



Handreiking

Evaluatieonderzoek ex post



Handreiking

Evaluatieonderzoek ex post

*Een praktisch handvat voor de opzet en uitvoering
van evaluatieonderzoek ex post*

'If you don't know how things went, how will you ever know if things can get better?'

Inhoud

Voorwoord	4
1 Inleiding	6
1.1 Belang van beleidsinformatie	6
1.2 Doel en doelgroep van deze handreiking	6
1.3 Leeswijzer	7
2 Evaluatieonderzoek ex post: wat, wanneer en hoe?	9
2.1 Evaluatieonderzoek ex post: wat is dat eigenlijk?	9
2.2 Waar richt evaluatieonderzoek ex post zich op?	10
2.3 Wanneer moet evaluatieonderzoek worden uitgevoerd?	12
2.4 Aan welke kwaliteitseisen voldoet goed evaluatieonderzoek?	14
3 De opzet en uitvoering van evaluatieonderzoek ex post: een stappenplan	16
3.1 De eerste stap: het formuleren van de probleemstelling en vragen van het onderzoek	16
3.2 De tweede stap: de keuze voor een onderzoeksopzet	24
3.4 De derde stap: het hoofdonderzoek: gegevensverzameling en -verwerking	30
3.4 De vierde stap: de rapportage en follow-up	38
4 Begrippenlijst	43
5 Een overzicht van aanbevolen literatuur	46
Bijlage 1 Overzicht stappen	48
Bijlage 2 Stappen in proces van uitbesteding van evaluatieonderzoek	49
Bijlage 3 Terms of Reference (TOR)	52

Voorwoord

Elk jaar worden er binnen de rijksoverheid honderden evaluatieonderzoeken uitgevoerd. Het doel is daarbij steeds gelijk: het verkrijgen van inzicht. Door middel van evaluatieonderzoek probeert de overheid inzicht te krijgen in de doeltreffendheid en doelmatigheid van het gevoerde beleid en/of bedrijfsvoering. Concreet gaat het dan om de vraag of beleid het beoogde effect heeft en of het goed wordt uitgevoerd. Met andere woorden: levert de overheid waar voor haar geld?

Beleidsvaluaties zijn daarmee onlosmakelijk verbonden met het resultaatgerichte sturingsmodel. Dit sturingsmodel vormt de kern van de operatie ‘Van Beleidsbegroting Tot Beleidsverantwoording’ (vbtv). In de begroting en het jaarverslag worden, naast financiële informatie, steeds meer gegevens opgenomen over wat de rijksoverheid doet, respectievelijk heeft gedaan en welke effecten daarmee worden beoogd, respectievelijk zijn gerealiseerd. Daarbij geldt dat het vaak niet op voorhand duidelijk is, wat precies de bijdrage van beleid aan het oplossen – of beheersen – van problemen is geweest. Evaluatieonderzoek ex post (letterlijk: achteraf) vormt zo het sluitstuk van een resultaatgerichte overheid. Een overheid die verantwoording af wil leggen over wat ze doet en waartoe dat leidt. Een overheid ook, die wil leren hoe haar beleid – of de uitvoering van dat beleid – kunnen worden verbeterd.

Evaluatieonderzoek is dus een belangrijk instrument. Maar het is ook kostbaar en moet goed worden toegepast. Deze handleiding wil mensen die een evaluatieonderzoek (laten) uitvoeren een praktische handreiking bieden hoe dit het beste en met het meeste effect kan gebeuren. Hiertoe bevat de handleiding een kader (wat is beleidsevaluatie? wanneer moet beleid eigenlijk geëvalueerd worden?), maar ook een praktisch stappenplan voor het opzetten en (doen) uitvoeren van goed onderzoek. Comptabele voorschriften worden daarbij zoveel als mogelijk vertaald in concrete checklists. Jargon wordt uitgelegd in een verklarende woordenlijst. Voor de echte liefhebber zijn verwijzingen naar vakliteratuur en websites toegevoegd.

Diverse departementen en de Algemene Rekenkamer hebben commentaar geleverd op een concept van voorliggende handreiking. Dit voorwoord is een goede plaats daarvoor dank uit te spreken. Graag nodigen we de gebrui-

kers van de handreiking uit hun ervaringen met ons te delen. Door middel van praktijkervaring kunnen we de handreiking immers verder verbeteren.

Uw opmerkingen of suggesties kunt u richten aan de afdeling Beleids-evaluatie en -Instrumentatie van de directie Begrotingszaken van het ministerie van Financiën¹. Hetzelfde geldt voor verzoeken om praktische ondersteuning bij de opzet of uitvoering van evaluatieonderzoek.

Ik hoop dat de handleiding slaagt in haar opzet.

André de Jong

Directeur-generaal Rijksbegroting

¹ Afdeling Beleidsevaluatie en -Instrumentatie, Anke Kijzers, telefoon: (070) 342 73 17; fax: (070) 342 79 62; e-mail: bz.beleidsevaluatie@minfin.nl.

1 Inleiding

1.1 Belang van beleidsinformatie

Beleidsinformatie speelt een belangrijke rol bij het nemen van besluiten over het te voeren beleid. De begrotingscyclus, maar ook de departementale jaarplancyclus, vormen daarbij een centrale rol. Een goede verdeling van middelen kan niet buiten goede beleidsinformatie. Beleidsinformatie die relevant is, betrouwbaar en tijdig beschikbaar. Daarnaast dient er (in het kader van de begrotingscyclus) verantwoording te worden afgelegd over het gevoerde beleid, de gevoerde bedrijfsvoering en de daarmee bereikte resultaten. Ook dat vergt beleidsinformatie die aan bepaalde kwaliteitseisen voldoet.

Sinds 1999 wordt aan het streven van de rijksoverheid om resultaatgericht te sturen en te verantwoorden, een belangrijke impuls gegeven met het project ‘Van Beleidsbegroting Tot Beleidsverantwoording’ (vBTB). Essentie van vBTB is dat financiële informatie wordt gekoppeld aan beleidsinformatie. De doelstellingen van beleid worden zoveel mogelijk concreet vertaald in ‘te bereiken effecten’. Naast informatie over de beoogde doelen (met streefwaarden, doelgroepen en termijnen) wordt in de vBTB-begroting tevens informatie opgenomen over de kosten en kwaliteit van de door de overheid geleverde prestaties: de in het kader van beleidsontwikkeling, -aansturing en/of – uitvoering geleverde producten en diensten.

Duidelijk is dat met vBTB het belang van beleidsinformatie in de begrotingscyclus sterk is toegenomen. Dit onderstreept het belang van kwalitatief hoogwaardige prestatiegegevens, waaronder met name de uitkomsten van evaluatieonderzoek. Veel beleidsinformatie ten behoeve van de begrotings- en verantwoordingsstukken vergt evaluatieonderzoek. Zo kan de vraag of de gerealiseerde effecten vallen toe te rekenen aan het gevoerde beleid (doeltreffendheid van beleid), in de meeste gevallen slechts met behulp van diepgaand onderzoek worden beantwoord.

1.2 Doel en doelgroep van deze handreiking

Het doel van deze handreiking is een praktisch handvat te bieden voor de opzet en uitvoering van evaluatieonderzoek ex post, met inachtneming van de bijbehorende kwaliteitseisen. Dit aan de hand van een aantal te doorlopen stappen. De handreiking bevat ook informatie voor procesbewakers (zoals medewerkers van FEZ of auditors).

Achterliggend doel is de kwaliteit van evaluatieonderzoek, en – langs die weg – van beleidsprocessen, te vergroten. In het verleden is er, met name vanuit de Algemene Rekenkamer, veel kritiek geweest op de evaluatiefunctie van de rijksoverheid. Dit is aanleiding geweest om de comptabele voorschriften te verduidelijken. De Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid vormt hierbij het overkoepelend kader. Deze ministeriële regeling (hierna te noemen: evaluatieregeling), die vanaf 1 januari 2002 van kracht is, kent op het terrein van beleidsevaluatie ex post de volgende bepalingen:

- 1 alle doelstellingen van beleid worden in de regel ten minste eens in de vijf jaar (nader) geëvalueerd;
- 2 evaluatieonderzoek voldoet aan kwaliteitseisen op het vlak van validiteit, betrouwbaarheid en bruikbaarheid;
- 3 de kwaliteit, alsmede de onafhankelijkheid waarmee een onderzoek is opgezet en/of uitgevoerd, moet achteraf kunnen worden beoordeeld, en;
- 4 de ambtelijke en politieke top wordt volgens een standaardprocedure op de hoogte gebracht van de uitkomsten van evaluatieonderzoek.

De handreiking is nadrukkelijk als handvat bedoeld en niet als blauwdruk. In geval van twijfel is de tekst van de evaluatieregeling leidend.

1.3 Leeswijzer

In hoofdstuk 2 zal eerst worden stilgestaan bij het begrip evaluatieonderzoek ex post: wat houdt evaluatieonderzoek ex post in (*paragraaf 2.1*), op welke aspecten richt evaluatieonderzoek ex post zich (*paragraaf 2.2*), wanneer moet evaluatieonderzoek ex post worden uitgevoerd (*paragraaf 2.3*) en aan welke eisen moet goed evaluatieonderzoek ex post voldoen (*paragraaf 2.4*)?

Vervolgens zal in hoofdstuk 3 worden ingegaan op de te doorlopen stappen bij de opzet en uitvoering van evaluatieonderzoek ex post. De volgende vier stappen kunnen in deze handreiking onderscheiden worden:

- | | |
|---------------|---------------------------------------------------------------------------------------------|
| Stap 1 | Het formuleren van de probleemstelling en vragen van het onderzoek (<i>paragraaf 3.1</i>) |
| Stap 2 | De keuze voor een onderzoeksopzet (<i>paragraaf 3.2</i>) |
| Stap 3 | Het hoofdonderzoek: gegevensverzameling en -verwerking (<i>paragraaf 3.3</i>) |
| Stap 4 | De rapportage en follow-up (<i>paragraaf 3.4</i>) |

In *bijlage 1* wordt een samenvattend overzicht gegeven van de vier stappen inclusief tussenstappen.

In hoofdstuk 4 wordt een lijst gegeven van de meest gebruikte begrippen in deze handreiking. Tot slot treft u in hoofdstuk 5 een overzicht aan van aanbevolen literatuur en relevante websites omtrent evaluatieonderzoek.

2 Evaluatieonderzoek ex post: wat, wanneer en hoe?

2.1 Evaluatieonderzoek ex post: wat is dat eigenlijk?

Strikt genomen behelst de term ‘evalueren’ niets meer of minder dan het waarderen van toestanden of processen. Als het gaat om beleidsevaluatie is dat waarderen minimaal gekoppeld aan de normen doelmatigheid en doeltreffendheid van beleid, maar ook de ‘algemene wenselijkheid’ van – doelstelling van – beleid zullen vaak worden beoordeeld. Hiermee wordt de klassieke vraag beantwoord: ‘Doet de overheid de goede dingen en doet ze die dingen goed?’ Het verschil tussen louter waarderen en evalueren schuilt in de wijze waarop tot een oordeel wordt gekomen. Aan evaluatieonderzoek mogen eisen worden gesteld: beleidsevaluatie vergt een zorgvuldige analyse van op systematische wijze verzamelde gegevens. Pas dan kunnen op verantwoorde wijze uitspraken worden gedaan over de doelmatigheid of doeltreffendheid van beleid.

Een belangrijk onderscheid is dat tussen evaluatieonderzoek *ex ante* (vooraf) en evaluatieonderzoek *ex post* (achteraf). Evaluatieonderzoek *ex ante* is toekomstgericht. Het doel is om de kosten en baten van alternatieve beleidsvoornemens in kaart te brengen en tegen elkaar af te wegen. Voor de opzet en uitvoering van evaluatieonderzoek *ex ante* wordt verwezen naar de handreiking evaluatieonderzoek *ex ante*. Deze handreiking behandelt evaluatieonderzoek *ex post*.

Evaluatieonderzoek *ex post* kan worden gedefinieerd als: systematisch onderzoek naar de effecten van bestaand beleid, de wijze waarop beleid wordt uitgevoerd en/of de kosten en kwaliteit van geleverde producten en diensten.

Beleidsvaluatie is onderdeel van het beleidsproces. Het is goed te bezien in welke fase van het beleidsproces evaluatieonderzoek *ex post* relevant is. In een beleidsproces kunnen vijf fasen worden onderscheiden:

- 1 **de beleidsvoorbereiding:** het formuleren van het beleidsprobleem, het verzamelen en analyseren van informatie over beleidsalternatieven, het opstellen van adviezen over het te voeren beleid;
- 2 **de beleidsbepaling:** het nemen van beslissingen over het te voeren beleid en de daarbij in te zetten instrumenten;

- 3 *de beleidsuitvoering*: het feitelijk toepassen van de gekozen middelen en instrumenten, resulterend in beleidsprestaties (in goed Nederlands: *output*) en beoogde effecten en neveneffecten (*outcome*);
- 4 *de evaluatie achteraf van het beleid*: het beoordelen van de beleidsinhoud, het uitvoeringsproces, de geleverde prestaties en de – daarmee bereikte – effecten van het beleid;
- 5 *de terugkoppeling*: het bijsturen van het beleid op basis van informatie over de inhoud, het proces, de prestaties en de effecten van het beleid.

Schema 1: De beleidscyclus



2.2 Waar richt evaluatieonderzoek ex post zich op?

Zoals aangegeven richt evaluatieonderzoek ex post zich primair op de doelmatigheid en doeltreffendheid van beleid. Daarbij wordt ervan uitgegaan dat evaluatieonderzoek kan voortbouwen op informatie die reeds beschikbaar is over de mate waarin doelen en voorgenomen beleidsprestaties zijn gerealiseerd (ook wel de mate van doelbereiking genoemd). Als deze informatie ontbreekt, zal dit aspect ook onderdeel moeten uitmaken van de vraagstelling van het onderzoek.

Kort samengevat staan bij het evalueren van beleid binnen de rijksoverheid vier vragen centraal:

1 Is er sprake van doelbereiking?

Te vertalen als: is het doel bereikt? Dit is veelal bekend, maar toch is het een goede vraag om te stellen. De vraag zou zich ook kunnen richten op in *hoeveer* er sprake is van doelbereiking.

2 Is het beleid doeltreffend?

Te vertalen als: in welke mate is het realiseren van de doelstellingen van beleid toe te schrijven aan het gevoerde beleid en de daarbij ingezette instrumenten?

3 Is het beleid doelmatig (kosteneffectief)?

Te vertalen als: als we weten in welke mate beleid heeft geleid tot effect, hoe kan de verhouding tussen gemaakte kosten en de gerealiseerde beoogde effecten worden beoordeeld? Hadden de doelstellingen van beleid bijvoorbeeld niet met minder middelen gerealiseerd kunnen worden of hadden er niet meer resultaten behaald kunnen worden met dezelfde inzet van middelen?

4 Is het beleid doelmatig uitgevoerd?

Te vertalen als: hoe kan de verhouding tussen de kosten en de kwaliteit van de – in het kader van beleidsontwikkeling, -aansturing en -uitvoering – door de overheid geleverde prestaties, producten en diensten worden beoordeeld?

Inzicht in de (mate van) *doelbereiking* bestaat in veel gevallen al. Of er sprake is van doelbereiking is veelal bekend. Evaluatieonderzoek zou nog kunnen uitwijzen *in hoeverre* de doelstelling is bereikt.

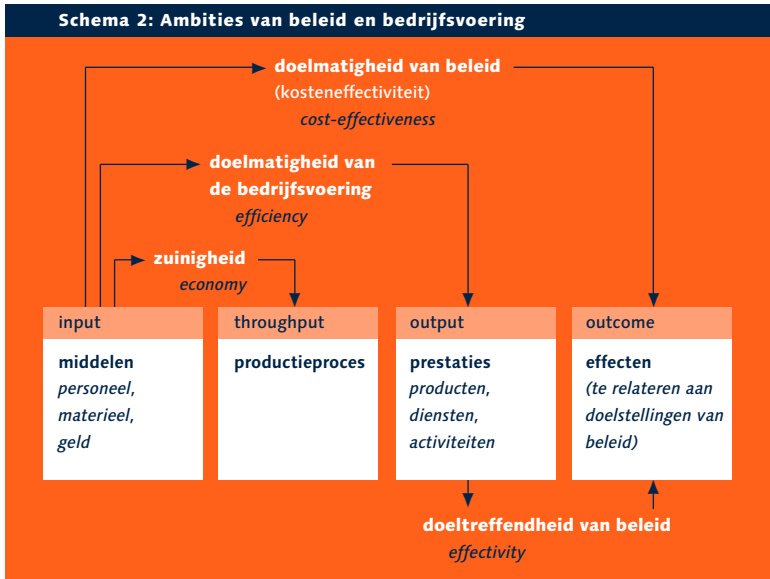
De *tweede vraag* – in hoeverre de gerealiseerde effecten vallen toe te rekenen aan het gevoerde beleid – (en ook de *derde vraag*) kan in de meeste gevallen slechts met behulp van evaluatieonderzoek ex post worden beantwoord.

Koppeling tussen beleid en resultaten

Het is niet eenvoudig om de doeltreffendheid van beleid aan te tonen. Hoe laat je bijvoorbeeld zien dat een verhoogde aanschaf van energiezuinige producten daadwerkelijk het resultaat is van subsidie (lees: beleid) van de overheid? Zou men ook zonder die subsidie niet zijn overgegaan tot de aanschaf van zuinigere apparaten? De koppeling tussen beleid en resultaten moet in dit geval worden gelegd door evaluatieonderzoek. Het is te kort door de bocht om aan te nemen dat het overheidsbeleid hier alleen aan heeft bijgedragen.

Er wordt onderscheid gemaakt tussen *doelmatigheid van het beleid* en *doelmatigheid van de bedrijfsvoering*. De doelmatigheid van de bedrijfsvoering geeft de relatie tussen middelen en prestaties (output) weer en beoordeelt daarmee de mate waarin een dienst doelmatig met de middelen omgaat. Had met minder middelen dezelfde prestaties (output) kunnen worden geleverd, of met evenveel middelen meer prestaties? De doelmatigheid van beleid behelst meer dan dat. Hier worden de middelen tegen de effecten afgezet. Voordeel van dit inzicht is dat de vraag gesteld kan worden of de effecten wel opwegen tegen de kosten. Bovendien kan ook bekeken worden of de juiste prestaties worden geleverd. Leiden andere prestaties/activiteiten wellicht tot meer effecten? Inzicht in de doelmatigheid van de uitvoering van beleid bestaat in veel gevallen al. Doelmatigheidsgegevens zijn vaak al beschikbaar door de aanwezigheid van reguliere prestatiegegevens over de kosten en kwaliteit van prestaties. Evaluatieonderzoek ex post kan dus worden beschouwd als ‘terugvaloptie’ c.q. aanvulling op reguliere prestatiegegevens over de doelmatigheid.

Kortom, evaluatieonderzoek ex post richt zich voornamelijk op de realisatie van de doelstellingen als gevolg van het gevoerde beleid. In onderstaand schema wordt deze relatie nog eens weergegeven.



2.3 Wanneer moet evaluatieonderzoek worden uitgevoerd?

Evaluatieonderzoek ex post is alleen zinvol als de resultaten ook gebruikt worden. Een goede timing van evaluatieonderzoek is daarbij van belang. De norm is dat het beleid dat moet leiden tot de realisatie van de in de begroting en jaarverslag opgenomen algemene en/of nader geoperationaliseerde doelstellingen van beleid in de regel ten minste eens in de vijf jaar (nader) wordt geëvalueerd aan de hand van evaluatieonderzoek ex post (RPE, bepaling 4.1). Indien wenselijk kan er uiteraard voor worden gekozen om eerder/vaker te evalueren. Reden hiervoor kan zijn een (voorgenomen) aanpassing of wijziging van het beleid. Een dergelijke (tussen)evaluatie zal zich moeten richten op de tot dan toe verwachte effecten of prestaties. Uiteraard heeft een dergelijke evaluatie alleen nut indien er al resultaten, in de zin van prestaties of effecten, te verwachten zijn. Het is bij het evaluatiemoment vooral van belang om een logisch moment uit te kiezen.

In twee gevallen kan worden afgeweken van de norm 'ten minste eens in de vijf jaar':

- de resultaten of effecten van beleid zijn niet binnen vijf jaar vast te stellen. Het is niet mogelijk of zinvol om (tussengelegen) doelstellingen te evalueren;
- de kosten van evaluatieonderzoek ex post staan in een negatieve verhouding tot de baten ervan (baten in termen van: extra inzicht in doelmatigheid of doeltreffendheid).

Evaluatieonderzoek en informatie over prestaties en effecten die op gezette tijden beschikbaar komt, moeten elkaar goed aanvullen. Het streven is om te komen tot een optimale inzet van bestaande, op reguliere basis beschikbare prestatiegegevens en periodiek uit te voeren evaluatieonderzoek. Als vuistregel geldt dat evaluatieonderzoek zich in eerste aanleg richt op die informatiebehoeften, waarin *niet* door monitorsystemen wordt voorzien.

Ideale VBTV-situatie	'De realiteit'
1 Monitoring systeem van reguliere prestatiegegevens (kwantitatief) voor zowel bedrijfsvoering als effecten, aangevuld met;	1 'Slechts' reguliere indicatoren over doelmatigheid van de bedrijfsvoering, aangevuld met;
2 evaluatieonderzoek (kwantitatief en kwalitatief).	2 evaluatieonderzoek naar 'van alles en nog wat', hopelijk ook naar het behalen van (gereconstrueerde) doelstellingen.

Bovenstaande box geeft de situatie weer zoals die in de ideale VBTV-situatie zou moeten zijn en de 'realiteit' waar men vaak tegenaan loopt. De 'realiteit' komt vaak niet overeen met de ideale situatie omdat het gros van het huidige beleid is opgezet in het pré VBTV-tijdperk. Veelal is bestaand beleid in een *www*-jasje gegoten, maar echt passen doet het nog niet. In het volgende hoofdstuk wordt aandacht besteed aan o.a. hoe 'oud' beleid (pré VBTV) toch op de 'nieuwe' wijze (RPE) te evalueren.

Verder is het nuttig om bij het opstellen van de evaluatieprogrammering na te gaan:

- welke instrumenten en doelstellingen van beleid geëvalueerd *moeten* worden (bijvoorbeeld in het kader van een wet, een verdrag, toezeggingen aan Kamer, EU-verplichtingen);
- of het wellicht verstandig is beleidsdoelstellingen en -instrumenten op een ander moment te evalueren (te denken valt aan financieel belang, ernstige twijfel aan effectiviteit of efficiëntie van beleid of het bestaan

van een politiek of maatschappelijk ‘momentum’ – bijvoorbeeld als er nieuwe grote nota’s of internationale afspraken in het verschiep liggen);

- of niet alle evaluaties in eenzelfde periode zijn gepland. (Dit geldt met name voor evaluaties van operationele doelstellingen. Vanwege de impact en het beslag op de organisatie die zo’n evaluatie kan hebben, verdient het de voorkeur deze evaluaties, en daarmee ook de evaluaties van instrumenten onder de operationele doelstellingen, te spreiden.)

Schema 3: Een overzicht van te evalueren aspecten, gekoppeld aan systemen van reguliere prestatiegegevens en periodiek evaluatieonderzoek

Beoordelings-norm	Te evalueren aspect	Meest logische informatiebron	Frequentie (afh. van opname in begroting en jaarverslag)	Voorbeeld prestatiegegevens (ontleend aan inburgeringsbeleid)
Doelbereiking	doelstelling gerealiseerd?	systemen van reguliere prestatiegegevens (terugval: evaluatie-onderzoek)	jaarlijks (terugval: meerjarig)	aantal oud-deelnemers aan inburgeringscursus met een betaalde baan
Doeltreffendheid van beleid	doelstelling gerealiseerd dankzij beleid?	periodiek evaluatie-onderzoek	minimaal eens in de vijf jaar	aantal oud-deelnemers met een betaalde baan <i>dankzij</i> de inburgeringsprogramma's
doelmatigheid van beleid	doelstellingen te realiseren met minder middelen/meer effecten met dezelfde middelen	periodiek evaluatie-onderzoek	minimaal eens in de vijf jaar	uitkomsten van een vergelijking van sobere en meer luxe inburgeringsprogramma's
doelmatigheid van de bedrijfsvoering (de ontwikkeling, aansturing en uitvoering van beleid)	kosten en kwaliteit van de geleverde producten en diensten	systemen van reguliere prestatiegegevens (terugval: periodiek evaluatie-onderzoek)	jaarlijks (terugval: meerjarig)	de kostprijzen van de op jaarbasis opgeleide cursisten (waarbij aan bepaalde kwaliteitseisen wordt voldaan)

2.4 Aan welke kwaliteitseisen voldoet goed evaluatieonderzoek?

Uitkomsten van evaluatieonderzoek ex post moeten bijdragen aan de kwaliteit van de besluitvorming over beleid. Parallel kunnen de uitkomsten bijdragen aan de kwaliteit van de besluitvorming over de bedrijfsvoering (e.g. het besluit tot een andere inrichting van de organisatie om de doelstellingen

van het beleid beter te realiseren). De deugdelijkheid van deze informatie is daarbij een belangrijke voorwaarde.

De evaluatieregeling schrijft voor dat evaluatieonderzoek ex post dient te voldoen aan ‘de kwaliteitseisen die gelden voor de voor beleidsevaluatie geëigende methoden en technieken van sociaal-wetenschappelijk onderzoek’. Dit zijn:

- a validiteit (de algemene geldigheid van de onderzoeksopzet en wijze waarop tot conclusies wordt gekomen);
- b betrouwbaarheid en nauwkeurigheid (van de gebruikte meetmethoden), en;
- c bruikbaarheid (van de uitkomsten van systeem of onderzoek).

Daarnaast bevat de evaluatieregeling de eis dat ‘de uitkomsten van evaluatieonderzoek ordelijk en controleerbaar tot stand komen’. Daarbij moet het mogelijk zijn om aan de hand van een onderzoeksdossier de kwaliteit en de onafhankelijkheid van evaluatieonderzoek vast te stellen.

Het onderzoeksdossier kan – in aanvulling op de informatie in het onderzoeksrapport – de volgende informatie bevatten:

- het onderzoeksplan (probleemstelling, opzet, bemensing, beoogd gebruik). Bij extern onderzoek: offerteaanvraag, offertebeoordeling, contract, déchargebeslissing. Bij intern onderzoek met name informatie over déchargebeslissing;
- de wijze waarop de (onderzoeks)activiteiten organisatorisch zijn ingebed (o.a. opdrachtgever, opdrachtnemer, etc);
- informatie over (intern en extern uitgevoerde) kwaliteitstoetsen, bijvoorbeeld notities die ten grondslag liggen aan uitgevoerd onderzoek.

Een onafhankelijke voorbereiding en/of uitvoering van evaluatieonderzoek of het op andere wijze waarborgen van de onafhankelijkheid waarmee de merites van beleid worden beoordeeld, zijn van groot belang voor de (gepercipieerde) betrouwbaarheid van evaluatieonderzoek. Als methoden om de onafhankelijkheid van beleidsevaluatie te vergroten valt te denken aan het werken met onafhankelijke begeleidingscommissies of onafhankelijke comités van toezicht.

Het bovenstaande komt wellicht nog wat abstract over. Om deze reden bevat het volgende hoofdstuk een aantal praktische stappen, waarin de genoemde doelen van evaluatieonderzoek en de te hanteren kwaliteitscriteria in de vorm van checklisten zijn opgenomen.

3 De opzet en uitvoering van evaluatieonderzoek ex post: een stappenplan

Een stappenplan is vaak een praktisch hulpmiddel bij het opzetten en (dan) uitvoeren van een evaluatieonderzoek. De volgende stappen zijn te onderkennen:

Stap 1	Het formuleren van de probleemstelling en vragen van het onderzoek
Stap 2	De keuze voor een onderzoeksopzet
Stap 3	Het hoofdonderzoek: gegevensverzameling en -verwerking
Stap 4	De rapportage en follow-up

Per stap wordt ingegaan op het doel, eventuele relevante tussenstappen en de bijbehorende kwaliteitscriteria (waar moet op worden gelet?). Het doorlopen van de stappen kan een herhalend proces zijn.

3.1 De eerste stap: het formuleren van de probleemstelling en vragen van het onderzoek

Het komen tot een heldere, volledige en juiste probleemformulering en – daaraan gekoppeld – een goede probleemstelling en vragen voor het onderzoek is één van de meest cruciale fasen in een evaluatieonderzoek ex post: het hele onderzoek wordt er in feite mee gestuurd. De eerste stap betreft daarom het uitwerken van de informatiebehoefte tot een duidelijke doelstelling en een onderzoekbare probleemstelling met onderzoeksvragen.

Vaak leven er bij aanvang van een evaluatieonderzoek allerlei verwachtingen van ‘state of the art’-beleid (e.g. heldere doelstellingen, aandacht voor causale verbanden). Omdat deze verwachtingen vaak niet worden uitgesproken, is het van belang om deze verwachtingen realistisch te maken. Om misverstanden te voorkomen, is het goed de verwachtingen expliciet te maken met behulp van de volgende checklist:

CHECKLIST

- Is de aanleiding tot het evaluatieonderzoek helder omschreven? Is er bijvoorbeeld sprake van een wettelijke verplichting, behoefte aan onafhankelijke beoordeling van beleidsuitvoering of resultaten of heeft men behoefte aan inzicht in resultaten?
- Wat is de informatiebehoefte of vraagstelling vanuit het 'beleid'?
- Is het doel van het evaluatieonderzoek helder omschreven? Is het probleem voldoende ingeperkt zodat het hanteerbaar is voor de onderzoeker?
- Zijn de doelen van het beleid helder aangegeven? Zijn de doelstellingen waar mogelijk gekwantificeerd (dat wil zeggen voorzien van streefwaarden, eindtermijnen en informatie over de doelgroep van het beleid)? Drukken de doelstellingen uit waar het om gaat? Zijn ze juist gekozen en niet manipulatief of symbolisch?
- Zijn de causale verbanden tussen de inzet van beleidsinstrumenten en het realiseren van effecten realistisch, gefundeerd, consistent, volledig? Mag worden verwacht dat de inzet van de instrumenten en de geleverde prestaties bijdragen aan het bereiken van de gestelde doelen?

Evaluatieonderzoek ex post zal in veel gevallen gericht zijn op de *effecten van beleid*. Een goede beleidsdoelstelling bevat informatie over de nagestreefde effecten en de manier waarop die worden gerealiseerd. Het evalueren van beleid heeft dan betrekking op *doelbereiking, doeltreffendheid, doelmatigheid van beleid/kosteneffectiviteit en doelmatigheid van bedrijfsvoering* (zie ook paragraaf 2.2). In wezen is daarmee de cirkel rond. Tegenvallende effecten zullen vaak aanleiding zijn om de beleidsuitvoering en/of de beleidsinhoud aan een nadere evaluatie te onderwerpen.

Het richten van een evaluatieonderzoek door te vertrekken vanuit 'de beleidstheorie'

Een beleidstheorie kan worden begrepen als 'het geheel van veronderstellingen dat aan beleid ten grondslag ligt'. In veel begrotingen en steeds meer beleidsnota's is zo'n theorie goed zichtbaar. De drie w-vragen uit vBtB vormen in wezen de essentie van zo'n beleidstheorie:

- wat willen we bereiken? (doelstellingen van beleid)
- wat gaan we daarvoor doen? (welke acties en instrumenten zullen naar verwachting bijdragen aan het bereiken van die doelstellingen?)
- wat gaat dat kosten?

Bovenstaand punt geldt voornamelijk voor (relatief) nieuw beleid. Uiteraard is er nog veel bestaand beleid dat is bedacht en opgezet in het pre-vBtB tijdperk. Dit is niet in wv-terminen opgezet, waardoor de beleidstheorie

erachter weinig (expliciet) zal zeggen over het behalen van doelstellingen. Voor zulk beleid is het dan ook niet altijd eenvoudig te achterhalen wat de aanvankelijke doelstellingen waren. Met de intrede van VBTB zou dit echter steeds minder frequent moeten voorkomen waardoor evaluatieonderzoek ex post zich in de toekomst meer zal kunnen richten op het beantwoorden van de HHH-vragen, ook op instrumenteel niveau:

- hebben we bereikt wat we wilden bereiken?
- hebben we gedaan wat we daarvoor zouden gaan doen?
- heeft dat gekost wat we zeiden dat het zou gaan kosten?

Intermezzo: wat te doen met 'slecht' (pre-VBTB) beleid?

Het komt dus nog steeds voor dat beleidsdoelstellingen en de veronderstellingen achter de inzet van instrumenten niet duidelijk zijn. In deze situatie spreken wij van 'slecht' beleid. Hier is niet duidelijk wat men wil bereiken óf niet duidelijk waarom bepaalde instrumenten worden ingezet voor bepaalde doelstellingen. Bij zulk beleid is het zaak om – voorafgaand aan een evaluatieonderzoek – de beleidstheorie te reconstrueren. Bij dit 'achteraf opstellen' van een beleidstheorie wordt de bijdrage die een beleidsvoerder impliciet verwacht van zijn instrumenten en werkzaamheden om een bepaald probleem aan te pakken, expliciet of zichtbaar gemaakt. Een reconstructie van een beleidstheorie kan op verschillende manieren plaatsvinden. Bestudering van beleidsnota's en parlementaire stukken waarin beleidsmaatregelen worden aangekondigd is de meest logische weg. Een andere mogelijkheid is het interviewen van degenen die bij het beleid zijn betrokken of betrokken zijn geweest ten tijde van instelling daarvan. In de meeste gevallen zal nu wel duidelijk zijn welke doelstellingen het beleid beoogde te behalen. Indien hier nog steeds geen eenduidige beleidsdoelstelling uitkomt, kan, uit praktisch oogpunt, de doelstelling van het bovengelegen operationele/algemene doel uit de begroting worden overgenomen. Op deze wijze richt de evaluatie zich op de *bijdrage* van het beleid of instrument aan de hogere doelstelling. Deze bijdrage wordt toch al enigszins verondersteld door de indeling van de begroting. Continuering van 'slecht' beleid ligt niet voor de hand als na beleidsreconstructie niet duidelijk is welke doelstelling wordt nagestreefd en geen aannemelijke bijdrage aan het bovengelegen doel is gebleken.

Hier dient wel vermeld te worden dat niet kost wat het kost gestuurd moet worden op prestaties en, in het verlengde daarvan, de evaluatie zich daar op moet richten. Indien prestaties weinig concreet zijn, kan worden teruggevallen op beheersing van het proces en/of de input. Een evaluatie moet in dit geval uitwijzen of het proces en de input optimaal en kwalitatief goed zijn én natuurlijk of het doel ook daadwerkelijk bereikt wordt.

Ook de Rekenkamer gaat er bij doelmatigheidsonderzoek vaak toe over de beleidstheorie te reconstrueren. Om de doelen van de beleidsmakers te reconstrueren maakt de Rekenkamer gebruik van een zogenaamde 'doelenboom'. Als onhelder is wat de officiële beleidsdoelen zijn, of welke beleidstheorie verondersteld wordt werkzaam te zijn, probeert de Rekenkamer in kaart te brengen welke relaties er (verondersteld worden te) bestaan tussen de instrumenten, de subdoelen en einddoelen van het beleid. Uiteraard moet zo'n doelenboom vervolgens worden besproken met de beleidsvoerder of 'doelstellingseigenaar'. Het gaat dan om vragen als: zijn alle doelen meegenomen? Kloppen de gesuggereerde veronderstelde relaties?

Net als bij beleidsontwikkeling zal er soms een versimpeling moeten plaatsvinden van de 'complexe werkelijkheid'. Dit om problemen hanteerbaar te maken voor onderzoekers. Een maatschappelijk probleem (meestal breed en ongreepbaar, zoals het werkloosheidsprobleem) dient allereerst 'gereduceerd' te worden tot een beleidsprobleem:

Voorbeeld: 'Is de huidige aanpak van de jeugdwerkloosheid de meest effectieve?'

Een beleidsprobleem dient vervolgens op zijn beurt weer te worden teruggebracht tot een onderzoeksprobleem, uit te werken in een probleemstelling van het onderzoek:

Voorbeeld: 'Hoeveel werkloze jongeren worden dankzij het huidige beleid aan een vaste baan geholpen?'

Met andere woorden: het beleidsprobleem is een specificatie en inperking van het maatschappelijk probleem; het onderzoeksprobleem is een verbijzondering van het beleidsprobleem.

Voorbeeld inburgeringsbeleid

Algemene beleidsdoelstelling: maatschappelijke integratie van nieuwkomers in Nederlandse samenleving

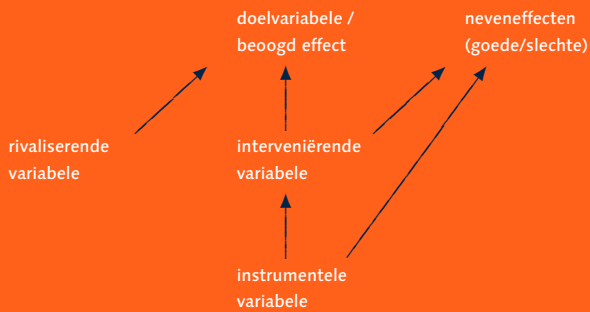
Nader geoperationaliseerde doelstelling: % van nieuwkomers met betaald werk

Onderzoeksprobleemstelling: welk % van nieuwkomers heeft na 2 jaar betaald werk en in hoeverre is dat toe te schrijven aan de inburgeringsprogramma's?

Analyseschema: een goed hulpmiddel

Een instrument dat behulpzaam kan zijn bij het formuleren van probleemstelling en onderzoeksvragen is het **analyseschema**. Een analyseschema vormt als het ware een 'plattegrond': het beleidsprobleem kan helder in kaart worden gebracht door de relatie tussen instrumenten, tussendoelen en einddoel en door de relatie tussen externe factoren, doelen en neveneffecten in kaart te brengen. Bij evaluatieonderzoek naar de effectiviteit van beleid staat de relatie tussen een instrument (instrumentele variabele) en het beoogde effect (doelvariabele) centraal. Echter, op deze relatie werken veelal andere factoren in. In een evaluatieonderzoek zal zelden sprake zijn van een eenvoudige causale relatie tussen instrumentele en een doelvariabele. Vaak verloopt het bereiken van effecten langs zogeheten 'interveniërende variabelen': tussengelegen effecten die onmisbare schakels vormen naar het uiteindelijke effect. Ook kan van andere variabelen dan de instrumentele variabele een zelfstandige invloed uitgaan op de doelvariabele (rivaliserende variabele; externe factoren).

Schema 4: Analyseschema



In eerder genoemd voorbeeld over inburgeringsbeleid worden inburgeringsprogramma's de instrumentvariabele genoemd en het percentage nieuwkomers met betaald werk de doelvariabele. Een interveniërende variabele in dit voorbeeld kan zijn: kennis bij nieuwkomers van de Nederlandse taal en cultuur. Ook kan er van andere variabelen een zelfstandige invloed uitgaan op de doelvariabele. Het is bijvoorbeeld denkbaar dat de krapte op de arbeidsmarkt een zelfstandige invloed heeft op het percentage nieuwkomers met betaald werk. Een neveneffect – als gevolg van de instrumentele en interveniërende variabelen – kan zijn het percentage nieuwkomers dat in Nederland blijft.

Schema 5: Analyseschema voor inburgeringsbeleid



Over het algemeen zijn de doelstelling en probleemstelling van het onderzoek nog op een tamelijk hoog abstractieniveau geformuleerd. De volgende tussenstap is daarom het 'vertalen' van probleem- en doelstelling in concrete onderzoeksvragen.

De onderzoeksvragen bij evaluatieonderzoek ex post volgen logischerwijs het doel van het onderzoek. Afhankelijk van dit doel kunnen zij betrekking hebben op één of meerdere van de volgende onderwerpen:

- **de uitvoering van het beleid:** Op welke wijze is de uitvoering verlopen? Welke knelpunten deden zich daarbij voor? Hoe kan de doelmatigheid van de uitvoering worden beoordeeld?
- **de doeltreffendheid of effectiviteit van het beleid:** In welke mate worden de doelstellingen van het programma bereikt? Is in de huidige situatie de gekozen beleidsmaatregel c.q. het gekozen beleidsalternatief ook de meest effectieve?
- **de doelmatigheid of kosten-effectiviteit van het beleid:** Hadden met dezelfde middelen meer gewenste effecten bereikt kunnen worden? Was hetzelfde resultaat mogelijk geweest met minder middelen?

- *de neveneffecten van het beleid*: het gaat hierbij om het optreden van effecten van het beleid (positief of negatief) die niet waren beoogd.

Extra stap bij uitbesteding van evaluatieonderzoek: opdrachtformulering

NB Een heldere opdrachtformulering is niet alleen noodzakelijk bij uitbesteding van het evaluatieonderzoek, maar tevens zeer nuttig wanneer het onderzoek intern wordt uitgevoerd! De opdrachtformulering kan het startdocument vormen voor uitvoering van de evaluatie.

De keuze om een evaluatieonderzoek uit te besteden, is afhankelijk van verschillende factoren. In de eerste plaats van het al dan niet aanwezig zijn van voldoende personeel. Met name het verzamelen en analyseren van gegevens is een tijdrovende bezigheid. Ten tweede kan men besluiten om onderzoek uit te besteden aan een onderzoeksbureau dat een bepaalde deskundigheid heeft opgebouwd. Voor het uitvoeren van een evaluatieonderzoek is onderzoekskundigheid en -ervaring een vereiste. Door uitbesteding van onderzoek kunnen capaciteit en deskundigheid 'ingehuurd' worden. Een derde reden voor uitbesteding kan zijn dat hiermee de onafhankelijkheid kan worden vergroot.

'De slager die zijn eigen vlees keurt'

Ook bij het kiezen van een onderzoeksbureau moet de onafhankelijkheid in acht worden genomen. Er is een voorbeeld bekend van een onderzoeksbureau dat onderzoek had gedaan naar het functioneren van een dienst en ter verbetering aanbevelingen had voorgesteld. Een aantal jaren later is hetzelfde bureau ingehuurd om de effecten van de aanbevelingen te evalueren. Hier kan de onafhankelijkheid van het onderzoeksbureau in twijfel worden getrokken, aangezien ze de effecten van hun eigen aanbevelingen beoordelen.

Voor de onderzoeksopzet maakt het in principe niet zoveel uit of een evaluatieonderzoek *intern* wordt uitgevoerd of wordt *uitbesteed*: het denkwerk (doel- en probleemstelling, onderzoeksvragen) moet toch binnen een beleidsorganisatie gebeuren. Voorwaarde voor interne uitvoering van evaluatieonderzoek is de beschikbaarheid over voldoende ervaring/deskundigheid en capaciteit. Voordelen van 'zelf doen' zijn: inwerktijd, kennis 'ter zake', 'grip' op de uitkomsten en leervermogen. Nadeel van 'zelf doen' is zoals eerder gezegd het beslag op de organisatie. Voordelen van uitbesteden zijn: neutraliteit, status/gewicht en (hopelijk!) de aanwezigheid van onderzoeksexperts. Nadelen van uitbesteden zijn: hoge kosten, het risico van

standaardadviezen en – iets dat nogal eens wordt onderschat – het proces dat leidt tot uitbesteden. Dit proces vergt het meeste denkwerk en kost veel tijd. Niet alleen het opstellen van het aanbestedingsdocument vergt veel tijd, maar ook het beoordelen van de offertes is een traject wat niet onderschat moet worden. Boven een bepaalde grens² moet er Europees worden aanbesteed. Hiervoor zijn speciale regels opgesteld, onder andere is er een minimum aantal dagen dat een opdracht gepubliceerd moet staan voor de deadline waarna overgegaan mag worden tot beoordeling en opdrachtverlening. Hier moet zeker rekening mee worden gehouden in de planning van het geheel³. Een ander nadeel van uitbesteding is dat inhoudelijke, gedetailleerde kennis van de uitvoering van het beleid en de uitwerking ervan op de doelgroep niet binnen de eigen organisatie wordt opgebouwd, maar buitenshuis.

Over het algemeen is een aantal *fasen/stappen* te onderscheiden in het proces van uitbesteding van een evaluatieonderzoek. Deze worden in *bijlage 2* nader beschreven.

In internationaal perspectief wordt het opstellen van *Terms of Reference* (ToR) als een vitale stap in het proces van uitbesteding van een evaluatieonderzoek beschouwd⁴. De Terms of Reference is een standaard onderzoeksopzet, bestaande uit 11 stappen, en vormt het basisdocument voor de uitvoering van een evaluatieonderzoek. Bij uitbesteding van het evaluatieonderzoek dient de ToR, samen met de goedgekeurde offerte en opdrachtbrief, als contract. In *bijlage 3* wordt kort de standaardopzet van de Terms of Reference weergegeven.

Enkele aandachtspunten bij evaluatie van uitvoeringsprocessen

Centraal bij de evaluatie van de uitvoering van beleid staat de vraag: zijn er knelpunten in de uitvoering aanwijsbaar, op grond waarvan de conclusie getrokken kan worden dat het effect van het beleid groter had kunnen zijn? Deze knelpunten hebben in veel gevallen te maken met:

- de (onvoldoende) kennis en deskundigheid van personeel
- de capaciteit en kwaliteit van betrokken uitvoeringsorganisaties
- de kwaliteit van ondersteunende faciliteiten (e.g. ICT, gegevensbeheer, communicatie)

2 2003: € 162.293,- voor diensten.

3 Alle aanbestedingseisen zijn terug te vinden op de site van het Ministerie van Economische Zaken (www.ez.nl).

4 Bron: MEANS (Means for Evaluating Actions of a Structural Nature), een evaluatieprogramma van DG Regionaal Beleid binnen de Europese Commissie.

- het bestaan van verborgen agenda's van de betrokken actoren
- onvoldoende bekendheid van de doelgroep met het beleid (e.g. subsidiemogelijkheden).

CHECKLIST KWALITEITSCRITERIA

Bij het doorlopen van de eerste stap kan de volgende checklist worden gehanteerd:

- Is er een onderzoeksprobleemstelling dan wel -doelstelling geformuleerd en is deze op een goede manier uitgewerkt in onderzoeksvragen? (Dat wil zeggen: is de vertaalslag naar concrete onderzoeksvragen gemaakt en richten zij zich op het doel van het onderzoek?)
- Worden de te onderzoeken begrippen op de juiste manier geoperationaliseerd? Dat wil zeggen: zijn de te onderzoeken eenheden helder gedefinieerd en zijn zij representatief voor de probleemstelling, doelstelling en vragen van het onderzoek?
- Sluit de opdrachtformulering aan op de onderzoeksprobleemstelling/-doelstelling en de onderzoeksvragen (indien aanwezig d.w.z. bij uitbesteding van onderzoek)?

In de vakliteratuur worden bovenstaande criteria ook wel *validiteitscriteria* genoemd. Validiteit (meet ik wat ik wil meten?) heeft betrekking op de algemene geldigheid van een evaluatieonderzoek.

AANVULLENDE VBTB-CRITERIA

- Wordt vermeld welke algemene en/of nader geoperationaliseerde (VBTB-)doelstellingen van beleid in het evaluatieonderzoek zijn betrokken?
- Wordt ingegaan op de beleidstheorie: welke causale relaties mogen verwacht worden? Wat zijn de veronderstelde probleem-oorzaak-gevolg-relaties en doel-middel-relaties?
- Wordt in de onderzoeksopzet rekening gehouden met/voortgeborduurd op reeds beschikbare prestatiegegevens?

3.2 De tweede stap: de keuze voor een onderzoeksopzet

Als eenmaal de doel- en probleemstelling en vragen van het onderzoek zijn vastgesteld, moeten keuzes gemaakt worden met betrekking tot het volgende element in evaluatieonderzoek: *de onderzoeksopzet*. Aangezien evaluatieonderzoek ex post in de meeste gevallen gericht zal zijn op het bepalen van de doeltreffendheid en doelmatigheid van beleid – en daarmee op *effecten* – zijn deze en de volgende daarop gericht.

NB *Voor uitvoeringsonderzoek of onderzoek naar de doelmatigheid van de bedrijfsvoering is de hieronder beschreven tweede stap niet relevant.*

Doel van de onderzoeksopzet is: het zoveel als mogelijk afbakenen van de invloed van het gevoerde beleid van andere factoren die van invloed zijn op de mate van doelbereiking. Het gaat er immers om de causale relatie tussen oorzaak (beleid) en gevolg (effect) vast te stellen. Er bestaan daarbij verschillende mogelijkheden. De bij aanvang van het beleid gewenste situatie (het beleidsdoel) kan, al dan niet volledig, bereikt zijn. Als dit zo is, kan dat al dan niet het gevolg zijn van het beleid. Maar ook als de nagestreefde situatie niet (volledig) optreedt, kan dat liggen aan het gevoerde beleid. De onderzoeksopzet geeft aan welke gegevens moeten worden verzameld en hoe die moeten worden geanalyseerd om op deze vragen antwoord te krijgen. Er bestaat een aantal typen onderzoeksopzetten waarmee de effectiviteit van beleid kan worden nagegaan. Mede afhankelijk van de beschikbaarheid van gegevens kan uit een van deze typen onderzoeksopzetten een keuze worden gemaakt:

- a een opzet gebaseerd op een *nameting*;
- b een opzet gebaseerd op een *voor- en nameting*;
- c een opzet gebaseerd op een *met- en zonder meting*, en;
- d een opzet gebaseerd op een combinatie van een *voor-na* en een *met-zonder meting*.

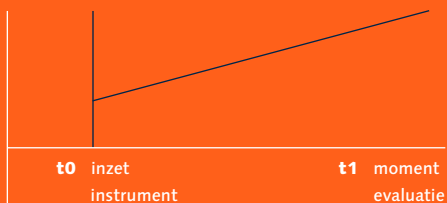
A *Opzet gebaseerd op nameting*

Kenmerkend voor dit type onderzoeksopzet is dat slechts gegevens beschikbaar zijn over de ontwikkeling die heeft plaatsgevonden nadat nieuw beleid is geïntroduceerd of nadat een nieuw beleidsinstrument is ingezet. Bij deze methode ligt de nadruk meer op het plausibel redeneren dan op het empirisch vaststellen.

Bijvoorbeeld

Vastgesteld wordt dat sinds de invoering van een subsidieregeling het aantal winkels in kleine dorpen stijgt, gelijk blijft of daalt.

Ontwikkeling van de doelvariabele



De kracht van evaluatieonderzoek dat louter gebaseerd is op nameting hangt mede af (1) van de volledigheid waarmee potentiële rivaliserende verklaringen kunnen worden geïnventariseerd en (2) van de mate waarin deze alternatieve verklaringen kunnen worden ontkracht of juist bevestigd. De methode is dus niet volledig betrouwbaar: harde conclusies over doelmatigheid en doeltreffendheid zijn uitgesloten. De methode levert niet meer op dan een plausibele indruk van de effecten van een ingezet instrument. Dat is niet erg; niet zelden is ‘beredeneerde plausibiliteit’ het hoogst haalbare.

Vergroten van zeggingskracht door alternatieve verklaringen in kaart te brengen

Voor de eerste drie beschreven modellen (nameting, voor- en nameting, met en zonder meting) geldt dat deze beredeneerde plausibiliteit vergroot kan worden door alternatieve verklaringen uit te sluiten. Dit is mogelijk door de volgende vragen te stellen:

- Het inventariseren welke factoren *naast* het beleid (instrument) in principe van invloed zijn op de doelvariabele. In ons voorbeeld is het de vraag welke factoren *naast* de subsidieregeling de aanwezigheid van winkels in kleine dorpen nog meer hadden kunnen beïnvloeden? Mogelijk antwoord: de economische ontwikkeling of de belastingdruk.
- Het beredeneren of van deze factoren een positieve of negatieve invloed uitgaat op de doelvariabele. Mogelijk antwoord in voorbeeld: economische groei leidt tot meer winkels, economische stagnatie tot minder.
- Het nagaan – door bestudering van literatuur of door interviews – of er aanwijzingen zijn voor de veronderstelde relaties. Mogelijk antwoord in ons voorbeeld: er is de laatste tijd sprake van een ongewijzigde economische situatie en er zijn evenmin wijzigingen in het belastingklimaat. Verondersteld (niet: bewezen!) mag dan worden dat de stijging van het aantal duurzame winkels in kleine kernen waarschijnlijk wordt veroorzaakt door de subsidieregeling.

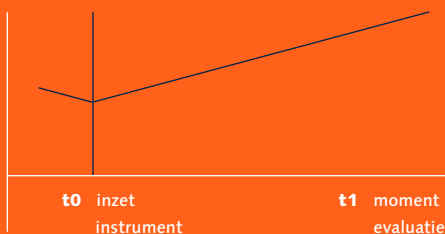
B *Opzet gebaseerd op voor- en nameting*

Kenmerkend voor een opzet die is gebaseerd op een voor- en een nameting is dat een ontwikkeling zowel voor en na de inzet van een beleidsinstrument wordt gemeten. Uitgegaan wordt van verschillen tussen de situatie voorafgaande aan de inzet van het instrument en de situatie na inzet. Het verschil in ontwikkeling dat men ziet, de feitelijke verandering, is veelal het uitgangspunt voor een nadere analyse. Deze analyse richt zich op het zo goed mogelijk ‘isoleren’ van de effecten van het beleid. Dit is nodig omdat het verschil ook andere oorzaken kan hebben; oorzaken die dan wel ‘toeval- lig’ samenvallen met het moment waarop het beleid werd ingezet. De hierboven beschreven vragen naar ‘alternatieve verklaringen’ zijn om die reden ook bij de onderzoeksopzet met voor- en nameting van toepassing.

In ons voorbeeld

Als uit een voor- en nameting blijkt dat er na invoering van een subsidiemaatregel meer winkels dan voorheen zich hebben gevestigd in kleine dorpen, is het op het eerste gezicht plausibel dit toe te schrijven aan het gevoerde beleid. Economische groei en de daarmee gepaard gaande vraag, die wellicht samenviel met het moment waarop de subsidieregeling van kracht werd, kan een alternatieve verklaring zijn.

Ontwikkeling van de doelvariabele



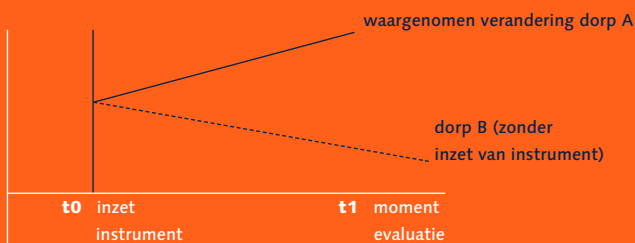
C *Opzet gebaseerd op met- en zonder-meting*

In geval van de met- en zonder-meting wordt de situatie die op een bepaald terrein is ontstaan als gevolg van de inzet van het beleidsinstrument vergeleken met de situatie die in een 'controlegroep' is ontstaan (dat wil zeggen: een vergelijkbare groep, populatie of domein waar geen beleidsinstrument is ingezet). In tegenstelling tot de eerder beschreven modellen wordt er in deze benadering expliciet naar gestreefd om de bijdrage van het beleidsinstrument te isoleren. Met andere woorden: bij de met- en zonder-meting wordt tegelijkertijd gekeken naar een groep die is blootgesteld aan het beleid (de doelgroep) en naar een groep die er niet aan is blootgesteld. Veel medische experimenten kennen deze opzet. Voorzichtigheid is echter ook hier geboden: conclusies over de doeltreffendheid of effectiviteit van het beleidsinstrument – of het geneesmiddel – kunnen alleen getrokken worden als beide groepen identiek zijn. Het is in de praktijk vaak moeilijk een controlegroep te vinden. Ook het beginsel van rechtsgelijkheid verzet zich vaak tegen het werken met grootschalige controlegroepen.

In ons voorbeeld

Uit een met- en zonder-meting kan naar voren komen dat er in dorpen waar de subsidiemaatregel is ingevoerd zich meer winkels hebben gevestigd dan in vergelijkbare dorpen waar geen maatregel is genomen. Ook hierbij is het op het eerste gezicht plausibel dit toe te schrijven aan het gevoerde beleid. De vergelijkbaarheid is vaak echter een probleem. Door gericht vragen te stellen aan betrokkenen over de motieven van vestiging e.d., kan aan geloofwaardigheid en zeggingskracht worden gewonnen.

Ontwikkeling van de doelvariable



D *Opzet gebaseerd op zowel voor- en nameting als met- en zonder-meting*

De enige opzet waarmee de doeltreffendheid van een bepaalde maatregel vrijwel zeker kan worden vastgesteld, is die waarbij zowel een voor- als nameting wordt gehouden als waarbij gewerkt wordt met een experimentele en een controlegroep.

Keuze onderzoeksopzet

Opgemerkt kan worden dat de keuze voor één van bovenstaande onderzoeksopzetten niet wordt bepaald door alleen strikt methodologische overwegingen (als dat wel het geval zou zijn, verdient opzet d de voorkeur!). De gegevens waarover men kan beschikken, zijn bepalend voor de keuze van een bepaalde opzet. Soms zal men alleen in staat zijn tot een nameting, in andere gevallen beschikt men over omvangrijke tijdreeksen. Informatie over controlegroepen is zeer zeldzaam: vaak gaat hier een echte experimentele onderzoeksopzet aan vooraf (waarmee deze opzet veelmeer een ex ante karakter zal hebben). Daarnaast zullen in veel gevallen de beschikbare tijd, het onderzoeksbudget, de tijdsdruk om het probleem op te lossen en de kosten van onderzoek van grote invloed zijn. Ieder evaluatieonderzoek vergt maatwerk.

Kwantitatief en kwalitatief evaluatieonderzoek

In kwantitatief onderzoek wordt gebruik gemaakt van de verschillende statistische technieken die voorhanden zijn. Het komt echter regelmatig voor dat het niet mogelijk is met kwantitatieve gegevens te werken, bijvoorbeeld wanneer men onvoldoende inzicht heeft in de variabelen die voor het onderzoek van belang zijn of wanneer men niet over de juiste gegevens beschikt.

NB *Wie de strekking van het bovenstaande – men heeft geen inzicht in de variabelen die voor het onderzoek van belang zijn – op zich in laat werken, zal moeten erkennen dat de kans groot is dat men ook geen inzicht zal hebben in de variabelen die voor het bereiken van beleidsdoelstellingen relevant zijn.*

In dergelijke gevallen wordt kwalitatief onderzoek verricht. Kwalitatief en kwantitatief onderzoek sluiten elkaar niet uit; eerder zijn ze complementair. Zo kan met kwantitatief onderzoek het effect van beleid in kaart worden gebracht, terwijl vervolgens met een kwalitatief onderzoek wordt onderzocht waarom een bepaald effect (niet) optreedt. Met kwantitatief onderzoek wordt nagegaan in hoeverre relaties *statistisch generaliseerbaar* zijn: in hoeverre geldt wat men heeft gevonden ook voor andere dan de onderzochte gevallen? Met kwalitatief onderzoek kan inzicht verkregen worden in de *validiteit* van relaties: wat zijn de oorzakelijke relaties tussen de verschillende variabelen en hoe beïnvloeden ze elkaar?

Het onderscheid tussen kwantitatief en kwalitatief evaluatieonderzoek wordt in de praktijk nogal eens vertroebeld doordat onderzoekers kwalitatieve studies proberen 'hard te maken' aan de hand van fraaie statistieken en grafieken. Gewaakt moet worden voor een 'kwantitatieve saus' bij een rapport dat vooral verslag doet van een kwalitatief onderzoek! Een voorbeeld is het zeer uitvoerig presenteren van de mening van respondenten over de kwaliteit van genoten inburgeringscursussen. Hoe relevant dergelijke meningen ook zijn, door het uitvoerig opnemen van allerhande kwantitatieve informatie wordt vaak ten onrechte de schijn gewekt dat de 'doelmatigheid' of 'doeltreffendheid' van beleid goed te meten is.

'Stakeholders' evaluation'

Een belangrijke trend is om ook de belanghebbenden bij -de uitvoering van- beleid in het evaluatieonderzoek mee te nemen. Daarbij kan het gaan om de doelgroep van beleid, maar ook om instanties die belast zijn met de uitvoering van bepaalde maatregelen. In de angelsaksische landen heeft deze zogeheten 'stakeholders' evaluation' geleid tot expliciete aandacht voor het perspectief en de belangen van groepen die door – de uitvoering van – beleid geraakt worden. Impliciet gevolg van een dergelijke insteek is dat het centrale perspectief van 'de beleidsmaker' wordt aangevuld met het perspectief van een beleidsnetwerk.

CHECKLIST KWALITEITSCRITERIA

Bij het doorlopen van de tweede stap kan de volgende checklist worden gehanteerd:

- Wordt er een beargumenteerde keuze gemaakt voor een specifiek onderzoeksontwerp en de te hanteren onderzoeksmethode(n)?

In de vakliteratuur wordt bovenstaand criterium ook wel een *validiteitscriterium* genoemd. Validiteit (meet ik wat ik wil meten?) heeft betrekking op de algemene geldigheid van een evaluatieonderzoek. Een goed onderzoeksontwerp is gericht op die geldigheid.

3.3 De derde stap: het hoofdonderzoek: gegevensverzameling en -verwerking

De derde stap betreft het *hoofdonderzoek: het verzamelen en verwerken van onderzoeksgegevens*. Het doel is het kiezen van een geschikte onderzoeksmethode en wijze waarop gegevens worden verwerkt en geanalyseerd. 'Geschikt' wil vooral zeggen:

- Biedt de methode antwoord op de onderzoeksvragen?
- Wordt er door de methode optimaal gebruik gemaakt van beschikbare gegevens?
- Is er – gegeven de gekozen methode – voldoende aandacht voor de aan beleidsinformatie te stellen kwaliteitseisen?

Er zijn diverse technieken om de benodigde evaluatiegegevens te verzamelen. Wij noemen: de documentenanalyse, het interview, de enquête, participerende observatie en de gevalsstudie of case-study. Afhankelijk van de aard van het evaluatieonderzoek en de gegevensbronnen wordt een keuze gemaakt uit de technieken. Nadat de gegevens zijn verzameld, kunnen de data-analyses beginnen. Opgemerkt moet worden dat het trekken van een steekproef vaak vooraf gaat aan een serie interviews, een enquête of een experiment.

Welke van de genoemde bronnen te gebruiken, hangt af van:

- de overwegend kwantitatieve of kwalitatieve aard van het evaluatieonderzoek c.q. de onderzoeksopzet;
- de aanwezigheid en beschikbaarheid van bestaand materiaal;
- de bereidheid van relevante respondenten tot medewerking;
- de randvoorwaarden in tijd en arbeidskracht van de ‘onderzoeker’.

NB *Als het goed is, zullen de bestaande systemen van reguliere prestatiegegevens een belangrijk deel van de benodigde gegevens leveren. Zo is bij het inburgeringsbeleid op basis van bestaande administraties vaak reeds bekend hoeveel mensen een cursus hebben gevolgd, hoeveel mensen zijn geslaagd en – door raadpleging van arbeidsbureaus – hoeveel van hen geschikt werk hebben gevonden.*

A Documentenanalyse

Een methode om relevante informatie te verzamelen is het analyseren van documenten. Gedacht kan worden aan beleidsstukken, begrotingsstukken, rapportages, overzichten, vergaderverslagen, statistieken e.d. Bij de inventarisatie en selectie van documenten is het van belang na te gaan of men alle relevante documenten heeft. Documentenanalyse is zowel een kwantitatieve als een kwalitatieve gegevensverzamelingsstechniek. Het kan immers gaan om het analyseren van statistische gegevens, maar ook van kwalitatieve beweegredenen/argumenten.

B Interview

Het interview (m.n. het diepte-interview) is een kwalitatieve gegevensverzamelingsstechniek. Zij past het beste bij een kwalitatieve onderzoeksopzet maar wordt ook vaak gebruikt als aanvulling op kwantitatieve gegevens. Er zijn verschillende interview-vormen. Hier worden onderscheiden het open, het semi-gestructureerde en het gestructureerde interview. Bij het *open interview (diepte-interview)* zijn de vragen niet van tevoren vastgesteld. Dit geeft de mogelijkheid om diep in te gaan op verscheidene onderwerpen. Deze vorm van interviewen kan goed gebruikt worden wanneer het van belang is inzicht te verkrijgen in ervaringen en meningen van personen (deskundigen). Het *semi-gestructureerde interview* houdt in dat een deel van de vragen vooraf is vastgesteld. Voordeel is dat een zekere mate van vergelijkbaarheid van de gegevens uit de interviews aanwezig is. Voorts blijft voldoende ruimte voor de geïnterviewde over om eigen ervaringen en meningen te vertellen. Het *gestructureerde interview* bestaat uit vooraf opgestelde vragen. Hierdoor worden gegevens op eenzelfde manier verkregen hetgeen de objectiviteit ten goede komt. De verkregen gegevens zijn ook in hoge mate

vergelijkbaar. Nadeel is dat weinig ruimte wordt gelaten voor de ervaringen van de geïnterviewde. Goed interviewen is erg moeilijk en vergt training!

C *Enquête*

De enquête kan meer beschouwd worden als een kwantitatieve gegevensverzamelingsstechniek en past daarom vaak beter bij een kwantitatieve onderzoekszopzet (zoals de ‘voor-na’-meting en de ‘met-zonder’-meting). De schriftelijke enquête is een lijst met opgestelde vragen die ter beantwoording aan personen wordt voorgelegd. Een vragenlijst bestaat meestal grotendeels uit gesloten vragen. Door middel van een aantal open vragen kan de gelegenheid gegeven worden eventuele ervaringen te vermelden. In schriftelijke (w.o. elektronische) enquêtes kunnen complexere vragen worden gesteld dan bij telefonische enquêtes. Een nadeel van schriftelijke enquêtes is dat niet goed valt te controleren wie het enquêteformulier invult; ook de aanwezigheid van derden kan bij de invulling van invloed zijn. De mondelinge enquête vindt veelal telefonisch plaats. De vragen kunnen zowel open als gesloten zijn. Het voordeel van de telefonische enquête is dat in korte tijd zeer veel informatie kan worden verzameld. Bovendien zal de invloed van anderen op de beantwoording van vragen meestal gering zijn. Voorts is de respons meestal hoger dan bij een schriftelijke enquête. Een nadeel is dat telefonische enquêtes in het algemeen niet te lang mogen duren. Evenals het houden van interviews is het maken van vragenlijsten vooral een ‘kunde’. Gezond verstand is hierbij cruciaal. Van groot belang is dat een vragenlijst wordt uitgetest en dat interviewers en enquêteurs worden getraind.

Intermezzo: het trekken van een steekproef

Als het gaat om interviews, enquêtes of experimenten, is het veelal niet mogelijk de totale doelgroep waarin men is geïnteresseerd in een onderzoek te betrekken. Er moet dan een keuze worden gemaakt welke eenheden in het onderzoek moeten worden betrokken. Dit kan gebeuren door het trekken van een steekproef. Het bepalen van een steekproef is een stap die onderdeel kan uitmaken van zowel kwantitatieve als kwalitatieve onderzoeksmethoden.

Een steekproef kan op verschillende manieren worden getrokken. Zo wordt in verkennend onderzoek vaak expliciet op zoek gegaan naar bepaalde onderzoekseenheden. Men is dan geïnteresseerd in onderzoekseenheden met bepaalde eigenschappen en streeft niet direct representativiteit na. Ten behoeve hiervan vindt soms een gerichte selectie van deze eenheden plaats; zo worden bijvoorbeeld extreme, afwijkende of typische onderzoekseenheden eruit gelicht (e.g. uitvallers bij inburgeringscursussen). Juist door analyse van die eenheden verwacht men een inzicht in het probleem te krijgen.

Er wordt ook wel onderscheid gemaakt tussen de *aselect getrokken* steekproef (waarbij iedere onderzoekseenheid uit de doelgroep een even grote kans heeft om deel uit te maken van de steekproef) en de *gestratificeerde* steekproef (waarbij binnen de doelgroep verschillende groepen worden onderscheiden).

Indien men een beeld van de totale doelgroep wil hebben, zijn bij het trekken van een steekproef de volgende vragen relevant:

- om welke populatie gaat het (de 'doelpopulatie' van de steekproef)?
- op basis van welk steekproefkader (lijst, bevolkingsregister, databank etc) kan de steekproef worden getrokken?
- wat is de omvang van de steekproef? De grootte van de steekproef wordt bepaald door de nauwkeurigheid en de betrouwbaarheid waarmee men uitspraken wil doen over de totale populatie (gegeven de spreiding binnen de populatie), de verwachte non-respons en het beschikbare budget.
- is de steekproef representatief? Een steekproefresultaat heeft in veel gevallen pas betekenis indien de steekproef beschouwd kan worden als een afspiegeling van de aspecten van de populatie of het proces waarin men geïnteresseerd is.

D *Participerende observatie*

Participerende observatie is een kwalitatieve gegevensverzamelingsmethode en past het beste bij een kwalitatieve onderzoeksopzet (zoals de nameting). Participerende observatie is tot nu toe bij evaluatieonderzoek weinig toegepast. Het centrale kenmerk ervan is dat de onderzoeker zich het 'lidmaatschap' probeert te verwerven van de groep of institutie die hij wil onderzoeken, d.w.z. deelgenoot probeert te worden van hun cultuur, dan wel daar zo sterk mee in contact te komen dat hij zich een 'insiders view' kan verwerven. Participerende observatie wordt gebruikt als gegevens alleen maar zijn te verkrijgen door middel van waarneming van gedrag. Het kan bijvoorbeeld inzicht geven in verschillende implementatiestijlen van ambtenaren bij regeltoepassing. Een probleem dat bij participerende observatie kan optreden, is dat de onderzochte groep een ander gedrag kan vertonen door de aanwezigheid van de onderzoeker. Ook kan de onderzoeker selectief waarnemen of interpreteren, bijvoorbeeld omdat hij zich te zeer identificeert met zijn onderzoeksobject. Met name als het gaat om langdurige 'participerende observatie' van als problematisch ervaren doelgroepen van beleid zoals daklozen, is de kans niet gering dat waarneming, interpretatie en conclusies daardoor 'gekleurd' raken.

E Case-study

De case-study is eveneens een kwalitatieve gegevensverzamelingsstechniek en past het beste bij een kwalitatieve onderzoeksopzet (zoals de nameting). Een case-study is een empirisch onderzoek:

- waarbij een probleem/verschijnsel in zijn complexe en concrete context wordt bestudeerd;
- waarbij de grenzen tussen het verschijnsel en zijn context niet altijd duidelijk zijn; en
- waarbij verschillende gegevensbronnen worden gebruikt.

Case-studies zijn er in verschillende soorten. Er kan naar één case worden gekeken of naar verschillende tegelijkertijd. Indien in eenzelfde studie naar meer cases wordt gekeken, kan dit de generaliseerbaarheid van de resultaten ten goede komen. Het is dan namelijk mogelijk de cases onderling te vergelijken. Belangrijke gegevensbronnen voor case-studies zijn: open interviews, documenten, archieven, statistieken, participerende observatie en de uitkomsten van eerder onderzoek.

'Triangulatie': gebruik van verschillende gegevensverzamelingsstechnieken

De keuze voor een bepaalde dataverzamelingstechniek zal in het algemeen ingegeven worden door enerzijds de mogelijkheden en beperkingen van de onderzoeker (bijvoorbeeld budget) en anderzijds door de mogelijkheden en beperkingen van de techniek. Geleit kan worden voor het tegelijkertijd toepassen van verschillende technieken -kwalitatief en kwantitatief – om dezelfde vragen te beantwoorden. Dit wordt 'triangulatie' genoemd, naar de techniek waarbij landmetingen worden verricht vanuit meerdere standpunten. Een voorbeeld van triangulatief onderzoek is het naast kwantitatieve analyses van de effecten van cameratoezicht op crimineel gedrag bevragen van 'criminele jongeren' (e.g. houden jullie als je op straat bent rekening met het feit dat er cameratoezicht is?).

Gegevensverwerking

Bij het verwerken van gegevens draait het om het beantwoorden van de onderzoeksvragen: kunnen op basis van de gegevens conclusies worden getrokken over de doelmatigheid en doeltreffendheid van beleid? Punten die bij elke verwerking van gegevens centraal staan zijn:

- Hoe sterk is het gevonden causale verband (tussen bijvoorbeeld de groep die aan beleid 'onderhevig' is geweest en de controle-groep)?
- Hoe kan in een analyse de invloed van het beleidsinstrument gescheiden worden van de rivaliserende verklarende factoren?

- Hoe kunnen gegevens over de onderzochte gevallen op zodanige wijze gepresenteerd worden, dat duidelijk is welke algemene uitspraken mogelijk zijn?

De eerste stap bij gegevensverwerking is vaak het ordenen van de gegevens per variabele, bijvoorbeeld in een tabel.

Voorbeeld: Wet Ketenaansprakelijkheid (1)

Door het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het Ministerie van Justitie is in het kader van de evaluatie van de Wet Ketenaansprakelijkheid een onderzoek gedaan naar de ontwikkeling van het aantal – malafide – faillissementen. De bedoeling van dit deelonderzoek was om na te gaan of door de wet inderdaad het 'bewust niet betalen van belastingen en premies bij onderaanneming en het ter beschikking stellen van arbeidskrachten' zou worden tegengegaan. Met andere woorden, de volgende vraag stond centraal: wordt het beoogde effect door de maatregel verwezenlijkt? Het ordenen van gegevens per variabele kan in dit voorbeeld plaatsvinden door het aantal faillissementen per maand (of jaar) en/of het aantal faillissementen per bedrijfstak aan te geven. Een eventuele trendbreuk wordt op deze wijze zichtbaar. Wanneer men met betrekking tot het laatste het aantal malafide faillissementen uitdrukt in een percentage van het totaal aantal faillissementen per bedrijfstak, worden verschillen tussen bedrijfstakken al snel duidelijk.

Het werken met *percentages* veronderstelt dat de achterliggende aantallen relatief groot zijn. Over het algemeen kan aangeraden worden geen percentages te berekenen als het totaal aantal gevallen waarop het percentage is gebaseerd minder is dan circa 50. Gegevens kunnen ook nog anders gepresenteerd worden. Zo kan het *rekenkundig gemiddelde* worden berekend.

Voorbeeld: Wet Ketenaansprakelijkheid (2)

In genoemd voorbeeld kan het rekenkundig gemiddelde aangegeven worden door te zeggen hoeveel malafide faillissementen er 'gemiddeld' per bedrijfstak zijn.

Bij analyse van twee of meer variabelen gaat het niet alleen om de doelvariabele en de instrumentvariabele(n), maar wordt vooral ook gekeken naar de invloed van concurrerende variabelen.

Voorbeeld: Wet Ketenaansprakelijkheid (3)

Als naar het eerder genoemde voorbeeld van de Wet Ketenaansprakelijkheid wordt gekeken, is een belangrijke effectiviteitsvraag of sinds de invoering van deze wet het aantal (malafide) faillissementen is verminderd. In een analyseschema:

Wet ketenaansprakelijkheid → Vermindering van het aantal faillissementen

De vraag is of er in de ontwikkeling van het aantal faillissementen een trendbreuk is opgetreden. Hiervoor kan een tijdreeks worden opgesteld. Ook kunnen er andere factoren zijn dan de Wet Ketenaansprakelijkheid die van invloed zijn op de ontwikkeling van het aantal faillissementen. Het zou bijvoorbeeld kunnen zijn dat de conjuncturele ontwikkeling van invloed is geweest. In analyseschema:

Wet ketenaansprakelijkheid ←→ Vermindering van het aantal faillissementen
↑
Conjuncturele ontwikkeling

Hoe is nu de invloed van de Wet Ketenaansprakelijkheid te onderscheiden van de invloed van de conjuncturele ontwikkeling? Dit kan gedaan worden door de ontwikkeling in een bepaalde groep (doelgroep) te vergelijken met de ontwikkeling in een aanverwante groep, een controle-groep (waar de Wet Ketenaansprakelijkheid niet of minder wordt toegepast). Van beide groepen kunnen twee trendlijnen geconstrueerd worden (één van de ontwikkeling van de faillissementscijfers vóór het van kracht worden van de wet en één van de ontwikkeling van de faillissementscijfers na het van kracht worden van de wet). Beide groepen zijn aan dezelfde rivaliserende factoren, hier de conjuncturele ontwikkeling, onderhevig. Indien blijkt dat een bepaalde trend aanwezig is bij de doelgroep en niet bij de controle-groep, kan gesteld worden dat het denkbaar is dat de ontwikkeling in de doelgroep (mede) kan zijn veroorzaakt door de Wet Ketenaansprakelijkheid.

Bij het doorlopen van de derde stap kan de volgende checklist worden gehanteerd:

- Wordt duidelijk aangegeven hoe de onderzoeksgegevens worden verzameld (opzet en onderzoeksmethoden)? Berust het onderzoek bijvoorbeeld op primaire bronnen/veldwerk (systematische verzamelde gegevens direct bij doelgroep) of op secundaire bronnen/bureauwerk (literatuurstudies, enquêtes onder deskundigen)?
- Zijn de eenheden op een adequate wijze geselecteerd en verzameld, gezien het type uitspraken dat men in het onderzoek wil doen (bijvoorbeeld op willekeurige wijze getrokken en een steekproef van voldoende omvang)?
- Zijn de kenmerken van de onderzochte eenheden op een adequate wijze gemeten en op een valide en achteraf controleerbare wijze vastgesteld, d.w.z. betrouwbaar (bij herhaling hetzelfde resultaat?) en valide (is onderzocht wat beoogd was te onderzoeken)?
- Wordt zoveel als mogelijk gebruik gemaakt van onafhankelijke gegevensbronnen?
- Is de onafhankelijkheid en/of professionaliteit van de partij die het evaluatieonderzoek feitelijk uitvoert, gewaarborgd? Een onafhankelijke voorbereiding en/of uitvoering van evaluatieonderzoek of het op andere wijze waarborgen van de onafhankelijkheid waarmee de merites van – het ontwikkelen, aansturen en uitvoeren van – beleid worden beoordeeld zijn van groot belang voor de (gepercipieerde) betrouwbaarheid van evaluatieonderzoek. Als methoden om de onafhankelijkheid van beleidsevaluatie te vergroten valt te denken aan het werken met onafhankelijke begeleidingscommissies of onafhankelijke comités van toezicht.
- Integriteit: is de informatie die het evaluatieonderzoek oplevert in overeenstemming met de werkelijkheid (geen informatie ten onrechte achtergehouden of verdwenen)?

Bovenstaande criteria worden ook wel **betrouwbaarheids-/nauwkeurigheidscriteria** genoemd. De uitkomsten van de fase van gegevensverzameling zijn betrouwbaar (kom ik bij herhaling van meting tot dezelfde conclusies?) als de 'buitenwacht' ervan op aan kan dat de meetresultaten kloppen. Een voorwaarde om deze betrouwbaarheidseis te kunnen controleren, is dat de verzameling van empirische gegevens op transparante en gedocumenteerde wijze plaatsvindt. Dit vergt als gezegd de aanwezigheid van een adequaat onderzoeks dossier. In het verlengde daarvan dienen (bron)gegevens beschikbaar en toegankelijk te zijn en dient de (wijze van) analyse traceerbaar te zijn.

- Worden de met het onderzoek verkregen gegevens op de juiste wijze geanalyseerd? (= validiteit van de gevolgtrekkingen!)

3.4 De vierde stap: de rapportage en follow-up

Nadat de gegevens zijn verzameld en analyse van de gegevens heeft plaatsgevonden, is het mogelijk om conclusies te trekken en aanbevelingen te doen. De vierde stap betreft de *rapportage van de uitkomsten* van evaluatieonderzoek. De daadwerkelijke benutting van – de uitkomsten van – evaluatieonderzoek, maar ook van reguliere prestatiegegevens, geldt als één van de kritische succesfactoren voor de begroting en verantwoording nieuwe stijl. Het doen van evaluatieonderzoek en het hebben van prestatiegegevens zijn immers geen doelen op zich. Het gaat er uiteindelijk om dat de aldus gegenereerde informatie wordt *benut* en zo de kwaliteit van de besluitvormings- en verantwoordingsprocessen helpt verbeteren. Het doel van de vierde stap is zo ervoor te zorgen dat de uitkomsten van evaluatieonderzoek ex post kunnen worden benut door politieke en bestuurlijke besluitvormers.

Van belang voor benutting zijn de volgende kenmerken van de *uitkomsten* van evaluatieonderzoek ex post:

- informatieve waarde: ‘bieden de uitkomsten iets nieuws?’
- herkenbaarheid: passen de uitkomsten in het beeld dat de beleidsvoerder en anderen van het beleidsveld hebben?
- toepasbaarheid: zijn de uitkomsten (aanbevelingen) praktisch van aard?
- mate van bedreiging: in hoeverre kunnen de uitkomsten beleidsvoerders, politici of andere betrokkenen zoals uitvoerders in een lastig parket brengen?

NB *Hierbij dient opgemerkt te worden dat een goed rapport altijd bedreigend kan zijn. De verplichte melding van de resultaten aan de politieke en ambtelijke top moet er toe leiden dat er ook daadwerkelijk acties worden ondernomen naar aanleiding van de uitkomsten.*

Om de kwaliteit en daarmee de benuttingskans te verhogen, is het zaak dat een evaluatierapport aan bepaalde eisen voldoet. Veel van die eisen hangen samen met de hierboven beschreven inhoudelijke eisen, zoals een valide opzet en interne consistentie. Maar een evaluatierapport wordt uiteindelijk geschreven voor een lezerspubliek: opdrachtgever, belanghebbenden, de Kamer. Voor al deze potentiële lezers moet het rapport toegankelijk zijn.

- Het rapport bevat een adequate, beknopte management samenvatting;
- Illustratieve grafieken, modellen en afbeeldingen maken de tekst toegankelijk;
- De tekst zelf is niet te technisch (bevat zo min mogelijk jargon);
- Het rapport is niet te dik (het werken met bijlagen kan behulpzaam zijn);
- Het rapport verschijnt op tijd.

Ten slotte geldt dat de uitkomsten van evaluatieonderzoek ex post bekend moeten zijn om benut te kunnen worden. Dit betekent dat de ambtelijke en politieke top op de hoogte moeten worden gebracht van het feit dat een evaluatieonderzoek is afgerond en van de uitkomsten op hoofdlijnen.

Meer in het algemeen geldt dat de uitkomsten van evaluatieonderzoek relevant zijn voor het beantwoorden van de drie VBTB-vragen uit de begrotingen en jaarverslagen. Met name de vraag 'wat hebben we dankzij het beleid bereikt' kan meestal niet met reguliere prestatiegegevens worden beantwoord. De uitkomsten van evaluatieonderzoek maken deel uit van de begrotingscyclus en worden om die reden opgenomen in de departementale begroting en, vooral, jaarverslagen. Als het goed is, zijn de begrotingscyclus en de beleids- en bedrijfsvoeringscycli aan elkaar gekoppeld (bijvoorbeeld door een op de begroting aansluitende jaarplancycclus). Langs die weg krijgt de ambtelijke en politieke top vaak al 'automatisch' inzicht in de uitkomsten van evaluatieonderzoek ex post. Echter, aangezien de doorlooptijden van deze cycli soms niet goed op elkaar aansluiten en de uitkomsten van evaluatieonderzoek op de vaste momenten van de cyclus al weer enkele maanden oud kunnen zijn, schrijft de RPE voor dat de ambtelijke en politieke top volgens een *standaardprocedure* op de hoogte dient te worden gebracht van de uitkomsten van evaluatieonderzoek (bepaling 6.1). De standaardprocedure wordt door de departementen zelf bepaald en vastgelegd.

Sinds 2003 worden de uitkomsten van evaluatieonderzoek ook ontsloten met behulp van het Evaluatieoverzicht Rijksoverheid. Dit is een integraal overzicht van de toepassing van evaluatieonderzoek en prestatiegegevens binnen de rijksoverheid, dat via het rijksbrede intranet RYX wordt gevuld en kan worden ingezien. Met behulp van handige zoekfuncties en sleutels (zoals de doelstellingen van beleid, begrotingsartikelen), is zeer snel over elk denkbaar onderwerp informatie op te vragen over afgeronde, maar ook lopende en geplande evaluatieonderzoeken.

Benutting: een kwestie van cultuur?

Naast het onderzoek en het rapport zelf en de wijze waarop e.e.a. samenhangt met de begrotingscyclus, hangt veel af van wat wel de 'cultuur' van een organisatie genoemd wordt. Zo zijn de volgende kenmerken van *gebruiker(s)* van de uitkomsten van evaluatieonderzoek van belang:

- Resultaatgerichtheid in denken en doen: is men bereid om open te zijn over ambities en rekenschap te geven over geleverde prestaties en bereikte effecten;
- Houding ten opzichte van leren: staat men open voor 'kritiek', streeft men naar verbetering;
- Belangstelling: bestaat er interesse voor wat achter de 'waan van de dag' schuilgaat.

Onderstaand zijn een paar voorbeelden opgenomen van benutting van uitkomsten van evaluatieonderzoek.

Uit een persbericht van de ministerraad d.d. 4 oktober 2002:

'De ministerraad heeft op voorstel van staatssecretaris Odink van Landbouw, Natuurbeheer en Visserij ingestemd met een tijdelijke aanpassing van enkele mestregels. De aanpassing vloeit voort uit de Evaluatie Meststoffenwet 2002. Deze evaluatie is in april van dit jaar naar de Tweede Kamer gezonden. Het toenmalige, demissionaire kabinet verbond daaraan geen conclusies meer. Het kabinet heeft nu besloten de aanscherping van enkele verliesnormen in het mineralen-aangiftesysteem voor 2003 met één à twee jaar te temporiseren. Daarnaast zijn maatregelen genomen, zodat minder melkveehouders onnodig contracten hoeven af te sluiten voor mest die ze niet daadwerkelijk hoeven af te voeren. Het kabinet is van mening dat deze aanpassingen verantwoord zijn, gelet op de te bereiken milieudoelen en de inmiddels behaalde milieuresultaten. De voorgestelde wijzigingen maken voor de agrarische sector de gestelde doelen op termijn beter haalbaar'.

Uit NRC Handelsblad d.d. 8 oktober 2002:

'Legalisering van prostitutie mislukt: de prostitutiebranche is twee jaar na afschaffing van het 'bordeelverbod' nog ver verwijderd van de status van reguliere bedrijfstak. Belastingdienst en Arbeidsinspectie tonen nauwelijks belangstelling voor prostituees. Gemeenten lopen achter bij de afgifte van vergunningen en naleving van voorschriften. Banken en andere zakelijke dienstverleners zijn huiverig om zaken te doen met de branche. Minderjarigheid en illegaliteit zijn nog steeds een veelvuldig voorkomend verschijnsel. Dat concludeert het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het ministerie van Justitie in het evaluatieonderzoek 'Het bordeelverbod opgeheven' dat vanochtend is gepubliceerd. Daarbij werden enkele honderden prostituees, exploitanten, politiemensen en andere sleutelfiguren geïnterviewd. De Tweede Kamer had bij afschaffing van het Bordeelverbod om die evaluatie gevraagd om in kaart te brengen of de situatie op het gebied van uitbuiting en vrouwenhandel en de controle daarop zijn verbeterd. Zonder landelijke normen voor minimale, overal geldende handhavingscriteria zal het niet mogelijk zijn de doelstellingen van de wetwijziging van twee jaar geleden in de praktijk te realiseren, is dan ook een van de belangrijkste aanbevelingen in de evaluatie'.

Bij het doorlopen van de laatste stap kan de volgende checklist worden gehanteerd:

Bevindingen/conclusies

- Worden de onderzoeksprobleemstelling en centrale vraagstelling beantwoord door de conclusies en aanbevelingen?

- Sluiten de bevindingen, conclusies en eventuele aanbevelingen op elkaar aan?

(Is er sprake van interne consistentie?)

In de vakliteratuur worden bovenstaande criteria ook wel *validiteitscriteria* genoemd.

Validiteit (meet ik wat ik wil meten?) heeft betrekking op de algemene geldigheid van een evaluatieonderzoek.

Gebruiksdoel

- Is het gebruiksdoel van het evaluatieonderzoek omschreven (als onderdeel van de doelstelling van het onderzoek) en sluit het aan op de informatiebehoeften van de ambtelijke en politieke besluitvormers?
- Sluiten de opdrachtformulering en de onderzoeksvragen aan bij het gebruiksdoel van de evaluatie (d.w.z. dat de wijze waarop de opdracht tot het onderzoek (inclusief gegevensverzameling, gegevensanalyse, etc.) is geformuleerd 'logisch' voortvloeit uit het doel waarvoor het onderzoek wordt verricht)?
- Sluiten de uitkomsten en conclusies/aanbevelingen aan bij het gebruiksdoel van de evaluatie (d.w.z. zijn de conclusies/aanbevelingen toegesneden op het doel waarvoor het onderzoek wordt uitgevoerd)?
- Zijn de uitkomsten van het evaluatieonderzoek tijdig beschikbaar: sluiten ze aan op de – aan de begroting en jaarverslag gekoppelde – cycli van bestuurlijk-politieke besluitvorming?

Presentatie

- Worden de uitkomsten van het evaluatieonderzoek op een heldere en toegankelijke wijze gepresenteerd? (is het evaluatierapport goed leesbaar?)
- Zijn er beleidsaanbevelingen gedaan?
- Zijn de aanbevelingen concreet genoeg geformuleerd en technisch uitvoerbaar? Houden de aanbevelingen rekening met de politiek/maatschappelijke/financiële haalbaarheid?
- Worden de budgettaire gevolgen van de aanbevelingen aangegeven?
- Bevat het evaluatierapport een samenvatting en is deze representatief van aard? (geeft de samenvatting de hoofdlijnen van het rapport op evenwichtige wijze weer?)

In de vakliteratuur worden bovenstaande criteria ook wel *bruikbaarheidscriteria* genoemd. De validiteit, betrouwbaarheid en nauwkeurigheid van evaluatieonderzoek zijn in belangrijke mate bepalend voor de uiteindelijke bruikbaarheid van de uitkomsten ervan.

AANVULLENDE VBTB-CRITERIA :

- Worden de uitkomsten van evaluatieonderzoek in de begroting opgenomen (in de toelichting van het betreffende beleidsartikel)? Wordt daarbij ingegaan op de mogelijke verklaring van eventuele verschillen tussen beoogde en behaalde resultaten van beleid?

Een positief gebruik van evaluatieonderzoek ex post is mede afhankelijk van dingen die je als onderzoeker of opdrachtgever zelf in de hand hebt. Deze factoren raken enerzijds de politiek en het beleid en anderzijds het evaluatieonderzoek. Factoren die bijdragen aan beter gebruik van evaluatieonderzoek ex post zijn (onder meer):

- Een nauwkeurige beschrijving van het probleem dat wordt onderzocht; indien een vraagstelling van één van de belanghebbenden wordt verwaarloosd kan dat een negatieve invloed hebben op het gebruik van evaluatieonderzoek;
- Gedurende het gehele evaluatieonderzoek moet tussen de verschillende partijen een open communicatie bestaan en dient er betrokkenheid te zijn. Politici en beleidsvoerders moeten open staan voor onderzoek en onderzoekers voor de problemen van de politiek en het beleid;
- Duidelijkheid voor politici, beleidsvoerders en onderzoekers over de beperkingen en grenzen van een evaluatieonderzoek;
- De inbedding van evaluatieonderzoek in besluitvormingsprocedures die vantevoren vastliggen, waardoor gebruik moet worden gemaakt van de onderzoeksresultaten;
- De instelling van een begeleidingscommissie die er onder andere zorg voor draagt dat met de hier genoemde factoren en de eerder genoemde kwaliteitscriteria rekening wordt gehouden.

4 Begrippenlijst

Algemene beleidsdoelstelling = einddoel = beoogd finaal effect

Effect dat men met behulp van beleid in de samenleving of overheidsorganisatie wil bereiken.

Beleidsinformatie

Gestructureerde en nader toegelichte gegevens over (1) de doelstelling(en) van beleid, (2) de beoogde en/of gerealiseerde effecten van beleid, (3) de daartoe te leveren en/of geleverde prestaties en (4) de daarmee gemoeide kosten.

Beleidsinstrumenten

Instrumenten – zoals subsidies, voorlichting en regelgeving – die worden ingezet om bepaalde (beoogde) doelstellingen te realiseren.

Doelbereiking

De mate waarin een – algemene en/of nader geoperationaliseerde – doelstelling van beleid, al dan niet dankzij de inzet van beleidsinstrumenten, wordt gerealiseerd (uitgedrukt in beoogde effecten).

Doelmatigheid van bedrijfsvoering

De kosten en kwaliteit van de prestaties, die in het kader van de ontwikkeling, aansturing en uitvoering van beleid worden geleverd.

Doelmatigheid van beleid = kosteneffectiviteit

De beoogde effecten van het beleid gerelateerd aan de ingezette middelen. Meer effecten met dezelfde middelen of dezelfde effecten met minder middelen.

Doeltreffendheid van beleid

De mate waarin een – algemene en/of nader geoperationaliseerde – doelstelling van beleid dankzij de ingezette beleidsinstrumenten wordt gerealiseerd.

Effecten (of finale effecten)

De uiteindelijk beoogde en/of gerealiseerde gevolgen van beleid in maatschappij en/of overheidsorganisatie.

Evaluatieonderzoek ex ante

Het op een systematische en methodologisch verantwoorde wijze in kaart brengen van de verwachte kosten en baten van beleidsalternatieven (kosten en baten in termen van respectievelijk inzet van middelen en verwachte positieve en negatieve effecten in de maatschappij of overheidsorganisatie).

Evaluatieonderzoek ex post

Onderzoek waarin – gedurende de uitvoering van beleid of achteraf- beleid en/of bedrijfsvoering wordt getoetst en beoordeeld op doelbereiking, doeltreffendheid en doelmatigheid óf: het op een systematische en methodologisch verantwoorde wijze onderzoeken en beoordelen van (1) de effecten van bestaand beleid in de samenleving of overheidsorganisatie, (2) de wijze waarop beleid wordt uitgevoerd en/of (3) de kosten en kwaliteit van – in het kader van bestaand beleid – geleverde producten en diensten.

Externe factoren

Factoren die op het einddoel in werken, maar niet door het beleid te beïnvloeden zijn.

Intermediaire effecten

De tussengelegen, stuurbare effecten die een bijdrage leveren aan de realisatie van de (finale) effecten.

Kwalitatief evaluatieonderzoek = constructivistisch evaluatieonderzoek

Evaluatieonderzoek waarbij de nadruk ligt op het achterhalen en interpreteren van overtuigingen, beweegredenen en waardeoordelen van bij beleid betrokken actoren.

Kwantitatief evaluatieonderzoek = rationeel-analytisch evaluatieonderzoek

Evaluatieonderzoek dat zich primair richt op kwantificeerbare beleidsinformatie (e.g. gegevens over de kosten en kwaliteit van in het kader van beleid geleverde prestaties, de aan beleid causaal toe te schrijven effecten van beleid).

Nader geoperationaliseerd doel = tussendoel

Beoogd effect dat een bijdrage levert aan of een indicatie geeft van het bereiken van het beoogd finaal effect.

Neveneffecten

Gevolgen in maatschappij of overheidsorganisatie die weliswaar voortvloeien uit de inzet van beleidsinstrumenten, maar die niet primair zijn beoogd.

Verifiërend evaluatieonderzoek

Onderzoek gericht op het vaststellen of gekozen effect-indicatoren of intermediaire effecten een adequate afspiegeling bieden van – c.q. adequate tussengelegen effecten benoemen voor – de nagestreefde finale effecten van beleid.

5 Een overzicht van aanbevolen literatuur

Beleidsevaluatie, J.Th. A. Bressers en A. Hoogerwerf, Samsom H. D. Tjeenk Willink, Alphen aan den Rijn, 1995

Beleidsonderzoek, het ontwikkelen en beoordelen van beleidsmaatregelen en -projecten, Ministerie van Financiën, Sdu Uitgeverij, Den Haag, 1993 (3e druk)

Evaluatiemethoden, een introductie, Ministerie van Financiën, Sdu Uitgeverij, Den Haag, 1992 (4e druk)

Evaluating socio-economic programmes; evaluation design and management, volume 1 MEANS collection, European Commission, European Communities, 1999.

Evalueren, P.G. Swanborn, Boom, Amsterdam, 1999

Fourth Generation Evaluation, E.S. Guba en Y.S. Lincoln, Sage, Newbury Park, 1989

Heeft beleid effect?- Enkele benaderingen voor de evaluatie van bestaand beleid, Afdeling Beleidsevaluatie en -Instrumentatie, Ministerie van Financiën, 1999

International Atlas of Evaluation, Jan-Eric Furubo, Ray C. Rist, Rolf Sandahl, Transaction Publishers, New Brunswick/New Jersey, 2002

Policy evaluation, a guide for managers, HMSO Publications Centre, 1988

Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid, Ministerie van Financiën, Drukkerij Van Deventer, 's-Gravenzande, 2002

Relevante websites

www.minfin.nl/vbtb

www.oecd.org/puma

www.beleidsonderzoek.nl

www.eval.org

www.worldbank.org/oed

www.europeanevaluation.org

http://europa.eu.int/comm/budget/evaluation/index_en.htm

Bijlage 1 Overzicht stappen

Stap 1 Het formuleren van de probleemstelling en vragen van het onderzoek

- Tussenstap 1A* het expliciteren van verwachtingen
- Tussenstap 1B* het 'richten' van een evaluatieonderzoek ex post
- Tussenstap 1C* het vertalen van een beleidsvraag/-doelstelling naar een doelstelling en probleemstelling van evaluatieonderzoek ex post
- Tussenstap 1D* het uitwerken van de onderzoeksdoelstelling en -probleemstelling in onderzoeksvragen
- Tussenstap 1E* het uitwerken van een opdrachtformulering

Stap 2 De keuze voor een onderzoeksopzet

Stap 3 Het hoofdonderzoek: gegevensverzameling en -verwerking

Stap 4 De rapportage en follow-up

- Tussenstap 4A* aandacht voor de uitkomsten zelf
- Tussenstap 4B* het opstellen van een evaluatierapport ('...een goede rapportage vergroot de kans op benutting...')
- Tussenstap 4C* het terugkoppelen van de uitkomsten van het evaluatieonderzoek ex post

Bijlage 2 Stappen in proces van uitbesteding van evaluatieonderzoek

1 **Vorbereiding en probleemstelling**

Het is aan te bevelen vooraf een onderzoeksplan op te stellen op basis waarvan offertes moeten worden uitgebracht. Belangrijk in het kader van deze handreiking is dat de opdrachtformulering aansluit op de onderzoeksprobleemstelling/-doelstelling en de onderzoeksvragen en op het gebruiksdoel van het evaluatieonderzoek. De wijze waarop de opdracht tot het onderzoek is geformuleerd moet immers ‘logisch’ voortvloeien uit het doel van het onderzoek.

Het is van belang dat alle bij het evaluatieonderzoek betrokken actoren de probleemstelling, doelstelling en onderzoeksvragen onderschrijven. Dit om discussie daarover in de loop van het evaluatieonderzoek te vermijden.

2 **Offertes vragen**

Hierbij gaat het om de selectie van te benaderen bureaus en de aanschrijving van geselecteerde bureaus die offerte mogen uitbrengen. De bureaus moeten meestal een periode van enkele weken worden gegund om een offerte op te stellen.

Globaal gesproken zijn drie vormen van uitbesteding van evaluatieonderzoek te onderscheiden:

- a De eerste vorm is dat *een beperkt aantal bureaus* gericht wordt gevraagd een offerte uit te brengen. Dit is aan te bevelen als men een aantal bureaus op het oog heeft, maar niet weet welk bureau het meest geschikt is.
Tip: drie bureaus is een mooi aantal! Bij de voorselectie van bureaus moet worden gekeken naar aanwezige deskundigheid en ervaring, het liefst blijkend uit eerdere onderzoeksrapporten;
- b De tweede vorm is dat *één bureau* wordt benaderd. Verwacht wordt dat dit bureau als enige gespecialiseerd en deskundig is om het onderzoek te verrichten. Zo'n bureau wordt gevraagd een offerte uit te brengen;
- c De derde vorm is de *openbare aanbesteding*. Men heeft geen idee welke onderzoeksbureaus het onderzoek aan zouden kunnen. Via een advertentie worden alle bureaus die interesse hebben, uitgenodigd hun belangstelling kenbaar te maken. Daarna volgt de voorselectie van bureaus.

NB *Als de geraamde waarde van een mogelijke opdracht voor het uitvoeren van een evaluatieonderzoek het drempelbedrag⁵ overschrijdt, dan moet er, vallend onder de richtlijn Diensten, Europees worden aanbesteed. Verder wordt verwezen naar de site van het Ministerie van Economische Zaken (www.ez.nl) en andere handleidingen voor openbare aanbesteding.*

3 Offertes beoordelen

Hierbij wordt een stramien aangegeven op basis waarvan de prijs/kwaliteit-verhouding van de offerte kan worden beoordeeld en wordt aangegeven hoe het stramien toegepast moet worden.

NB *Onderstaande criteria helpen bij de ordening van de kwaliteit van de offertes en de bureaus.*

Een volgende stap is dat de offertes door de opdrachtgever en de begeleidingscommissie beoordeeld worden. Op onder andere de volgende aspecten kan beoordeeld worden:

- de uitwerking van de probleemstelling en de operationalisering van de onderzoeksvragen;
- de wijze waarop de gegevens worden verzameld en een eventuele steekproef wordt getrokken;
- de wijze waarop men de verzamelde gegevens zal analyseren;
- de wijze waarop bestaande prestatiegegevens en evaluaties worden gebruikt;
- de deskundigheid van het bureau en de onderzoekers;
- de wijze waarop men (tussentijds) zal rapporteren;
- de wijze waarop de verzamelde data ter beschikking van de opdrachtgever worden gesteld (*tip: maak afspraken over het vrijgeven van het bronnenmateriaal! De aanwezigheid hiervan wordt voorgeschreven!*);
- de planning en de begroting van het onderzoek.

Door vergelijking van de offertes op de verschillende criteria kan tenslotte een keuze worden gemaakt.

4 Opdrachtverlening

Hierbij gaat het om een beargumenteerd voorstel (aan gemandateerde) over welk bureau het evaluatieonderzoek moet gaan uitvoeren. Een opdrachtbrief wordt opgesteld met precieze opdrachtformuleringen en voorwaarden.

5 2003: € 162.293,- voor diensten

5 Bewaking en begeleiding van het evaluatieonderzoek

Op basis van periodiek overleg en tussenproducten kan tussentijds bijgestuurd worden. Voor een goede voortgangscategorie is het zaak dat opdrachtgever en opdrachtnemer regelmatig contact hebben. Dit ook om het onderzoeksproces tijdig te kunnen bijsturen. Afhankelijk van de duur van het onderzoek zal dit maandelijks of zelfs wekelijks moeten plaatsvinden. Als vuistregel geldt dat – na de startbijeenkomst – op zo’n drie momenten overleg wenselijk is. Het is verstandig af te spreken op welke facetten wanneer en in welke vorm wordt teruggekoppeld.

6 Beoordeling en acceptatie eindrapport

De opdrachtgever beoordeelt (en accepteert) de kwaliteit van het aangeboden evaluatierapport.

NB In de praktijk wordt hiernaar toegewerkt.

7 Gebruik van resultaten/follow-up

Op basis van het evaluatierapport kunnen conclusies worden getrokken t.a.v. het nut of de noodzaak van beleidswijzigingen.

Bijlage 3 Terms of Reference (ToR)

Standaard onderzoeksopzet

- 1 Voorwaardenscheppend kader
- 2 Bereik van de evaluatie
- 3 Wie gaat de evaluatie gebruiken? En waarvoor?
- 4 Onderzoeksvragen en criteria
- 5 Kennis en beschikbare gegevens
- 6 Methode (richtlijnen)
- 7 Onderzoekscontext
- 8 Planning
- 9 Budget
- 10 Vereiste kwalificaties
- 11 Structuur van het voorstel

Voor meer informatie over de ToR wordt verwezen naar:

Evaluating socio-economic programmes; evaluation design and management, volume 1 MEANS collection, European Commission, European Communities, 1999.

Colofon

Ministerie van Financiën

Directie Begrotingszaken

Postbus 20201

2500 EE Den Haag

Internet www.minfin.nl/vbtb

Oplage 1000

november 2003

022okt2003

vormgeving Studio Tint, Den Haag

druk Drukkerij van Deventer, 's-Gravenzande



Postbus 20201
2500 EE Den Haag
Telefoon (070) 342 75 42
Internet www.minfin.nl

november 2003

Voorwaardenscheppend kader

Wat is een "voorwaardenscheppend kader"? Een terechte vraag. In wezen wordt het "formele kader" bedoeld: wat is de formele context van het onderzoek?

Het gaat daarbij om drie zaken. Het gaat daarbij in de eerste plaats om de **formele aanleiding** van het onderzoek, zoals een aankondiging in een wet of – meer algemeen – de bepaling uit de RPE dat alle beleid in beginsel minimaal eens in de vijf jaar moet worden geëvalueerd. In de tweede plaats moet duidelijk zijn wie de **opdrachtgever** is: wie initieert het onderzoek, wie formuleert de opdracht, wie stelt de onderzoeksvoorwaarden op. Ten slotte moet duidelijkheid bestaan over het **algemene doel van de evaluatie**: gaat het bijvoorbeeld om het afleggen van verantwoording of dient er een knelpunt in de uitvoering.

Behorende bij de uitgave:

Handreiking Evaluatieonderzoek ex post, november 2003. Ministerie van Financiën
Bijlage 3 pagina 52