

## Voorwoord

Dit nummer van *Justitiële verkenningen* gaat over beleidsevaluatie. We hebben het dan over een uit de twintigste eeuw stammend fenomeen, maar vanzelfsprekend evalueren mensen al eeuwenlang hun activiteiten. Zo zouden al in 2200 v. Chr. evaluaties van persoonsselectie hebben plaatsgevonden in China (Shadish, Cook e.a., 1995, p. 21). Op zoek naar een effectief middel tegen scheurbuik ontwikkelde de Britse scheepschirurgijn James Lind in 1747 iets wat leek op een zogeheten experimenteel design. Op het schip waar hij diende, selecteerde hij twaalf mannen die aan scheurbuik leden en verdeelde hen in zes paren. Elk paar kreeg dagelijks een andere aanvulling op het basisdieet als mogelijke 'remedie': cider, een elixer, zeewater, een mix van knoflook, mosterd en radijs, azijn, en het laatste paar kreeg dagelijks twee sinaasappelen en een citroen. De twee mannen die het citrusfruit aten, herstelden opmerkelijk snel (Lind, 1753). Hoewel er naar huidige maatstaven wel wat aan te merken valt op zijn onderzoeksopzet, had Lind zo de superioriteit van citrusvruchten als remedie aangetoond.

De aanleiding voor dit themanummer is niet alleen de explosieve toename van het aantal evaluaties in Nederland en andere landen. Ook de recente discussies over de methoden van evaluatieonderzoek en het maken van syntheses van bestaande evaluaties bieden meer dan genoeg stof. Een interessante ontwikkeling is de opleving van het experiment als evaluatiemethode. Amerikaanse sociaal wetenschappers werkten al experimenteel in de jaren twintig en dertig, de Britten deden er in de jaren vijftig ervaring mee op. De samenleving mag dan geen laboratorium zijn, het voorbeeld van de medische wetenschap met haar uitgebreide gecontroleerde testen van medicijnen en andere behandelmethoden is toch voor menig sociaal wetenschapper een inspiratiebron, zoals bijvoorbeeld blijkt uit het inmiddels klassieke werk van Donald Campbell en Julian Stanley, *Experimental and Quasi-Experimental Designs for Research* (1966). Dat het experiment als onderzoeksmethode weer terug is van weggeweest, werd afgelopen november duidelijk op de conferentie '(Quasi) Experimental Evaluations and Dutch Society; trends and developments in criminal justice and policies, social integration and education'. De artikelen in dit nummer zijn bijna allemaal

gebaseerd op presentaties die daar zijn gehouden. De conferentie was georganiseerd door het WODC in samenwerking met het ministerie van Financiën en de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Een onderwerp dat tijdens de discussies op de conferentie wél aan de orde kwam, maar in deze aflevering van JV nauwelijks, is het tegenstrijdige fenomeen dat evaluaties een bescheiden rol spelen in het beleidsproces. Tegenstrijdig, omdat er wél keer op keer opdrachten voor evaluatie worden verstrekt en er dus veel geld aan wordt besteed. Afgezien van de soms gebrekkige kwaliteit en ongelukkige timing van evaluaties speelt hier iets heel anders: de kloof tussen onderzoekers enerzijds en beleidsambtenaren en politici anderzijds, botsende denkwerelden, logica's, belangen en loyaliteiten. Misschien komen we daar nog eens op terug in een ander thema-nummer, want het is een onderwerp op zich.

In Nederland begon de evaluatie van overheidsbeleid rond 1980. Sindsdien is het aantal evaluaties gestaag toegenomen. Op het terrein van rechtshandhaving bijvoorbeeld werden in 1995 door het ministerie van Justitie achttien evaluaties uitgevoerd. In 2005 zijn dat er 37.

Evaluatie-‘societies’ van onderzoekers en beleidsmakers kwamen aanvankelijk alleen voor in de VS en Canada, maar figureren nu in vrijwel elk Europees land en in verscheidene Afrikaanse en ‘Australaziatische’ landen. Ook supranationale instituties zoals de Europese Unie en de Wereldbank evalueren hun beleid. Alleen al op het terrein van ontwikkelingshulp lopen de aantallen evaluaties in de duizenden binnen een tijdsbestek van enkele jaren.

Een mogelijke verklaring voor deze groei is dat overheden, maar ook bedrijven en maatschappelijke organisaties tegenwoordig transparant moeten zijn en verantwoording moeten kunnen afleggen over hun beleid en de wijze waarop geld wordt besteed. In de jaren negentig van de vorige eeuw was dit idee sterk gelieerd aan de New Public Management-filosofie die populair was in veel westerse samenlevingen.

Deze ontwikkelingen ontmoeten ook veel kritiek. Er wordt wel gesproken van een ‘evaluatie-industrie’, een ‘audit society’ of ‘controlemaatschappij’. Meer evaluatie, monitoring en controle leiden niet automatisch tot succesvoller beleid. Het kan zelfs de efficiency en effectiviteit van beleid en organisaties schaden, omdat er teveel

hopen dat dit type onderzoek ook in andere landen, waaronder Nederland, opgang zal maken.

Pawson en Klein Haarhuis beschrijven vervolgens de theoriegestuurde aanpak van evaluatieonderzoek. Zij laten zien hoe de theorie van een programma kan worden achterhaald. Zo wordt duidelijk of de veronderstelde onderliggende mechanismen die het beoogde effect moeten bewerkstelligen, daadwerkelijk optreden. Zij illustreren het nut van deze aanpak met een aansprekend voorbeeld: de evaluatie van 'Megan's law', een regeling in de VS waarbij buurten worden ingelicht als een uit de gevangenis of inrichting ontslagen seksueel delinquent zich daar vestigt.

Experimentele evaluaties komen in Nederland nog nauwelijks voor. Hoewel er methodologisch het nodige is aan te merken op bestaande evaluaties, kunnen deze toch bruikbaar zijn, zo betogen Klein Haarhuis, Van Ooyen-Houben, Kleemans en Leeuw in hun bijdrage. Hierin wordt een synthese van 31 onderzoeken op het terrein van rechtshandhaving opgesteld. Van belang is dat onderzoeken een gedegen beschrijving bevatten van de onderzoeksopzet, de meetmethode en de eventuele selectiemethode. Ook het gebruik van meerdere informatiebronnen draagt bij aan de bruikbaarheid. Voorts moet duidelijk zijn of beleid daadwerkelijk is geïmplementeerd en moet aandacht zijn besteed aan de vooronderstellingen die ten grondslag liggen aan de interventie. De inhoudelijke synthese wijst onder meer uit dat projecten in de rechtshandhaving vaak bestaan uit verhoogde input van mankracht, expertise en middelen, en uit intensivering van samenwerking. Daarnaast noemen de auteurs een aantal systematische knelpunten.

Webbink borduurt in zijn bijdrage voort op de moeizame relatie tussen experiment en sociale wetenschappen. Het gerandomiseerd toewijzen aan een interventie- of een controlegroep wordt vaak bemoeilijkt door ongewilde doch onvermijdelijke selectieprocessen. Verschillende recente onderzoeken maken tegenwoordig gebruik van natuurlijke experimenten om de causale effecten van beleidsinterventies te bepalen. Het gaat dan om situaties die toevallig zijn ontstaan door natuurlijke processen of door overheidsbeleid en die lijken op een omgeving waarbinnen een gecontroleerd experiment zou kunnen worden uitgevoerd. De auteur geeft verscheidene voorbeelden.

Bijl bespreekt hoe evaluatieonderzoek een rol zou kunnen spelen in het integratiebeleid. Een complicerende factor daarin is dat meerdere en naar hun aard verschillende doelstellingen worden nagestreefd. Besproken wordt wat men verstaat onder het begrip integratie alsmede hoe het beleid op zijn kwaliteiten zou kunnen worden beoordeeld. Positief is de grotere beschikbaarheid van kwantitatieve gegevens en de ruimere mogelijkheden om databestanden van uiteenlopende organisaties aan elkaar te koppelen, meent de auteur. Dat biedt nieuwe mogelijkheden voor goede beleidsevaluaties. De Jong en Kabel, ten slotte, stellen dat evaluatie onontbeerlijk is als men antwoord wil geven op de vraag of de belastingbetaler waar krijgt voor zijn geld. Het aantal evaluaties in Nederland is explosief gestegen, maar de kwaliteit ervan is niet navenant toegenomen. Een verbetering kan worden bereikt door meer effectevaluaties, duidelijker beleidsdoelstellingen, meer garanties voor de onafhankelijkheid van evaluaties en betere verslaglegging. Hoewel consensus lijkt te bestaan over de potentieel grote waarde van evaluaties, is de invloed ervan op het beleidsproces doorgaans marginaal. De auteurs, werkzaam bij het ministerie van Financiën, kondigen aan dat een strategie wordt ontwikkeld ter bevordering van ex ante beleidsevaluaties en effectevaluaties. Ook zou de focus van evaluatieonderzoek moeten verschuiven van instrumenten naar concrete beleidsterreinen.

C. Klein Haarhuis  
F.L. Leeuw  
M.P.C. Scheepmaker

### Literatuur

**Campbell, D.T., J. Stanley**

*Experimental and quasi-experimental designs for research*  
Chicago, Rand McNally, 1966

**Lind, J.**

*A treatise of the scurvy in three parts; containing an inquiry into the nature, causes and cure of that disease, together with a critical and chronological view of what has been published on the subject*

London, A. Millar, 1753

**Shadish, W.R., T.D. Cook e.a.**

*Foundations of program evaluation, theories of practice*  
London, Sage, 1995

**Thiel, S. van, F.L. Leeuw**

*The performance paradox in the public sector*

Public performance & management review, 25e jrg., nr. 3, 2002, p. 267-281