

ARTIKELEN

Veiligheidsbeleid: onderbouwd en effectief?

De meerwaarde van beleidstheorieën voor beleid en beleidsevaluatie *

Peter van der Knaap

Dit themanummer richt zich op de effecten van het sociaal en fysiek veiligheidsbeleid. Over het normatieve uitgangspunt dat beleid effectief moet zijn en dat de positieve effecten groter moeten zijn dan de negatieve effecten, bestaat wel consensus. Maar hoe kan de kans op effectief beleid vooraf worden vergroot? Wanneer is beleid goed onderbouwd en wat telt als bewijs of 'evidence' dat beleid zal kunnen werken of heeft gewerkt? Welke beleidsinformatie hebben we daarbij nodig?

Centraal in dit artikel staat de bijdrage van beleidstheorieën aan het ontwikkelen en evalueren van beleid en het belang van kwalitatief voldoende beleidsinformatie daarbij. In het artikel wordt ingegaan op recent onderzoek van de Algemene Rekenkamer naar (1) de onderbouwing van en (2) de beschikbare informatie over het veiligheidsbeleid van het derde en vierde kabinet-Balkenende. Betoogd wordt dat een theoretische onderbouwing van beleidsmaatregelen en een evaluatief gebruik van doelstellingen en indicatoren kunnen bijdragen aan beleid dat niet alleen effectief is, maar waarbinnen ook geleerd kan worden.

1 Beleidstheorieën: rationaliteit en evidence als basis bij beleidsontwikkeling

Beleid en bestuur kennen vele 'als, dan'-redeneringen. Als we X doen, dan leidt dat tot Y. Als we meer toezicht houden, dan neemt de kans op ongewenst gedrag af. Als we de dijken hoger maken, dan overstroomt het land minder snel. Als we de taximarkt liberaliseren, dan zal de verhouding tussen prijs en kwaliteit verbeteren. De voorbeelden zijn legio. Deze redeneringen vormen elk een 'beleidstheorie': een stelsel van veronderstellingen en verwachtingen over het verband tussen de inzet van beleidsinstrumenten en het bereiken van doelstellingen (vergelijk Leeuw 1983; Hoogerwerf 1984).

Vanuit een rationeel perspectief op beleid biedt een beleidstheorie een bruikbaar hulpmiddel voor het ontwikkelen van beleid. Zo kan een beleidstheorie zorgen

* Dr. P. van der Knaap is directeur doelmatigheidsonderzoek bij de Algemene Rekenkamer. Dit artikel schreef hij op persoonlijke titel. Hij bedankt Rudi Turksema, Marieke Gorree, Erik Israël, Mariëlle van den Dongen, Jooske de Sonnaville en Jeroen Kerseboom voor hun bijdragen aan dit artikel.

voor focus in het vergelijken van alternatieve maatregelen: welke aannamen gelden bij elk van de opties? Welk alternatief biedt daarbinnen de beste kans dat een doelstelling binnen de geldende randvoorwaarden gehaald wordt? Welke informatie is beschikbaar over de kans dat interventie X zal kunnen leiden tot uitkomst Y? Pawson (2002) vat die richtinggevende werking van een beleidstheorie mooi samen: 'A theory tells us where to look for evidence, theories drive learning. Evaluation tests programmes, programmes are theories, evaluation is policy-testing.'

Vooraf uitgevoerd (*ex ante*) evaluatieonderzoek heeft een besluitvormingonderbouwende functie. Het doel is de kwaliteit van de besluitvorming over te voeren beleid te verbeteren door een gerichte inbreng van informatie over de verwachte geschiktheid en (kosten)effectiviteit van beleidsalternatieven. Om de beste maatregelen te kunnen kiezen, is immers vooraf inzicht nodig in de vraag welke interventies het hoogste rendement kunnen opleveren tegen welke risico's. Op basis van *ex ante* evaluatie kan het beste alternatief worden geselecteerd: het alternatief waarbij de verhouding tussen verwachte (maatschappelijke) kosten en verwachte (maatschappelijke) baten het gunstigste is (zie Ministerie van Financiën 2003a).

Als het gaat om dit vooraf vergelijken en waarderen van beleidsalternatieven, kent Nederland een lange traditie (Van Hoesel e.a. 2005). Ook het ministerie van Justitie heeft in het verleden waardevolle initiatieven ondernomen. Het werk van het – inmiddels uitgeplaatste – Expertisecentrum Rechtspleging en Rechtshandhaving (bijvoorbeeld de *ex ante* handhaaf- en uitvoerbaarheidstoets en de Tafel van Elf) en van het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) zijn daarvoor illustratief.¹ De Algemene Rekenkamer hanteert als norm dat de overheid bij de totstandkoming van beleid niet alleen nagaat of het beleid uitvoerbaar en handhaafbaar is, maar ook of – en zo ja, in welke mate – het beleid kán bijdragen aan het bereiken van de doelstellingen: 'Daartoe expliciteren organisaties de beleidstheorie waarop het beleid gebaseerd is en toetsen ze de veronderstellingen over de uitvoering van het beleid' (Algemene Rekenkamer 2005b). Als het goed is, maakt de overheid vanuit de verwachte effectiviteit een weloverwogen keuze voor de in te zetten beleidsinstrumenten (*ibid.*).

Internationaal gezien heeft de rationele benadering van beleid sinds de jaren tachtig een stevige impuls gekregen door toepassing van methoden uit de school van het *New Public Management* (NPM), zoals *logic modeling*. In ons land is de operatie Van Beleidsbegroting tot Beleidsverantwoording (VBTB) van grote invloed geweest. De drie vragen die daarbij centraal staan – wat willen we bereiken, wat gaan we daarvoor doen en wat mag het kosten – leiden als het ware vanzelf tot een beleidstheorie. Doelstellingen en indicatoren moeten daarbij zo concreet mogelijk zijn. Daarnaast worden ook andere kwaliteitseisen aan beleidsinformatie gesteld (zie par. 4).

1 Het Expertisecentrum Rechtspleging en Rechtshandhaving (ERR) van het ministerie van Justitie is vanaf 1 november 2009 verder gegaan als de afdeling Nalevingsexpertise van het Centrum voor Criminaliteitspreventie en Veiligheid (CCV).

Het laatste decennium is daaraan toegevoegd dat beleid niet alleen in theorie mag werken, maar *evidence-based* moet zijn. Paradoxaal genoeg wordt ook van nieuwe maatregelen verwacht dat er 'bewijs' is dat ze het gewenste resultaat zullen hebben.² Strikt genomen zijn het dan ook de resultaten van wetenschappelijk onderzoek die in een *evidence-based* beleidsontwikkeling de vorm en inhoud van beleidsmaatregelen zouden mogen bepalen. In Nederland, maar zeker in het Verenigd Koninkrijk en de Verenigde Staten, gaat *evidence-based* beleid en de roep om meer kwantitatief, experimenteel onderzoek vaak hand in hand. In de Verenigde Staten werkt de *Council for Excellence in Government* bijvoorbeeld actief aan het promoten van zuiver experimenteel, kwantitatief beleidsonderzoek (zie www.excelgov.org/evidence). Ook in Nederland klinkt vanuit onder meer het ministerie van Financiën de roep om een bredere toepassing van experimenteel onderzoek. In met name het onderwijsbeleid en het criminaliteitsbeleid is sprake van een toename van op beleidstheorieën toegesneden, experimenteel kwantitatief beleidsonderzoek (Van Hoesel e.a. 2005).

2 Beleidsevaluatie: feedback, verantwoording en leren

Naast beleidsevaluatie *ex ante*, draagt de Nederlandse rijksoverheid ook achteraf uitgevoerde beleidsevaluaties hoog in het vaandel (beleidsevaluatie *ex post*; zie bijvoorbeeld Van der Knaap 2000 en Leeuw 2009). Niet lang geleden schreef de minister van Financiën in de Miljoenennota daarover:

'Het moment om te bezien wat er van de kabinetsdoelstellingen terecht is gekomen, is het Jaarverslag. Daarnaast zal – zowel vooraf als achteraf met evaluatieonderzoek – op zijn minst aannemelijk moeten worden gemaakt, maar bij voorkeur aangetoond, dat de acties van de overheid bij zullen dragen c.q. hebben bijgedragen aan het politiek geformuleerde doel in de beleidsprogramma's en de beleidsbegroting.' (Ministerie van Financiën 2003c)

Het begrip 'beleidstheorie' is in ons land al langere tijd verbonden aan evaluatieonderzoek *ex post* (zie bijvoorbeeld Leeuw 1983). Onderzoeken of beleid effect heeft, geldt van oudsher als de centrale opdracht van de beleidsevaluatie. Het gaat vaak om het analyseren van de mate waarin beleid heeft geresulteerd in het realiseren van de beoogde beleidsdoelstellingen. De hamvragen zijn in eerste instantie geënt op de beleidstheorie: klopten de veronderstellingen? Werkte het beleid? Leidde het verhogen van de dijken bijvoorbeeld inderdaad tot minder overstromingen?

Evaluatieonderzoek *ex post* heeft daarmee in eerste instantie een feedbackfunctie: de inzet van instrument X leidt in die en die mate tot het bereiken van doelstel-

2 De staatssecretaris van Onderwijs sprak bijvoorbeeld daarover op 29 november 2007 op de Conferentie over onderwijssociologie 'Kennis en beroepsonderwijs': 'Het is belangrijk om nieuwsgierig te zijn, te willen weten wát beter kan en hóe het beter kan, maar: scholen zijn geen laboratoria waar ruimte is voor mislukte experimenten. Daarom moet je evidence-based te werk gaan.'

ling Y. Deze functie van beleidsevaluatie past naadloos binnen het rationele beleidsmodel. Volgens dit model moet immers actief worden gestuurd op prestaties en de effecten van die prestaties, en moet daar achteraf ook verantwoording over worden afgelegd. Naast informatie die afkomstig is uit monitoringsystemen (bijvoorbeeld het aantal jeugdige delinquenten dat wacht op behandeling c.q. is behandeld in inrichtingen), wordt met behulp van evaluatieonderzoek gezien hoe doeltreffend de behandelwijzen zijn geweest in relatie tot het terugdringen van crimineel gedrag. Het parlement verkrijgt volgens dit model inzicht in de doeltreffendheid en doelmatigheid van de bestedingen en kan – als ook de rechtmatigheid op orde is – decharge verlenen. De verantwoordingscyclus is rond wanneer de beleidsverantwoordelijke instantie zich heeft verantwoord over inzet van middelen en de daarmee bereikte resultaten en eventueel beterschap heeft beloofd ten aanzien van ‘verbetermogelijkheden’.

Daarmee kom ik bij de tweede functie die beleidsevaluatie van oudsher wordt toebedacht: het leveren van een bijdrage aan een lerende overheid. Op basis van de uitkomsten van evaluaties kunnen immers verbeteringen worden aangebracht, waardoor beleid doelmatiger en doeltreffender kan worden uitgevoerd. Het antwoord op de vraag in welke mate overheidsbeleid het gewenste resultaat heeft opgeleverd, kan daarbij aanleiding zijn om terug te keren tot de oorspronkelijk gehanteerde veronderstellingen. Perfect uitgevoerd, maar op verkeerde assumpties gestoeld beleid zal vaak immers weinig of zelfs niet effectief zijn. Een tweede oogmerk van theoriegerichte beleidsevaluatie is dan ook een beter begrip van de achterliggende mechanismen. De vragen die dan gesteld moeten worden, zijn: *Waarom* klopten de veronderstellingen wel of niet? *Waarom* werkte het beleid (niet)? *Waarom* leidde het verleggen van dijken (niet) tot minder (kans op) overstromingen? Welke andere effecten traden op en hoe kunnen we die waarderen?

De antwoorden op deze vragen kunnen verstrekkende gevolgen voor het beleid hebben. Als immers aangetoond is dat beleid niet werkt, moeten de veronderstellingen waarop beleid stoelt worden herzien (vergelijk Van der Knaap 2006). Is het nog steeds een goede doelstelling om helemaal geen overstromingen te hebben of moeten rivieren juist meer ruimte hebben om hoogwater te kunnen afvoeren?

Het is de bedoeling dat al het beleid van de rijksoverheid periodiek wordt geëvalueerd (Ministerie van Financiën 2006). Voor beleid dat met doelstellingen in de rijksbegroting is opgenomen, gebeurt dat met behulp van ‘beleidsdoorlichtingen’ (ibid.). Kenmerkend voor een beleidsdoorlichting is de systematische vraagstelling naar nut en noodzaak van het beleid: expliciet wordt aandacht besteed aan de probleemanalyse, de actualiteit van de doelstelling en de argumentatie voor de rol van de rijksoverheid (waarom moet de rijksoverheid dit oplossen en niet de markt of een andere overheid). Vragen die bij elke doorlichting worden gesteld zijn:

- Welke doelstelling heeft de overheid geformuleerd voor de oplossing van het probleem?
- Welke instrumenten (zullen) worden ingezet?
- Hoe is de samenhang tussen de instrumenten?
- Is er sprake van overlap?

- Wat is bekend over de uitvoering van beleid?
- Dragen de instrumenten bij aan de geformuleerde doelstellingen (oplossing van het probleem)?
- Wat zijn belangrijke positieve en negatieve neveneffecten?

De beleidsdoorlichtingen zijn door het ministerie van Financiën in 2006 ingevoerd. Sindsdien zijn er enkele tientallen uitgevoerd, zoals Diversiteitbeleid bij het Rijk, Energiemarkt, Handelspolitiek, Slachtofferzorg, Re-integratiebeleid, Benutting binnenvaart en vaarwegen en de Nota Ruimte. De ministeries voeren de beleidsdoorlichtingen zelf uit. Het ministerie van Financiën is verantwoordelijk voor de coördinatie. De uitkomsten van eerder uitgevoerde evaluaties – bijvoorbeeld van afzonderlijke regelingen of instrumenten – worden in een beleidsdoorlichting meegenomen (vergelijk Van der Knaap 2009).

Ook de Algemene Rekenkamer kijkt in haar doelmatigheidsonderzoek specifiek naar de mate waarin politiek-bestuurlijke verwachtingen over de werking van beleid uitkomen en naar verklaringen voor tegenvallende resultaten (Algemene Rekenkamer 2005a). Zo heeft de Rekenkamer recentelijk onderzocht of het bieden van meer vrijheid bij de besteding van specifieke uitkeringen inderdaad (zoals de beleidstheorie wil) leidt tot meer maatwerk, integrale beleidsvoering en slagvaardigheid bij decentrale overheden en welke verklaringen er zijn voor succes en falen (Algemene Rekenkamer 2005c). Naast de Rekenkamer gaan ook onderzoeksorganisaties als het WODC en het Sociaal en Cultureel Planbureau (SCP) bij hun evaluaties frequent uit van de gehanteerde beleidstheorie. De huidige directeur van het WODC noemt het gegeven dat in de meeste onderzoeken van dat centrum ruim aandacht is voor de theorie achter de te onderzoeken maatregel als een van de drie elementen die bepalend zijn voor de mate waarin onderzoek benut wordt (Klein Haarhuis e.a. 2009).

3 De voor- en nadelen van een theoriegerichte aanpak samengevat

Het vertrekken vanuit een beleidstheorie bij het ontwikkelen, uitvoeren en evalueren van beleid biedt voordelen. Deze laten zich langs de fasen van de beleidscyclus samenvatten (waarbij beleidsontwikkeling logischerwijs de basis legt voor de uitvoering en evaluatie van beleid) (tabel 1).

Naast voordelen kent het hanteren van een beleidstheoretische invalshoek ook risico's. Het eerste risico is de dwingende en soms ook simplificerende werking die van een beleidstheorie en de bijbehorende prestatie- en effectindicatoren kan uitgaan. De kans is niet denkbeeldig dat alles wat niet in theorieën, prestatiegegevens en doelstellingen valt uit te drukken, zich aan de waarneming van besluitvormers zal onttrekken (De Bruijn 2006). Een beleidstheorie stuurt niet alleen de ontwikkeling van beleid of zijn uitvoering, maar ook de beleidsevaluatie achteraf. Dat geldt vooral wanneer de uitgangspunten van beleid stevig gearticuleerd zijn: wanneer consensus tussen betrokken partijen groot is en/of een centrale actor

Tabel 1: Functies van beleidstheorieën bij het ontwikkelen, uitvoeren en evalueren van beleid en daarbij passende operationele vragen

Fase	Functie beleidstheorie	Operationele vragen
Bij het ontwikkelen van beleid	Beleidstheorie als middel om focus te verkrijgen en systematisch te zoeken naar onderbouwing van de verwachte werking van beleidsmaatregelen	Van welke instrumenten mag verwacht worden dat zij zullen bijdragen aan het realiseren van doelstellingen? Welk 'bewijs' bestaat daarvoor?
Bij het uitvoeren van beleid	Beleidstheorie als een (afgesproken) richtlijn voor handelen en bijsturen	Wat zou er wanneer door wie moeten gebeuren? Welke bijsturing is mogelijk op basis van tussentijds beschikbare monitorinformatie?
Bij het evalueren van beleidsprogramma's	Beleidstheorie als ijkpunt voor beoordeling van succes versus inhoud	Hoe kan de mate waarin doelstellingen zijn gerealiseerd, worden beoordeeld? Welke verklaringen moeten worden gezocht voor tegenvallende beleidsresultaten? Welke onverwachte effecten zijn er opgetreden en hoe kunnen die worden beoordeeld?

haar beleidstheorie krachtig neerzet, kan dit leiden tot tunnelvisie, rigiditeit en angst voor vernieuwing (Van der Knaap 2006). Maar ook als een beleidstheorie niet meer echt 'leeft' en achteraf gereconstrueerd moet worden (bijvoorbeeld Justitie 2008), kan dit risico zich manifesteren.

Een tweede risico is de kwaliteit van de bewijsvoering. Ik merkte al op dat in het Verenigd Koninkrijk, de Verenigde Staten en ook in ons land de invoering van *evidence-based* beleid gepaard gaat met een roep om streng, op experimentele leest geschoeid onderzoek. In de praktijk zijn onder de noemer *evidence-based* beleid ook pragmatische combinaties waar te nemen. Daarbij worden niet alleen de uitkomsten van onderzoek, pilots en/of eerdere beleidsevaluaties gebruikt bij het onderbouwen van beleidskeuzes, maar bijvoorbeeld ook ervaringen in het buitenland of zelfs de mening van enkele experts over de mogelijke effectiviteit van maatregelen. Hoewel de kern steeds is dat voor een bepaalde (beleids)theoretische aanpak van een probleem – de mogelijke interventie – systematisch bezien wordt welk bewijs beschikbaar is voor de werkzaamheid ervan, verschilt het bewijs in kwaliteit of 'hardheid'.³

Lam wijst op deze relatieve eenvoud van het centraal stellen van beleidstheorieën in beleidsevaluatie als een van de oorzaken van het succes van de beleidsgerichte benadering: "Theory-based evaluation and objective-based evaluation have gained tremendous popularity in the last decade because of its power to augment an evaluation's capacity to assess program impact without strong experimental designs"

3 Niet zelden wordt gesteld dat bij het streven naar *evidence-based* beleid gezocht wordt naar een *onderbouwing* van de verwachte effectiviteit van maatregelen. Vanuit het gegeven dat bij beleidsontwikkeling gewerkt wordt met theoretische hypothesen over de effectiviteit van interventies, valt evenveel te zeggen voor het omverwerpen van claims op effectiviteit.

(Lam 2002, 1). Ook het woord *evidence* heeft – zoals veel woorden die uit het Verenigd Koninkrijk of de Verenigde Staten overwaaien om innovaties een naam te geven – vooral positieve connotaties. Dat geldt zeker wanneer het in combinatie met beleid en beleidspraktijk wordt gebruikt. De vraag is nu: welk bewijs, welke *evidence* telt bij het onderbouwen van beleidsmaatregelen en het aantonen van effectiviteit?

4 Inzicht in onderbouwing en effectiviteit van veiligheidsbeleid: onderbouwing en kwaliteit van beleidsinformatie

Hoe staat het met de toelichting van beleidsambities en -maatregelen bij het veiligheidsbeleid van het huidige kabinet? Hoe zijn de maatregelen onderbouwd, welke rol spelen beleidstheorieën daarbij en wat kan gezegd worden over de kwaliteit van de beleidsinformatie?

In 2007 presenteerde het vierde kabinet-Balkenende zijn beleidsprogramma *Samen werken, samen leven*. Het programma bestaat uit 74 doelstellingen en 10 projecten en geeft aan waar het kabinet zich gedurende vier jaar voor wil inzetten. De doelstellingen in het beleidsprogramma zijn een uitwerking van de thema's die in het coalitieakkoord waren opgenomen. Het bestaat uit zes pijlers:

- een actieve internationale en Europese rol;
- een innovatieve concurrerende en ondernemende rol;
- een duurzame leefomgeving;
- sociale samenhang;
- veiligheid, stabiliteit en respect;
- overheid en dienstbare publieke sector.

De Algemene Rekenkamer voert jaarlijks onderzoek uit naar de kwaliteit van beleidsinformatie zoals die is opgenomen in de departementale jaarverslagen. Het veiligheidsbeleid is zowel in 2006 als in 2008 in dat onderzoek onderwerp van analyse geweest. In 2008 heeft de Algemene Rekenkamer gekeken naar de volgende doelen van het veiligheidsbeleid:

- een reductie van de criminaliteit met 25% in 2010 ten opzichte van 2002;
- 100.000 minder gestolen fietsen in 2010 ten opzichte van 2006.

Als het kabinet duidelijk wil maken waarom juist de gekozen instrumenten voor het bereiken van het gestelde doel ingezet (moeten) worden, zou het de Tweede Kamer moeten voorzien van een onderbouwing dat het aannemelijk is dat er een causale relatie is tussen de ingezette middelen, de daarmee te leveren prestaties en de verbeterde situatie rond criminaliteit en fietsendiefstal die daarna ontstaat.

4.1 Onderbouwing van het beleid: beschikbaarheid van valide en betrouwbare informatie

Als het gaat om het aantonen dat beleid effect heeft (dat interventie X leidt tot uitkomst Y), is er feitelijk maar één manier om dat te doen: de experimentele methode. In essentie komt het erop neer dat een experimentele, uit *random* gese-

lecteerde eenheden bestaande doelgroep wel, en een identieke, eveneens *random* geselecteerde controlegroep niet aan een bepaalde interventie wordt blootgesteld. Bij absoluut gelijke omstandigheden en populaties van voldoende omvang kunnen verschillen na toepassing van de interventie dan worden toegeschreven aan de interventie (Ministerie van Financiën 2003b; zie voor een uitvoerige beschrijving Swanborn 1991).

In de praktijk van overheidsbeleid is een dergelijke ‘gouden standaard’ niet altijd realistisch. Doordat er sprake kan zijn van een lange beleidsketen en doordat er ook externe factoren van invloed kunnen zijn op de te bereiken doelen, is het vaak moeilijk om het causale verband tussen prestaties en effecten volgens de experimentele methode te ‘bewijzen’. Toch zijn er alternatieven voor de gouden standaard. Zo worden in ‘quasi-experimenten’ bestaande verschillen tussen bijvoorbeeld jeugdige delinquenten benut (vergelijk ook Oosterbeek 2001 of Wittebrood & Van Dijk 2007). Ook kunnen trendbreuken op effectief beleid duiden, vooral als betrokkenen en experts het verschil toeschrijven aan een beleidsmaatregel en kansrijke alternatieve verklaringen beargumenteerd uitsluiten. Ook kan – zeker als het gaat om het vooraf inschatten van het potentieel van een bepaalde maatregel – gebruik worden gemaakt van systematische *reviews* van eerder en/of elders uitgevoerde analyses van werkzaamheid van interventiemechanismen (vergelijk Leeuw 2008).

In het rapport *Staat van de beleidsinformatie 2006* constateerde de Algemene Rekenkamer dat de ministers van BZK en van Justitie geen samenhangende onderbouwing hadden van de veronderstelde werkzaamheid van de 150 maatregelen uit het toenmalige Veiligheidsprogramma (Algemene Rekenkamer 2006). Hoe deze maatregelen – al dan niet in samenhang – van invloed zouden zijn op de omvang van de criminaliteit, konden de ministers niet goed aangeven. Conform de aanbeveling van de Rekenkamer hebben beide bewindspersonen het SCP vervolgens gevraagd een (re)constructie te maken van de beleidstheorie van het beleid van de vorige kabinetten.

In mei 2008 was de uitkomst hiervan nog niet beschikbaar, zodat de conclusie in de *Staat van de beleidsinformatie 2008* was, dat het veronderstelde causale verband tussen instrumenten en effecten niet onderbouwd was (Algemene Rekenkamer 2008). Het beste zou zijn, zo concludeerde de Algemene Rekenkamer, om wanneer nieuw beleid wordt ingezet, meteen een theoretische onderbouwing op te stellen over de vraag waarom maatregelen werkzaam kunnen of zullen zijn. Dit staat het snel nemen van maatregelen – bijvoorbeeld in crisissituaties – overigens niet in de weg: het dwingt beleidsmakers wel om verwachtingen te expliciteren en stelt hen vervolgens in staat gericht op zoek te gaan naar informatie die iets zegt over het veronderstelde causale verband tussen maatregel en effect. De Algemene Rekenkamer beval de ministers dan ook aan om zo veel als mogelijk in de begrotingen te specificeren welke (extra) kosten gemoeid zijn met de realisatie van de beleidsdoelen criminaliteitsreductie en het terugbrengen van het aantal fietsendiefstallen. Daarbij zouden de berekeningen van de kosten van criminaliteit, die

het WODC van het ministerie van Justitie jaarlijks maakt, gebruikt kunnen worden voor het maken van een kosten-batenanalyse.

Inmiddels heeft het SCP in zijn rapport *Sociale Veiligheid ontsleuteld* een analyse gemaakt welke onderbouwing er reeds beschikbaar is voor welke (type) beleidsmaatregelen (Van Noije & Wittebrood 2008). In deze studie maakt het SCP inzichtelijk welke beleidsmaatregelen werkzaam zijn of kunnen zijn, welke dat niet zijn en van welke maatregelen de kennis nog ontbreekt. Het SCP bevestigde dat het Veiligheidsprogramma geen geëxpliciteerde beleidstheorie kende, maar meer een stapeling van noties en maatregelen betrof (Van Noije & Wittebrood 2008). De reconstructie van de beleidstheorie van het SCP is door het kabinet overgenomen (Ministerie van BZK 2008; zie ook par. 5 van dit artikel).

Als we geloofwaardig willen aantonen dat een verandering op een doelvariabele wordt veroorzaakt door een bepaalde interventie, dan zal het onderzoeksontwerp tot valide uitspraken moeten leiden en zullen metingen betrouwbaar moeten zijn (Swanborn 1991, 322 e.v.). Begrippen moeten *valide* gemeten worden: zodanig geoperationaliseerd dat je inderdaad meet wat je meten wilt. *Betrouwbaarheid* houdt in dat het gebruikte meetinstrument zorgvuldig wordt toegepast en dat herhaalde metingen steeds hetzelfde resultaat opleveren (Algemene Rekenkamer 2005b).⁴

Op het vlak van de betrouwbaarheid van de informatie heeft de Algemene Rekenkamer in het onderzoek van 2006 vastgesteld dat de betrouwbaarheid van de gegevens over overlastreductie in probleemwijken onzeker is, 'omdat informatie hierover ontbreekt' (Algemene Rekenkamer 2006, 45). In het onderzoek van 2008 zijn – behoudens de consistentie van de 25%-doelstelling (zie hierna) – geen specifieke problemen geconstateerd. In haar reguliere rechtmatigheidsonderzoek, waarover de Algemene Rekenkamer rapporteert in haar rapporten bij de jaarverslagen, is wel een algemene tekortkoming geconstateerd rond de betrouwbaarheid van de geaggregeerde beleidsinformatie.⁵

4.2 Overige kwaliteit van de beleidsinformatie: concreetheid, relevantie en consistentie

Concreetheid

Naast eisen op het vlak van validiteit en betrouwbaarheid stelt de Algemene Rekenkamer ook dat doelstellingen en indicatoren concreet (geformuleerd) moeten zijn, dat de beleidsinformatie een relevant beeld geeft van de geleverde prestaties en de mate waarin effecten zijn bereikt, en dat de onderliggende gegevens consistent zijn (Algemene Rekenkamer 2005a; vergelijk Maso 1987). Een beoogd

4 Smaling vat dit kernachtig samen als de afwezigheid van 'onsystematische fouten' (1992, 307).

5 Wel is voor het jaarverslag over 2008 bij Justitie vastgesteld dat de totstandkoming van de niet-financiële informatie een aantal tekortkomingen kent die veelal samenhangen met het onttrekken van informatie aan decentrale systemen en het aggregeren ervan naar centrale (management)informatie: 'De betrouwbaarheid van het aggregatieproces is onvoldoende gewaarborgd. Veel onderdelen van het ministerie zijn echter reeds geruime tijd bezig om hun informatiesystemen te vernieuwen en de tekortkomingen op te lossen' (Algemene Rekenkamer 2009, 38).

effect is concreet geformuleerd als het zelf meetbaar, tijdgebonden en richtinggevend is of wanneer het met een indicator meetbaar is gemaakt, tijdgebonden is geworden en van een streefrichting is voorzien (vergelijk Algemene Rekenkamer 2005b, 14). Een prestatie is concreet als óf de beleidsactiviteiten óf de inzet van een instrument, zelf dan wel via een specifieke indicator, meetbaar is gemaakt, tijdgebonden is en een streefrichting bevat (ibid.).

Het reduceren van de criminaliteit met 25% is een algemeen geformuleerde, hogere doelstelling van het kabinet. De doelen waarin deze doelstelling is uitgewerkt (afname van geweldsdelicten, afname van vermogensdelicten, daling van de criminaliteit tegen ondernemingen, verbetering van het ophelderingspercentage en een daling van de recidive), zijn evenwel concreet geformuleerd: ze zijn overwegend meetbaar, tijdgebonden en bevatten een streefrichting (Algemene Rekenkamer 2006; 2008). In het beleidsprogramma van het kabinet worden voorbeelden genoemd van nieuwe maatregelen, zoals de inzet van vijfhonderd extra wijkagenten en de inzet van vijfhonderd extra forensisch assistenten. Ook deze prestaties zijn concreet geformuleerd, waardoor het mogelijk is om later verantwoording af te leggen over het al dan niet realiseren ervan (Algemene Rekenkamer 2008).

Relevantie

Voor een goede uitvoering van, bijsturing van en verantwoording over beleid moeten relevante gegevens verzameld worden. Relevant betekent in elk geval dat alle informatie die voor een beleidsvoerder, gelet op de beleidsdoelen, beschikbaar zou moeten zijn om te sturen, te leren of te verantwoorden, ook beschikbaar is. Doelstellingen en indicatoren moeten voor de doelgroepen en uitvoerders van beleid de essentie van een beleidsopgave uitdrukken. Verder geldt dat de aanwezige informatie actueel moet zijn en dat de benodigde informatie op tijd beschikbaar dient te zijn (Algemene Rekenkamer 2005b).

De relevantie van informatie over geleverde prestaties hangt samen met het beleidsdoel dat daarmee gediend moet worden. In 2008 waren de vragen welke prestaties zouden moeten leiden tot de beoogde afname van criminaliteit en fietsendiefstal, en hoe deze prestaties gemeten en verantwoord zullen worden, nog niet uitgewerkt (Algemene Rekenkamer 2008). De Algemene Rekenkamer stelde in 2006 wel vast dat de ministeries van BZK en Justitie via enquêtes controleren of de effectindicatoren die zij voor het veiligheidsbeleid hanteren voor de doelgroepen van het beleid werkelijk relevant zijn. Daarbij concludeerden de onderzoekers dat meten altijd maatwerk is: lokale verschillen bepalen in sterke mate de relevantie van doelstellingen en indicatoren (Algemene Rekenkamer 2006, 40-41). De Algemene Rekenkamer constateerde instemmend dat de nieuwe Veiligheidsmonitor in dat maatwerk kon voorzien.

Consistentie

Als de definitie en operationalisatie van waarop gestuurd wordt of wat gemeten wordt frequent verandert, kun je de gegevens van verschillende jaren niet verge-

lijken. Gegevens moeten daarom in de tijd vergelijkbaar zijn en wijzigingen in eerder verstrekte gegevens moeten worden geëxpliciteerd (vergelijk Smaling 1992). Verder moeten gegevens uit verschillende bronnen over hetzelfde onderwerp onderling consistent zijn. Dat wil zeggen dat gegevens die uit verschillende bronnen afkomstig zijn met elkaar vergeleken moeten worden en transparant, bijvoorbeeld naast elkaar, moeten worden gepresenteerd. Als gegevens strijdig zijn, moet dit worden aangegeven en – beter nog – worden verklaard.

De doelstelling om de criminaliteit met 25% te reduceren is in 2002 geformuleerd in de beleidsnota *Naar een veiliger samenleving* (Ministerie van Justitie en Ministerie van BZK 2002). Er werd een Veiligheidsprogramma opgestart, dat van de Tweede Kamer de status van groot project toegewezen kreeg. Het vierde kabinet-Balkenende heeft de 25%-doelstelling in zijn beleidsprogramma als gezegd overgenomen. Met behulp van de Veiligheidsmonitor Rijk wordt sinds een aantal jaren in beeld gebracht hoe het slachtofferschap van criminaliteit, de onveiligheidsgevoelens van burgers, buurtproblemen en meningen over het functioneren van de politie zich ontwikkelen. De Veiligheidsmonitor Rijk wordt jaarlijks uitgebracht door het Centraal Bureau voor de Statistiek (CBS) en is ontstaan uit een integratie van drie eerdere afzonderlijke reguliere onderzoeken op dit terrein (CBS 2008).

Daarbij bleek het berekenen van de al geboekte en nog te realiseren afname van de criminaliteit lastig, doordat in 2005 de meetsystematiek door het CBS was gewijzigd (Algemene Rekenkamer 2008). Na deze wijziging waren de criminaliteitscijfers fors omhoog gegaan. Vanuit BZK en Justitie is het CBS gevraagd om te analyseren waar dit grote verschil aan te wijten was: was de criminaliteit echt toegenomen óf vielen de cijfers alleen maar hoger uit, doordat het onderzoek anders werd uitgevoerd? Het CBS kon deze vraag niet met zekerheid beantwoorden. De conclusie was wel dat de cijfers van voor en na de trendbreuk van 2005 niet te vergelijken waren. Desondanks hebben de ministers van BZK en van Justitie in 2006 wél conclusies getrokken over de ontwikkeling van de criminaliteit. Conclusies, die op basis van het beschikbare cijfermateriaal van het CBS volgens de Algemene Rekenkamer niet getrokken hadden mogen worden (Algemene Rekenkamer 2008, 82). Voor fietsendiefstal geldt vorenstaand probleem overigens niet, aangezien voor de aanpak van dit probleem de situatie in 2006 als uitgangspunt is gekozen.

4.3 Slotsom: nog geen goed bewijs mogelijk voor effectiviteit

In de brief bij de tweede voortgangsrapportage over het project *Veiligheid begint bij Voorkomen* schrijven de ministers van Justitie en van BZK dat Nederland veiliger is geworden:

‘De structurele daling van de criminaliteit, die zich al sinds de start van het vorige veiligheidsprogramma in 2002 voordoet, heeft zich ook in 2008 voortgezet. Het aantal door burgers ondervonden delicten is verder afgenomen. Eind 2008 gaf één op de vier burgers aan de afgelopen twaalf maanden slacht-

offer geweest te zijn van een misdrijf. In 2006 was dit nog 30%.' (Ministerie van Justitie 2009a)

De ministers schrijven dat deze tendens zich voordoet bij nagenoeg alle soorten criminaliteit, maar vooral bij vandalisme-, gewelds- en vermogensdelicten, waaronder de fietsendiefstal.

Ten aanzien van dat laatste punt, de fietsendiefstal, meldde het kabinet op Verantwoordingsdag 2009 dat, hoewel de doelstelling van 100.000 minder gestolen fietsen 'ambitieuw' was, het aantal van 110.000 minder gestolen fietsen in 2008 al bereikt was (110.000 minder fietsen gestolen ten opzichte van 2006). Daarmee, zo stelde het kabinet tevreden vast, ligt de doelstelling niet alleen op koers, maar is het beoogde resultaat voor 2010 bereikt (www.regering.nl/verantwoordingsdag). Een halfjaar later blijkt het aantal gestolen fietsen in 2008 met 38.000 te zijn gedaald ten opzichte van 2007 (Ministerie van Justitie 2009b, 5). Voor de volledigheid: nergens in de voortgangsrapportage worden deze resultaten expliciet toegeschreven aan het gevoerde beleid.

Dit laatste is in lijn met de conclusies van het Rekenkameronderzoek uit 2006 en 2008. Op zich waren die conclusies niet negatief: de ministers van BZK en van Justitie spannen zich in om te komen tot afrekenbaar beleid. Tegelijkertijd zijn er op diverse punten verbeteringen mogelijk. Problematisch achtte de Algemene Rekenkamer vooral dat de ministers door een tussentijdse wijziging van de manier waarop criminaliteit wordt gemeten, niet meer kunnen aangeven in hoeverre de problemen sinds 2002 zijn afgenomen. Strikt genomen betekent dit dat er geen goede bewijsvoering mogelijk is dat de bestede middelen en de daarmee genomen maatregelen daadwerkelijk hebben geresulteerd in het realiseren van de doelstellingen.

De ministers reageerden overwegend instemmend op de aanbevelingen van de Rekenkamer om (1) uit te werken welke prestaties moeten leiden tot de afname van criminaliteit en fietsendiefstal en (2) duidelijk te maken hoe deze prestaties gemeten en verantwoord zullen worden. Tegelijkertijd gaven de ministers aan dat er grenzen zijn aan het een-op-een koppelen van prestaties aan (beoogde) effecten, zeker in een complex beleidsterrein als criminaliteit (Algemene Rekenkamer 2008).

5 Conclusies en perspectieven

In 2006 konden de verantwoordelijke ministers niet goed aangeven hoe beleidsmaatregelen zouden moeten bijdragen aan meer veiligheid. Sindsdien is het nodige gebeurd. De door het SCP gereconstrueerde beleidstheorie is door het kabinet overgenomen. De verantwoordelijke ministers gaven in hun reactie aan om in toekomstige beleidsplannen op het gebied van veiligheid de beleidstheorie veel nadrukkelijker te gaan expliciteren (Van Noije & Wittebrood 2008; Ministerie van BZK 2008, 3).

Anno 2010 wordt er gestuurd op de doelstellingen en indicatoren van het veiligheidsbeleid en wordt er systematisch verantwoording afgelegd over de voortgang. Daarbij is het positief dat in de voortgangsrapportage onderscheid gemaakt wordt tussen de bereikte resultaten tot en met 2006 enerzijds, en de voortgang op de herijkte doelstellingen vanaf dat jaar anderzijds. De bronnen van de beleidsinformatie worden aangegeven. Doelen zijn veelal gekwantificeerd en worden gevolgd via de veiligheidsmonitor en de monitor criminaliteit bedrijfsleven. Daarnaast bevat de voortgangsrapportage diverse verwijzingen naar beleidsevaluaties, -onderzoek en -analyses (bijvoorbeeld de tussenevaluatie van de Wet tijdelijk huisverbod). Ook wordt ingegaan op nieuw beleid, zoals de stallingsmogelijkheden voor fietsen waar gemeenten een actieve rol bij spelen. Bij de bestrijding van ernstige criminaliteit wordt gewerkt met proeftuinen om, al lerende, robuuste hindernissen voor criminele activiteiten neer te zetten waar *best practices* worden toegepast (Ministerie van Justitie 2009b, 45). Het WODC monitort en evalueert daartoe deze experimenten.⁶

Dat dit nodig is, blijkt uit de rapportage van het SCP. Twee derde van de maatregelen die op effectiviteit konden worden beoordeeld, zijn volgens de onderzoekers als 'effectief' of 'veelbelovend' te bestempelen. Daar staat tegenover dat bijna een derde als 'niet-effectief' werd beoordeeld, terwijl een aantal maatregelen onder bepaalde omstandigheden zelfs als 'contraproductief' te kwalificeren valt (Van Noije & Wittebrood 2008).

Al met al is het veiligheidsbeleid anno 2010 een treffend voorbeeld van hoe in Nederland gestreefd wordt naar beleidstheoretisch onderbouwd, *evidence-based* beleid. Het onderzoek van het SCP en het WODC is illustratief voor de meerwaarde van het gebruik van beleidstheorieën bij beleidsontwikkeling en evaluerend onderzoek. Het biedt de mogelijkheid om het vizier actief te richten op die facetten die kenmerkend zijn voor een goede beleidsvoorbereiding en -uitvoering. Aan de voorkant, bij het ontwikkelen van beleid en het onderbouwen van budgetten bij nieuwe maatregelen, biedt het expliciet maken van een beleidstheorie en het verschaffen van informatie over de verwachte effectiviteit houvast voor mensen die uiteindelijk over voorstellen moeten beslissen. Met name kan worden beoordeeld in welke mate een beleidsmaker zijn verwachtingen rond de beleidstheorie geëxpliciteerd heeft. Op basis daarvan kan vervolgens de onderbouwing van de verwachting, dat de inzet van instrumenten zal bijdragen aan het bereiken van gewenste effecten, worden beoordeeld.

Tegelijkertijd geeft het woord 'theorie' al aan, dat op voorhand nooit met absolute zekerheid te garanderen is welke effecten een maatregel zal kennen. Het is dus van belang dat goed wordt nagegaan in hoeverre de in de beleidstheorie veronderstelde gedragsmechanismen bij de doelgroep van het beleid overeenkomen met eerdere ervaringen en kennis (vergelijk Leeuw 2008). Ook de samenhang tussen causale ketenveronderstellingen waaruit een beleidstheorie is opgebouwd, is relevant: vertonen deze een logische samenhang of is er sprake van onderlinge strij-

6 Zie www.wodc.nl/onderzoeksdatabase/om-proeftuinen.aspx.

digheden? Ten slotte is het – zeker wanneer belangen groot zijn en er bijvoorbeeld ingrijpende maatregelen op het gebied van privacy worden voorgesteld – soms nodig om de in een beleidstheorie vervatte veronderstelde causale relaties te toetsen – van *evidence* te voorzien – met behulp van pilots, proeftuinen en (quasi)experimenten.

Om goede lessen uit beleidsonderzoek te trekken, is het zaak dat de bewijsvoering van voldoende kwaliteit is. Ook aan de beleidsinformatie, die bij wijze van verantwoording aan de Tweede Kamer wordt gezonden, moeten kwaliteitseisen worden gesteld: het vormt immers de munitie om een beleidsdebat goed geïnformeerd te kunnen voeren. Door duidelijk aan te geven hoe een monitorsysteem is opgezet of hoe een evaluatieonderzoek – *ex ante* dan wel *ex post* – is uitgevoerd, kan duidelijk worden gemaakt hoe gewerkt wordt aan validiteit en betrouwbaarheid. Bij de vraag of alle relevante beleidsinformatie beschikbaar is, is het verstandig om te letten op de ‘toereikendheid’ van informatie. Deze vraag is vooral boeiend, omdat het ons bewust maakt van de valkuil van perfectionisme: de neiging om te veel te willen meten. Dit wordt nog versterkt door de mogelijkheden van de toepassing van ICT.

Tot slot: de belangrijkste waarde van het bewust hanteren van beleidstheorieën bij het ontwikkelen en evalueren van beleid is mijns inziens dat gericht kan worden bijgedragen aan het verbeteren van beleid. Goed beleid doet recht aan zich wijzigende omstandigheden. Voor het beoordelen van relevantie is het niet alleen zaak om te kijken naar beleidsdoelen, maar zeker ook de vraag te stellen of indicatoren en gegevens voor de doelgroepen en uitvoerders van beleid inderdaad – en nog steeds! – de essentie van een beleidsissue uitdrukken. Welke doelstellingen en indicatoren kunnen we loslaten, bijvoorbeeld omdat het probleem niet langer bestaat of geen prioriteit meer heeft? Welke nieuwe doelstellingen en indicatoren zijn relevant? Daarbij geldt dat van de meeste beleidstheorieën – en dus doelstellingen en indicatoren – anno 2009 geen universele of oneindige geldigheid wordt verwacht. De relevantie, maar ook effectiviteit van beleidsprogramma’s hangt sterk af van de lokale, regionale of sectorale context waar het beleid wordt toegepast (vergelijk Pawson & Tilley 1997).

Effectief beleid vergt vaak maatwerk. Goede beleidsevaluaties gaan dan ook niet alleen over gerealiseerde prestaties of bereikte effecten. Goede beleidsevaluaties gaan ook niet alleen over de waardering van beleid door belanghebbenden. Een goede beleidsevaluatie doet beide en zoekt tegelijkertijd, rekening houdend met de specifieke context waarin een maatregel wordt toegepast, naar *verklaringen* voor (gebrek aan) effectiviteit en waardering. Vanuit deze drieslag bevat een goed evaluatierapport zinvolle aanbevelingen en gaan voorstellen voor nieuw beleid in op de gekozen remedies voor geconstateerde tekortkomingen.

Een goede evaluatieonderzoeker richt daarbij het vizier ook op de door de overheid gehanteerde beleidstheorie zelf. Beleidstheorieën hebben een sturende uitwerking op de vormgeving, uitvoering en evaluatie van beleid. Soms is dat wenselijk: het leidt tot focus, gedeelde referentiekaders en ijkpunten voor beoordeling

van succes en falen. In andere gevallen leidt het tot simplificatie, tunnelvisie, defensief gedrag en leerbarrières. De waardering van deze effecten zal van belanghebbende tot belanghebbende verschillen: wat voor de één focus is, zal door de ander als inperking worden gevoeld. Daarbij speelt dat het onmogelijk is om iedere verschillende interpretatie of elke uitzonderingssituatie in een beleidstheorie of beleidsprogramma te verdisconteren. Theoriegerichte en tegelijkertijd responsieve evaluaties, waarbij de evaluator expliciet openstaat voor veranderingen in context of de mening van betrokkenen, kunnen helpen om in dit hiaat te voorzien.

Literatuur

- Algemene Rekenkamer (2005a) *Staat van de beleidsinformatie 2005*. Den Haag.
- Algemene Rekenkamer (2005b) *Essentialia van goed openbaar bestuur*. Den Haag.
- Algemene Rekenkamer (2005c) *Beleidsvrijheid en specifieke uitkeringen*. Den Haag.
- Algemene Rekenkamer (2006) *Staat van de beleidsinformatie 2006*. Den Haag.
- Algemene Rekenkamer (2008) *Staat van de beleidsinformatie 2008*. Den Haag.
- Algemene Rekenkamer (2009) *Rapport bij het jaarverslag 2008 – Ministerie van Justitie 2008*. Den Haag.
- Bruijn, H. de (2006) *Prestatiemeting in de publieke sector – Tussen professie en verantwoording*. Utrecht: Lemma.
- Centraal Bureau voor de Statistiek (2008) *Veiligheidsmonitor Rijk, Landelijke rapportage*. Den Haag: CBS.
- Dunn, W.N. (1994) *Public policy analysis*. Englewood Cliffs: Prentice Hall.
- Hoesel, P. van, F. Leeuw & J. Mevissen (2005) *Beleidsonderzoek in Nederland – Kennis voor beleid; ontwikkeling van een professie*. Assen: Van Gorcum.
- Hoogerwerf, A. (1984) Beleid berust op veronderstellingen. De beleidstheorie. *Acta Politica*, 14(4), 493-531.
- Klein Haarhuis, C.M. & L.L.C. Hagen (m.m.v. M.P.C. Scheepmaker) (2009) *Toetsen en verbinden – Over het WODC in de tijd en de relatie tussen onderzoek en beleid in het bijzonder*. Den Haag: Boom Juridische uitgever.
- Knaap, P. van der (2000) Performance management and policy evaluation in the Netherlands: towards an integrated approach. *Evaluation*, 6(3), 335-350.
- Knaap, P. van der (2006) Responsive evaluation and performance management – Overcoming the downsides of policy objectives and performance indicators. *Evaluation*, 12(3), 278-293.
- Knaap, P. van der (2009) Beleidsdoorlichtingen en doelmatigheidsonderzoek onder de loep. *TPC*, december 2009, 7-11.
- Lam, T.C.M. (2002) *Theory-based evaluation and objective-based evaluation: An integration of the two approaches*. Paper presented at the 2002 European Evaluation Society conference, Sevilla.
- Leeuw, F.L. (1983) *Overheid en bevolkingsgroei: een evaluatie van beleidstheorieën*. Leiden: Rijksuniversiteit Leiden.
- Leeuw, F.L. (1993) Institutionaliseren van evaluatieonderzoek bij de overheid. In: M. Donker & J. Derks (red.) *Rekenschap – Evaluatieonderzoek in Nederland: de stand van zaken*. Utrecht: Nederlands Centrum Geestelijke Volksgezondheid.
- Leeuw, F.L. (2008) *Gedragsmechanismen achter overheidsinterventies en rechtsregels*. Oratie. Maastricht: Universiteit van Maastricht.

- Leeuw, F.L. (2009) Evaluation policy in the Netherlands. In: W.M.K. Trochim, M.M. Mark & L.J. Cooksy (red.), *Evaluation policy and evaluation practice*. Chicester: Wiley.
- Maso, I. (1987) *Kwalitatief onderzoek*. Meppel: Boom.
- Ministerie van BZK (2008) *Beleidsreactie SCP-rapport 'Sociale Veiligheid ontsleuteld'*. Den Haag.
- Ministerie van Financiën (2003a) *Handreiking beleidsevaluatie ex ante*. Den Haag.
- Ministerie van Financiën (2003b) *Handreiking beleidsevaluatie ex post*. Den Haag.
- Ministerie van Financiën (2003c) *Miljoenennota 2004*. Den Haag: Sdu Uitgevers.
- Ministerie van Financiën (2005) *Evaluatie ex post onderzoek – Bijlage bij het eindrapport VBTB-evaluatie*. Den Haag.
- Ministerie van Financiën (2006) *Regeling periodiek evaluatieonderzoek en beleidsinformatie*. Den Haag.
- Ministerie van Justitie (2008) *Evaluatie Justitie in de buurt nieuwe stijl – Verbindende netwerken in de veiligheidshuizen*. Den Haag.
- Ministerie van Justitie (2009a) *Aanbiedingsbrief bij tweede voortgangsrapportage 'Veiligheid begint bij voorkomen'*. Den Haag, 30 oktober 2009.
- Ministerie van Justitie (2009b) *Tweede voortgangsrapportage 'Veiligheid begint bij voorkomen'*. Den Haag.
- Ministerie van Justitie en Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (2002) *Naar een veiliger samenleving*. Den Haag.
- Noije, L. van & K. Wittebrood (2008) *Sociale veiligheid ontsleuteld: veronderstelde en werkelijke effecten van veiligheidsbeleid*. Den Haag: Sociaal en Cultureel Planbureau.
- Oosterbeek, H. (2001) *Voortschrijdend inzicht*. Rede uitgesproken bij de aanvaarding van het ambt van hoogleraar in de Onderwijseconomie. Amsterdam: Universiteit van Amsterdam.
- Pawson, R. (2002) *Nothing as practical as a good theory*. Lezing tijdens de 5^{de} tweejaarlijkse conferentie van de European Evaluation Society. Sevilla.
- Pawson, R. & N. Tilley (1997) *Realistic Evaluation*. Londen: Sage Publications.
- Smaling, A. (1992) Objectiviteit, betrouwbaarheid en validiteit. In: G.J.H. Bruinsma & N.A. Zwanenburg, *Methodologie voor bestuurskundigen – Stromingen en methoden*. Muidersberg: Coutinho.
- Swanborn, P.G. (1991) *Methoden van sociaal-wetenschappelijk onderzoek*. Meppel: Boom.
- Wittebrood, K. & T. van Dijk (2007) *Aandacht voor de wijk: effecten van herstructurering op de leefbaarheid en veiligheid*. Den Haag: Sociaal en Cultureel Planbureau.