uit 1996 die de Pekingrapportages invoerde. Het Instituut voor de Gelijkheid van Vrouwen en Mannen bereidde een wetsontwerp voor ter vervanging van de wet van 1996 wat uiteindelijk resulteerde in de wet van 12 januari 2007.<sup>75</sup>

Deze wet voorziet in een duidelijke identificatie van de overheidskredieten die bedoeld zijn voor de bevordering van de gelijkheid van mannen en vrouwen en legt zodoende de wettelijke basis voor gendergevoelig budgetteren. Voorts legt het een “gendertoets” op voor elke nieuwe beleidsmaatregel en dient de genderdimensie beter verankerd te worden in de federale beleidslijnen. De wet verplicht de federale regering om bij het begin van de legislatuur, in het kader van de regeringsverklaring, voor het geheel van de beleidsdomeinen, strategische doelstellingen vast te leggen voor het bevorderen van de gelijkheid van mannen en vrouwen. Het gaat daarbij vooral om het verwezenlijken van doelstellingen zoals die in het Peking Actieplatform werden bepaald. De wet voorziet tevens in de opvolging en evaluatie van deze acties. Ministers dienen relevante genderindicatoren vast te leggen die toelaten om te meten in welke mate de vooropgestelde strategische doelstellingen gehaald worden. Ze dienen ook jaarlijks te rapporteren over de acties, maatregelen en projecten die moeten bijdragen tot de realisatie van hun strategische doelstellingen. De oorspronkelijk jaarlijkse verslaggeving over de opvolging van de resoluties van het Peking Actieplatform wordt als gevolg daarvan teruggebracht tot een tweejaarlijkse rapportering. Eventuele knelpunten bij de verwezenlijking van de strategische doelstellingen en hoe ze te overkomen moeten vermeld worden in de tussentijdse evaluaties. In de evaluatie op het einde van de legislatuur moet de situatie van mannen en vrouwen vergeleken worden met die aan het begin van de legislatuur. Zo wordt nagegaan in welke mate de initieel geformuleerde strategische doelstellingen ook effectief gehaald werden.

---

<sup>75</sup> Wet van 12 januari 2007 strekkende tot controle op de toepassing van de resoluties van de wereldvrouwenconferentie die in september 1995 in Peking heeft plaatsgehad en tot integratie van de genderdimensie in het geheel van de federale beleidslijnen, BS 13-02-2007.



In vergelijking tot de wet van 1996 biedt de nieuwe wet op twee vlakken een aanzet tot het voeren van een substantiëler gender mainstreamingsbeleid: er is een concretere doelstelling waardoor de wet het louter technisch-formalistische niveau overstijgt en er zijn duidelijkere uitvoeringsmodaliteiten.

De nieuwe wet voorziet niet enkel in een opvolging van de doelstellingen van het Peking Actieplatform, maar ook in een *'integratie van de genderdimensie in het geheel van haar beleidslijnen, maatregelen, begrotingsvoorbereidingen of acties en dit met het oog op het vermijden of corrigeren van eventuele ongelijkheden tussen vrouwen en mannen'* (art. 2 § 1). Ook moet, in navolging van het pilootproject Cel Gender Mainstreaming, aan het begin van de legislatuur elke minister strategische doelstellingen vastleggen. Men zou kunnen argumenteren dat de resoluties van het Peking Actieplatform op zich al doelstellingen aangeven. Dat zijn er echter veel en in dit opzicht is het niet slecht dat de regering aan het begin van de legislatuur kleur bekent met betrekking tot haar prioriteiten. Dit biedt aan het einde van de legislatuur een toetssteen voor het gevoerde beleid. Alles staat of valt echter met de definitie van de strategische doelstellingen. De ervaring van de Cel Gender Mainstreaming heeft uitgewezen dat dit geen evidente opdracht is. Het Instituut voor de Gelijkheid van Vrouwen en Mannen, dat belast wordt met de begeleiding en ondersteuning van de integratie van een genderdimensie, en de functie vervult die de Cel Gender Mainstreaming had in het pilootproject, kan theoretisch gezien een belangrijke rol spelen op dit vlak. De vraag is echter in welke mate het die in de praktijk kan waarmaken (Celis & Meier, 2006).

In tegenstelling tot de wet van 1996 bevat die van 2007 een aantal aanzetten tot het concreet implementeren van een gender mainstreamingstrategie. Zo moet ten eerste elke minister genderindicatoren goedkeuren om de integratie van een genderdimensie en de realisatie van de strategische doelstelling te meten. Ten tweede is er een soort emancipatie-effect-rapportage voor elk wet- en regelgevend initiatief. Het model voor die gendertoets moet echter nog vastgelegd worden bij Koninklijk Besluit en zoals bij de invulling van de strategische doelstellingen hangt ook hier veel af van de wijze waarop wetgevende of reglementaire project doorgelicht zullen worden op hun gendereffecten en vooral wat de gevolgen van die doorlichting zijn, met name of er in het geval men genderongelijkheid detecteert ook effectief de wet of het reglement bijstuurt. Dat zal bepalen of de nadruk zal komen te liggen op beleidsvorming (met beleidsevaluatie als middel) dan wel op beleidsevaluatie zonder terugkoppeling naar de beleidsvorming.

Ten derde besteedt de wet van 2007 ook aandacht aan de middelen. Deze inbedding van gendergevoelig budgetteren in een algemeen gender mainstreamingsbeleid wordt in de literatuur beschouwd als een belangrijke bron van succes (Holvoet, 2007; Sharp & Broomhill, 2002). De voor de verwezenlijking van de strategische doelstellingen voorziene kredieten worden toegelicht in een gendernota die bij het ontwerp van de algemene uitgavenbegroting gevoegd wordt. Het voorstel voor de gendernota van het Instituut voor de Gelijkheid van Vrouwen en Mannen omvat twee luiken: een overzicht van de budgetten voorzien voor specifieke acties ten voordele van vrouwen en een genderanalyse van de budgetten. Deze analysefase voorziet eerst in een "quick scan", een vlugge toetsing waarbij de dossierbeheerder

evalueert of de maatregel, actie of project al dan niet op korte termijn een effect zal hebben op gender. De quick scan wordt uitgevoerd op basis van de reeds gekende feiten (wetenschappelijk analyse of statistische informatie) of op basis van expertise binnenshuis of na consultatie van het Instituut voor de Gelijkheid van Vrouwen en Mannen. Indien hieruit blijkt dat het beleid een genderimpact heeft dient men over te gaan op een diepgaandere analyse. Deze “gender impact assessment” dient in kaart te brengen in welke mate mannen en vrouwen voordeel halen uit de overheidsuitgave of -inning en wat de impact is op de socio-economische situatie van mannen en vrouwen. Deze analyse dient te gebeuren aan de hand van kwalitatieve en kwantitatieve genderindicatoren. Dit voorstel roept twee kritische bedenkingen op. Ten eerste blijft de bekommernis die eerder werd geformuleerd naar aanleiding van de bespreking van het onderdeel van het pilootproject gender mainstreaming dat betrekking had op gendergevoelig budgetteren: in welke mate zal de analyse die in hoofdzaak door ambtenaren zal worden uitgevoerd een overheidskritische analyse kunnen realiseren die ook macro-economische assumpties inzake gender in vraag stelt? Veel valt met andere woorden af van de genderindicatoren die men in beschouwing neemt. Ten tweede blijft het heikele punt dat niet wordt bepaald wat de consequentie is van een negatieve gender impact assessment.

Ten vierde zet de wet van 2007 de verschillende ministers ertoe aan ervoor te zorgen dat de verzamelde of geproduceerde statistieken naar sekse worden uitgesplitst. Veel zal echter afhangen van de wijze waarop de wet naar de praktijk wordt vertaald. Het genereren van volgens sekse gesegregeerde data mag bij voorbeeld geen doelstelling op zich worden en de toekomst moet uitwijzen in welke mate deze data ingezet worden in het beleid om te komen tot een meer egalitaire samenleving.

#### **4. Verklaringen voor gender mainstreaming als een proces van meten en weten**

De beschrijving van de initiatieven die in België betreffende gender mainstreaming werden genomen, duidt er inderdaad op dat er een sterke neiging aanwezig is om het gelijke kansenbeleid te reduceren tot een formalistische oefening van meten en weten. De praktijk van de Pekingrapportages wijst uit dat de betrokken beleidsmakers deze beleidsstrategie niet aanwendden als een procedure om binnen hun beleidsdomein te werken aan effectieve gendergelijkheid. Ze lijken ze eerder te ervaren als een formele vereiste waaraan voldaan moet worden waarbij niet zelden bestaande beleidsinitiatieven, die soms ver verwijderd zijn van de gender mainstreaminglogica, naar voren werden geschoven. Het federale pilootproject gender mainstreaming had duidelijk de intentie om te komen tot een gender mainstreaming van beleid maar de daartoe bedachte formule was heel formalistisch en nam in veel gevallen de bovenhand op een inhoudelijke focus. De Vlaamse EER en LEER als meer dwingende en gerichte instrumenten om gendergelijkheid te realiseren misten dan weer de structurele inbedding (en een verplichting) om effect te ressorderen. De latere Pekingdecreten en -ordonnantie (2002 en 2006) en de federale gender mainstreamingwet (2007) combineren een doelgerichte aanpak en een structurele inbedding, maar ook hier dienen we te concluderen dat er tot nog

toe weinig tekenen zijn dat gender mainstreaming als hefboom wordt aangewend voor een substantieel gelijke kansenbeleid gericht op het realiseren van effectieve gendergelijkheid. Ook de formule voor de nieuwe Pekingrapporten blijft getuigen van een wens om in eerste instantie te voldoen aan formele vereisten. Het is uiteraard nog te vroeg om de federale gender mainstreamingwet te evalueren. De centrale vraag hier is of de strakke fasering en ingebouwde meetmomenten aan de hand van genderindicatoren de betrokken beleidsactoren al dan niet zal dwingen om een beleidsproces te doorlopen en bij te sturen met meer gendergelijkheid als doel en uitkomst. De ervaring met de gender mainstreaminginstrumenten als de Pekingrapporteringen en het pilootproject in het bijzonder voeden de vrees dat de gender mainstreamingwet van 2007 wel eens in eerste instantie de momenten vermenigvuldigt waarop beleidsactoren zich in regel moeten stellen met de formele vereisten betreffende rapportering en meting. De Vlaamse OCM lijkt deze vallen te vermijden, maar ook hier is het te vroeg om te stellen dat de beleidsevaluatie en -monitoring die erin vervat zit ook op langere termijn zullen resulteren in een beleid dat daadwerkelijk gendergelijkheid realiseert.

De vraag die zich vervolgens opdringt is waarom gender mainstreaming in de praktijk vaak verwordt tot een formalistisch proces waarbij het meten van gelijkheid en het monitoren en evalueren van beleid een doel op zich wordt en niet langer een middel om meer gelijkheid te bewerkstelligen. We zien hiervoor een viertal redenen. Twee ervan zijn terug te brengen op de implementatie van de beleidsstrategie in een daartoe "onaangepaste" beleidscontext; een derde en vierde verklaring zijn terug te brengen op het concept gender mainstreaming zelf, en meer bepaald op de afwezigheid van een concrete invulling van gelijkheid en op de nadruk op specifieke beleidsinstrumenten.

Gender mainstreaming vergt een implementatie in de reguliere beleidsprocessen en dat transversaal in verschillende beleidsdomeinen. Deze horizontale of transversale dynamiek van gender mainstreaming dient met andere woorden geoperationaliseerd te worden in traditioneel veeleer verticale beleidsstructuren. Om een houvast te bieden aan de actoren in deze beleidsstructuren die weinig ervaring en affiniteit hebben met een gelijke-kansenbeleid wordt sterk gefocust op het uittekenen van procedures die gevolgd moeten worden om tot het gewenste resultaat te komen. Dit blijkt duidelijk uit de ervaring van de federale Cel Gender Mainstreaming en de eerste Pekingrapporten. Zo diende de door de Cel Gender Mainstreaming uitgewerkte procedure in eerste instantie om in gendervraagstukken onervaren beleidsmensen te helpen met het realiseren van hun beleidsdoelstellingen. Ter begeleiding van deze ambtenaren werd sterk de nadruk gelegd op het rationele verloop van een beleidscyclus en op de interne coherentie van het beleid (met name de afstemming op elkaar van doelstellingen, doelgroepen en ingezette middelen). Ook andere instrumenten ter bevordering van gender mainstreaming (zoals de EER, LEER en de OCM) zijn instrumenten en processen die reguliere beleidsactoren in staat zouden moeten stellen om een gender mainstreamingsbeleid te voeren. Om deze beleidsprocedures en instrumenten die in het kader van gender mainstreaming werden uitgewerkt ook effectief in te zetten voor het bereiken van een ruimere doelstelling, namelijk gendergelijkheid realiseren (wat bijvoorbeeld verschilt van niet discrimineren), is evenwel genderexpertise vereist. Wanneer het gender mainstreamingsbeleid dat gevoerd wordt zich situeert

in het expert-democratisch model, dient die genderexpertise opgebouwd te worden bij de reguliere beleidsactoren. De genderexpertise wordt niet aangeleverd door de procedures en de instrumenten zelf, maar ze moet wel aanwezig zijn voor een succesvolle toepassing ervan. Genderdeskundigheid moet via training en vorming bij de politici en/of ambtenaren opgebouwd worden of door externen aangeleverd worden (meestal dient het tweede om te komen tot het eerste). In de praktijk blijkt die transfer en verinnerlijking van genderdeskundigheid niet gemakkelijk te verlopen. Wat bijgevolg na een eventuele fase van expertise-opbouw overblijft, zijn vaak enkel de procedures en instrumenten zonder een voldoende hoge mate van genderexpertise. Dat leidt dan uiteraard tot een toepassing van de procedures en instrumenten en het voldoen aan de minimale vereisten zonder het oogmerk van een streven naar reële gendergelijkheid.

Een tweede verklaring voor de vaststelling dat gender mainstreaming toegepast in de beleidspraktijk de neiging heeft om zich te beperken tot een focus op het monitoren en evalueren van het beleid is eveneens terug te brengen op de beleidscontext. Door het transversale opzet van gender mainstreaming zijn diegenen die gender gelijkheid bezielen niet diegenen die ze horen te bevorderen. De verantwoordelijkheid voor het beleid ligt dus bij andere diensten en personen dan diegenen die het beleid dienen te implementeren en vaak is er ook geen duidelijke hiërarchische relatie tussen beiden. De OCM is daar een duidelijk voorbeeld van. Bijgevolg zijn het instellen van procedures en het opleggen van het gebruik van bepaalde beleidsinstrumenten ook ingegeven door een wens om resultaten te kunnen afdwingen. Zo werd de procedure van de Cel Gender Mainstreaming deels ingesteld om enige greep te hebben op de actoren en het verloop van het project. Het stappenplan had niet enkel de bedoeling om de ambtenaren te begeleiden in het proces van gender mainstreaming, maar ook om hen te dwingen om het proces te doorlopen en om effectief stappen te zetten. Hetzelfde geldt voor de Pekingrapporten en de OCM. Ook hier was de formalistische procedure een controlemiddel om beleidsactoren, die misschien niet allen op eenzelfde lijn zaten wat betreft prioriteit verlenen aan het gendervraagstuk, toch tot actie aan te zetten. Waar de reële doelstelling van het beleid evenwel geen breed draagvlak heeft, wordt een minimaal naleven van die procedures echter de maximale inspanning voor de beleidsactoren en dus een doel op zich. Als deze procedure geen eisen oplegt met betrekking tot het effectief bereiken van gendergelijkheid (bijvoorbeeld door consequenties te koppelen aan negatieve evaluaties naar het verderzetten van het beleid en naar het formuleren van nieuwe doelstelling toe) wordt het meten van gendergelijkheid en het toepassen van de gender mainstreamingsinstrumenten een doel op zich.

Al bij al kunnen we stellen dat het meten van beleid niet tot doel verheven wordt door gender mainstreaming, maar dat de rationalistische en intentionele assumpties die in de strategie van gender mainstreaming vervat zitten, daar in de praktijk wel toe bijdragen. De rationalistische kijk op het beleidsproces en het beleid op zich verklaart de procedurele en formalistische instrumenten. Immers, als gender gelijkheid het doel is en gender deskundigheid gegeven is, dan kunnen dergelijke instrumenten een proces structureren. Dat was ook hun initiële bedoeling. Als een of beide echter niet gegeven zijn, dan komt het inhoudsloze van deze instrumenten aan de oppervlakte. Ze bieden dan een gemakkelijke oplossing om niet aan gender mainstreaming te moeten doen, om niet het doel van gendergelijkheid te moeten nastreven maar wel

in regel te zijn met de verplichtingen. Zoals eerder gesteld worden de intentionele en rationele assumpties die aan gender mainstreaming ten grondslag liggen recent in vraag gesteld, onder andere door actoren die zelf mee aan de basis gelegen hebben van rationalistische instrumenten zoals de EER. In de beleidspraktijk wordt echter nog sterk voortgebouwd op die rationele en intentionele veronderstellingen, onder andere door de hegemonie van de door de Raad van Europa gegeven invulling aan gender mainstreaming. De sterke nadruk die de Raad van Europa legt op het inzetten van de gender mainstreamingsstrategie laat weinig ruimte om de strategie en de instrumenten die ze impliceert in vraag te stellen.

Een derde verklaring voor de vastgestelde beleidspraktijken is o.i. terug te brengen op de definitie van gender mainstreaming en vooral op de afwezigheid van een helder geformuleerde doelstelling van de strategie. Gender en gendergelijkheid zijn complexe en geen eenduidige concepten. De gender mainstreamingsstrategie vermijdt uitspraken hierover en laat de invulling ervan als het ware over aan de betrokken beleidsactoren. Deze beleidsautonomie heeft aanzienlijke voordelen zoals het kunnen afstemmen van het concrete beleidsdoel op het beleidsdomein waarin het ingepast wordt en het ontwikkelen van “issue-ownership”, namelijk een sterk gevoel van betrokkenheid wanneer de beleidsactoren voldoende inspraak hebben in de concrete invulling van de beleidsdoelen en –processen. Maar het brengt ook het grote nadeel met zich mee dat gender mainstreaming een veeleisende strategie is die zich moeilijk naar de beleidspraktijk laat vertalen. Gender, zijnde de maatschappelijke constructie van sekse, leidt tot complexe gehelen aan ongelijkheid die zich moeilijk laten vatten in eenvoudige schema’s, gezien verschillende dimensies van ongelijkheid elkaar vorm geven en beïnvloeden. Een evolutie op een van de dimensies kan gevolgen hebben voor alle andere. Dit geheel aan evoluerende dynamieken schematisch vatten is quasi onmogelijk, laat staan dat het tot een gemakkelijk te hanteren beleidsinstrument zou leiden. Daarnaast is het evenmin gemakkelijk om de middelen vast te leggen om tot gender gelijkheid te komen. Net deze complexiteiten met betrekking tot gender als concept en gendergelijkheid als doel hebben er, naast politieke motieven, wellicht toe geleid dat het inhoudelijke uit de instrumenten bleef. En waar het erin verwerkt werd, werd het gereduceerd tot heel gemakkelijke stellingen of kwam de kritiek dat het instrument niet haalbaar was.

Ten vierde is gender mainstreaming als beleidsstrategie ook deels verantwoordelijk voor haar eigen falen door de klemtoon die gelegd wordt op een aantal randvoorwaarden, zoals de nood aan naar sekse uitgesplitste statistieken. De rationalistische assumpties omtrent de beleidscyclus en het beleid leidden bij voorbeeld mede tot een nadruk op statistieken om correcte beleidsbeslissingen te kunnen treffen. Daar waar dergelijke gegevens ontbreken, evenals misschien de politieke wil om effectief een beleid te voeren en de middelen schaars zijn, kan het voeren van een gender mainstreamingbeleid blijven hangen op het voorbereidende niveau van datavergaring.

## 5. Gender mainstreaming: een meerwaarde voor het realiseren van gendergelijkheid?

Deze bijdrage analyseerde een aantal pogingen tot introductie en implementatie van gender mainstreaming in het Belgische beleid sinds het midden van de jaren 1990. Het door de VN en Europa gepromote gender mainstreamingsbeleid heeft het Belgische gelijke kansenbeleid duidelijk beïnvloed. In navolging van het Peking Actieplatform zetten de verschillende overheden de eerste stappen in het voeren van een gender mainstreamingsbeleid, zowel op het vlak van strategische inbedding in diverse beleidsdomeinen als betreffende het formuleren van doelstellingen die reële gelijkheid nastreven en het voorzien van de nodige middelen. Onze verkennende analyse van deze initiatieven toont aan dat alvast aanvankelijk de neiging sterk aanwezig was om de beleidsevaluatie en -monitoring centraler te stellen dan het eigenlijke doel, met name beleidsvorming ter realisatie van gendergelijkheid. De verklaringen hiervoor zijn veelvuldig en hebben te maken met het implementeren van een nieuwe strategie in beleidsdomeinen waar weinig of geen genderexpertise aanwezig was. Om houvast te bieden aan de beleidsactoren en om hen te dwingen om acties te ondernemen inzake gendergelijkheid werd de nadruk sterk gelegd op de te volgen procedures en aan te wenden beleidsinstrumenten. De inhoudelijke component –met name wat gender is, hoe genderdynamieken resulteren in genderongelijkheid en wat gendergelijkheid precies inhoudt– werd minder precies bepaald. De manier waarop gender mainstreaming van meet af aan door de Raad van Europa werd geconcipeerd, gaf daar ook aanleiding toe. Gender mainstreaming wordt naar voren geschoven als een uitgesproken intentioneel en rationalistisch beleidsproces opgebouwd rond doelstellingen, probleemanalyse en inzet van specifieke beleidsinstrumenten.

Een aantal nieuwere initiatieven zoals de federale gender mainstreamingwet van 2007 en de Vlaamse OCM bevatten elementen die het potentieel hebben om tot een meer substantieel gelijke kansenbeleid te komen, maar veel hangt af van de precieze invulling die er in de toekomst aan gegeven zal worden. Echter, ook al wordt de doelstelling van gendergelijkheid centraal gesteld, dan nog zullen de strategieën die voorliggen een aantal rationalistische en intentionele assumpties volgen die het realiseren van gender gelijkheid zoniet blokkeren dan toch niet noodzakelijk bevorderen. Niettemin vestigt deze evolutie in gunstige zin onze aandacht op het belang van het tijdsperspectief. De eerste, eerder formalistische fase waarin beleidsevaluatie en –monitoring een doel op zich waren, kan gezien worden als een noodzakelijke eerste stap die een basis legde voor performantere vormen van gender mainstreaming die we nu zien verschijnen. Op hun beurt geven deze misschien de aanzet om tot de implementatie te komen van gender mainstreaming initiatieven die nog sterker gericht zijn op het realiseren van gendergelijkheid. In elk geval biedt de tijd de mogelijkheid tot het ontwikkelen van genderexpertise, die, zoals we herhaaldelijk aangaven, een belangrijke voorwaarde is voor het inzetten van de gender mainstreamingsprocedures en instrumenten in het kader van een gelijke-kansenbeleid dat reële maatschappelijke impact resorteert.

De opbouw van de nodige expertise is echter een continu en omvattend project dat enkel volledig gerealiseerd zal worden indien gendergelijkheid en gender mainstreaming aanvaard wordt als behorend tot de “core business” van de



besluitvormers en beleidsactoren. Er dient immers niet alleen algemene kennis en expertise opgebouwd te worden inzake gender, genderdynamieken en hoe beleid genderongelijkheid creëert, bestendigd en kan verhelpen. Een dergelijke vorming kan eenmalig georganiseerd worden en hiervoor kunnen bestaande vormingspakketten en consultants ingeschakeld worden. Daarnaast is er echter ook kennis nodig over hoe dit alles zich concreet vertaalt in de respectievelijke beleidsdomeinen die aan constante evoluties onderhevig zijn. Deze naar het beleidsdomein geconcretiseerde en evoluerende kennis is veel moeilijker te garanderen. Het vergt eigen studiewerk van beleidsactoren en studiediensten of bestellen van externe expertenstudies. Dat laatste is duidelijk een ambitieuzere aangelegenheid die een hoge inzet van personele en financiële middelen vergt.

Naast verbeteringen aan de beleidscontext, zou de blik ook gericht kunnen worden op het concept gender mainstreaming zelf om in de toekomst te vermijden dat de gender mainstreaming in de praktijk vervalt in beleidsevaluaties en –monitoring als doel op zich. Zo kan men zich de vraag stellen of een invulling van de definitie van gender en van gendergelijkheid een oplossing zou kunnen bieden. Zodoende zou de strategie zelf een (deel van de) probleemanalyse bevatten evenals een meer concreet einddoel aangeven. Dat zouden dan ideologische keuzes zijn die politici dienen te maken wanneer ze gender mainstreaming adopteren. Dit zou impliceren dat besluitvormers het politieke debat voeren over de gewenste man-vrouwverhoudingen waarbij ze dan bijvoorbeeld expliciet uitspraken doen over of ze het ideaal nastreven van een gelijke verdeling van betaalde arbeid buitenshuis en onbetaalde arbeid in de familie tussen man en vrouw dan wel het ideaal van een potentieel ongelijke verdeling, mits deze geen ongelijke behandeling genereert (bijvoorbeeld inzake pensioensrechten).

Voor het concreet invullen van gendergelijkheid zou ook een beroep gedaan kunnen worden op externe actoren uit het middenveld. Dat zou een valorisatie van het eerder vermelde participatief-democratische model van gender mainstreaming inhouden. Vooral het expert-bureaucratische model werd in Europa geïmplementeerd, niet zelden in het kader van administratieve hervormingen met als doel het beleid efficiënter en performanter te maken. Het heeft de neiging om introspectief te worden en zodoende het ruimer kader uit het oog te verliezen. Zoals we aangaven, biedt de inschakeling van “buitenstaanders”, i.c. betrokken burgers en het middenveld, zoals dat voorzien is in het participatief-democratisch model, de mogelijkheid om meer kritische analyses te maken omdat deze personen buiten het dominante denkkader staan. De invulling van gendergelijkheid en de gewenste man-vrouwverhoudingen zouden dan in samenspraak met het betrokken middenveld en burgers tot stand kunnen komen in een proces van consultatie en overleg. Het legt ook expliciet de link tussen beleid en samenleving, iets wat, zoals deze bijdrage uitgebreid aantoonde, wel eens uit het oog wordt verloren door de uitgesproken intentionele en rationalistische aard van gender mainstreaming. Een belangrijk nadeel is evenwel dat het beleidsvoorbereidend proces hierdoor verlengd en verzaamd wordt.

We willen geenszins beweren dat deze twee suggesties soelaas bieden voor de problemen die we vaststellen bij de implementatie van een veeleisende strategie als gender mainstreaming. Wel willen we de aandacht erop vestigen dat gender

mainstreaming onderwerp moet zijn van evaluatie en bijsturing. Op dit ogenblik lijkt het immers weinig waarschijnlijk dat de piste van gender mainstreaming zal verlaten worden. Integendeel, gezien de recente tendens om het gelijke-kansenbeleid niet langer uitsluitend op gender te focussen maar om gelijktijdig ook andere achterstelligheidsdynamieken te viseren (onder meer etnisch-culturele achtergrond, seksuele geaardheid, fysieke capaciteiten), lijkt het er eerder op dat de mainstreamingstrategie een uitbreiding zal kennen. Bijvoorbeeld in Noord-Ierland en Schotland heeft het mainstreamingsbeleid naast mannen en vrouwen onder meer ook betrekking op personen met verschillende geloofs- en politieke overtuigingen, etnisch-culturele minderheidsgroepen, leeftijdsgroepen, mindervaliden, holebi's, taalgroepen, klassen, en personen met en zonder anderen ten laste (Celis & Cortier, 2008). Het komt er dus op aan om de mainstreamingsstrategie te verbeteren en te maken dat de inbedding ervan aan de hand van beleidsevaluatie en -monitoring geoptimaliseerd wordt en dat erover gewaakt wordt dat ze blijvend ingezet worden voor de realisatie van gelijkheid.



## Bibliografie

Behning, U. & Serrano Pascual, A. (2001). *Gender Mainstreaming in the European Employment Strategy*. Brussels: ETUI.

Beveridge F. & Nott S. (2002). Mainstreaming: A Case for Optimism and Cynicism. In: *Feminist Legal Studies (Special Issue)*, 10 (3), 299-331.

Beveridge, F., Nott, S. & Stephen, K. (2000). Mainstreaming and the engendering of policy-making: a means to an end? In: *Journal of European Public Policy*, 7 (3), 385-405.

Booth, C. & Bennett, C. (2002). Gender Mainstreaming in the European Union. Towards a New Conception and Practice of Equal Opportunities? In: *European Journal of Women's Studies*, 9 (4), 430-446.

Budlender, D. (2000). The Political Economy of Economics in the South. In: *World Development*, 28 (7), 1365-1378.

Celis, K. & Cortier, E. (2008). *De implementatie van gender mainstreaming. Implementatie in België, buitenlandse best practices en een checklist van stimulerende en remmende factoren voor de implementatie van gender mainstreaming. Literatuurstudie uitgevoerd in het kader van het onderzoeksproject 'De structurele implementatie van gendermainstreaming in de institutionele en administratieve mechanismen van het Brussel Hoofdstedelijk Gewest'*. Onuitgegeven onderzoeksrapport. Brussel: Ministerie van het Brussels Hoofdstedelijk Gewest.

Celis, K. & Meier, P. (2006). *De macht van het geslacht. Gender, politiek en beleid*. Leuven: Acco.

Celis, K. & Meier, P. (2007a). State feminism and women's movements in Belgium: complex patterns in a multi-level setting. In: J. Kantola and J. Outshoorn (eds.). *Changing State Feminism: Women's policy agencies confront shifting institutional terrain*. New York: Palgrave, 62-81.

Celis, K. & Meier, P. (2007b). *De L'Europe aux Régions et Communautés: une politique d'égalité à plusieurs visages?* Paper gepresenteerd op de Journée d'étude du groupe 'Genre et politique' de l'ABSP-CF, ULB, 7 december 2007.

Cel Gender Mainstreaming (2003). *Eindrapport en evaluatie van de door de federale regering opgerichte cel Gender Mainstreaming*. Brussel: FOD Werkgelegenheid, Arbeid en Sociaal Overleg.

Cortier, E. & Celis, K. (2008). *De implementatie van gender mainstreaming. Screening van de institutionele en administratieve mechanismen in het Brussels Hoofdstedelijk Gewest. Screening uitgevoerd in het kader van het onderzoeksproject 'De structurele implementatie van gendermainstreaming in de institutionele en administratieve mechanismen van het Brussel Hoofdstedelijk Gewest'*. Onuitgegeven

- onderzoeksrapport. Brussel: Ministerie van het Brussels Hoofdstedelijk Gewest. Council of Europe (1998). *Gender Mainstreaming. Conceptual framework, methodology and presentation of good practice*. Strasbourg: Council of Europe.
- Daly, M. (2005). Gender Mainstreaming in Theory and Practice. In: *Social Politics*, 12 (3), 433-450.
- Hafner-Burton, E. & Pollack, M. (2002a). Mainstreaming Gender in Global Governance. In: *European Journal of International Relations*, 8 (3), 339-73.
- Hafner-Burton, E. & Pollack, M. (2002b). Gender mainstreaming and Global Governance. In: *Feminist Legal Studies*, 10 (3), 285-298.
- Holvoet, N. (2007). Gender budgeting in Belgium: findings from a pilot project. In: *European Societies*, 9 (2), 275-300.
- Liebert, U. (2002). Europeanising Gender Mainstreaming: Constraints and Opportunities in the Multilevel Euro-Polity. In: *Feminist Legal Studies*, 10 (special issue), 241-256.
- Lombardo, E. (2005). Mainstreaming Gender Equality in the European Constitution-Making Policy. In: Schonlau, J. e.a. (eds.). *The Making of the European Constitution*. Houndmills: MacMillan.
- Lombardo, E. & Meier, P. (2006). Gender mainstreaming in the EU: incorporating a feminist reading? In: *European Journal of Women's Studies*, 13 (2), 151-166.
- Lombardo, E. & Meier, P. (2008). What do they say about power? Policy frames on gender inequality in politics in the Netherlands and Spain. In: *Journal of Women, Policies and Politics* (in druk).
- Lombardo, E., Meier, P. & Verloo, M. (red.) (2009). *The Discursive Politics of Gender Equality. Stretching, Bending and Policy-Making*. New York: Routledge.
- Mazey, S. (2001). *Gender mainstreaming in the EU: Principles and Practice*. London: Kogan Page Publishers.
- Meier, P. (2006). Implementing gender equality: gender mainstreaming or the gap between theory and practice. In: Hellsten, S.K. Holli, A.M. & Daskalova, K. (eds.). *Women's Citizenship and Political Rights*. New York: Palgrave.
- Meier, P. & Woodward, A. (1997). *De Emancipatie-EffectRapportage: Een instrument ter bevordering van gelijke kansen op alle beleidsterreinen*. Brussel: Ministerie van de Vlaamse Gemeenschap.
- Meier, P., Meyntjens, M. & Vander Steene, A. (1999). *LEER: een instrument voor lokaal beleid op maat van mannen en vrouwen*. Brussel: Ministerie van de Vlaamse Gemeenschap.

Meier, P. e.a. (2000). *De eer van ons beleid: Emancipatie-Effectrapportage. Een instrument ter bevordering van gelijke kansen op alle beleidsterreinen*. Brussel: Ministerie van de Vlaamse Gemeenschap.

Meier, P. e.a. (2007). The Pregnant Worker and Caring Mother: Framing Family Policies across Europe. In: Verloo, M. (ed.). *Multiple Meanings of Gender Equality. A critical frame analysis of gender policies in Europe*. Budapest: CEU Press, 109-140.

Mertens, T. & Steegmans, N. (2006). 'M/V United'. *Genderjaarboek 2006. Deel 3 MV United in beleid*. Antwerpen/Brussel: Steunpunt Gelijkekansenbeleid/ESF Agentschap Vlaanderen.

Morre, J. (2008). *Gender budgeting, een onmisbaar instrument voor de implementatie van de 'wet gender mainstreaming' van 12 januari 2007*. Presentatie op het seminarie gender mainstreaming georganiseerd door het Instituut voor de Gelijkeheid van Vrouwen en Mannen, Brussel, 16 april 2008.

Mosesdottir, L. & Erlingsdottir, R.G. (2005). Spreading the Word Across Europe. Gender Mainstreaming as a Political and Policy Project. In: *International Feminist Journal of Politics*, 7 (4), 513-531.

Oldersma, J. (2000). Het moderne feminisme en het sociaal kapitaal van Nederland. In: Hooghe, M. (red.). *Sociaal kapitaal en democratie, verenigingsleven, sociaal kapitaal en politieke cultuur*. Leuven: Acco, 197-223.

Pollack, M.A. & Hafner-Burton, E. (2000). Mainstreaming Gender in the European Union. In: *Journal of European Public Policy*, 7 (3), 432-456.

Rees, T. (1998). *Mainstreaming Equality in the European Union: Education, Training, and Labour Market Policies*. New York: Routledge.

Rees, T. (2005). Reflections on the Uneven Development of Gender Mainstreaming in Europe. In: *International Feminist Journal of Politics*, 7 (4), 555-574.

Rubery, J. e.a. (2004). The ups and downs of European gender equality policy. In: *Industrial Relations Journal*, 35 (6), 603-628.

Sharp, R. & Broomhill, R. (2002). Budgeting for equality: the Australian Experience. In: *Feminist Economics*, 8 (1), 25-47.

Squires, J. (2005). Is Mainstreaming Transformative? Theorising Mainstreaming in the Context of Diversity and Deliberation. In: *Social Politics*, 12 (3), 366-388.

Steegmans, N. (2007). *M/V United. Genderjaarboek 2007/ Deel 2: Onderwijsmonitor (Deel 1)*. Antwerpen: Steunpunt Gelijkekansenbeleid/ESF Agentschap Vlaanderen.

Stratigaki, M. (2004). The Cooptation of Gender Concepts in EU Policies: The Case of 'Reconciliation of Work and Family'. In: *Social Politics*, 11 (1), 30-56.

Stratigaki, M. (2005). Gender Mainstreaming versus Positive Action: An on-going Conflict in EU Gender Equality Policy. In: *European Journal of Women's Studies*, 12 (2), 165-186.

True, J. & Mintrom, M. (2001). Transnational Networks and Policy Diffusion: The Case of Gender Mainstreaming. In: *International Studies Quarterly*, 45 (1), 27-57.

United Nations (1995). *Beijing Declaration and Platform for Action. Fourth World Conference on Women*. New York: UN.

Verloo, M. (2000) *Gender mainstreaming: Practice and Prospects*. Report prepared for the Council of Europe, EG (99)13.

Verloo, M. (2001). *Another Velvet Revolution? Gender mainstreaming and the politics of implementation*. IWM Working Paper No. 5/2001. Vienna: IWM.

Verloo, M. (red.) (2007). *Multiple Meanings of Gender Equality. A critical frame analysis of gender policies in Europe*. Budapest: CEU Press.

Verloo, M. & Roggeband, C. (1994). *Emancipatie-effectrapportage: theoretisch kader, methodiek en voorbeeldrapportages*. Den Haag: VUGA.

Woodward, A. (2003). European Gender Mainstreaming: Promises and Pitfalls of Transformative Policy. In: *Review of Policy Research*, 20 (1), 65-88.

# Hoofdstuk 9: Een Dorpshuis als Ontmoetingsplaats: Ontwerp van Beleidsevaluatie

Steven Lannoo & Dries Verlet

<b>Inleiding</b> .....	<b>208</b>
<b>1. De Interventie</b> .....	<b>208</b>
1.1. Een dorpshuis als ontmoetingsplaats .....	208
1.2. De stakeholders .....	209
1.3. De beleidstheorie: Sociale Verbondenheid en Lokale Faciliteiten .....	210
<b>2. Methodologisch Raamwerk</b> .....	<b>213</b>
2.1. Types beleidsevaluatieonderzoek .....	213
2.2. Ex-post evaluatie .....	216
2.3. Methoden: kwalitatieve versus kwantitatieve methoden: twee werelden van waarheden? .....	218
2.4. Interviews en Focusgroepen .....	219
<b>3. Het Concrete Beleidsevaluatieontwerp</b> .....	<b>223</b>
3.1. Input .....	223
3.2 Throughput .....	223
3.3 Output .....	224
3.4. Outcome .....	224
<b>4. Uitleiding</b> .....	<b>226</b>
<b>Bibliografie</b> .....	<b>227</b>

## Inleiding

In deze bijdrage wordt een ontwerp van beleidsevaluatie voorgesteld. Het onderwerp van de evaluatie is een beleidsinterventie in de gemeente Assenede, meer bepaald in het dorp Boekhoute, waar een dorpshuis werd gecreëerd ten einde het aanbod aan lokale dienstverlening te versterken, maar voornamelijk ook om de sociale verbondenheid in het dorp te verstevigen. De centrale vraag in deze bijdrage is hoe men deze onderzoeksvraag het best kan aanpakken, rekening houdend met de specifieke context waarin het onderzoek dient te gebeuren (de betrokken actoren, inputs, processen, output, beoogde outcome, ...). Een fundamentele vraag terzake is welk type evaluatieonderzoek het meest aangewezen is om de centrale onderzoeksvraag te beantwoorden. De ultieme doelstelling is te komen tot een accurate meting die niet wordt vertekend omwille van de keuze van het onderzoekspopzet. Daarnaast moeten er ook fundamentele keuzes worden gemaakt wat betreft het soort methode die we hanteren om data te verzamelen en welke afwegingen hier een doorslaggevende rol in spelen. Deze keuze hangt natuurlijk in belangrijke mate samen met het soort informatie dat we willen verkrijgen. Het is in deze context dat we hier stuiten op het verschil tussen de nood aan eerder kwalitatieve dan wel eerder kwantitatieve data. Hoewel het soort data niet ipso facto samenhangt met het soort methode dat we hanteren (eerder kwantitatief, dan wel kwalitatief),<sup>76</sup> liggen beide elementen in elkaars verlengde. Bepalend in deze keuze is natuurlijk de concrete onderzoeksvraag. Dit komt aan bod in de eerste paragraaf, net zoals het ruimere kader waarbinnen het vooropgestelde onderzoek zich situeert.

## 1. De Interventie

### 1.1. Een dorpshuis als ontmoetingsplaats

De beleidsinterventie waarvoor in deze bijdrage een evaluatie wordt ontworpen, situeert zich in de gemeente Assenede, meer bepaald in het dorp Boekhoute. Assenede is een fusiegemeente van de oude entiteiten Oosteeklo, Bassevelde, Assenede en Boekhoute, gelegen in het noorden van Oost-Vlaanderen (Het Meetjesland), grenzend aan Nederland. Boekhoute ligt in het *Krekengebied*, een open polderstreek. Via de Braakman, een zijarm van de Westerschelde was het dorp vroeger verbonden met de zee. Boekhoute was dan ook een echt vissersdorp dat leefde van de mosselpluk en garnaalvangst. Sinds de afsluiting van de Braakman in 1952 is daaraan een einde gekomen. De vissersgeschiedenis is echter overal in het dorp aanwezig (Hasquin e.a., 1980). Het dorpshuis, waar deze studie over gaat, is dan ook gelegen in de zolderverdieping van het Vissermuseum.

In 2001 werd op initiatief van de vzw Samenlevingsopbouw in samenwerking met de gemeente Assenede de bevraging “Dorp In Zicht” georganiseerd. Dorp In Zicht is

---

<sup>76</sup> Zo kunnen we een kwantitatieve methode (zoals survey), perfect hanteren om eerder kwalitatieve data te genereren (in casu bijvoorbeeld door middel van open vragen).

een van oorsprong Britse methodiek met een dubbel doel. Ten eerste wil men via een bevraging van, door en voor de dorpsbewoners peilen naar het dorpsleven. Er worden vragen gesteld over lokaal bestuur, verenigingsleven, dienstverlening, enzovoort. De exacte vragen die in het onderzoek worden gebruikt worden gekozen door de *bewonersgroep*, een groep dorpsbewoners die de organisatie van de bevraging op zich nemen.<sup>77</sup> De informatie die op die manier wordt verzameld moet dorpsbewoners ondersteunen die actie voor hun gemeenschap ondernemen. De belangrijkste doelstelling van Dorp In Zicht is echter om een dynamiek los te maken in het dorp die aan het dorpsleven nieuwe impulsen geeft. Dorp In Zicht is in Vlaanderen een methodiek die door de vzw Samenlevingsopbouw wordt ondersteund en gepromoot, met de medewerking van de Vlaamse overheid (Dorp In Zicht, 2008).

Uit de bevindingen van de bevraging bleek dat er in het dorp een sterke vraag bestond naar meer en betere communicatie met de gemeentelijke overheid, naar een aantal voorzieningen die niet in het dorp aanwezig waren en naar algemene sociale verbondenheid onder de dorpsbewoners. Zowel de bewonersgroep, de initiatiefnemende vzw als de gemeente Assenede waren ervan overtuigd dat de oprichting van een dorpshuis wel eens het ideale antwoord op die vragen zou kunnen vormen. In 2002 werd een extra bevraging rond het oprichten van een dorpshuis georganiseerd. Daaruit bleek dat er een groot enthousiasme bij de bevolking bestond voor het project. Een projectgroep voor de oprichting van het dorpshuis werd opgestart vanuit het team dat de organisatie van Dorp In Zicht op zich had genomen, aangevuld met enkele bewoners en vertegenwoordigers van verenigingen. Het gemeente- en OCMW-bestuur werd van het begin bij de oprichting van het Dorpshuis betrokken. Zij beslisten om een deel van het oude stadhuis ter beschikking te stellen en de kosten van het gebruik van gas, water, elektriciteit en onderhoud te betalen. Financiële steun werd verder verkregen van de Europese Unie (in het kader van het Europees plattelandsontwikkelingsprogramma PDPO) en de Provincie Oost-Vlaanderen. Het dorpshuis werd zodanig ingericht dat het een combinatie werd van een ontmoetingsruimte waar men gezellig een kop koffie kan komen drinken, informatiekanaal waarbij computers met toegang tot internet, infomappen en brochures bezoekers inlichtingen moeten verstrekken en dienstverlening, daar in het dorpshuis verschillende zitdagen van de gemeente, politie, enzoverder worden gehouden. Naar mate de tijd vorderde kreeg het dorpshuis extra functies mee: er werd een sociaal restaurant ingericht, bijkomende zitdagen georganiseerd, een toeristisch infokantoor georganiseerd en vorming aangeboden. In het dorpshuis zijn twee mensen werkzaam onder een sociaal tewerkstellingsstatuut. Zij moeten aanwezig zijn tijdens de openingsuren, koffie aanbieden aan bezoekers en werken verder aan het bedenken en organiseren van activiteiten en de uitbouw van de werking. Bij de aanvang van het project werd de inhoudelijke werking gestuurd door Samenlevingsopbouw Oost-Vlaanderen. Tegenwoordig heeft de vzw Meetjeslandse Leerwerkbedrijven (die al van het begin de werknemers leverde), deze taak overgenomen. Het dorpshuis is nog steeds gevestigd in het oude stadhuis. Tegenwoordig betreft het de zolderverdieping, boven het Visserijmuseum. Omdat dat museum zal uitbreiden wordt binnenkort een nieuwbouw voorzien (Lannoo, 2008).

---

<sup>77</sup> De methodiek situeert zich in de sfeer van het participatief onderzoek waarbij de onderzoekseenheden zelf een actieve rol spelen in het tot stand komen en uitvoeren van het onderzoek. Zoals we later zullen aantonen zijn deze ideeën ook aanwezig in het veld van het beleidsevaluatieonderzoek.



## 1.2. De stakeholders

Er werd getracht de verschillende stakeholders in het project te identificeren, ten einde te weten waarop de evaluatie zich moet richten. Ten eerste zijn de dorpsbewoners zelf uiteraard stakeholders. De dorpsbewoners van Boekhouten vormen immers de betrokken doelgroep. Van hen wordt er verondersteld dat ze behoefte hebben aan meer dienstverlening en meer sociale samenhang in het dorp. Binnen de dorpsbewoners kunnen we nog een onderscheid maken tussen enerzijds diegenen die vaak gebruik maken van de infrastructuur en dienstverlening van het dorpshuis en diegenen die dat niet of weinig doen.

Een tweede categorie van stakeholders vormen de verantwoordelijken binnen Samenlevingsopbouw en Meetjeslandse leerwerkbedrijven die het initiatief tot oprichting van het dorpshuis hebben genomen. Verder zijn er de politieke verantwoordelijken, met name de burgemeester en het schepencollege. Het is dan ook de uitvoerende macht die de beslissing om een dorpshuis te steunen en te financieren heeft genomen. In tweede instantie zijn ook de raadsleden en de politieke partijen betrokken. De raadsleden moeten het plan immers goedkeuren, en de partijen zijn vaak de organen waarbinnen de houding van de verschillende groepen binnen de politieke elite ten opzichte van een beleidsinitiatief worden bepaald.

Ten slotte is ook de administratie een belangrijke betrokken partij. De administratie zal hier niet enkel moeten instaan voor enkele praktische elementen van de steun en financiering, maar is zelf ook onderwerp van de beleidsinterventie. Door het invoeren van het dorpshuis zullen een aantal ambtenaren zitdagen moeten houden in het dorpshuis. Zij zullen in een nieuwe omgeving terecht komen en zich aan die nieuwe omgeving moeten aanpassen. De betrokken personen zijn hier in eerste instantie het personeel dat in het dorpshuis werkt en voor de coördinatie instaat (bij Samenlevingsopbouw en de Leerwerkbedrijven) en ten tweede het personeel van de gemeente, het OCMW en de politiezone die er zitdagen organiseren.

## 1.3. De beleidstheorie: Sociale Verbondenheid en Lokale Faciliteiten

Het centrale concept in de voorliggende beleidsevaluatie is *sociale verbondenheid*. In dit hoofdstuk herconstrueren we de beleidstheorie die aan de basis ligt van de onderzochte interventie. Om de waarde van deze theorie te toetsen gaan we op zoek naar steun voor de uitgangspunten in de wetenschappelijke literatuur.

De beleidsmakers wilden met de creatie van een dorpshuis twee doelstellingen realiseren. Ten eerste wilden zij hiermee de vraag naar meer (lokale) dienstverlening bij de dorpsbewoners opvangen. Ten tweede wilden zij de behoefte aan meer sociale samenhang bevredigen. Via de creatie van een dorpshuis zou aan deze twee doelstellingen moeten worden voldaan. Men gaat ervan uit dat het dorpshuis gebruikt zal worden omdat het aan een bestaande behoefte voldoet. Op die manier zal rechtstreeks aan de eerste doelstelling worden voldaan. Men gaat er hier ten tweede van uit dat het dorpshuis een ontmoetingsplaats voor de dorpsbewoners kan vormen. De aanwezigheid van het dorpshuis en de activiteiten die er georganiseerd worden zouden eveneens een sociale dynamiek in het dorp moeten doen ontstaan.

Het aanbieden van deze ontmoetingsplaats zou mensen meer samenbrengen, waardoor de sociale samenhang in het dorp wordt versterkt. De beleidstheorie wordt schematisch weergegeven in onderstaand veranderingsmodel.

**Figuur 9.1: Veranderingsmodel**



Wanneer we spreken over sociale samenhang, dan botsen we in de literatuur op de concepten sociaal kapitaal en sociale cohesie. De laatste twee decennia hebben zowel onderzoekers als beleidsmakers steeds meer aandacht gegeven aan die concepten. Robert Putnam (1993, 1995, 2000) is in deze context één van de belangrijkste auteurs. Vooral de publicatie van *Bowling Alone: America's Declining Social Capital* (1995), waarin de auteur stelt dat sinds de jaren '70 het sociaal kapitaal in de Amerikaanse samenleving dramatisch is afgenomen, heeft wetenschappers en politici aangezet om te zoeken naar manieren om de sociale verbondenheid tussen mensen te vergroten.

Het belang dat gehecht wordt aan sociaal kapitaal komt voort uit de idee dat sociale verbondenheid, veralgemeend vertrouwen en normen van wederkerigheid (drie centrale concepten in de notie van het concept), zeer veel voordelen opleveren voor de bredere samenleving (Coffé & Geys, 2006). Sociaal kapitaal kan nochtans ook een keerzijde hebben. Zoals Narayan (1999, 10) het zegt: *'the same ties that bind also exclude'*. Een sterke verbondenheid kan resulteren in een sterke sociale controle, wat op zijn beurt kan voorkomen dat mensen opkomen tegen bestaand onrecht en ongelijkheid. Op die manier kan sociale verbondenheid de individuele vrijheid beperken.

Toch is in de literatuur sociale verbondenheid het meest geassocieerd met positieve effecten. Mensen die veel participeren aan sociale activiteiten zullen doorgaans ook meer participeren aan politieke activiteiten (Van Deth, 1992). Goed geïntegreerde individuen zijn vaak minder ethnocentrisch (Jakobs e.a., 2001). Sociaal kapitaal wordt geassocieerd met een betere gezondheid (Latkin & Curry, 2003; Cattell, 2001, ...), economische ontwikkeling (Woolcock, 1998), minder criminaliteit (Rice & Sumberg, 1997), enzoverder. Wat misschien nog het allerbelangrijkste is: sociale integratie lijkt de meest doorslaggevende factor te zijn in het voorspellen van het subjectief welzijn (Verlet & Devos, 2008). Met andere woorden, goed geïntegreerd zijn in het sociale

weefsel maakt mensen gelukkig. Dit is niet echt verassend, aangezien mensen in de eerste plaats sociale wezens zijn. Daarom hebben ze nood aan contacten met anderen. Hebben mensen de indruk dat ze onvoldoende positieve interacties met anderen hebben, dan voelen ze zich ongelukkig, boos en ongezond. Ondanks het feit dat sociale cohesie soms ook negatieve effecten kan hebben, moeten we het in zijn geheel dan ook als een positief gegeven beschouwen. Want volkomen vrij, maar sociaal geatomiseerd zijn is een situatie die slechts voor zeer weinig mensen bevredigend kan zijn.

Er dient hier te worden opgemerkt dat de meting van concepten zoals sociaal kapitaal en sociale verbondenheid uiterst moeilijk is. Van Deth (2003) merkt terecht op dat die moeilijkheden voornamelijk voortkomen door conceptuele problemen, eerder dan door methodologische. In de literatuur rond sociale verbondenheid is er sprake van een veelzijdigheid aan theoretische invalshoeken die allemaal eigen dimensies van het concept identificeren, wat dan zijn gevolgen heeft voor het meten. Zo concentreren sommige auteurs zich op netwerken en resources (Granovetter, 1973; Wilson & Musick, 1997), anderen dan weer op attitudes (Stolle, 2000). Sommigen benaderen verbondenheid vanuit het individu (van der Gaag, 2005), anderen hebben een meer collectieve benadering (Putnam, 2000). Het is dus aan de onderzoeker om in het labrynt van invalshoeken en perspectieven een keuze te maken en die te expliciteren. Wij zullen in ons onderzoek kiezen voor een individueel perspectief waarbij we aandacht hebben voor de netwerken van mensen en de *resources* die ze uit die netwerken halen. We zullen ons daarbij concentreren op de netwerken die zich op het dorp zelf richten omdat de doelstelling van de interventie er net in bestond om die netwerken te verstevigen.

De meeste onderzoekers, beleidsverantwoordelijken en beleidsondersteuners zijn het er mee eens dat sociale samenhang iets positiefs is, waardoor ze zoeken naar manieren om die samenhang in de maatschappij te verhogen. Daarom is het belangrijk om te weten welke indicatoren het *niveau van sociale verbondenheid* van mensen kunnen voorspellen. Heel veel factoren op het individuele niveau werden reeds geïdentificeerd. Tv-kijkgedrag, religieuze en levensbeschouwelijke betrokkenheid, leeftijd, opleidingsniveau (Hooghe, 1999), arbeidsomstandigheden (Vézina e.a., 2004, Lindström, 2006), enzoverder zijn allemaal factoren die een deel van de variantie in sociale verbondenheid verklaren. Naast individuele factoren, is ook de invloed van de woonomgeving onderzocht. Verschillende studies toonden aan dat er verschillen in sociale verbondenheid tussen buurten bestaan, die niet verklaard kunnen worden door de verschillen tussen de individuen die er wonen (bv. Hart & Dekker, 2003). Binnen de traditie van de Chicago-school, worden die verschillen verklaard door een concentratie van achterstelling, zoals armoede, werkloosheid, verloederding van de buurt en culturele heterogeniteit. In Vlaanderen toonden Coffé en Geys (2006) reeds aan dat sterk heterogene gemeenten een lagere graad van sociaal kapitaal bezaten. Een ander buurtkenmerk dat verschillen in sociale cohesie kan verklaren, is de hoeveelheid en de kwaliteit van de faciliteiten in de buurt. In het Verenigd Koninkrijk hebben Bowling en Stafford (2007) aangetoond deze faciliteiten een positieve invloed hadden op het sociaal functioneren van ouderen. De auteurs geloven dat de aanwezigheid van sociale faciliteiten en diensten de mogelijkheid van mensen om elkaar te ontmoeten vergroot. Of zoals Witten e. a. het zien: die

faciliteiten *'provide informal meeting places, outside home and work, where social relationships can be formed and maintained'* (Witten e. a., 2001, 308).

We kunnen concluderen dat de ideeën die in de beleidstheorie naar voren komen rond de positieve perceptie van sociale verbondenheid en rond de mogelijkheid om die te versterken via de creatie van faciliteiten/ontmoetingsplaatsen ook in de wetenschappelijke literatuur aanwezig is. In beperkte mate zijn de positieve effecten van ontmoetingsplaatsen ook al vastgesteld. Het is echter nog de vraag of die positieve effecten ook in deze concrete situatie zullen optreden.

## 2. Methodologisch Raamwerk

In wat volgt worden enkele belangrijke methodologische twistpunten besproken die op de voorliggende case van toepassing zijn. Wij behandelen achtereenvolgens de verschillende theoretische invalshoeken rond evaluatieonderzoek, het specifieke karakter van een ex-post evaluatie, de discussie rond kwantitatieve en kwalitatieve methoden en enkele beschouwingen rond de specifieke onderzoekstechnieken die in dit ontwerp aan bod komen.

### 2.1. Types beleidsevaluatieonderzoek

De definities en omschrijvingen inzake beleidsevaluatie zijn divers (cf. bijdrage Verlet en Devos). Niettemin, komen –aldus Dahler-Larsen– vier elementen telkens terug: systematisch vastleggen; waarde toekennen, praktische relevantie en een evaluandum. Dezelfde auteur onderscheid in diezelfde analyse drie benaderingen in het beleidsevaluatieonderzoek: een doelgerichte benadering, een theoriegebaseerde benadering en een responsieve/participatorische benadering (Dahler-Larsen, 2005). Deze drie benaderingen worden hier kort omschreven.

In de doelgerichte, “goal-oriented”, benadering worden activiteiten geëvalueerd op grond van de mate waarin ze formeel vastgelegde doeleinden helpen bereiken. De terminologie input/processes/output/outcome, die overigens uit de private sector komt overwaaien, treedt op de voorgrond in deze benadering. De mate van doelbereiking staat voorop en wordt gezien als het ultieme criterium voor de beleidsevaluatie. Deze benadering stookt met een representatieve theorie, waarbij een publieke actor uiteindelijk wordt afgerekend op het halen van vooropgestelde en gelegitimeerde doelstellingen. Het is duidelijk dat deze benadering ook schatplichtig is aan het New Public Management-denken dat een shift bepleitte van de aandacht voor regels en procedures naar output/outcome (Dahler-Larsen, 2005).

Niettemin zijn er volgens Dahler-Larsen (2005) een aantal belangrijke kritieken te formuleren op deze benadering. Zo faalt deze benadering in het bereiken van de vooropgestelde causaliteit. Hoewel men vaak veronderstelt dat de outcome de effecten zijn van publieke activiteiten, is deze causaliteit zelden methodologisch hard te maken. Een andere kritiek houdt in dat een dergelijke benadering vaak weinig zoden aan de dijk brengt wat betreft aanbevelingen teneinde de performantie van de overheid te verbeteren. Men geeft wel aan of en in welke mate doelstellingen worden

bereikt, maar over het “hoe” is men minder duidelijk. Een andere belangrijke kritiek in de context van de voorliggende SVR-studie is dat (positieve dan wel negatieve) neveneffecten worden over het hoofd gezien. Het doet daarmee afbreuk aan het complex geheel waarin de geëvalueerde outcome wordt bekeken. In dezelfde sfeer kunnen we de kritiek situeren dat de strikte focus op outcome kan leiden tot veranderingen in de activiteiten die het voorwerp zijn van evaluatie. In die context kunnen we verwijzen naar o.a. de performance paradox (cf. supra). Een laatste kritiek op deze benadering is dat het op zich heel moeilijk is om outcome te meten.

Een alternatieve benadering van beleidsevaluatie zijn de zogenaamde “theory-based approaches”. Bij een dergelijke benadering wordt het evaluandum geplaatst binnen een ruimer geheel veronderstellingen. Deze veronderstellingen kunnen we samenvatten onder de noemer beleidstheorie of zoals Dahler-Larsen (2005) het noemt in zijn analyse een “program theory”. Een dergelijke theorie geeft aan waarom een betrokken activiteit zou leiden tot de gestipuleerde outcome en wat er verder nog zou gebeuren. In deze context kunnen we onder meer de “theory-driven evaluation” van Chen (1990) situeren net zoals ook de “realistic evaluation” van Pawson & Tilly (1997).

Door de evaluatie te baseren op een programmatheorie komen een aantal voordelen naar voor in vergelijking met de doelgerichte evaluatie. Zo kan een theoriegebaseerde evaluatie een sterkere empirische basis zijn voor beleidsaanbevelingen. Een dergelijke benadering laat ook beter toe om maatstaven voor outcome te plaatsen in een ruimer geheel. Indicatoren worden immers maar al te vaak ontwikkeld in een theoretisch vacuüm. Ten derde laat een theoriegebaseerde evaluatie ook makkelijker toe om (positieve dan wel negatieve) neveneffecten te integreren in het evaluatieonderzoek. Een ander voordeel is dat de theoriegebaseerde evaluatie toelaat om kennis te accumuleren doorheen verschillende evaluaties. Ten slotte schuwt een dergelijke benadering ook geen controversie inzake het duiden van causale verbanden. Het leidt vaak tot technisch jargon, maar niettemin draagt het bij tot een open en kritische discussie van de veronderstellingen op grond waarvan beleid (veelal) wordt gebaseerd (Dahler-Larsen, 2005).

De mate waarin een dergelijke theoriegebaseerde benadering van evaluatie ingang en verspreiding kent is evenwel een ander paar mouwen. Theorieconstructie wordt in de dagdagelijkse praktijk binnen de publieke sector als iets vaags beschouwd en in de praktijk is er weinig ruimte voor een dergelijke benadering van evaluatie. Een ander controversieel gegeven is dat men bij theoriegebaseerde evaluatie vaak veronderstelt dat er überhaupt een programmatheorie is. Verschillende actoren hebben vaak hun eigen theoretische veronderstellingen. De evaluator construeert in dat geval vaker theoretische veronderstellingen, eerder dan ze te reconstrueren. Een meer fundamentele kritiek op deze benadering houdt in dat dergelijke evaluaties niet altijd gebaseerd zijn op een goed begrip tussen enerzijds het theoretische causale denken enerzijds en het normatieve/politieke denken anderzijds. De normatieve/politieke dimensie wordt vaak onderschat bij een dergelijke benadering. Een andere fundamentele kritiek is dat theorieën per definitie een reductie zijn van een complexe maatschappelijke realiteit. Het is nog maar de vraag in welke mate deze complexiteit te vatten is in een theorie (Dahler-Larsen, 2005).

Een derde groep benaderingen zijn deze die Dahler-Larsen (2005) de “responsive and participatory approaches” noemt. Eigen aan deze benaderingen is het ontkennen van de relevantie van externe constructen zoals “doelstellingen”, “outcomes” of “theorieën” met betrekking tot een specifieke evaluatiesituatie. In de plaats daarvan worden dergelijke benaderingen gekenmerkt door ‘*an immersion into a case context, a dialogue with local stakeholders, and/or a sharing of responsibility for the evaluation with local people opens the evaluation towards the complexities of the local setting, identification of local criteria, and an evaluation process which sometimes benefits the involved partners both directly and indirectly.*’ (Dahler-Larsen, 2005, 632). Daarmee delen beide benaderingen alvast de bezorgdheid voor participatie en invloed als democratische waarden. Een achterliggend idee bij deze benadering van evaluatie is dan ook het dichten van de spreekwoordelijke kloof tussen de burger en de politiek (Abma, 2004). Het is in deze context dat Stake (2004) “responsive evaluation” plaatste tegenover “priordinate evaluation” (waarbij a priori een selectie en meting van enkele outcome criteria plaatsvindt). Bij een responsieve evaluatie staat de case study voorop als methode om een beeld te krijgen van de eigenheid van een lokale situatie. Op die manier kan men een goed begrip ontwikkelen van de aspecten eigen aan een lokale situatie, aspecten die niet op voorhand worden geselecteerd, maar het resultaat zijn van een vooronderzoek, interviews en –vooral– observatie. Stake, grondlegger van het concept, benadrukt in zijn werk dat responsieve evaluatie op zich geen participerende evaluatie is, maar participerende evaluatie omvat veelal een graad van responsiviteit.

Bij een participerende evaluatie worden verschillende deelnemers uitgenodigd om verantwoordelijkheid te delen voor het evaluatieproces. Voor dergelijke vorm van evaluatie wordt er gekozen teneinde contextueel betekenisvolle kennis op te bouwen. Met deze benadering komen we in de buurt van wat de “fourth generation of evaluation” wordt genoemd (cf. Guba en Lincoln, 1998).<sup>78</sup> In een dergelijke evaluatie hebben alle stakeholders een stem in het uitwisselen van ideeën, visies en bezorgheden omtrent een programma. In die zin is de benadering eerder “constructivistisch” te noemen. Diverse theorieën kunnen we onderbrengen onder de koepel van participerende evaluaties. Gemeenschappelijk hebben ze alvast hun oog voor de bruikbaarheid van de evaluatie en het democratisch karakter van het proces (Dahler-Larsen, 2005). Ze maakten van evaluatie alvast meer dan een technisch gegeven, maar een praktijk waarin sociale, ethische en politieke aspecten een rol krijgen (Abma, 2004). Een belangrijk element bij dit type evaluatie is dat men niet alleen oog heeft voor aspecten die de kwaliteit van het beleid als dusdanig ten goede kunnen komen, maar dat men ook aandacht schenkt aan de kwaliteit van de dialoog en overleg terzake (Van der Knaap, 2006). Binnen een complexe politieke en organisatorische setting waarin vele stakeholders diverse visies hebben, is een participerende evaluatie zeker op zijn plaats. Het laat ook toe om een genuanceerder beeld te krijgen, waarin de mening van alle belanghebbenden aan bod komt. Voor sommige overheidsmanagers is dit de uitgelezen manier om de vinger aan de publieke pols te houden, terwijl anderen de mogelijke discordantie in

---

<sup>78</sup> Wel is het zo dat Guba en Lincoln de focus verschuiven van overleg tussen stakeholders naar het vinden van consensus. Een stap die door de meeste bepleiters van responsieve evaluatie niet wordt gezet (Abma, 2004).



de meningen van de verschillende stakeholders vrezten. Daarenboven is er de kans dat meer expliciete politieke agenda's van participerende evaluatie in conflict komen met de prioriteiten van overheidsmanagers (Dahler-Larsen, 2005). Het toepassen van deze benadering impliceert alvast dat de evaluatoren wat van hun controle op het evaluatieproces uit handen geven en openstaan voor een zekere mate van ambiguïteit. Een andere mogelijke barrière voor het hanteren van een dergelijke methode is de hardnekkige perceptie dat kwalitatieve methoden, die nauwer aansluiten bij een dergelijke benadering, vaak als minder sterk worden beschouwd (Abma, 2004).

De kritieken op deze groep benaderingen zijn divers, net zoals de types benaderingen onder deze koepel. Zo komt soms te sterk de nadruk te nadruk te liggen op de bruikbaarheid van de evaluaties in hoofde van de opdrachtgevers. Daarnaast wordt ook de politieke agenda eigen aan dergelijke evaluaties bekritiseerd. Soms ontbreekt bij participerende evaluatie ook de expertise aan evaluatie gezien men te veel vastklampt aan de specificiteit van wat men wil evalueren. Controle delen over het evaluatieproces is vaak veel meer controversieel dan participerende evaluatoren willen toegeven. Het komt de participerende beleidsevaluatie dan ook ten goede om, naast het succes, ook oog te hebben voor mindere successen in de participerende evaluatie (Dahler-Larsen, 2005).

Wij zijn ervan overtuigd dat de verschillende benaderingen elk hun positieve en negatieve kanten hebben. In het kader van het vermijden van perverse effecten van beleidsevaluatie lijkt het ons dan ook aangewezen zich niet op één benadering blind te staren maar het beste van de verschillende werelden te combineren. Zo zijn wij in de eerste plaats geïnteresseerd in het meten van effecten van een interventie en houden we rekening met het onderscheid tussen de concepten input, throughput, output, en outcome (cf. de doelgerichte benadering). Tegelijkertijd besteden wij aandacht aan het wetenschappelijk-theoretisch werk dat reeds rond ons onderzoeksonderwerp werd gemaakt en waar wij met ons onderzoek aan willen bijdragen (cf. de theoriegerichte benadering). Ten slotte is er in ons onderzoek veel aandacht voor de ideeën en concepties van de participanten aan het onderzoek (cf. de participatieve benadering).

## **2.2. Ex-post evaluatie**

Traditioneel wordt beleidsevaluatie gesitueerd op het einde van de beleidscyclus. Een bepaalde interventie is doorgevoerd, en achteraf wordt nagegaan of ze haar doel heeft bereikt of niet, en welke middelen daarvoor nodig zijn geweest. Een meer moderne visie gaat er echter van uit dat evaluatieonderzoek overall in de beleidscyclus zijn plaats kent. In deze context worden drie vormen van evaluatieonderzoek onderscheiden. Ex-ante evaluatie vindt plaats in de beleidsvoorbereidende fase. Het gaat om een ondersteunende analyse bij het ontwerpen van de interventie, waarbij onder andere de kosten en baten van bepaalde alternatieve oplossingen van beleidsprobleem worden afgewogen. Tussentijdse (ex-durante) evaluatie vindt plaats tijdens de implementatie van de interventie. Men onderzoekt hierbij of de evaluatie wordt doorgevoerd zoals gepland, of er geen onverwachte obstakels optreden, enzovoort. De bedoeling van een dergelijke evaluatie is vaak bijsturen



van de interventie en het vroegtijdig opsporen van problemen. Soms wordt ook nagegaan of bepaalde tussentijdse doelstellingen wel worden gehaald. De laatste vorm van evaluatieonderzoek is ex-post evaluatie. Ex-post evaluatie is eigenlijk de klassieke vorm van evaluatieonderzoek, en situeert zich dan ook aan het einde van de beleidscyclus. Het zwaartepunt van dit soort onderzoek ligt meestal bij het evalueren van de effecten van het beleid. Voor de beleidsmakers betekent dit een mogelijkheid om verantwoording af te leggen voor het beleid aan hogere overheden, drukingsgroepen, kiezers, of de samenleving in het algemeen. Ex-post evaluatie moet beleidsmakers ook de mogelijkheid geven tot beleidsleren, zodat het beleid onder meer efficiënter en effectiever kan verlopen (De Peuter e.a., 2007).

Het is duidelijk dat we in deze casus te maken hebben met een ex-post evaluatie. De interventie, i.e. de creatie van een dorps huis, is enkele jaren terug doorgevoerd en zou dus zijn effecten al moeten gegenereerd hebben. Het is de bedoeling van het onderzoek om na te gaan of de doelstellingen van het beleid, namelijk de creatie van verbondenheid en het voldoen aan een behoefte aan meer lokale dienstverlening, werden ingevuld. De interesse ligt dus duidelijk op de beleidseffecten en minder op de beleidsinhoud.

Bij een ex-post evaluatie kan men het onderscheid maken tussen een onderzoek naar doelbereiking dan wel naar effectiviteit. Wanneer het onderzoek eerder doelbereiking voor ogen heeft, dan gaat men in feite onderzoeken of het beleid het product heeft afgeleverd dat men wilde afleveren. In ons geval zou dat betekenen dat we onderzoeken of het gecreëerde dorps huis de functies en middelen heeft zoals men die oorspronkelijk had voorzien, en of het ook daadwerkelijk door de bezoeker werd bezocht. Het gaat hier met andere woorden om het meten van de *output* van het beleid. Bij een evaluatie die gericht is op effectiviteit gaat men na of de maatschappelijke doelstellingen die werden geponeerd ook effectief zijn bereikt en of dat ook effectief aan het gevoerde beleid te danken is. Men richt zich hierbij dus op de *outcome* van het beleid (De Peuter e.a., 2007). In ons geval willen we dan onderzoeken of dat de creatie van het dorps huis ook effectief aanleiding heeft gegeven tot het stijgen van de verbondenheid in het dorp en tot het bevredigen van een bestaande behoefte aan meer lokale dienstverlening. Of men zich al dan niet richt op de evaluatie van de output dan wel van de outcome, hangt uiteraard af van de onderzoeksvraag. Daar wij in de eerste plaats geïnteresseerd zijn in de gecreëerde effecten zullen we ons daarom ook richten op de outcome. Het mag echter duidelijk zijn dat een interpretatie van gegevens over effecten heel dikwijls gepaard zal moeten gaan met gegevens over de output van het beleid. Een afwezigheid van effect kan namelijk zowel te wijten zijn aan de inadequaatheid van een maatregel als het niet implementeren van de interventie zoals die werd gepland. Daarom zullen we ons eveneens richten op de output, dus op het dorps huis zelf.

Het onderzoek naar effectiviteit brengt ons echter bij de zeer netelige kwestie van causaliteit (cf. bijdrage Verlet & Devos). Het vaststellen dat de doelstellingen van een beleid werden bereikt is immers niet voldoende om de conclusie te mogen trekken dat het beleid ook effectief is geweest. Het is namelijk mogelijk dat een externe oorzaak het gewenste effect te weeg heeft gebracht. Omgekeerd is het mogelijk dat een gewenst effect dat effectief door de interventie werd gecreëerd negatief werd gecompenseerd door een externe oorzaak. In dat geval is het beleid effectief,

ondanks het feit dat de doelstellingen niet werden gehaald. Over wat causaliteit nu precies is zijn bibliotheken volgeschreven. In de sociale wetenschappen, voornamelijk in kwantitatief onderzoek, volgt men meestal de zogenaamde “Probabilistische Causaliteitstheorie” van Suppes (1970): A is de oorzaak van B als (1) er een (probabilistische) relatie is tussen A en B (A gaat meestal samen met een vermindering of vermeerdering in B), (2) A in de tijd voor B komt en (3) er geen derde variabele C bestaat die de oorzaak is van zowel A als B (Tacq, 1984).

De eerste twee voorwaarden zijn makkelijk vast te stellen. Het is echter zeer moeilijk om aan de derde voorwaarde te voldoen. In theorie bestaat de mogelijkheid van een experimentele setting<sup>79</sup> waarbij men op basis van het toeval onderzoekenheden toebedeeld aan een experimentele conditie (waarin de interventie wel wordt doorgevoerd) en een controleconditie (waarin de interventie niet wordt doorgevoerd). Het verschil tussen beide condities geeft het effect van de interventie weer. In de praktijk is het echter meestal onmogelijk een experimenteel opzet te creëren. In ons geval zou een experimenteel opzet willen zeggen dat we een grote groep dorpen per toeval indelen in een groep waar wel een dorps huis wordt gecreëerd en één waar er geen wordt gecreëerd. In alle dorpen zou dan een nulmeting en een meting na de interventie moeten plaatsvinden. Dit zou het onderzoek echter onbetaalbaar maken. Bovendien is het niet evident om dorpen die per toeval aan een conditie zijn toegewezen te verplichten of te verbieden om een dorps huis te creëren. Ten slotte is het ook niet mogelijk om een nulmeting te organiseren, aangezien de interventie reeds werd doorgevoerd. De meeste vormen van quasi-experimenteel onderzoek kunnen dus ook niet worden toegepast (Loosveldt, 2003). Ten slotte moeten we nog opmerken dat het kwantitatief meten van een variabele als verbondenheid heel erg moeilijk, intensief en vooral ook zeer duur zou zijn. Daarom zijn we voornamelijk aangewezen op het gebruik van alternatieve kwalitatieve methoden. In de volgende paragraaf gaan we daar dieper op in.

### ***2.3. Methoden: kwalitatieve versus kwantitatieve methoden: twee werelden van waarheden?***

Met het onderscheid tussen kwalitatieve dan wel kwantitatieve onderzoeksmethoden zijn we terecht gekomen bij een belangrijke tweespalt binnen het beleidsevaluatieonderzoek meer specifiek en binnen de sociale wetenschappen meer in het algemeen. Het doornemen van de literatuur terzake heeft vaak iets van een discussie tussen believers en non-believers, al is dit een schijn discussie. Beide types methoden hebben immers elk hun verdienste en zijn elk in hun opzicht dienstig in beleidsevaluatieonderzoek.<sup>80</sup> Het debat over het gebruik van kwalitatieve dan wel kwantitatieve methoden wordt niet alleen weerspiegeld in het beleidsevaluatieonderzoek, het was ook een belangrijke drive in de uitbouw ervan.

---

<sup>79</sup> Men kan erover discussiëren of een puur experimentele setting in de sociale wetenschappen überhaupt mogelijk is. Randomisering van de groepen garandeert namelijk niet dat alle factoren in het onderzoek onder controle worden gehouden. Op die discussie gaan we hier echter niet verder in.

<sup>80</sup> De literatuur omtrent de voor- en nadelen van kwalitatieve en kwantitatieve methoden is immense. Degelijke overzichten zijn onder meer te vinden in Bryman, 2004; Grix, J., 2004; Mannheim e.a., 2002; Sarantakos, 2005.

Eerder kwalitatieve methoden (zoals bijvoorbeeld elite-interviews, directe observatie en focusgroepen) zijn de uitgelezen methoden om een beeld te krijgen van subjectieve ervaringen en de betekenis die men hecht aan deze ervaringen (Davies, 2007). Daarnaast laten dergelijke methoden ons ook toe om een zicht te krijgen van processen inzake bijvoorbeeld besluitvorming. Hoewel binnen de publieke sector vele processen ook geformaliseerd zijn in regels, is het essentieel om ook aandacht te hebben voor de informele dimensie. Daarnaast kunnen kwalitatieve methoden ons ook helpen om de context waarin een specifiek fenomeen zich voordoet te begrijpen. De term context is hier ruim te interpreteren en kan gaan van de individuele biografie van een betrokken actor tot de bredere sociale setting (Verlet, 2007).

De kritieken op eerder kwalitatieve methoden zijn divers (zie bijvoorbeeld Grix, 2004; Manheim e.a., 2002; Sarantakos, 2005).<sup>81</sup> Voorwerp van discussie is onder meer de betrouwbaarheid en de representativiteit bij het gebruik van deze methoden. De beperkte herhaalbaarheid en vergelijkbaarheid van kwalitatief onderzoek staat veralgemeningen veelal in de weg. Daarnaast zijn vele kwalitatieve methoden gevoelig aan reactiviteit bij de dataverzameling. Vaak is er immers een interactie tussen de onderzoeker en het onderzoeksobject. De kritiek uit zich dan ook niet zelden op de validiteit of geldigheid van de data verkregen door kwalitatieve methoden. Natuurlijk gelden bovengenoemde kritieken niet voor alle specifieke kwalitatieve methoden. Daarnaast is het gratis om aan te nemen dat kwantitatieve methoden ongevoelig zijn voor de genoemde problemen.

Niettemin worden kwantitatieve methoden (zoals surveyonderzoek, geaggregeerde data-analyse en inhoudsanalyse) geprezen omwille van hun wetenschappelijke credits die ze verzamelen door bijvoorbeeld vergelijkingen en veralgemeningen makkelijker te maken. Het is alvast zo dat doorgaans het aantal cases dat men kan bestuderen groter is. Daarnaast laat de systematiek van kwantitatieve methoden inderdaad makkelijker veralgemeningen toe.<sup>82</sup> Kwantitatieve methoden maken het ons makkelijker om op zoek te gaan naar regelmatigheden in bijvoorbeeld het gedrag van individuen of maatschappelijke gebeurtenissen. Op grond van die regelmatigheden tracht met het verleden en het heden beter te kunnen verklaren en –voor zover als haalbaar– voorspellingen te maken over de toekomst (zie bijvoorbeeld Grix, 2004; Manheim e.a., 2002; Sarantakos, 2005).

Kwantitatieve methoden worden evenwel verweten te falen in het ontsluiten van betekenissen en het tegen het licht houden van slechts een beperkt facet van de maatschappelijke werkelijkheid. Zo zou men bijvoorbeeld te weinig oog hebben voor de rol die de habitat van individuen en organisaties speelt. Daarnaast wordt ook bij de rapportering het soms te technische jargon en de misleidende weergave van data op de korrel genomen. Vaak suggereert men ook causaliteit, terwijl hooguit een samenhang empirisch kan worden aangetoond. In het onderzoek door middel van kwantitatieve methoden komt eveneens vaak de nadruk al te sterk te liggen

---

<sup>81</sup> Hiermee hebben we natuurlijk niet gezegd dat het systematisch tewerk gaan bij kwalitatieve methoden onbelangrijk is, zelfs integendeel. Ook kwalitatieve methoden ontleen in belangrijke mate hun wetenschappelijke waarde door een systematische analyse van data.

<sup>82</sup> Daarvoor kunnen we onder meer verwijzen naar Carton & Van Geel (2001).

op gemakkelijk meetbare fenomenen (die niet altijd het meest relevant zijn) en propageert men –voorbarig– een waardevrijheid (Bryman, 2004).

Het is duidelijk dat beide types onderzoeksmethoden zo elk hun eigen mogelijkheden en beperkingen hebben. Het is niet aangewezen om een strikte scheiding te maken tussen kwantitatieve dan wel kwalitatieve methoden. In het slechtste geval zou dit er toe leiden dat beide types elkaar uitsluiten en niet te combineren vallen, terwijl net de rijkdom ligt in de combinatie van beiden. Cruciaal in deze is dat de keuze voor een specifieke methode voor een (deel van de) analyse in eerste instantie dient gebaseerd te zijn op de onderzoeksvraag in kwestie.

Hierbij is het ook belangrijk om aan te geven dat ze op diverse vlakken ook gelijkenissen vertonen. Zo zijn beiden gericht op een analyse van de empirische werkelijkheid. Beiden zijn dan ook gebaseerd op een gelijkaardige logica, met name op grond van empirisch vaststellingen komen tot veralgemeningen en veronderstellingen. Men hanteert gekende feiten om te komen tot kennis over datgene we nog niet weten. Alleen bestuderen ze beiden een ander aspect van deze maatschappelijke werkelijkheid. De ene methode is dan ook niet beter of slechter dan de andere, alleen is in een aantal gevallen de ene methoden meer geschikt dan de andere. Het meest ideaal is dan ook de combinatie van beide methoden, “triangulation” zoals dit ook wel eens wordt genoemd. De slotsom is dan ook dat kwantitatieve en kwalitatieve methode het best worden gecombineerd in een beleidsevaluatieonderzoek (Grix, 2004; Manheim e.a., 2002). Afhankelijk van het onderwerp zal een eerder kwalitatieve meting noodzakelijk zijn, dan wel een kwantitatieve analyse gebaseerd op strikt geformuleerde indicatoren.

## **2.4. Interviews en Focusgroepen**

Het is weinig verwonderlijk dat we hier opteren voor een combinatie van kwantitatieve en kwalitatieve methoden teneinde een antwoord te bieden op de vooropgestelde onderzoeksvraag.

Om te weten hoe een grote groep respondenten denkt of handelt is een veel beproefde methode surveyonderzoek. Bij de praktische uitwerking van een survey zullen we in de bijdrage niet stilstaan.<sup>83</sup> Het is een methode van dataverzameling waarbij de informatie op een directe wijze wordt verkregen van individuele personen die worden geselecteerd teneinde de basis te vormen voor veralgemeningen tot ruimere populaties. Het is dan ook de meest ontwikkelde kwantitatieve methode binnen de sociale wetenschappen (Manheim e.a., 2002). Naast het specificeren van surveydesign (qua doelstelling, type survey en organisatie doorheen de tijd) is het natuurlijk belangrijk om de juiste vragen te stellen aan uw respondenten. Eigen aan survey is dat de bevraging gestructureerd gebeurt en dat er in deze weinig ruimte is voor eventuele bijsturing. Bij een grondige pretest kunnen al veel problemen worden vermeden, maar dan nog is een goede voorbereiding op de uiteindelijke bevraging essentieel. Hierbij kunnen we gebruik maken van andere methoden om

---

<sup>83</sup> Daarvoor kunnen we onder meer verwijzen naar Carton & Van Geel (2001).

te komen tot een goede vraagstelling. Zo kunnen we bijvoorbeeld gebruik maken van elite-interviews en focusgroepen teneinde een goed inzicht te verwerven in het voorwerp van de studie, met name het inplanten van een dorps huis teneinde de dienstverlening en de sociale verbondenheid te optimaliseren.

Interviews zijn een beproefde en populaire methode om data te verzamelen. In eerste instantie denken we dan spontaan aan diepte- of zogenaamde elite-interviews, maar natuurlijk kunnen we evengoed in het kader van face-to-face en telefonische survey spreken van interviews. Essentieel verschil tussen beide types is de mate waarin ze gestructureerd of gestandaardiseerd zijn. Beide types kunnen dan ook op een continuüm geplaatst worden (Verlet, 2007).

Bij gestandaardiseerde of gestructureerde interviews wordt elke vraag op dezelfde manier gesteld, veelal met pasklare antwoordmogelijkheden. Dit is bijvoorbeeld het geval bij een survey waarin gesloten vragen worden gesteld. De voordelen van dergelijke interviews is dat deze eenvoudig te kwantificeren zijn, in principe minder interviewvaardigheden vergen en makkelijk vergelijkbaarheid mogelijk maken.

Aan de andere kant van het continuüm vinden we de ongestandaardiseerde of ongestructureerde interviews. De bedoeling van dergelijke interviews is dat men kwalitatieve data wil genereren. Deze data is veelal meer gedetailleerd, men hanteert een meer flexibel interviewschema en kan beter de achterliggende betekenissen ontwaren. Zeggen dat een dergelijk interview helemaal niet gestructureerd zou zijn, is evenwel de waarheid geweld aan doen. Men kan namelijk argumenteren dat alleen al een lijstje met vooropgestelde vragen een vorm van structurering inhoudt. Een dergelijk type interview is bijvoorbeeld interessant aan het begin van een onderzoek om een duidelijk inzicht te verwerven in de problematiek. Het voornaamste knelpunt is echter de vergelijkbaarheid die heel gebrekkig is (Sarantakos, 2005). Dit kan men dan evenwel ondervangen door bijvoorbeeld in een andere fase van het onderzoek meer gestructureerde interviews (bijvoorbeeld survey) te hanteren.

Eigen aan een continuüm is natuurlijk dat er ook tussenvormen zijn, net zoals de eindpunten zelden in hun pure vorm voorkomen.<sup>84</sup> Deze tussenvormen kunnen we labelen als semi-gestructureerde interviews. Het is een combinatie van een eerder kwantitatieve gestructureerde benadering en eerder kwalitatieve ongestructureerde interviews. Dit type vinden we onder meer terug bij zogenaamde focusgroepen. Dergelijke (als kwalitatief geklasseerde) methode houdt in dat men diepgaand gaat discussiëren over een specifiek thema onder leiding van een moderator. Niettemin wil men bij het gebruik van deze methode vaak ook enige vergelijkbaarheid introduceren door bijvoorbeeld focusgroepen te gaan houden over een gelijkaardig thema maar bij verschillende stakeholders in een beleidsevaluatie. Op zich kunnen ook elite-interviews als semi-gestructureerd worden beschouwd, indien men bijvoorbeeld dezelfde vragen gaat stellen bij diverse actoren. Op die manier genereert men

---

<sup>84</sup> Zoals aangegeven, ook ongestructureerde interviews, zoals elite-interviews, hebben iets van structuur, net zoals ook bij bijvoorbeeld survey ongestructureerde elementen een plaats kunnen krijgen (bijvoorbeeld door middel van open vragen).

immers vergelijkbaarheid. Dit is evenwel geen evidente oefening, gezien men geen al te strak raamwerk aan het interview mag opdringen. Het is op zich een moeilijke evenwichtsoefening waarbij men een evenwicht moet vinden tussen enerzijds het behandelen van wat men als onderzoeker belangrijk vindt qua thema's en anderzijds moet men de respondenten toelaten om zelf thema's naar voor te brengen zonder te vervallen in voor het onderzoek irrelevante onderwerpen.

Gezien het onderwerp van deze evaluatie zouden we een lans willen breken voor het hanteren van focusgroepen. Zoals aangegeven betreft het een eerder kwalitatieve methodologie waarbij men een nauwkeurig geselecteerde groep mensen bijeenbrengt om onder leiding van een moderator diepgaand te discussiëren over een bepaald onderwerp (Manheim e.a., 2002).

Natuurlijk kan men deze methode hanteren als centrale methode om de data te verzamelen, maar evengoed valt deze methodiek toe te passen in andere fasen van het onderzoeksproces. Zo kunnen we deze methode hanteren om de onderzoeksvraag verder te verfijnen, om indicatoren voor bepaalde fenomenen te selecteren of de interpretatie van data uit bijvoorbeeld survey te vergemakkelijken. Op zich kunnen we hier nog de vraag stellen waarom dan geen interview zoals bij een enquête of elite-interview. Probleem met survey is dat we mogelijkerwijze de verkeerde vragen stellen, wat moet worden vermeden rekening houdend met de tijd en kostprijs van een dergelijke (doorgaans grootschalige) bevraging. Ook al hanteert men open vragen, dan nog kan de manier van vragen stellen en andere effecten omwille van de interactie tussen interviewer en geïnterviewde een rol spelen. Ook de interpretaties van vragen en antwoorden die de onderzoeker dan wel de geïnterviewde geeft kan wel eens verschillen. Maar het meest essentiële element is dat opinies van mensen in belangrijke mate worden beïnvloed door reacties en opinies van anderen. Gezien elite-interviews en survey in een geïndividualiseerde context worden afgenomen, heeft het gebruik van focusgroepen hier een belangrijke meerwaarde. Het centrale kenmerk van focusgroepen is dan ook dat deze gestoeld zijn op de interactie tussen de deelnemers teneinde een zicht te krijgen op een bepaald fenomeen. Focusgroepen zijn dan ook de uitgelezen manier om te peilen naar zogenaamde "audience reception", hoe de betekenis van iets vorm krijgt en waarom mensen een bepaalde houding hebben. Het is een methode die ook beter dan andere toelaat om het belang van aspecten van een fenomeen in te schatten. Daarnaast is het eveneens een vaststelling bij het gebruik van deze methode dat deelnemers aan een focusgroep spontaan hun mening beargumenteren en verdedigen, dit in tegenstelling tot bijvoorbeeld de situatie binnen een klassieke interviewsituatie. Deze kwaliteiten maken het in vele gevallen een geschikte methode om te hanteren in het kader van beleidsvoorbereiding en -evaluatie (Bryman, 2004; Manheim e.a., 2002).

Daarnaast heeft ook het gebruik van focusgroepen haar beperkingen (Manheim e.a., 2002; Sarantakos, 2005). Een eerste mogelijks probleem is dat van de subjectieve interpretaties. Net als de andere kwalitatieve methoden, zijn de verkregen data gevoelig aan de manier waarop men als onderzoeker de discussie interpreteert. Ten tweede is er de beperkte representativiteit. Het gaat over relatief kleine groepen,

discussies met een 6- tot 10-tal deelnemers waarbij men dan nog de kans loopt dat het een atypische groep is. Daarom is het op zich aangewezen om meerdere sessies te zien voor gelijkaardige groepen, wat nog niet wegneemt natuurlijk dat veralgemeningen tot ruime populaties niet aangewezen is. Ten derde moeten we er rekening mee houden dat het een artificiële setting is en blijft. Zo kan de samenstelling van zo'n groep sterk verschillen van de groepen waarmee men dagdagelijks in contact komt. Binnen elke focusgroep ontwikkelt zich eveneens een specifieke dynamiek. Meerdere focusgroepsessies en een zorgvuldige selectie van de deelnemers kunnen al veel problemen verhelpen. Ten slotte moeten we ook hier, net zoals bij andere methoden overigens, rekening houden met methode-effecten. Zo kunnen vooroordelen van de moderator de discussie sturen. De mate waarin de hierna genoemde problemen zich stellen, hangt natuurlijk sterk af van de samenstelling, de doelstelling en de kwaliteit van de voorbereiding van dergelijke focusgroepen (Manheim e.a., 2002).

Focusgroepen lijken ons bijvoorbeeld een aangewezen methode om de interpretatie van de verschillende stakeholders met betrekking tot het functioneren van het dorps huis na te gaan. We stellen vier focusgroepen op: een groep van dorpsbewoners die geen of weinig gebruik maken van het dorps huis, een groep van bewoners die er veel gebruik van maakt, een groep van verantwoordelijken en personeelsleden en een groep van de beleidsmakers. In de vier groepen houden we een discussie waarin we dezelfde vragen aanbrengen. Voorbeelden zijn: wat is de belangrijkste meerwaarde van het dorps huis, wat zijn de belangrijkste minpunten, heeft het dorps huis Boekhoutte meer leven ingeblazen, was het dorps huis nodig, brengt het mensen dicht bij elkaar, et cetera. Ten einde atypische discussies en dito conclusies te vermijden, doen we er goed aan om verschillende focusgroepen te organiseren per subgroep.

Probleem blijft dat we pas ex post een meting uitvoeren. Dit maakt het bijzonder lastig om effecten in kaart te brengen door een ex ante en ex post analyse met elkaar te vergelijken. Dit zou op zich de meest ideale onderzoekssituatie zijn om het effect van een beleidsinterventie, met name het inplanten van een dorps huis, na te gaan. Als de vergelijkingsmogelijkheid doorheen de tijd wegvalt kunnen we eventueel door middel van een vergelijking in de "ruimte" doorvoeren door het opzetten van een comparatief onderzoek (Burnham e.a., 2004). Als methode kunnen we het comparatief onderzoek plaatsen naast een experimentele<sup>85</sup> of statistische<sup>86</sup> methode (cf. Lijphart). Het komt er op neer dat we doormiddel van een vergelijkend onderzoek de "isolatie" van bepaalde factoren trachten te bewerkstelligen. De vraag is natuurlijk welke dorpen men dan moet gaan vergelijken. De basisoptie hier is een most-similar dan wel most-different design voor de vergelijkende analyse. In het eerste geval opteren we voor gelijkaardige dorpen wat de omgevingsvariabelen variabelen betreft, maar verschillen wat betreft de onafhankelijke variabelen (i.c. het

---

<sup>85</sup> Binnen evaluatieonderzoek wordt er bijzonder weinig of niet gebruik gemaakt van experimenten.

<sup>86</sup> Wat in essentie een vergelijkend onderzoek weerhoudt om statistisch te zijn is het aantal cases in de analyse.



al dan niet aanwezig zijn van een dorpshuis). Dit vanuit de redenering dat op die manier de gemeenschappelijke kenmerken constant worden gehouden. Bij een most-different design, kiezen we net sterk verschillende dorpen qua omgevingsvariabelen, terwijl de onafhankelijke variabelen gelijk zijn (bijvoorbeeld allen een dorpshuis). De redenering hier is dat wanneer men dan toch gelijkenissen vaststelt tussen die dorpen, men de verschillen tussen de dorpen niet kan aangrijpen ter verklaring. Dit vereist natuurlijk op zich een grondige basiskennis van specifieke dorpen die we willen opnemen in het comparatief onderzoek. Sowieso is het aangewezen om ook in deze context meervoudige indicatoren te gebruiken.

### 3. Het Concrete Beleidsevaluatieontwerp

In wat volgt zullen we de methodologische theorie uit het vorige hoofdstuk toepassen op onze concrete casus. Ter structurering gebruiken we de indeling input/troughput/output/outcome die we aan de doelgerichte benadering hebben ontleend.

Er dient opgemerkt dat aan de effecten die men in kaart wil brengen uiteraard ook een tijdsdimensie vasthangt. Het zou kunnen zijn dat de creatie van het dorpshuis in de beginfase een impuls geeft aan het dorpsleven, maar geen blijvend effect ressorteert. Omgekeerd kan het zijn dat het dorpshuis eerst uit de startblokken moet schieten, vooraleer het zijn positieve resultaten bereikt. We moeten dan ook rekening houden met het moment waarop de evaluatie plaatsvindt.

#### 3.1. Input

Hoewel de nadruk ligt op het evalueren van de maatschappelijke effecten, of de outcome van de interventie, zal ook een evaluatie van de input worden uitgevoerd. Die gegevens moeten ons later in staat stellen de uiteindelijke resultaten van het onderzoek beter te kunnen interpreteren. In dit deel van het onderzoek willen we een antwoord vinden op onder andere de volgende vragen: hoe werd de keuze gemaakt voor de inplanting van het dorpshuis, hoeveel mensen werken er in het dorpshuis, wie werkt er, welke taken voeren ze uit? Hoeveel mensen werden betrokken bij de opzet van het dorpshuis en kreeg het project al dan niet een hoge prioriteit? Hoeveel financiële middelen vloeien naar het project en waar worden die allemaal aan besteed? Dergelijke vragen kunnen het best worden aangepakt met een analyse van beleidsdocumenten. Heel wat gegevens omtrent oprichting, financiering en dergelijke worden immers neergeschreven in schriftelijke verslagen van gemeenteraden, collegevergaderingen, vergaderingen van vzw's et cetera. Wat staat neergeschreven vormt echter niet altijd een getrouwe weergave van de werkelijkheid. Daarom zal dit deel van het onderzoek worden aangevuld met interviews van enkele sleutelpersonen uit de gemeentelijke politiek en administratie (burgemeester, gemeentesecretaris en enkele betrokken lagere ambtenaren) en verantwoordelijken van Samenlevingsopbouw en Meetjeslandse Leerwerkbedrijven.

In deze fase wordt ook onderzocht of er wel degelijk een behoefte bestond aan een dergelijke dienstverlening. Daarvoor wordt een beroep gedaan op de resultaten

van een analyse van Dorp in Zicht (2008). Die vraag zal ook in de georganiseerde focusgroepen aan bod komen (cf. infra).

### 3.2 Throughput

Om te kunnen begrijpen waarom een bepaalde beleidsinterventie wel of niet werkt is het zeer belangrijk om een zicht te krijgen op de processen die zich afspelen bij de beleidsdoorvoering. We willen met andere woorden een zicht krijgen op vragen zoals: Hoe zijn de verschillende verantwoordelijkheden over verschillende mensen en organisaties verdeeld? Hoe zijn de persoonlijke verhoudingen tussen de verschillende mensen die bij het functioneren van het dorps huis betrokken zijn? Wat verwachten bezoekers van het dorps huis? Hoe verlopen de interacties tussen professionelen, vrijwilligers, bezoekers en beleids mensen?

Dergelijke vragen kunnen slechts worden beantwoord door diepgaande gesprekken te voeren met de mensen in kwestie. Eerder kwalitatieve technieken zoals diepte-interviews en focusgroepen lijken ons hier de beste methode. Personen die zeker dienen te worden geïnterviewd zijn de werknemers van het dorps huis, hun leidinggevend en als ook de gebruikers. Om gebruikers te identificeren zullen we gebruik maken van de gegevens die via het gebruikers monitoringssysteem (cf. infra) worden verzameld.

### 3.3 Output

Een derde element dat we dienen te onderzoeken vooraleer we aan de analyse van de effecten kunnen denken, is de output die door de beleidsinterventie werd gerealiseerd. Wat de uitvoering betreft nemen we enerzijds het gebruik van het dorps huis onder de loep, en gaan we anderzijds ook het functioneren van de werking na. Wat het gebruik betreft willen we nagaan hoeveel bezoekers het dorps huis krijgt, wie die bezoekers precies zijn (komt bijna iedereen langs, gaat het telkens om dezelfde individuen, ...) en wat ze er precies komen doen. Informatie hieromtrent kunnen we verkrijgen door een bezoekers monitoringssysteem te organiseren. Het is namelijk zo dat het dorps huis geen systematische gegevens bijhoudt over wie van de diensten en infrastructuur gebruik maakt, of wie aan de activiteiten deelneemt. Een monitoringssysteem dat gegevens omtrent de bezoeken bijhoudt moet ons in staat stellen een beter inzicht te krijgen in het gebruik van het huis. Een dergelijk systeem kan eenvoudig ontwikkeld worden door bezoekers/gebruikers een formulier te laten invullen en deze gegevens vervolgens in een databank op te slaan voor analyse.

Naast de gebruikers aantallen willen we ook informatie verzamelen over het functioneren van het dorps huis zelf. We willen te weten komen welke dienstverlening effectief wordt aangeboden. Hier zullen we de koppeling maken naar de informatie over voorgaande behoeften die we uit de "Dorp in zicht"-studie verkregen hebben. Informatie moet ook verkregen worden over de openingsuren, het aantal personeelsleden dat meestal aanwezig is, enzovoort. Deze informatie willen we

verkrijgen via het uitpluizen van interne documenten (dienstregeling, ...), aangevuld door informatie uit diepte-interviews met de werknemers, leidinggevenden en enkele bezoekers.

De informatie die hier werd verzameld moet worden vergeleken met de oorspronkelijke plannen en doelstellingen. Op die manier willen we te weten komen of beleidsinterventie effectief werd doorgevoerd zoals ze werd gepland, wat de precieze afwijkingen van de oorspronkelijke plannen zijn, en waarom die afwijkingen gebeurd zijn. Dit laatste moet vooral via diepte-interviews en focusgroepen aan het licht komen.

### 3.4. Outcome

#### 3.4.1. Premissen

Hoewel de informatie die we via bovenvermelde technieken zullen verzamelen ook op zichzelf een grote waarde heeft, staat ze in dit onderzoek voornamelijk in het teken van de outcome-evaluatie. Met dit deel van de evaluatie willen we nagaan of dat de maatschappelijke veranderingen die als doelstelling van de interventie werden geponeerd ook daadwerkelijk hebben plaatsgevonden.

Zoals we reeds eerder aanhaalden is het testen van effecten in geen sinecure. In de realiteit van het beleidsevaluatieonderzoek is het veelal onmogelijk om een causaal verband methodologisch zuiver vast te stellen. Meestal baseren onderzoekers zich bij het leggen van verbanden dan ook op enkele premissen van epistemologische aard. Hoewel onderzoeksontwerpen of –rapporten geen filosofische traktaten hoeven te zijn, is het in het kader van de *intersubjectiviteit* van de wetenschap wel aangewezen om basisveronderstellingen bij het beschrijven van een methode te expliciteren.

In het onderzoeksontwerp kiezen wij uitdrukkelijk niet voor een comparatief uitgangspunt. Zoals we reeds eerder hebben vermeld, is een vergelijking in de tijd in onze situatie onmogelijk. We hebben immers te maken met een reeds doorgevoerde interventie, zodat het uitvoeren van een nulmeting niet meer tot de mogelijkheden behoort. Ook het onder sectie 2.4. aangehaalde alternatief om een vergelijking in de ruimte te maken via een “most similar design”, lijkt ons in deze context niet opportuun. Bij een dergelijk ontwerp zouden we twee of meer dorpen zoeken die op alle (of zoveel mogelijk) kenmerken gelijk zijn, behalve op de onafhankelijke variabele, i.e. de aanwezigheid van het dorpshuis. Daar bij het onderzoek naar sociale verbondenheid heel wat unieke en moeilijk kwantificeerbare kenmerken van dorpen, zoals de lokale geschiedenis,<sup>87</sup> de structuur van het verenigingsleven, et cetera. Het toeschrijven van verschillen tussen twee gelijkaardige dorpen aan de aanwezigheid van het dorpshuis zal onzes inziens dan ook altijd een stelling zijn die op los zand is gefundeerd.

Ten einde dan toch een effect te kunnen onderzoeken stellen wij een aanpak voor waarbij we volgende veronderstelling maken. Stel dat de introductie van het

---

<sup>87</sup> Denken we bijvoorbeeld aan het vrij unieke gegeven van Boekhoute als een verzand vissersdorp.

dorpshuis iets heeft veranderd aan de sociale verbondenheid aldaar, dan moet dat de dorpsbewoners zeker opgevallen zijn. De introductie van een dergelijk dorpshuis is immers een opvallende gebeurtenis waar iedereen van op de hoogte is. De dorpsbewoners zelf, evenals de andere personen die bij het project betrokken zijn (personeel, ...), zullen ons dan ook in staat stellen om informatie over het effect te verkrijgen. Een aanpak die in hoofdzaak gestoeld is op kwalitatieve interviews en focusgroepen, lijkt ons dan ook de meest aangewezen methode om onze onderzoeksvraag van een antwoord te voorzien.

Zoals dat met premissen altijd het geval is, kunnen we de onze niet bewijzen. We zijn er echter van overtuigd dat ze zeer aannemelijk is. In wat volgt zullen we in meer concrete termen een schets geven van hoe we een dergelijk project zouden aanpakken.

### **3.4.2. Doelstelling 1: Voldoen aan de Vraag naar Lokale Dienstverlening**

In eerste instantie zal de vraag naar het bereiken van dit doel worden aangepakt met de informatie die verzameld zal worden via de eerder besproken methodes. We hebben het over de informatie betreffende bezoekersaantallen, effectief aanbod van dienstverlening en het behoeftenonderzoek van Dorp in Zicht.

Deze informatie zal echter worden aangevuld via een schriftelijke enquête bij de voltallige bevolking van het dorp. Deze enquête zal gepost kunnen worden in een brievenbus aan het dorpshuis. In de enquête worden vragen opgenomen naar de tevredenheid over de dienstverlening van het dorpshuis, de noodzaak en het nut van het bestaan ervan, het gebruik en hoe en met wat de dienstverlening eventueel uitgebreid dient te worden. In de enquête zal voornamelijk van een gesloten vragenlijst gebruik gemaakt worden, om een snelle en kwantitatieve verwerking mogelijk te maken.

Deze kwantitatieve informatie dient om een eerste zicht te krijgen op de houding van de bewoners ten aanzien van het dorp. Die informatie dient om een zicht te krijgen op de houding van de bewoners van het dorp en zal uitgebreid worden met informatie verzameld uit focusgroepen en interviews. In een ideale situatie moeten wij per onderscheiden groep stakeholders verschillende focusgroepgesprekken organiseren. We zullen daarom telkens drie focusgroepen maken uit de populatie frequente gebruikers, niet-frequente gebruikers en niet-gebruikers van het dorpshuis, de politieke verantwoordelijken, de administratie en de initiatiefnemers van de vzw. We zullen de groepen eveneens diversifiëren naar geslacht, opleiding en herkomst (al dan niet opgegroeid in het dorp). Het voordeel van het gebruik van focusgroepen ten aanzien van interviews is hier dat sociale verbondenheid bij uitstek een groepsfenomeen is, waardoor inzichten met betrekking tot dit fenomeen dan ook het best via groepsgesprekken tot uiting komen. Dit voordeel weegt echter niet op tegen het nadeel van de te grote druk van leidinggevenden op hun onderschikten. Daarom zullen we de werknemers van het dorpshuis en hun leidinggevenden ook individueel interviewen.

### 3.4.3. Doelstelling 2: Verstevigen van de Sociale Verbondenheid

Teneinde het al dan niet bereiken van de tweede doelstelling te kunnen nagaan, wordt de enquête waarover sprake onder sectie 3.4.2 uitgebreid met vragen over de werking van het dorps huis als ontmoetingsplaats en over het eventuele effect op de sociale verbondenheid in het dorp. We vragen of de respondent reeds mensen ontmoet heeft in het dorps huis die hij anders minder zag, of de oprichting van het dorps huis een positieve verandering in het dorp teweeg gebracht heeft, of het dorps huis het gevoel geeft dat het dorp meetelt in de gemeente, of het dorps leven gestimuleerd wordt door het dorps huis, of de mensen meer met elkaar omgaan sinds de komst van het huis, et cetera.

De kwantitatieve analyse zal ook hier moeten worden aangevuld met kwalitatieve informatie. Zoals eerder vermeld worden verschillende focusgroepen ingericht bevolkt door de verschillende stakeholders teneinde te proberen ontdekken of het dorps huis wel degelijk iets veranderd heeft aan de verbondenheid in het dorp, aan het dorps leven, et cetera.

## 4. Uitleiding

In deze bijdrage werd een plan voor een ontwerp van beleidsevaluatie voorgesteld. De centrale beleidsinterventie die het onderwerp van de bijdrage vormt, betreft de creatie van een buurthuis in een dorp met een dubbel doel. Ten eerste moet het buurthuis in de behoefte aan meer lokale dienstverlening vanwege de overheid voorzien. Ten tweede moet het een ontmoetingsplaats vormen voor de buurtbewoners, en aldus de sociale samenhang in het dorp versterken. We hebben aangetoond dat de idee van lokale faciliteiten die als ontmoetingsplaatsen de sociale cohesie kunnen versterken, ook in de literatuur aanwezig is. In het methodologisch raamwerk dat we hebben geschetst, werd een poging ondernomen om de verschillende inzichten die bestaan rond invalshoeken, de afweging van kwalitatieve en kwantitatieve methoden en het hanteren van specifieke technieken te schetsen en aan te tonen welke gevaren gepaard gaan met het maken van dergelijke keuzes. In een laatste deel hebben we deze inzichten dan toegepast op onze concrete case, ten einde een ontwerp van beleidsevaluatieonderzoek op te stellen.

We willen deze bijdrage beëindigen met het aanstippen van enkele belangrijke aandachtspunten die doorheen de tekst naar voren zijn gebracht.

Ten eerste wensen wij aan te stippen dat de discussie tussen believers en non-believers van kwalitatieve en kwantitatieve methoden in onze ogen niet vruchtbaar is. De keuze tussen beide invalshoeken moet volledig gemaakt worden met het oog op de onderzoeksvraag en de concrete onderzoeksituatie. Uiteindelijk werd er geopteerd voor de combinatie van beiden. De nadruk komt evenwel te liggen op een eerder kwalitatieve benadering in onze concrete case.

Ook de keuze van een concrete methode moet uiteindelijk gestoeld zijn op het zo goed mogelijk beantwoorden van de onderzoeksvraag en de specifieke situatie waarin het onderzoek plaats vindt. Het best wordt dicht allicht geïllustreerd door de keuze tussen focusgroepen en diepte-interviews bij onze outcome-evaluatie.

Ten slotte willen we nog eens benadrukken dat het proberen meten van een effect een uiterst moeilijke zaak is, en dat het leggen van een attributie tussen maatschappelijke veranderingen en een interventie die een dergelijke verandering heeft beoogt op een methodologisch zuivere manier (i.e. via een experimentele setting) in de praktijk zo goed als onhaalbaar is. De onderzoeker zal zich daarom moeten baseren op andere premissen inzake causaliteit. Dit hoeft geen probleem te vormen, zolang die premissen maar expliciet worden gemaakt.

## Bibliografie

Abma, T., (2004). Responsive evaluation: the meaning and special contribution to public administration. In: *Public Administration*, 82 (4), 993-1012.

Anoniem (2008). *Rapport Dorp in Zicht Ursel*. Eeklo: Samenlevingsopbouw Oost-Vlaanderen. <http://www.dorpinzicht.be/rapporten/Uitgebreid%20rapport%20Ursel.pdf>, geraadpleegd op 15 mei 2008.

Anoniem, (2008). *Dorp In Zicht. Leefbaarheidsonderzoek in Plattelandskernen*. Eeklo: Samenlevingsopbouw Oost-Vlaanderen. [www.dorpinzicht.be](http://www.dorpinzicht.be), geraadpleegd op 31 augustus 2008.

Bryman, A. *Social research methods. Second edition*. Oxford/New York: Oxford University Press.

Burnham, P., Grant, W. & Gillant, K. (2004). *Research Methods in Politics*. Houndsmill/ New York: Pargrave Macmillan.

Carton, A. & Van Geel, H. (2001). *Kwaliteitsrichtlijnen bij het uitvoeren van surveyonderzoek*. Brussel: Ministerie van de Vlaamse Gemeenschap/Administratie Planning en Statistiek.

Cattell, V. (2001). Poor people, poor places, and poor health: the mediating role of social networks and social capital. In: *Social Science & Medecine*, 52 (10), 1501-1516.

Coffé, H. & Geys, B. (2006). Sociaal kapitaal in heterogene gemeenten. Een studie naar de relatie tussen een heterogene bevolkingssamenstelling en het sociaal kapitaal in Vlaamse gemeenten. In: *Tijdschrift voor sociologie*, 27 (1), 345-367.

Dahler-Larsen, P. (2005). Evaluation and public management. In: Ferlie, E., Lynn, L. & Pollitt, C. (eds.). *The Oxford handbook of public management*. New York: Oxford University Press, 622-636.

Davies, M.B. (2007). *Doing a succesful research project using qualitative or quantatieve methods*. Houndsmill/New York: Pargrave Macmillan.

De Peuter, B., De smet, J. & Bouckaert, G. (2007). *Handleiding Beleidsevaluatie. Deel 1: Evaluatiedesign en -management*. Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

Granovetter, M.S. (1973), The strenght of weak ties. In: *American Journal of Sociology*, 78 (6), 1360-1380.

Grix, J. (2004). *The foundations of Research*. Houndsmill/New York: Pargrave Macmillan.



Hasquin, H., Van Uyten, R. & Duvosquel, J.M. (1980). *Gemeenten van België. Geschiedkundig en administratief-geografisch woordenboek*. Brussel: Gemeentekrediet van België.

Jakobs, D., Abts, K. Phalet, K. & Swyngedouw, M. (2001). Verklaringen voor ethnocentrisme. De rol van sociaal kapitaal, sociaal-economische onzekerheid, sociale integratie en gevoelens van discriminatie. Een verkenning. In: (APS), *Vlaanderen gepeild! De Vlaamse overheid en burgeronderzoek 2001*. Brussel: Ministerie van de Vlaamse Gemeenschap, Administratie Planning en Statistiek.

Lannoo, S. (2008), *Niet gepubliceerd gesprek met Luc Joos en Tine Rollier* (vzw Samenlevingsopbouw).

Latkin, C.A. & Curry, A.D. (2003). Stressful Neighborhoods and Depression: A Prospective Study of the Impact of Neighborhood Disorder. In: *Journal of Health and Social Behavior*, 44 (1), 34-44.

Loosveldt, G. (2003), Experimentele designs. In: Billiet, J. & Waege, H., *Een samenleving onderzocht. Methoden van sociaal-wetenschappelijk onderzoek*. Antwerpen: De Boeck.

Manheim, J.B., Rich, R.C. & Willnat, L., (2002). *Empirical Political Analysis. Research Methods in Political Science. Fifth edition*. New York: Longman.

Putnam, R. (1993), *Making Democracy Work. Civic traditions in modern Italy*. Princeton: Princeton University Press.

Putnam, R. (1995). Bowling Alone: America's Declining Social Capital. In: *Journal of Democracy*, 6 (1), 65-78.

Putnam, R. (2000). *Bowling Alone: Collapse and revival of American community*. New York: Simon & Schuster.

Rice, T.W., & Sumberg, A. F. (1997). Civic culture and government performance in the American states. In: *The Journal of Federalism*, 27 (1), 99-114.

Sarantakos, S., *Social Research. Third edition*. Houndsmill/New York: Pargrave Macmillan.

Stolle, D. (2000), Onderzoek naar sociaal kapitaal. Naar een attitudinale benadering. In: Hooghe, M. *Sociaal Kapitaal en Democratie*. Leuven: Acco, 25-60.

Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam: North Holland Publishing Company.

Tacq, J. (1984). *Causaliteit in Sociologisch Onderzoek. Een beoordeling van Causale Analysetechnieken in het Licht van Wijsgerige Opvattingen over Causaliteit*. Deventer: Van Loghum Slaterus.

Van der Knaap, P. (2006). Responsive evaluation and performance management. Overcoming the downsides of policy objectives and performance indicators. In: *Evaluation*, 12 (3), 278-293.

Van Deth, J. W. (1992). De politieke betekenis van maatschappelijke participatie. In: *Acta Politica*, 27 (4), 28-53.

Van Herck, B. (2007). *Meetjesland 2020, een plan met toekomst*. Eeklo: Streekplatform+ Meetjesland.

Verlet, D. & Devos, C. (2008). The main determinants for subjective well-being: a quest for the Holy Grail? Can local governments enhance the perceived quality of life? In: Moller, V. & Huschka, D. (ed.). *Quality of Life and the Millennium Challenge. Advances in Quality-of-Life Studies, Theory and Research*. Social Indicators Research Series (volume 35). Springer, in press.

Verlet, D. (2007)., *Syllabus Onderzoeksmethoden 2*. Gent: Hogeschool Gent, Departement Handelswetenschappen en Bestuurskunde.

Woolcock, M. (1998). Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework. In: *Theory and Society*, 27 (1), 151-208.

Wilson, J. & Musick, M. (1997). Who Cares? Toward an Integrated Theory of Volunteer Work. In: *American Sociological Review*, 62 (5), 694-713.

## ***Epiloog: Beleidsevaluatie is een evidentie, beleid evalueren is dat niet***

Dries Verlet & Carl Devos

Studiedienst van de Vlaamse Regering & Universiteit Gent

Na lectuur van dit boek is men onmiskenbaar overtuigd dat evaluatieonderzoek een complex proces is, waar heel veel komt bij kijken. De verschillende auteurs geven elk een invulling aan het concept beleidsevaluatie, maar de grote lijnen komen vaak terug. Het zijn niet die grote lijnen, maar wel de operationalisering, de talrijke uitvoeringsmodaliteiten en uiteenlopende criteria en doelstellingen die voor grote heterogeniteit zorgen in het veld van de beleidsevaluatie.

Alle auteurs en onderzoekers benadrukken het grote belang van beleidsevaluatie. Hoe we dat precies moeten doen is veel minder evident. Zo blijkt het behoorlijk moeilijk te zijn om een goede, doordachte evaluatie in de dagdagelijkse beleidspraktijk om te zetten. Er bestaan vele invalshoeken, mogelijkheden en criteria. In dit relatief jong onderzoeksveld is er nog niet zoveel stroomlijning. De vraag is ook of dat wenselijk is, en of niet precies de vele vormen, methoden en invalshoeken een rijkdom vormen waaruit beleidsmakers kiezen om de gepaste evaluatie samen te stellen die overeenkomt met hun doelstellingen en mogelijkheden en rekening houden met de context waarin ze werken. Of zoals Wollmann (2007, 389) het verwoordde: *'it is widely agreed that there is no "king's road" in the methodological design of evaluation research, but, instead, one should depart from a "pluralism of methods". The selection and combination of the specific set and mix of methods depends on the evaluative question to be answered and on the time frame and financial and personnel resources available.'*<sup>88</sup>

Beleidsevaluatie bevat onvermijdelijk normatieve elementen. Het moet uiteindelijk, op een zo objectief mogelijke manier en op basis van zo duidelijk mogelijke en beargumenteerde criteria, beleid waarderen. Het evalueren van vooropgesteld, lopend of gevoerd beleid moet rekening houden met vele gevoeligheden, van beleidsmakers, van betrokken maatschappelijke actoren, van diegenen die bepaalde opties verdedigen in weerwil van het verzet van anderen, et cetera. De uitspraken die het evaluatieonderzoek doet, moeten antwoord geven op de gestelde vragen, maar moeten uiteraard ook rekening houden met het, al dan niet oneigenlijk, gebruik of zelfs misbruik van die conclusies. Dit is eigen aan het beoordelen van beleidsopties. Het formuleren van de probleemstelling en het vooropstellen van toetsingscriteria vereist uiteraard keuzes, en die keuzes impliceren dat alternatieve evaluatiecriteria dan weer niet aan bod komen. Dat onderzoek dient strikt wetenschappelijk te gebeuren: met een duidelijke onderzoeksvraag, een duidelijke theoretisering, glasheldere operationalisering en onderzoeksdesigns, transparantie inzake de dataverzameling en -analyse, alsook een duidelijke rapportering.

---

<sup>88</sup> Wollmann, H. (2007). Policy Evaluation and Evaluation Research. In: Fischer, F., Miller, G. & Sidney, M., Handbook of Public Policy Analysis. Boca Raton/London/New York: CRC Press.

Cruciaal voor evaluatieonderzoek is de terugkoppeling naar het beleid, anders verliest het onderzoek veel van haar zin. Beleidsevaluatie is in de eerste plaats al belangrijk omdat het inzicht geeft in de werking van de overheid, de relaties tussen actoren, de capaciteiten van de bestuurders, de controlesystemen, et cetera. Maar uiteindelijk dient dit onderzoek om nadien beter te doen dan voordien of om de beste beleidsmethoden te identificeren die nadien ook elders ingezet kunnen worden, gelet op de omstandigheden die ze vereisen. Het onderscheid tussen beleidsevaluatieonderzoek enerzijds en wat er uiteindelijk mee wordt gedaan anderzijds is niet evident in die gevallen waarin de evaluatie intern gebeurt. Het spanningsveld tussen het evaluandum, de evaluatoren, de opdrachtgevers en de beleidsverantwoordelijken speelt op de achtergrond een belangrijke rol.

In Vlaanderen wordt volop gewerkt aan de uitbouw van goede vormen van beleidsevaluatie. Nederland staat er voor gekend dat beleidsmakers aldaar, in vergelijking met Vlaanderen, beter kunnen terugvallen op een uitgebouwde traditie van monitoring en meting. Hoewel monitoring op zich nog geen evaluatie is (en omgekeerd), kan men wel stellen dat er in Nederland meer basismateriaal beschikbaar is, op grond waarvan men aan de slag kan gaan. Die informatie is noodzakelijk, maar onvoldoende voor een gezonde beleidsevaluatiecultuur. We moeten dus over de grenzen van overheden, beleidsectoren en zelfs nationale staten heen van elkaar leren om die rijkdom en valkuilen van beleidsevaluatie beter te leren kennen. Alleen op die manier kunnen we doordacht beleid evalueren en vermijden dat de mogelijke negatieve effecten die eraan verbonden zijn zich voordoen. De groei van de overheid in haar verschillende en diverse verschijningsvormen, van de behoefte aan publieke tussenkomst en regelcapaciteit, de toenemende verwachtingen t.a.v. het overheidsoptreden in ruime zin, de versterking van het vermogen om dat optreden te volgen en te analyseren en vele andere evoluties maken beleidsevaluatie onvermijdelijk. We moeten er alles aan doen om die zo slim mogelijk te organiseren. Dit boek heeft daartoe een bijdrage willen leveren.

## Over de auteurs

**Frank Bongers** is doctor in de sociale wetenschappen. Sinds kort is hij strategisch adviseur bij de provincie Noord-Brabant in 's-Hertogenbosch. Daarvoor werkte hij ruim acht jaar als senior beleidsonderzoeker en -adviseur bij Dialogic Innovatie & Interactie in Utrecht. Hij is gespecialiseerd in overheid & ICT, beleidsonderzoek (in het bijzonder evaluatie en effectmetingen) en innovatieonderzoek (metingen, kennisuitwisseling). Hij heeft in de afgelopen jaren vele evaluaties uitgevoerd in ondermeer het Nederlandse innovatie-, wetenschaps- en ICT-beleid. Zijn laatste evaluatieonderzoek betrof een effectmeting van het programma Collectief Onderzoek van IWT Vlaanderen. Voorts is hij lid van een lokale rekenkamer en daar medeverantwoordelijk voor evaluatie van gemeentelijke dossiers.

**Marleen Brans** (Lic. Politieke & Sociale Wet. K.U.Leuven, M.A. in Policy Processes (Hull University UK), PHD (Europees Universitair Instituut, Firenze) is hoofddocent aan het Instituut voor de Overheid en doceert beleidsanalyse, ontwerp&strategie, beleidsevaluatie en comparative public policy. Ze is coördinator en promotor van onderzoek over beleidsvoorbereiding, -implementatie, en -evaluatie, alsook over evidence-based en lerend beleid. Zij is lid van meerdere internationale professionele netwerken (EPAN, EAAPA) en (co)-promotor van verschillende internationale governance projecten.

**Karen Celis** is docent in de opleiding Bestuurskunde en Publiek Management van de Hogeschool Gent – Universiteit Gent. Karen Celis publiceert over en voert en begeleidt fundamenteel en beleidsondersteunend onderzoek naar de politieke vertegenwoordiging van maatschappelijke groepen en naar gelijke-kansenbeleid. Ze is lid van het internationale onderzoeksnetwerk “Research Network on Gender and the State” en het onderzoeksnetwerk “Gendered Ceremony and Rituals in Parliament”. Ze is co-convenor van de “ECPR Standing Group of Gender and Politics” en co-editor van ‘*The Oxford Handbook of Gender and Comparative Politics*’ (Oxford University Press, forthcoming).

**Bart De Peuter** (Licentiaat in de Politieke Wetenschappen en European Master in Public Administration, K.U.Leuven) is wetenschappelijk medewerker aan het Instituut voor de Overheid (K.U.Leuven). Zijn expertise situeert zich op de thema's monitoring, beleidsevaluatie, toekomstverkenning, lokale besturen en interbestuurlijke relaties. Hij is auteur van een vierdelige handleiding beleidsevaluatie en coördinator van de open postacademische vorming “beleidsevaluatie”. Hij is lid van het coördinatiecomité van het Vlaams Evaluatieplatform (VEP – [www.evaluatieplatform.be](http://www.evaluatieplatform.be)).

**Carl Devos** is hoofddocent aan de Vakgroep Politieke Wetenschappen van de Universiteit Gent en hoofd van de onderzoeksgroep GHIPS (Ghent Institute for Political Studies) en DEMAPOL (Research Centre on Decision Making in Politics and Policy). Hij doctorerde over de veranderde rol van vakbonden in een geglobaliseerde wereld. Vandaag doceert hij over Politicologie, Politieke Besluitvorming, en het Belgisch Federalisme. Zijn onderzoeksinteresse beslaat een brede waaier onderwerpen waaronder politieke ideologie, federaliseringsprocessen en sociaal overleg. Carl Devos is voormalig hoofdredacteur van 'Samenleving en Politiek' en is nu hoofdredacteur van Res Publica, het Vlaams-Nederlands wetenschappelijk tijdschrift voor politieke wetenschappen. Hij is columnist voor Het Nieuwsblad en voor de VRT nieuwssite deredactie.be.

**Amy-Jane Gielen** is oprichter van A.G. Advies, een onderzoeks- en adviesbureau op het gebied van jeugd, welzijn en integratie. Zij is cum laude afgestudeerd als politicoloog met accenten in bestuurskunde, beleidswetenschap, conflict studies en radicalisme. Amy-Jane Gielen heeft diverse antiradicaliseringsprojecten van het Slotervaartse actieplan geëvalueerd: de interreligieuze dialoog, de socratische gesprekken, de religieus-seculiere kring, religieuze en culture opvoedingsondersteuning en de training identiteit en weerbaarheid voor orthodoxe moslima's. Daarnaast adviseert zij diverse publieke organisatie op het gebied van antiradicalisering. Tot slot heeft zij een gastaanstelling voor onderzoek en onderwijs aan de afdeling Politicologie van de Universiteit van Amsterdam.

**Gerd Junne** is hoogleraar Internationale Betrekkingen aan de Universiteit van Amsterdam. Hij heeft politicologie, rechten en economie gestudeerd in Berlijn en Genève en als consultant voor internationale organisaties gewerkt en veel ontwikkelingsprojecten geëvalueerd. Zijn recent onderwijs en onderzoek richt zich met name op conflict preventie, conflict transformatie en de ontwikkeling van crisisgebieden na perioden van grootschalig geweld.

**Henk L. Klaassen** heeft ruimtelijke economie gestudeerd aan de Erasmus Universiteit Rotterdam en aansluitend Urban Planning aan de School of Architecture and Urban Planning aan de UCLA. In 1980 trad hij in dienst van het Secretariaat van de Commissie voor de Ontwikkeling van Beleidsanalyse. In 1983 werd hij senior projectleider bij de afdeling Beleidsanalyse van de Directie Begrotingszaken van het Ministerie van Financiën. In 1988 volgde een overstap naar de Opleiding Bestuurskunde van de Erasmus Universiteit Rotterdam. In 1995 promoveerde hij op een proefschrift over besluitvorming in de publieke sector en de rol van beleidsanalyse daarbij. Hij is nu universitair hoofddocent Economie bij de opleiding Bestuurskunde, waarbij hij zich bezighoudt met onderwijs en onderzoek op het terrein van bedrijfsvoeringstechnieken in de publieke sector, financieel management, verzelfstandiging en sturingsvraagstukken.

**Lannoo Steven** is als doctorandus verbonden aan de Vakgroep Politieke Wetenschappen van de Universiteit Gent en de onderzoeksgroep GHIPS (Ghent Institute for Political Studies). Onder het promotorschap van Prof. dr. Carl Devos doet hij onderzoek naar sociale verbondenheid en sociaal kapitaal op het platteland. Een deel van zijn onderzoek spitst zich toe op de werking en evaluatie van projecten die sociale verbondenheid moeten versterken.

**Petra Meier** is docente Politieke Wetenschappen aan de Universiteit Antwerpen. Haar onderzoeksinteresses betreffen democratie en vertegenwoordiging, kiesstelsels en instrumenten voor “groepsvertegenwoordiging”, feministische beleidsanalyse, de vrouwenbeweging, “state feminism” en de Belgische politiek. Recent publiceerde zij ‘*The Discursive Politics of Gender Equality. Stretching, Bending and Policymaking*’ (red. samen met Emanuela Lombardo en Mieke Verloo, 2009, Routledge) evenals ‘*Genre et politique transnationale: entre mouvements sociaux et politiques publiques*’ (red. samen met Bérengère Marques-Pereira en David Paternotte, 2009, Academia Bruylant).

**Valérie Pattyn** (Licentiaat in de Politieke Wetenschappen, K.U.Leuven & University of Exeter, Kandidaat in de Pedagogische Wetenschappen K.U.Leuven) is als assistent verbonden aan het Instituut voor de Overheid, waar zij betrokken is bij onderwijs en doctoraatsonderzoek rond het thema van beleidsevaluatie en evidence-based beleid.

**Walter Van Dongen** is doctor in de moraalwetenschappen (RUG, Gent) en licentiaat in de economische wetenschappen (KUL, Leuven). Hij werkt op de Studiedienst van de Vlaamse Regering (SVR) en doet interdisciplinair beleidsgericht onderzoek naar de tijdsverdeling van mannen en vrouwen binnen gezinnen en bedrijven en de wisselwerking met de arbeidsmarkt, het fiscale systeem en de publieke voorzieningen. Hij gaat uit van de positieve interactie tussen de conceptuele, normatieve en empirische dimensie, als basis voor toekomstgerichte beleidsmodellen, beleidsperspectieven en praktische beleidsinstrumenten.

**Frans van Nispen** heeft bestuurskunde en politicologie gestudeerd aan de Universiteit van Leiden. Hij is –na zijn militaire dienstplicht te hebben vervuld– zijn loopbaan begonnen bij het secretariaat van de Commissie voor de Ontwikkeling van Beleidsanalyse (1977-1980), dat was ondergebracht bij het Ministerie van Financiën te ‘s-Gravenhage. Op dit moment is hij universitair hoofddocent Bestuurskunde aan de Erasmus Universiteit Rotterdam waar hij tevens is gepromoveerd op een onderzoek naar het gebruik van de resultaten van de heroverweging van overheidsuitgaven (1993). Hij is daarnaast als Affiliated Professor of Public Policy verbonden aan George Mason University te Fairfax, Virginia. Hij heeft verscheidene publicaties op zijn naam staan over onderwerpen op het raakvlak van beleid, in het bijzonder de begroting en onderzoek in international vergelijkend perspectief. In opdracht van de EU, de OECD en WB heeft hij advies gegeven op het gebied van de openbare financiën.



**Dries Verlet** is doctor in de politieke wetenschappen (UGent, faculteit Politieke en Sociale Wetenschappen). Na een doctorassistentschap aan de Vakgroep Politieke Wetenschappen (UGent) is hij sinds november 2007 werkzaam op de Studiedienst van de Vlaamse Regering. Op de Studiedienst is hij adviseur beleidsevaluatie binnen het team kwaliteit, statistiek, survey en toekomstverkenningen. Vanuit deze functie is Dries betrokken bij een diversiteit aan beleidsevaluatieonderzoeken binnen de Vlaamse overheid. Daarnaast is hij gastdocent aan het Departement Handelswetenschappen en Bestuurskunde van de Hogeschool Gent waar hij onderzoeksmethoden doceert binnen de opleiding Bestuurskunde en Publiek Management.

**Ellen Wayenberg** is docent aan het Departement Handelswetenschappen en Bestuurskunde van de Hogeschool Gent en gastdocent aan het Departement Politieke Wetenschappen van de Universiteit Gent. Tot 1 oktober 2006 was Ellen werkzaam aan het Instituut voor de Overheid (KUL) waar zij in 2004 promoveerde met een proefschrift getiteld *'Vernieuwingen in de Vlaamse centrale-lokale verhoudingen: op weg naar partnerschap? Een kwalitatieve studie van de totstandkoming en uitvoering van het sociaal impulsbeleid'*. Ellen werkt en werkte aan diverse projecten en onderzoeken rond uiteenlopende facetten van het binnenlands beleid en bestuur. Ellen is lid van de Vlaamse Adviesraad voor Bestuurszaken en van diverse redactieraden. Tevens is ze medevoorzitster van de Permanent Study Group on Intergovernmental Relations in het kader van de European Group on Public Administration (EGPA).

## **Hoofdstuk 1: Beleidsevaluatie: gelijk hebben ≠ gelijk krijgen**

**Henk Klaassen & Frans van Nispen**

Al enige decennia is er een internationale trend waarneembaar in de richting van een steeds grotere prestatieoriëntatie als het gaat om overheidshandelen. Dit beperkt zich niet alleen tot het beleidsvormende deel van het overheidshandelen, maar heeft evenzeer betrekking op de wijze waarop verantwoording wordt afgelegd. Beleidsonderzoek ter facilitering van beleidsvormingsprocessen (beleidsevaluatie ex ante) heeft dan ook een enorme vlucht genomen. De focus op beleidsprestaties lijkt een onbedoeld neveneffect te hebben in die zin dat analyses ten behoeve van beleid steeds partiëler worden (de focus op beoogde effecten). Belangengroepen, stakeholders zien hun belangen minder goed vertegenwoordigd en komen met eigen beleidsonderzoeken. De institutionalisering van belangen die zich in de samenleving heeft voorgedaan, heeft zich daardoor onder andere vertaald in een proliferatie van beleidsonderzoek. Beleidsvoerders laten zich bijstaan door een interne of externe onderzoekers. Het streven is daarbij niet langer gericht op een integrale afweging van alle belangen, maar slechts op de onderbouwing van een al vooringenomen standpunt, dan wel de ondermijning van het standpunt van de tegenstanders in de politieke arena. Zij zijn tegelijkertijd tot elkaar veroordeeld, hetgeen noopt tot compromissen en “trade-offs”, waarbij de uitslag niet zozeer wordt bepaald door de macht van het argument, als wel door de macht van de sterkste. Het doel heiligt de middelen: “gelijk hebben” is ondergeschikt aan “gelijk krijgen”. Misschien wordt het tijd voor een terugkeer van de integrale beleidsevaluatie.

## **Hoofdstuk 2: Prestatiemeting: doel of middel? Over de niet-evidente meting van prestaties binnen de publieke sector**

**Dries Verlet & Carl Devos**

Hoewel beleidsevaluatie veel meer omvat dan het in kaart brengen van de prestaties binnen de publieke sector, is prestatiemeting een centraal gegeven binnen het onderzoeks- en toepassingsveld. In deze bijdrage gaat onze bijzondere aandacht gaat uit naar de potentiële onbedoelde negatieve effecten die een dergelijke meting kan hebben op het functioneren van een organisatie zelf. Hiermee willen we geenszins een afbreuk doen aan het belang van prestatiemeting. We willen enkel bijdragen tot een bedachtzame meting, die onontbeerlijk is indien we op een correcte manier o.a. de efficiëntie en effectiviteit van het gevoerde beleid in kaart willen brengen. Het meten van overheidsprestaties, waarover velen zich graag uitspreken, is immers geen evidente opdracht. Deze bijdrage bestaat uit twee delen. In eerste instantie staan we stil bij de uitdagingen van het meten van overheidsbeleid of -prestaties. Daarna komt het hoofdonderwerp aan bod: welke positieve en negatieve effecten roept publieke prestatiemeting op? Hoe kunnen we de negatieve effecten verklaren en voorkomen?

### **Hoofdstuk 3: Preventief gedrag bij beleidsevaluaties** **Frank Bongers**

Beleidsvoerders, ofwel politiek-ambtelijke leidinggevenden en hun ambtelijke staf, kunnen preventief gedrag gaan vertonen wanneer zij weten dat een beleid waar zij verantwoordelijk voor zijn later het onderwerp van een evaluatie wordt. Dit gedrag kan ongewenst zijn, bijvoorbeeld wanneer beleidsvoerders op basis van deze (voor)kennis besluiten doelstellingen minder concreet en minder ambitieus formuleren, zodat de kans op een positieve beoordeling van de resultaten en effecten van het gevoerde beleid achteraf toeneemt. Het is echter maar de vraag of beleidsvoerders zich zo zullen gedragen. Deze gedachtegang veronderstelt immers dat preventief gedrag vooral veroorzaakt wordt door beleidsevaluaties en dat dit gedrag altijd ongewenst is. De praktijk is echter weerbarstiger en kent voldoende factoren die preventief gedrag afzwakken. Dit hoofdstuk behandelt deze factoren en gaat ook kort in op het effect van beleidsevaluaties op het beleidsproces.

### **Hoofdstuk 4: Risico's op perverse effecten bij beleidsevaluatie? Een vergelijking vanuit mogelijke evaluatiemotieven tegen de achtergrond van de Vlaamse hervormingscontext** **Bart De Peuter, Valérie Pattyn & Marleen Brans**

Bestuurskundigen refereren vaak aan ongewenste neveneffecten verbonden aan de praktijk van prestatiemeting. Ook worden uiteenlopende motieven geïdentificeerd om beleid te evalueren. In dit hoofdstuk gaan de auteurs na in welke mate de verschillende motieven om beleid te evalueren meer of minder risico's op perverse effecten inhouden, met name in de vorm van distorties tijdens de processen van het meten zelf en/of het vormen van waardeoordelen over het beleid als evaluandum. Deze analyse wordt gebaseerd op de aandachtspunten die de literatuur aanreikt en in vergelijkend perspectief toegepast op mogelijke evaluatiemotieven tegen de achtergrond van recente hervormingstrends bij de Vlaamse overheid. Welke motieven prevaleren? En welke mogelijke neveneffecten dienen we hierbij in te calculeren? Een inzicht in evaluatiemotieven en hun vatbaarheid voor disfuncties brengt de auteurs tot het formuleren van een aantal pistes om de risico's op perverse effecten te verkleinen.

### **Hoofdstuk 5: De plaats en betekenis van beleidsevaluaties in de Vlaamse overheid –de domeinen Binnenlands Bestuur en Stedenbeleid verkend** **Ellen Wayenberg**

Dit hoofdstuk omvat een eerste verkenning van de plaats en betekenis van beleidsevaluaties in de Vlaamse overheid. Die verkenning werd beperkt tot de domeinen Binnenlands Bestuur en Stedenbeleid die op vlak van beleidsvoering doorgaans vlot mee-evolueren met gangbare Europese trends en ontwikkelingen. Is dit ook het geval in verband met die trend tot (meer) beleidsevaluatie die zich doorheen Europa sinds de eeuwwisseling manifesteerde? Om dit te achterhalen werd een eigen analyse van beleidsdocumenten in beide domeinen aangevuld met een bevraging van politieke en ambtelijke beleidsverantwoordelijken. Daaruit bleek dat beleidsevaluatie vandaag duidelijk haar plaats heeft verworven in beide domeinen

want de aandacht voor en de wil tot het (laten) evalueren van beleidsmaatregelen en –initiatieven is er onmiskenbaar. En terecht volgens de bevraagde beleidsverantwoordelijken die in evalueren duidelijk een weg erkenden tot het verbeteren van “hun” beleid. Zij erkenden echter ook dat die weg vandaag nog onvoldoende wordt bewandeld want resultaten uit evaluatiestudies werken nog te weinig door in de beleidsvoering. Daarmee sloegen ze zichzelf als beleidsverantwoordelijken op de borst en getuigden ze niet terug te schrikken van enige kritische zelfevaluatie. En precies zo’n attitude is onontbeerlijk om ervoor te zorgen dat beleidsevaluaties in de toekomst niet alleen een plaats hebben in de Vlaamse beleidsvoering maar ook voldoende betekenis.

## **Hoofdstuk 6: Evaluatie van antiradicaliseringsprojecten. Hoe meet je of radicalisering wordt tegengegaan?**

**Amy-Jane Gielen & Gerd Junne**

In november 2004 werd Nederland voor het eerst geconfronteerd met religieus gemotiveerd terrorisme op eigen bodem. Sindsdien zijn er op lokaal niveau allerlei initiatieven ontwikkeld om polarisatie en radicalisering tegen te gaan, bijvoorbeeld door het stimuleren van interetnisch en interreligieus contact of het vergroten van de weerbaarheid van (moslim)jongeren. Over het effect van deze initiatieven is echter maar weinig bekend. In toenemende mate is er belangstelling voor “wat werkt?”, maar een antwoord blijft om methodische redenen vaak uit. Hoe meet men immers of radicalisering is tegengegaan? Wij betogen dat traditionele evaluatiemethoden onvoldoende handvatten bieden. Door te kijken naar methoden die worden gebruikt om vredesopbouwprojecten te evalueren en vervolgens te bespreken hoe deze methoden ingezet kunnen worden om antiradicaliseringsbeleid te evalueren, probeert dit hoofdstuk een brug te slaan tussen het lokale en internationale niveau.

## **Hoofdstuk 7: Doelmatige publieke diensten in een democratische samenleving: een geïntegreerde benadering en praktische evaluatie-instrumenten**

**Walter Van Dongen**

De centrale vraag is in welke mate overheden en gesubsidieerde organisaties, als volwaardige productieve actoren in het marktsysteem, hun publieke diensten doelmatig voortbrengen en aanbieden aan de andere actoren. Uitgaande van een algemene benadering van de diverse soorten dienstverlening in het marktsysteem worden de begrippen effectiviteit en efficiëntie bepaald, met aandacht voor de eigenheid van de publieke sector en voor de betekenis van andere basiswaarden zoals vrijheid, gelijkheid en solidariteit en de doelstellingen die daaruit voortvloeien.

Het meten van de doelmatigheid van publieke diensten vereist in de praktijk een correcte meting/registratie van een aantal basisgegevens van de productieprocessen binnen de afgebakende diensten, aan de input- en outputzijde, uitgaande van de beleidsdoelen met betrekking tot beschikbaarheid, gebruik, bezettingsgraad, productiviteit, kostenefficiëntie en aandeel van de persoonlijke bijdrage. Er wordt gepleit voor de ontwikkeling van praktische evaluatie-instrumenten op organisatieniveau, als een kwantitatieve basis voor een brede beleidsevaluatie, op maat van de sectoren en organisaties.

Ten slotte wordt de evolutie van de dagopvang van kinderen jonger dan drie jaar in het Vlaamse Gewest in de periode 1990-2002 toegelicht aan de hand van enkele basisindicatoren.

## **Hoofdstuk 8: Als meten het doel wordt: de vertaling van gender mainstreaming naar de beleidspraktijk**

**Petra Meier & Karen Celis**

Sinds het midden van de jaren 1990 staat in het gelijke kansenbeleid het concept gender mainstreaming centraal: het bevorderen van gendergelijkheid in de verschillende beleidsdomeinen. Meer dan tien jaar later laten de resultaten volgens veel critici te wensen over. Deze bijdrage analyseert de verschillende pogingen tot implementatie van gender mainstreaming in het Belgische beleid. Hierbij wordt gefocust op de vraag of de beleidsevaluatie en -monitoring die vaak centraal staan in het gender mainstreamingsbeleid een middel zijn om gendergelijkheid te realiseren dan wel een doel op zich zijn geworden. Zowel het federale, als het Vlaamse, het Waalse en het Brusselse gelijke-kansenbeleid wordt vanuit die vraagstelling belicht. De conclusie is dat gender mainstreaming in België de neiging vertoont om te vervallen in een formalistisch proces van meten en weten, maar dat er niettemin een hoopvolle evolutie is vast te stellen naar meer output en outcome gericht beleid.

## **Hoofdstuk 9: Een Dorpshuis als Ontmoetingsplaats: Ontwerp van Beleidsevaluatie**

**Steven Lannoo & Dries Verlet**

In deze bijdrage wordt een ontwerp van evaluatie van een interventie in de gemeente Assenede voorgesteld. Het gaat om een plan waarbij een dorps huis wordt opgezet dat moet tegemoet komen aan een behoefte aan meer lokale dienstverlening, maar dat vooral ook de sociale samenhang in het dorp moet versterken. In een eerste deel van de bijdrage wordt het project dieper onder de loep genomen, worden de stakeholders geïdentificeerd en de beleidstheorie nader omschreven. Vervolgens worden het methodologische raamwerk waarbinnen het evaluatieontwerp zich situeert dieper uitgespit. In het laatste deel passen we de theorie toe op de case en stellen op die manier een ontwerp van beleidsevaluatie voor.

## Lijst met figuren

<b>Beleidsevaluatie: een inleidende situering .....</b>	<b>7</b>
Figuur 1: Het productieproces in de publieke sector .....	11
Figuur 2: Beleidsevaluatie geplaatst binnen de beleidscyclus .....	13
Figuur 3: Evaluatiecriteria afgeleid uit de beleids- en beheerscyclus als open systeemmodel .....	14
<b>Hoofdstuk 1: Beleidsevaluatie: gelijk hebben ≠ gelijk krijgen .....</b>	<b>20</b>
Figuur 1.1: Plaatsing van beleidsanalyse .....	23
Figuur 1.2: Het concept kosten-batenanalyse .....	25
Figuur 1.3: Het concept effectrapportage .....	28
Figuur 1.4: Het concept performance management .....	31
<b>Hoofdstuk 2: Prestatiemeting: doel of middel? Over de niet-evidente meting van prestaties binnen de publieke sector .....</b>	<b>40</b>
Figuur 2.1.: Dimensies bij beleids- of prestatie-meetsystemen .....	48
<b>Hoofdstuk 5: De plaats en betekenis van beleidsevaluaties in de Vlaamse overheid –de domeinen Binnenlands Bestuur en Stedenbeleid verkend ...</b>	<b>106</b>
Figuur 5.1: Modellen inzake het gebruik van onderzoek van overheidswege .....	117
<b>Hoofdstuk 6: Evaluatie van antiradicaliseringsprojecten. Hoe meet je of radicalisering wordt tegengegaan? .....</b>	<b>122</b>
Figuur 6.1: Het antiradicaliseringsbeleid van de Gemeente Amsterdam .....	126
Figuur 6.2: Splijtzwammen vs. verbindingsschakels: een kader voor conflicttransformatie .....	132
Figuur 6.3: Splijtzwammen vs. verbindingsschakels in internationale conflictgebieden .....	133

Figuur 6.4: Splitszwammen vs. verbindingsschakels: een kader voor conflictransformatie .....	134
Figuur 6.5: Analyse kader voor ontwerp en evaluatie van interventiestrategieën .....	136
Figuur 6.6: Vier delen van het Peace and Conflict Assessment framework .....	137
Figuur 6.7: Indicatoren voor de toetsing van antiradicaliseringsprojecten .....	140
Figuur 6.8: De projecten in Slotervaart in kaart gebracht .....	141
Figuur 6.9: Hypothesen t.a.v. het effect .....	147
<b>Hoofdstuk 7: Doelmatige publieke diensten in een democratische samenleving: een geïntegreerde benadering en praktische evaluatie-instrumenten .....</b>	<b>152</b>
Figuur 7.1: Een algemeen model voor menselijke activiteiten .....	154
Figuur 7.2: Verdeling van de dienstverlening onder de diverse aanbieders .....	158
Figuur 7.3: Breder model voor menselijke activiteiten en efficiëntie/effectiviteit .	163
<b>Hoofdstuk 9: Een Dorpshuis als Ontmoetingsplaats: Ontwerp van Beleidsvaluatie .....</b>	<b>207</b>
Figuur 9.1: Veranderingsmodel .....	211



## Lijst met tabellen

### **Hoofdstuk 4: Risico's op perverse effecten bij beleidsevaluatie? Een vergelijking vanuit mogelijke evaluatiemotieven tegen de achtergrond van de Vlaamse hervormingscontext ..... 77**

Tabel 4.1: Evaluatiemotieven vermeld in beleidsnota's van de Vlaamse Regering (2004-2009) ..... 86

Tabel 4.2: Vergelijkende inhoudsanalyse van 7 beheersovereenkomsten (Rekenhofstudie en eigen studie) ..... 93

### **Hoofdstuk 5: De plaats en betekenis van beleidsevaluaties in de Vlaamse overheid –de domeinen Binnenlands Bestuur en Stedenbeleid verkend ..... 106**

Tabel 5.1: Voorwerp van aangekondigde beleidsevaluaties in de domeinen Binnenlands Bestuur en Stedenbeleid (*en van extern uitgevoerd evaluatieonderzoek*) ..... 114

Tabel 5.2: Modellen van “research utilization” ..... 115

### **Hoofdstuk 7: Doelmatige publieke diensten in een democratische samenleving: een geïntegreerde benadering en praktische evaluatie-instrumenten ..... 152**

Tabel 7.1: Aantal volledige plaatsen per 1000 kinderen van drie maand tot drie jaar (Vlaams Gewest, 1990-2002) ..... 168

Tabel 7.2: Aantal gebruikers per 1000 kinderen van drie maand tot drie jaar (Vlaams Gewest, 1990-2002) ..... 169

Tabel 7.3: Volle verblijfsdagen per gebruiker per jaar en de benuttingsgraad in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002) ..... 169

Tabel 7.4: Aantal volle verblijfsdagen per gebruiker per week in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002) ..... 170

Tabel 7.5: Aantal volledige verblijfsdagen per plaats en bezettingsgraad in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002) ..... 170

Tabel 7.6: Aantal plaatsen per VTE in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002) .....	171
Tabel 7.7: Aantal volledige verblijfsdagen per VTE in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002) .....	171
Tabel 7.8: Kostprijs per volledige plaats in de gesubsidieerde opvang (Vlaams Gewest, 1990-2002) .....	172
Tabel 7.9: Kostprijs per volledige verblijfsdag in de gesubsidieerde dagopvang (Vlaams Gewest, 1990-2002) .....	173
Tabel 7.10: Kostprijs per verblijfsdag bij de opvang thuis (Vlaams Gewest, 1990-2002) .....	173

