

## CIP-GEGEVENS

ISBN 90 6092 475 4  
NUGI 654

© 1991 J.Th.A. Bressers en A. Hoogerwerf, Enschede

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikel 16b Auteurswet 1912 jº het Besluit van 20 juni 1974, Stb. 351, zoals gewijzigd bij Besluit van 23 augustus 1985, Stb. 471 en artikel 17 Auteurswet 1912, dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (postbus 882, 1180 AW Amstelveen). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezing, readers en andere compilatiewerken (artikel 16 Auteurswet) dient u zich te richten tot: Samsom H.D. Tjeenk Willink, postbus 4, 2400 MA Alphen aan den Rijn.

# Inhoud

Woord vooraf	11
Auteurs	13
<b>DEEL I</b>	
<b>INLEIDING</b>	
<b>1 Inleiding tot de beleidsevaluatie</b>	19
<i>prof.dr. J.Th.A. Bressers en prof.dr. A. Hoogerwerf</i>	
1.1 Opzet van het hoofdstuk	19
1.2 Evaluatie	21
1.3 Beleidsevaluatie	22
1.4 Wetenschappelijke beleidsevaluatie	26
1.5 Opzet van het boek	29
1.6 Samenvatting	31
Literatuur	31
Vragen en opgaven	32
<b>DEEL II</b>	
<b>STAPPEN IN HET EFFECTIVITEITSONDERZOEK</b>	
<b>2 De voorbereiding van evaluatie-onderzoek</b>	35
<i>dr. M.J. Arentsen</i>	
2.1 Inleiding	35
2.2 Het gebruik van evaluatie-onderzoek in het beleidsproces	36
2.3 Van bevraagingsvraag tot onderzoeksvraag	38
2.4 De keuze van het onderwerp en de evaluatiecriteria	42
2.5 De keuze van de onderzoeksmethoden	47
2.6 Samenvatting	48

6		Inhoud	
	Literatuur	48	
	Vragen en opgaven	49	
<b>3</b>	<b>Hoe zien het beleid en het beleidsveld er uit?</b>	50	
	<i>prof.dr. J.Th.A. Bressers en prof.dr. A. Hoogerwerf</i>		
	3.1 Inleiding	50	
	3.2 Het in kaart brengen van het beleid en het beleidsveld	51	
	3.3 De gebruiksmogelijkheden van een model van beleid en beleidsveld bij evaluatie-onderzoek	54	
	3.4 Beleidsevaluatie als toetsing van een beleidstheorie	60	
	3.5 Reconstructie en evaluatie van een beleidstheorie	62	
	3.6 Samenvatting	67	
	Literatuur	68	
	Vragen en opgaven	68	
<b>4</b>	<b>Het vaststellen van de mate van doelbereiking</b>	70	
	<i>drs. K.R.D. Lulofs en drs. J. Schuddeboom</i>		
	4.1 Inleiding	70	
	4.2 Het operationaliseren van de beleidsdoelen	71	
	4.3 De informatiebronnen en de dataverzameling	73	
	4.4 De interpretatie van de gegevens	79	
	4.5 Samenvatting	82	
	Literatuur	83	
	Vragen en opgaven	83	
<b>5</b>	<b>Wat is de bijdrage van het beleid aan de doelbereiking?</b>	85	
	<i>Deel 1: Kwantitatieve methoden</i>		
	<i>drs. F.H.J.M. Coenen</i>		
	5.1 Inleiding	85	
	5.2 Causaliteit	86	
	5.3 Quasi- en non-experimentele onderzoeksopzetten	88	
	5.4 Samenvatting	100	
	Literatuur	100	
	Vragen en opgaven	100	

7		Inhoud	
<b>6</b>	<b>Wat is de bijdrage van het beleid aan de doelbereiking?</b>	102	
	<i>Deel 2: Kwalitatieve methoden</i>		
	<i>prof.dr. J.Th.A. Bressers en prof.dr. M. Herweijer</i>		
	6.1 Inleiding	102	
	6.2 Het belang van kwalitatieve methoden naast kwantitatieve methoden	103	
	6.3 Kwalitatieve methoden samen met kwantitatieve methoden	105	
	6.4 De beslisboomanalyse als methode van effectiviteitsonderzoek	111	
	6.5 Verklaren met een case-study	114	
	6.6 Samenvatting	119	
	Literatuur	120	
	Vragen en opgaven	120	
<b>7</b>	<b>Hoe valt de effectiviteit van beleid te verklaren?</b>	122	
	<i>Deel 1: Empirisch onderzoek</i>		
	<i>prof.dr. J.A.M. Maarse</i>		
	7.1 Inleiding	122	
	7.2 De uitvoering en doorwerking van een beleid	123	
	7.3 Uitvoeringsonderzoek en de verklaring van de effectiviteit van beleid	125	
	7.4 De evaluatie van de beleidsuitvoering	128	
	7.5 Analyse van de doorwerking van een beleid	131	
	7.6 Samenvatting	134	
	Literatuur	134	
	Vragen en opgaven	134	
<b>8</b>	<b>Hoe valt de effectiviteit van beleid te verklaren?</b>	136	
	<i>Deel 2: Instrumententheorie</i>		
	<i>prof.dr. J.Th.A. Bressers en dr. P.-J. Klok</i>		
	8.1 Inleiding	136	
	8.2 Uitgangspunten voor een instrumententheorie	137	
	8.3 Beleidseffectiviteit en de gedragskeuze van de doelgroep	139	
	8.4 De toepassing van beleidsinstrumenten	143	
	8.5 Voorbeeld: het schone-autobeleid	147	
	8.6 Samenvatting	153	
	Literatuur	154	
	Vragen en opgaven	154	

<b>9</b>	<b>Hoe wordt het evaluatie-onderzoek gepresenteerd en gebruikt?</b>	<b>155</b>
	<i>dr. I.M.A.M. Pröpper</i>	
9.1	Inleiding	155
9.2	Dimensies en omschrijving van het gebruik van evaluatie-onderzoek	155
9.3	De presentatie van het evaluatie-onderzoek	158
9.4	De aansluiting tussen het evaluatie-onderzoek en de beleidspraktijk	162
9.5	Wegen waarlangs de bruikbaarheid van het evaluatie-onderzoek kan worden vergroot	168
9.6	Samenvatting	171
	Literatuur	172
	Vragen en opgaven	172

**DEEL III****ANDERE VORMEN VAN BELEIDSEVALUATIE**

<b>10</b>	<b>Kosten-batenanalyse en decisie-analyse</b>	<b>175</b>
	<i>prof.dr. N.P. Mol en dr. H.J. Blommestein</i>	
10.1	Economische evaluatie	175
10.2	Kosten-batenanalyse	176
10.3	Decisie-analyse	188
10.4	Samenvatting	191
	Literatuur	192
	Vragen en opgaven	192
<b>11</b>	<b>Doelmatig beheer en beleid</b>	<b>194</b>
	<i>dr. F.C.J. van der Doelen en dr. F.L. Leeuw</i>	
11.1	Inleiding	194
11.2	De recht- en doelmatigheidspiramide	195
11.3	Interne controle door departementale onderdelen: evaluatie-onderzoek	198
11.4	De centrale afdeling Financieel-Economische Zaken: de prestatiebegroting	200
11.5	De Inspectie der Rijksfinanciën: de heroverwegingsprocedure	203
11.6	De Algemene Rekenkamer: economy, efficiency, effectiviteit	205
11.7	Grenzen en mogelijkheden van doelmatigheidsonderzoek	209

11.8	Samenvatting	211
	Literatuur	211
	Vragen en opgaven	212
<b>12</b>	<b>Juridische evaluatie van beleid</b>	<b>214</b>
	<i>mr.dr. H.M. de Jong en mr.dr. R.J. Jue</i>	
12.1	Inleiding	214
12.2	Hoofdvormen van juridische evaluatie	214
12.3	Juridische evaluatie	216
12.4	Samenvatting	226
	Literatuur	227
	Vragen en opgaven	227
<b>13</b>	<b>Hoe valt het evaluatie-onderzoek zelf te beoordelen? (meta-evaluatie)</b>	<b>229</b>
	<i>prof.dr. A. Hoogerwerf en drs. D.C. Zoutendijk</i>	
13.1	Meta-evaluatie	229
13.2	Achtergronden en vormen van meta-evaluatie	229
13.3	Eisen vanuit de wetenschap te stellen aan een evaluatierapport	232
13.4	Eisen vanuit de beleidspraktijk te stellen aan een evaluatierapport	236
13.5	Resultaten van een meta-evaluatie	240
13.6	Samenvatting	245
	Literatuur	246
	Vragen en opgaven	247
	<b>Noten</b>	<b>249</b>
	<b>Zakenregister</b>	<b>269</b>

## 4 Het vaststellen van de mate van doelbereiking

DRS. K.R.D. LULOFS EN DRS. J. SCHUDDEBOOM

### 4.1 Inleiding

*Als u van vreugde opspringt, zie dan toe dat niemand de grond onder uw voeten weghaalt.  
(Stanislaw Jerez Lec)*

In dit hoofdstuk wordt de vraag behandeld op welke wijze de mate van doelbereiking van een beleid kan worden vastgesteld. Deze vraag kan worden beantwoord zonder de oorzaken van het eventuele bereiken van de doeleinden van het beleid aan de orde te stellen. In het ongewisse blijft dan in welke mate de geconstateerde veranderingen inderdaad aan de ingezette beleidsinstrumenten kunnen worden toegeschreven en in hoeverre deze worden veroorzaakt door andere, buiten het beleid gelegen, factoren.<sup>1</sup> De vraag naar de beleidseffectiviteit blijft dan buiten beschouwing. Het vaststellen van de mate van doelbereiking vormt in evaluatie-onderzoek wel veelal een voorbereiding op het bepalen en verklaren van de mate van beleidseffectiviteit. De in dit hoofdstuk geschetste aanpak van het vaststellen van doelbereiking houdt daarom zoveel mogelijk rekening met vervolgonderzoek naar de beleidseffectiviteit en de verklaring daarvan (hoofdstukken 6 t/m 8).

De wijze waarop het bepalen van de mate van doelbereiking wordt aangepakt is – zoals vele dingen in het leven – onderhevig aan de spanning tussen het wenselijke en het haalbare. Tot de algemene eisen die aan een goed evaluatie-onderzoek moeten worden gesteld behoren de uitvoerbaarheid, accuraatheid en bruikbaarheid.<sup>2</sup> Voor het bepalen van de mate van doelbereiking moeten in ieder geval de volgende drie deelstappen worden doorlopen:

- Stap 1. Het operationaliseren van de beleidsdoeleinden of andere evaluatiecriteria. Bij het bepalen van de mate van doelbereiking zijn het uiteraard de beleidsdoelen die als maatstaf worden geoperationaliseerd.
- Stap 2. Het verzamelen van de gegevens die noodzakelijk zijn om de

onderzoeksvraag (wordt aan de criteria voldaan?) te kunnen beantwoorden.

Stap 3. Het interpreteren van de gegevens (data) en het trekken van conclusies.

Deze deelstappen vormen het uitgangspunt voor de rest van dit hoofdstuk. In paragraaf 4.2 wordt de operationalisatie van begrippen behandeld, in paragraaf 4.3 worden de te gebruiken informatiebronnen en de dataverzameling besproken. De data-interpretatie en het trekken van conclusies komen in paragraaf 4.4 aan bod. Het hoofdstuk wordt afgesloten met een samenvatting (par. 4.5). Belangrijke aandachtspunten, potentiële moeilijkheden en mogelijke remedies worden voor elk van de drie deelstappen aangegeven. De potentiële moeilijkheden liggen op het terrein van de bruikbaarheid, de geldigheid en betrouwbaarheid van het onderzoek. De mogelijke oplossingen moeten worden gezocht in een zorgvuldige operationalisatie, een weloverwogen dataverzameling met gebruik van kwalitatieve en kwantitatieve gegevens uit verschillende bronnen, en een zorgvuldige data-interpretatie.

### 4.2 Het operationaliseren van de beleidsdoelen

Het ligt voor de hand dat de evaluator die de mate van doelbereiking wil vaststellen, in overleg met de opdrachtgever de beleidsdoelen als bron van evaluatiecriteria kiest. Dit lijkt erg eenvoudig maar in de praktijk zijn er enkele complicerende factoren. De officiële beleidsdoelen geven niet altijd een juist of bruikbaar beeld van wat de beleidsvoerders werkelijk willen bereiken. Dit pleit echter niet voor het verwerpen van het gebruik van officiële beleidsdoeleinden als evaluatiecriteria. Ook wanneer de beleidsvoerders iets anders nastreven dan de officiële doeleinden kunnen deze in belangrijke mate de legitimatie van het beleid in het democratisch proces vormen. Uit een oogpunt van democratische controle kunnen de officiële doeleinden uitstekend als maatstaf worden gebruikt.<sup>3</sup>

Een ander probleem voor de evaluator kan liggen in de wijze waarop de officiële beleidsdoeleinden zijn geformuleerd. Deze kunnen erg vaag geformuleerd, onrealistisch hoog of laag, veranderlijk of zelfs tegenstrijdig zijn. In deze paragraaf houden wij ons bezig met de wijze waarop men beleidsdoelen operationeel en dus meetbaar kan maken. Soms lukt dit probleemloos, bijvoorbeeld als de beleidsdoelen op een demate laag niveau van abstractie zijn geformuleerd dat zij direct meetbaar zijn. Voorbeelden daarvan zijn: 'een lagere gemiddelde snelheid op de autowegen' of 'een lager gasgebruik per eengezinswoning'. Vaker zal er echter sprake zijn van beleidsdoelen die geen betrekking hebben op direct observeerbare zaken.<sup>4</sup> Dit vraagt om een vertaling van de voor het evaluatie-onderzoek geëxpliciteerde doeleinden in meetbare termen. Wan-

neer een doel bijvoorbeeld luidt 'een goede kwaliteit van het fluviale milieu' moet bijvoorbeeld allereerst het begrip 'fluviaal milieu' worden verduidelijkt.

Meestal zal deze problematiek door de betrokkenen al bij de uitwerking van het evaluatieproject aan de orde zijn gesteld. Vervolgens moet de evaluator komen tot een omschrijving van de essentie van een 'goede kwaliteit van het fluviale milieu'. De meest relevante aspecten of onderdelen van het fluviale milieu zullen al bij de systematische beschrijving van het onderwerp van de evaluatie zijn geïnventariseerd (zie hoofdstuk 3). In essentie is de problematiek van de geldige operationalisering een vertaalprobleem. Het gaat om de vertaling van spreek- of schrijftaal in onderzoekstaal.<sup>5</sup> De evaluator ziet zich verder voor de taak gesteld het verschijnsel 'kwaliteit van het fluviale milieu' te indiceren in meetbare termen. Indiceren in meetbare termen wil zeggen het ontwikkelen van indicatoren die een zo adequaat en volledig mogelijke afbeelding van het originele begrip vormen en – in tegenstelling daarmee – meetbaar zijn.

Voor een beleidsdoel als 'een goede kwaliteit van het fluviale milieu' moet de mate van doelbereiking via deze indirecte weg worden gemeten. Van belang is in dit verband de opmerking van Boesjes-Hommes dat indicatoren niet met behulp van de regels van de logica uit de betekenis van de begrippen kunnen worden afgeleid. Noch biedt de opvatting dat er sprake is van een causale relatie tussen begrip en indicator perspectieven.<sup>6</sup> Het belang van een zorgvuldige operationalisatie staat echter buiten kijf wil men komen tot een goed evaluatieonderzoek. Drie vragen kunnen de evaluator bij de vervulling van zijn operationalisatetaak als hulpmiddel dienen:<sup>7</sup>

1. *Is er een voldoende sterke relatie tussen het te operationaliseren doel en de indicatoren?*

Zo is bijvoorbeeld de mate waarin leden van culturele minderheden werk hebben gevonden wellicht een redelijke indicator van de economische integratie van deze minderheden in de Nederlandse samenleving, maar niet voor hun politieke of culturele integratie.

2. *Wordt het te evalueren doel in al zijn dimensies omvat door de gekozen indicatoren?*

Het streven om tot meetbare indicatoren te komen leidt vaak tot een zekere beperking of zelfs verschuiving van het evaluatiecriterium (in casu meestal de beleidsdoelen). Vaak neemt men er genoegen mee dat slechts een deel van het evaluatiecriterium via de operationalisatie wordt gemeten. Om dit informatieverlies te compenseren kan worden gegrepen naar de constructie van een index. Een index is een schaal die is samengesteld uit een groep indicatoren waarop wordt gescoord. Men gaat er daarbij van uit dat als op al deze indicatoren te zamen wordt gelet, een redelijke afbeelding ontstaat van het evaluatiecriterium dat men wilde operationaliseren. Een voorbeeld van een complex beleidsdoel

is het verminderen van verzuring. De operationalisatie van dit beleidsdoel vraagt om inzicht in de stoffen die bijdragen aan de verzuring, het inwerkingsproces en de wijze waarop de verspreiding van de verzurende stoffen in zijn werk gaat.

3. *Zijn er interveniërende factoren die verandering aanbrengen in de afbeelding die de indicator maakt van het doel?*

Een voorbeeld van een dergelijke complicatie is de operationalisatie van het doel 'terugdringing van de milieucriminaliteit' in de vorm van de indicator 'aantal door de politie geconstateerde criminele daden met milieu-effecten'. Deze wijze van operationalisatie kan een ernstig gevaar voor de validiteit van de meting van de doelbereiking meebrengen. Dit is bijvoorbeeld het geval indien gelijktijdig met het criminaliteitbestrijdingsbeleid een intensivering van het handhavingsbeleid wordt doorgevoerd. De operationalisatie is dan waarschijnlijk niet meer valide als meting van de milieucriminaliteit, omdat een krachtiger politieoptreden in de regel zal leiden tot een toeneming van de geregistreerde criminaliteit. De statistieken geven dan niet alleen de eventueel toegenomen milieucriminaliteit weer, maar ook de toegenomen 'pakkans'.

#### 4.3 De informatiebronnen en de dataverzameling

Het meetinstrument omvat behalve de operationalisatie ook de wijze en de tijdstippen waarop de data worden verzameld. Er zijn geen universeel geldende uitspraken te doen over de vraag welke gegevens en hoeveel gegevens verzameld moeten worden in een evaluatieonderzoek. Bij het maken van de verschillende keuzes zullen onder meer de beschikbare tijd en het beschikbare geld een rol spelen. Wel is het steeds belangrijk om de algemene eisen van uitvoerbaarheid, accuraatheid en bruikbaarheid van het onderzoek in het oog te houden. Patton heeft deze eisen nader uitgewerkt en op deze wijze voor 'het verzamelen van data' enig houvast verschaft in de vorm van 'regels'. Een kanttekening is hierbij overigens op zijn plaats. Patton's uitwerkingen zijn met name gebaseerd op de Amerikaanse commerciële evaluatietraditie. De Nederlandse situatie, waar het accent meer ligt bij beleidsevaluaties, is daarmee niet geheel vergelijkbaar. Wij hebben daarom hieronder Patton's uitwerking enigszins aangepast voor de Nederlandse situatie. Dit resulteert in vijf vuistregels.

*De uitvoerbaarheidsregel ten aanzien van de te gebruiken methoden.*

De wijze waarop de evaluatie wordt aangepakt moet zo praktisch mogelijk zijn, zodat het bewerkstelligde ongerief voor hen die worden onderzocht minimaal is en toch de noodzakelijke informatie kan worden verzameld.

3) *De bruikbaarheidsregel ten aanzien van de omvang en de selectie van informatie.*

De dataverzameling moet zodanig doelgericht worden aangepakt dat de gestelde evaluatievragen over het onderwerp van evaluatie kunnen worden beantwoord. Daarbij moet rekening worden gehouden met de behoeften en interesses van de eventuele opdrachtgever en andere mogelijke gebruikers.

4) *De accuratesseregel ten aanzien van de informatiebronnen en de verzamelde data.*

De informatiebronnen waaruit wordt geput kunnen zodanig gedetailleerd worden beschreven dat de adequaatheid van de informatie kan worden bepaald. Verder moet de keuze van de informatiebronnen kunnen worden verdedigd.

5) *De validiteitsregel.*

De wijze waarop de informatie wordt verzameld en de wijze waarop de meetinstrumenten en de meetprocedures worden gekozen of ontwikkeld moeten de validiteit (geldigheid) van de uiteindelijke conclusies voor het beoogde gebruik verzekeren. Een meting (in casu het vaststellen van de mate waarin de doelen zijn bereikt) is valide (geldig) indien de evaluator met de gekozen werkwijze meet wat hij wilde meten.

6) *De betrouwbaarheidsregel.*

De instrumenten en procedures voor dataverzameling kunnen zodanig worden gekozen, ontwikkeld en uitgevoerd dat de informatie voldoende betrouwbaar is voor het beoogde gebruik. Een meting is betrouwbaar (consistent) indien bij herhaling van de meting in eenzelfde situatie in essentie dezelfde resultaten worden geproduceerd.

De bovenstaande vuistregels vormen de leidraad bij de beantwoording van de vraag hoe om te gaan met het verzamelen van data. De volgende vragen met een inventariserend karakter spelen hierbij een rol:

1. Uit welke bronnen kunnen de data worden geput?
2. Welke tijdstippen zijn relevant voor het verzamelen van data?
3. Welke technieken kan men gebruiken om data te verzamelen?

Naast een overzicht van mogelijke bronnen, relevante tijdstippen en technieken geven wij nu per vraag ook een schets van de implicaties van de verschillende keuzes die men maakt.

*Ad 1.* Bij het evalueren van beleid kan men de volgende relevante bronnen of waarnemingseenheden onderscheiden: deskundigen (of 'insiders'), de doelgroepen van het beleid, archieven en statistieken.

Deskundigen worden in de praktijk van het beleidsevaluatie-onderzoek vaak gebruikt als gegevensbron. Door bij deskundigen, die sleutelposities in het beleidsproces of in het beleidsveld bekleden, data te verzamelen wordt snel inzicht verkregen in verschillende aspecten van de beleidsdoelstellingen en van de doelbereiking in het beleidsveld. Het verzamelen van gegevens bij de doelgroep geeft vooral inzicht in de effecten van het beleid op de doelgroep. Het geformuleerde onderscheid dient slechts als globale indicatie voor het verschil in de gebruiksmogelijkheden van de bronnen. Ook bij de doelgroep is het mogelijk om inzicht in de beleidsdoelstellingen te krijgen. Omgekeerd geldt dat het ook bij de deskundigen mogelijk is om iets te kunnen achterhalen over effecten van het beleid op de doelgroep. Maar dit zijn dan indirecte metingen, hetgeen afbreuk kan doen aan de accurateheid van de gegevens.

Zowel bij de bron 'deskundigen' als bij de bron 'doelgroepen' geldt dat de waarnemingseenheden dermate talrijk kunnen zijn dat een steekproef noodzakelijk is. Dit houdt in dat men een deelverzameling kiest uit de gehele verzameling van potentiële waarnemingseenheden. Het trekken van een steekproef kan op verschillende manieren geschieden. Welke vorm men kiest hangt af van de vraag waar het accent van het onderzoek op ligt. Gaat het vooral om het beschrijven van de populatie, dan ligt een aselechte steekproef voor de hand. In een aselechte steekproef wordt gestreefd naar een zo getrouw mogelijke afspiegeling van de werkelijke populatie. Wil men bepaalde relaties aantonen dan is het belangrijk om de populatie in de in dit geval selecte steekproef zo homogeen mogelijk te houden en laat men alleen variatie bestaan op de variabelen waartussen men een relatie verwacht.<sup>8</sup>

Voorbeelden van relevante schriftelijke bronnen zijn archiefstukken en statistieken. Archieven kunnen met name een rol spelen bij het verkrijgen van inzicht in het verloop van het beleidsproces. Uit statistieken valt vergelijkbaar materiaal te halen over de stand van zaken ten aanzien van de verschillende indicatoren in het beleidsveld.

Welke bronnen men daadwerkelijk gaat gebruiken hangt af van de probleemstelling van het onderzoek, de beschikbare gegevens, de beschikbare middelen en de techniek die men wil toepassen om de gegevens te ordenen. Als men deze vraagpunten toepast op de verschillende bronnen van gegevens kan men ruwweg het volgende beeld schetsen. Statistieken en archieven bieden bij complexere vraagstellingen in het algemeen onvoldoende gegevens. Een volgend probleem is dat bij het gebruik van archieven en statistieken moet worden opgepast voor systematische vertekeningen. Wanneer men met bestaande gegevens werkt, is het van belang om na te gaan hoe deze gegevens werden verkregen, om te voorkomen dat er een grotere of geringere exactheid aan wordt toegeschreven dan ze verdienen en om de kans op een onjuiste interpretatie te verminderen.

Het gebruiken van deskundigen en doelgroepen als bron van gegevens kan een probleem zijn als gevolg van de beschikbare tijd en financiën. Het werken met een steekproef biedt hiervoor een oplossing. Het trekken van een steekproef moet dan wel met de nodige zorgvuldigheid geschieden, wil men de betrouwbaarheid van het onderzoek waarborgen. Ook de toe te passen techniek vraagt om de nodige aandacht bij het gebruiken van deskundigen en doelgroepen als waarnemingseenheden. Net als bij de bewerking van reeds bestaand materiaal is systematische vertekening van de resultaten mogelijk. Bij het behandelen van de verschillende technieken zullen we op dit punt terugkomen, omdat er een nauwe relatie is tussen de mogelijk toe te passen technieken en de bronnen die men kiest. Samengevat bepalen de volgende aspecten uit welke bronnen men gaat putten:

- a. de aanwezigheid en beschikbaarheid van bestaand materiaal;
- b. de bereidheid van potentieel relevante respondenten tot medewerking;
- c. de randvoorwaarden in tijd en arbeidskracht van de onderzoeker(s);
- d. de overwegend kwantitatieve of kwalitatieve aard van de evaluatiecriteria.

(Ad 2.) Idealiter worden in het doelbereikingsonderzoek en vooral in het effectiviteitsonderzoek op ten minste twee tijdstippen data verzameld: vóórdat de beleidsprikkel is toegediend en ná de toediening van de beleidsprikkel. In de praktijk ziet men dat onderzoekers dit onder meer trachten te bereiken door het opnemen van retrospectieve vragen. Voor de beantwoording van dergelijke vragen moet de respondent vanuit het moment van vraagstelling terugblikken en een beschouwing over de gang van zaken in het verleden geven. Bij het waarnemen en interpreteren van de verschillende waarnemingen moet dan wel rekening worden gehouden met een aantal mogelijke validiteitsproblemen. Met name bij de meting van de vóór het gevoerde beleid bestaande uitgangssituatie dient men attent te zijn op het gevaar dat retrospectieve informatie uit enquêtes (een veel gebruikte vorm van dataverzameling) vaak niet erg betrouwbaar is, zeker wanneer de respondenten zelf geen bronnen hebben waarin de toenmalige situatie is vastgelegd. Dit is nog sterker het geval indien een vraag een 'sociaal-wenselijk' antwoord kent. In zo'n geval kan de liefde voor de waarheid bij de respondent ernstig worden aangetast.

Schriftelijk enquêteren kan helpen de betrouwbaarheid te vergroten. Zo ook exact geformuleerde vragen en antwoordcategorieën. Ook het inbouwen van talrijke antwoordcategorieën die niet aan de norm voldoen kan helpen om vertekeningen op grond van sociaal-wenselijke antwoorden te voorkomen. Verder kan men projectieve vragen stellen zoals: 'Denkt u dat de mensen in uw omgeving minder milieudelicten begaan?' in plaats van 'Heeft uw bedrijf in het afgelopen jaar milieudelicten gepleegd?' Uiteraard kan dit gepaard gaan met interpretatieproblemen van de antwoorden.

Ad 3.) Grofweg kan men onderscheid maken tussen kwalitatieve en kwantitatieve technieken. In de gereedschapskist van de kwantitatief georiënteerde onderzoeker vinden we de schriftelijke en de gestructureerde mondelinge enquête, statistieken en secundaire analyse voor zover het kwantitatieve gegevens betreft. De in het beleidsevaluatie-onderzoek meest gebruikte en wellicht ook meest geschikte methode voor de dataverzameling is de schriftelijke enquête ('survey'). Indien er voor een mondelinge enquête wordt gekozen zal de kwantitatief georiënteerde onderzoeker kiezen voor gestructureerde interviews. De eigen dataverzameling kan meer gestructureerd zijn naarmate de evaluator vooraf een scherper en betrouwbaarder beeld heeft van de mogelijk relevante aspecten. Deze informatie kan de evaluator verkrijgen door het voeren van oriënterende gesprekken aan het begin van een onderzoek. Deze hebben in de regel een open en ongestructureerd karakter.

Daarmee zijn we aangekomen bij de kwalitatieve dataverzamelingstechnieken. De in hoofdstuk 3 beschreven methode waarmee het te evalueren beleid en het beleidsveld in kaart worden gebracht, maakt ook gebruik van deze benadering van dataverzameling. Het vaststellen van de mate waarin aan de gestelde beleidsdoelen is voldaan, is in de regel niet goed mogelijk zonder dat het beleid en het beleidsveld in kaart zijn gebracht. Andere technieken in deze stroom zijn alle vormen van observatie in het onderzoeksveld, open interviews en de analyse van allerlei documenten zoals vergaderverslagen, nota's en dergelijke. In hoofdstuk 6 wordt nader ingegaan op de kwalitatieve analyse.

Het verdient aanbeveling om in overleg met de opdrachtgever de evaluatiecriteria (beleidsdoelen) en hun operationalisaties vast te leggen. Deze voorbereiding vergt een kwalitatieve analyse van het beleid van waaruit vervolgens op zinnige wijze kan worden bepaald welke variabelen moeten worden onderzocht om de mate van doelbereiking te kunnen vaststellen. Wanneer dit zorgvuldig gebeurt zal het vaak mogelijk zijn om meer kwantitatieve onderzoeksmethoden te gebruiken om de situatie te onderzoeken. De evaluatiecriteria en hun operationalisaties kunnen op deze wijze zo nauwkeurig worden omschreven dat het vaak mogelijk is gestructureerde interviews of schriftelijke enquêtes met gesloten vragen te gebruiken.

Als aandachtspunt bij de dataverzameling kan nog worden genoemd dat het aanbeveling verdient om verschillende informatiebronnen (mondeling en schriftelijk) te gebruiken om de uitkomsten te kunnen vergelijken. In dit licht is het zinvol om kwantitatieve data over de mate van doelbereiking te onderwerpen aan een check met kwalitatieve instrumenten zoals de opinie van deskundigen en analyse van schriftelijke bronnen.

Naast het gebruik van bestaande gegevens is zoals reeds werd aangegeven de meest gebruikte en wellicht ook meest geschikte techniek bij het vaststellen

van de mate van doelbereiking de mondelinge of schriftelijke enquête. De interviews kunnen de vorm aannemen van min of meer informele gesprekken, thematisch gerichte vraaggesprekken en gesprekken op basis van een lijst van gestandaardiseerde open vragen. Interviews waarmee men data wil verzamelen voor het leggen van statische verbanden vereisen een zekere mate van standaardisatie om de vergelijkbaarheid van de uitspraken van de respondenten te vergroten. Dezelfde variëteit is in principe denkbaar bij schriftelijke enquêtes.

Bij het behandelen van de tijdstippen waarop men kan meten is reeds ingegaan op enkele aspecten die aandacht verdienen bij het toepassen van de enquête. De problemen die men kan verwachten in de vorm van 'sociaal-gewenste' antwoorden spelen ook een rol bij andere dan retrospectieve vragen. Een ander gevaar is een eventuele suggestieve vraagstelling: 'Vindt u "X" ook niet erg slecht?' in plaats van 'Wat vindt u van "X"?'. Een laatste validiteitsprobleem is het zogenaamde 'halo-' of 'uitstralings'-effect. Zowel interviewer als geïnterviewde plegen de scores consistent te maken als zij te maken hebben met verwante vragen. Deze gevaren zijn overigens zeker niet uitsluitend aan het gebruik van enquêtes verbonden.

Het gebruik van *verschillende* bronnen naast elkaar kan in een aantal gevallen de mogelijkheid tot metingen die zowel geldig als betrouwbaar zijn, vergroten.

Het werken met gesloten vragen, al dan niet schriftelijk gesteld, is aan te bevelen:

- ① bij dat deel van de vragen waarbij de mogelijke antwoordcategorieën volstrekt duidelijk zijn;
- ② wanneer men het onderzoek bewust wil beperken tot enkele mogelijke relaties tussen scherp gedefinieerde en weloverwogen geïndiceerde variabelen;
- ③ wanneer men wil beschikken over gegevens die onderling goed vergelijkbaar zijn, waardoor zij bij kwantitatieve analysetechnieken kunnen worden gebruikt. Op deze wijze kunnen bijvoorbeeld onderdelen worden getoetst van een beeld dat op grond van open interviews is verkregen.

Met name de laatste mogelijkheid maakt duidelijk dat bij vaststelling van doelbereiking een kwantitatieve analyse een belangrijke rol speelt om de nadelen van de in andere opzichten soms meer geschikte kwalitatieve benadering te ondervangen.

#### 4.4 De interpretatie van de gegevens

Tijdens de analysefase moet de onderzoeker de verzamelde data herleiden tot een *maat voor doelbereiking* om uitspraken te kunnen doen over de mate van doelbereiking. In feite is dit de vraag op welke wijze men de werkelijkheid met de gekozen maatstaven wil vergelijken. Het feit dat er verschillende vergelijkingsmaten zijn wordt vaak vergeten. Bij het meten van de doelbereiking is het niet alleen problematisch wat men als doeleinden (evaluatiecriteria) ziet, maar ook wat men als *bereiken* beschouwt. Het begrip 'doelbereiking' kent vele dimensies. Deze dimensies komen voort uit de verschillende vergelijkingsbases, waarmee de situatie die op een bepaald tijdstip is bereikt, wordt vergeleken. De periode waarin het beleid werd uitgevoerd en die in het onderzoek wordt geëvalueerd noemen wij de evaluatieperiode. De volgende variabelen kunnen in verschillende combinaties deze vergelijkingsbases met de bereikte situatie (S) vormen:<sup>9</sup>

1. de uitgangssituatie aan het begin van de evaluatieperiode (U);
2. de beoogde situatie voor het einde van de evaluatieperiode (B);
3. de beoogde situatie voor een later tijdstip (B');;
4. de volmaakt geachte situatie (B''), dat wil zeggen: meest gewenste situatie (bijvoorbeeld: geen werkloosheid, geen milieuverontreiniging, etc.).

Daarnaast kan de vergelijkingsmaat in absolute en in relatieve termen worden gesteld, ten minste wanneer de criteria in kwantitatieve termen kunnen worden geoperationaliseerd. Het doel kan een vermindering of een vermeerdering van iets inhouden. Wanneer we uitgaan van een vermindering als doel, dan kunnen voorbeelden van vergelijkingsmaten er als volgt uit zien:

1. bereikte verbetering: U-S.
2. bereikte procentuele verbetering:  $\frac{100 \cdot (U-S)}{U}$
3. verschil tussen de beoogde en de bereikte situatie: B-S;
4. procentueel verschil tussen de beoogde en de bereikte situatie:  $\frac{100 \cdot (B-S)}{B}$
5. procentueel verschil tussen de bereikte en de beoogde verbetering:  $\frac{100 \cdot (B-S)}{U-B}$



6. procentueel gedeelte van de beoogde verbetering dat is bereikt:  

$$\frac{100 \cdot (U-S)}{U-B}$$

Mutatis mutandis zijn dergelijke maten voor doelbereiking ook te construeren met in plaats van de beoogde situatie voor het einde van de evaluatieperiode (B) de beoogde situatie voor een tijdstip na de evaluatieperiode (B') of de meest gewenste situatie (B''). Al deze maten hebben voor- en nadelen. Alle voegen ze iets toe aan het beeld dat de onderzoeker zich vormt van de mate van doelbereiking.

De eerste en tweede maat zijn slechts gesteld in termen van de variabelen U en S, zij bevatten alleen feitelijke variabelen en geen normatieve variabelen. Alleen de richting van de gewenste verandering speelt mee doordat deze het teken, positief of negatief, van de gemeten doelbereiking bepaalt. Bij deze maten is de uitkomst dus niet afhankelijk van de gestelde doelen of van de oorspronkelijke discrepantie tussen de uitgangssituatie en de gewenste situatie (= U-B), maar is de uitkomst een mogelijk effect van het beleid, temidden van andere rivaliserende verklaringen.

De maten drie tot en met zes worden alle berekend met behulp van een variabele die niet alleen is afgeleid van de richting, maar ook van het ambitieniveau van het beleid en wel in de vorm van maatstaf B: de beoogde situatie aan het einde van de periode. Dit betekent dat in deze maten een waardering voor de geconstateerde effecten is opgenomen die mede de uitkomst van de maat bepaalt. Door een voorbeeld in te voeren in de genoemde maten kan gemakkelijk worden nagegaan hoe bij de maten drie tot en met zes bij gegeven waarden voor U en S, de uitkomst wordt bepaald door de hoogte van de gestelde doelen B.

Het ontbreken van gegevens over één of meer variabelen maakt het hanteren van één of meer maten onmogelijk. Het verdient daarom aanbeveling *om bij de opzet van de dataverzameling rekening te houden met de doelbereikingsmaten die men wenst te hanteren.*

Een tweede reden voor een vroegtijdige bezinning op de te gebruiken doelbereikingsmaten ligt in het feit dat niet alle maten zich evengoed lenen voor de status van afhankelijke variabele bij de beantwoording van vervolgvragen naar de *bijdrage van de beleidsuitvoering (de toepassing van de beleidsmiddelen)* aan de geconstateerde mate van doelbereiking (= effectiviteitsvraag). Wanneer de score op een doelbereikingsmaat ook afhangt van het ambitieniveau van de doelstelling, bestaat een deel van de verklaring niet uit de feitelijke situatie en de invloed van de beleidsuitvoering daarop. Omdat met de bijdrage van de beleidsuitvoering aan de doelbereiking steeds de invloed van de beleidsuitvoering op de feitelijke situatie wordt bedoeld, is het een nodeloze

complicatie wanneer ook andere zaken, zoals het ambitieniveau, de mate van doelbereiking mede bepalen.

Het is zelfs denkbaar dat men bij het verklaren van de mate van doelbereiking van het woonwagendstandplaatsenbeleid van twee provincies in het ene geval een negatieve score moet verklaren en in het andere geval een positieve, terwijl in beide provincies de situatie en de ontwikkeling daarvan, alsmede de invloed van het beleid daarop identiek waren en alleen het ambitieniveau verschilde.

De eerste twee maten, die alleen feitelijke variabelen en geen normatieve variabelen bevatten, zijn het meest geschikt om te gebruiken wanneer men ook de vervolgvragen naar de effectiviteit van het beleid wil beantwoorden.

Dit betekent niet dat alleen de eerste twee maten kunnen worden gebruikt. Bij het gebruik van de maten drie tot en met zes (en andere te bedenken maten) moet echter worden opgepast voor de betekenis van het normatieve aspect voor de interpretatie van de uitkomst. Zo kunnen de doelen bijvoorbeeld onredelijk hoog of laag zijn gesteld om politieke redenen. Toch zal de opdrachtgever vaak in de uitkomsten van deze maten zijn geïnteresseerd. Als aanvulling op de eerste twee maten en met gezond verstand gebruikt, is daar niets op tegen.

Aanvullend kan worden opgemerkt dat een doel niet altijd qua omvang of tijdsperiode is gekwantificeerd of gedurende het beleidsproces ongewijzigd blijft. Ook in deze gevallen is een vergelijking tussen een waargenomen ontwikkeling van een situatie en een doel wel mogelijk. Een niet in kwantitatieve termen geformuleerd doel kent immers wel altijd een richting, namelijk meer, even veel of minder van een bepaald verschijnsel. Dit maakt wel enkele maten van vergelijking onmogelijk (in ons voorbeeld de maten drie tot en met zes), maar maakt de vergelijking als zodanig aan de hand van de eerste twee maten niet onmogelijk.

Indien het doel gedurende het beleidsproces is gewijzigd moet het tijdstip waarop men de uitgangssituatie definieert zorgvuldig worden gekozen.

De betrouwbaarheid van het evaluatie-onderzoek en dus ook die van de vaststelling van de mate van doelbereiking betreft de herhaalbaarheid van het onderzoek. In de voorafgaande deelstappen is reeds gewezen op de testen die hiertoe als controle kunnen dienen. De betrouwbaarheid van de gevolgde analyse kan worden gecontroleerd door verschillende onderzoekers onafhankelijk van elkaar hetzelfde materiaal te laten analyseren.

#### 4.5 Samenvatting

Met behulp van een drietal deelstappen is in dit hoofdstuk verduidelijkt hoe men de mate van doelbereiking van een beleid kan meten. Bij elk van de deelstappen geldt dat de invulling op een dusdanige wijze moet geschieden dat er wordt voldaan aan de eisen van uitvoerbaarheid, accuraatheid en bruikbaarheid.

In de eerste stap staat de operationalisatie van de beleidsdoelen centraal. Veelal zal het hierbij om doeleinden gaan die niet betrekking hebben op direct observeerbare zaken. Dit vraagt om het indiceren van een doel in meetbare termen. Om te komen tot een zorgvuldige operationalisatie is het noodzakelijk dat:

- er een voldoende sterke relatie is tussen het te operationaliseren doel en de indicatoren;
- de gekozen indicatoren het te evalueren doel in al zijn dimensies omvatten;
- er geen interveniërende factoren zijn die een verandering aanbrengen in de betekenis van de indicator voor het doel.

In de tweede deelstap gaat het om het verzamelen van gegevens. Hierbij komen de volgende aspecten aan de orde: de keuze van de bronnen waaruit men de gegevens wil putten, de keuze van de tijdstippen waarop wordt gemeten en de te gebruiken technieken. Mogelijke bronnen van gegevens zijn: deskundigen, doelgroepen van beleid, archieven en statistieken. Idealiter zijn er twee tijdstippen waarop men meet: het moment waarop er sprake is van een Ausgangssituatie (het beleid is nog niet begonnen) en het moment waarop het beleid reeds enige tijd is uitgevoerd.

In de technieken die een rol kunnen spelen bij het verzamelen van de gegevens kan men een onderscheid aanbrengen tussen kwantitatieve en kwalitatieve methoden. Een combinatie van technieken die zowel kwantitatieve als kwalitatieve gegevens opleveren wordt gezien als de beste benadering om te komen tot een goed evaluatie-onderzoek.

Uit welke bronnen de gegevens worden geput hangt af van de mogelijkheden die er zijn. Hierbij kan men denken aan beschikbaarheid van het materiaal en de hoeveelheid middelen die de onderzoeker ter beschikking staat.

De laatste stap die gezet moet worden alvorens men het onderzoek naar de mate van doelbereiking kan afsluiten is het interpreteren van de data en het trekken van conclusies. In deze fase vraagt het meten van de mate van doelbereiking de nodige aandacht. Er zijn verschillende vergelijkingsbases mogelijk op grond waarvan men een uitspraak kan doen over 'de mate van' doelbereiking. Men kan kiezen voor een vergelijkingsbasis waarbij men zich beperkt tot de feitelijke variabelen. Een voorbeeld van een dergelijke constructie is de bereikte (procentuele) verandering tussen de Ausgangssituatie en de bereikte situatie.

Deze constructie biedt de beste mogelijkheden indien men na het meten van doelbereiking wil overgaan op de beantwoording van vervolgvragen naar de bijdrage van het beleid aan de geconstateerde mate van doelbereiking. Een andere mogelijkheid is het kiezen van een vergelijkingsbasis waarbij normatieve variabelen (in dit geval beleidsdoeleinden) een rol spelen. Bij deze keuze wordt de mate van doelbereiking beïnvloed door de gestelde doelen of door de discrepantie met de gewenste situatie.

In de fase van de dataverzameling is het gewenst om reeds duidelijkheid te hebben over de te kiezen vergelijkingsbasis voor de doelbereiking. Anders kunnen gegevens ontbreken die voor het hanteren van bepaalde maten onmisbaar zijn.

#### Literatuur

- R.W. Boesjes-Hommes, *De geldige operationalisering van begrippen*, Meppel, 1970.  
 J.Th.A. Bressers, *Beleidsdoelmatigheid en waterkwaliteitsbeleid*, Enschede, 1983.  
 J.Th.A. Bressers, *Beleidsbeoordeling en beleidsdoelmatigheid*, in: A. Hoogerwerf (red.), *Overheidsbeleid, vierde druk*, Alphen aan den Rijn, 1989.  
 A.D. de Groot, *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen*, 's-Gravenhage, 1961.  
 M.B. Miles en Huberman A.M., *Qualitative evaluation methods: A source book of new methods*, Beverly Hills/Londen, 1984.  
 M.B. Miles en Huberman A.M., *Drawing valid meaning from qualitative data: Toward a shared craft*, in: *Educational Researcher*, 1984.  
 M. Quinn Patton, *Qualitative evaluation methods*, Beverly Hills/ Londen, 1980.  
 M. Quinn Patton, *Evaluation Standards Committee, Standards for evaluation of educational programs, projects and materials*, New York, 1981.  
 P.G. Swanborn, *Aspecten van sociologisch onderzoek*, Meppel/Amsterdam, 1979.

#### Vragen en opgaven

1. Men wil weten in hoeverre de milieubelangen een volwaardige plaats krijgt in het besluitvormingsproces van grote projecten (projecten waarvoor een milieueffectrapportage moet worden opgesteld).
  - a. Operationaliseer het beleidsdoel 'milieu een volwaardige plaats geven in het besluitvormingsproces van grote projecten'.
  - b. Geef aan welke aspecten de aandacht vragen bij de operationalisatie van het beleidsdoel.
  - c. Op welke wijze en op welke tijdstippen kan informatie worden verzameld om te bepalen of het hierboven geformuleerde beleidsdoel al dan niet bereikt is?
  - d. Welke vergelijkingsmaten verdienen aanbeveling om de mate van doelbereiking te meten? Beargumenteer uw antwoord.

2. Een beleidsvoerder twijfelt bij een onderzoek zowel aan de betrouwbaarheid van de gekozen bronnen als aan de gehanteerde analysemethode.
  - a. Aan de hand van welke aandachtspunten kan de beleidsvoerder een goed beeld krijgen over de betrouwbaarheid?
  - b. Op welke wijze kan de beleidsvoerder de verrichte analyse controleren?
3. Hoe verhouden de richtlijnen van Patton zich tot de algemene eisen die men stelt aan een goed evaluatie-onderzoek?

## 5 Wat is de bijdrage van het beleid aan de doelbereiking?

### *Deel 1: Kwantitatieve methoden*

DRS. F.H.J.M. COENEN

#### 5.1 Inleiding

In het vorige hoofdstuk is het vaststellen van de mate van doelbereiking behandeld. De mate van doelbereiking zegt echter nog niets over de effectiviteit van het beleid. Effectiviteit van beleid is de mate waarin een bepaald beleid bijdraagt aan het bereiken van bepaalde doeleinden. Een beleid is slechts effectief voor zover de mate van doelbereiking aan dit beleid te danken is. Naast het gevoerde beleid kunnen ook andere verklaringen voor de mate van doelbereiking gelden, de zogenaamde alternatieve of rivaliserende verklaringen.

Door middel van het kiezen van een geschikte onderzoeksopzet is het in principe mogelijk rivaliserende verklaringen als oorzaak van de doelbereiking zo goed als uit te sluiten. Bij het toepassen van de onderzoeksopzetten blijven er echter altijd nog een of meer 'valkuilen' over. We spreken daarbij over de validiteit (geldigheid) van de onderzoeksresultaten: is er gemeten wat men wilde meten? Deze geldigheid wordt onder meer bepaald door de mate waarin de onderzoeksopzet het mogelijk maakt rivaliserende verklaringen uit te sluiten.

Hierbij speelt het begrip causaliteit een belangrijke rol. Het gaat immers om oorzaak-gevolgrelaties tussen het beleid en de beleidseffecten. In paragraaf 5.2 zullen we nader ingaan op het begrip causaliteit. In paragraaf 5.3 zal een aantal in de praktijk van het beleidsevaluatie-onderzoek meest gebruikte onderzoeksopzetten worden behandeld. We zullen deze onderzoeksopzetten illustreren met behulp van voorbeelden van Nederlandse beleidsevaluaties.