

KWALITEIT VAN BELEIDSINFORMATIE

ERVARINGEN MET 'EVIDENCE BASED POLICY AND PRACTICE' IN HET VERENIGD KONINKRIJK

Dr. P. van der Knaap

Peter van der Knaap is directeur bij de Algemene Rekenkamer. Deze bijdrage is geschreven op persoonlijke titel.

In het Verenigd Koninkrijk (VK) is de afgelopen vijf jaar op uitvoerige schaal ervaring opgedaan met 'evidence based policy and practice'. Een mooi, aansprekend concept, dat zich lastig laat vertalen. 'Bewezen beleid' komt misschien nog het dichtst in de buurt. Bij 'evidence based policy' draait het er om dat de uitkomsten van audits, inspecties en onderzoek (het 'bewijs') echt worden gebruikt bij het verbeteren van beleid en uitvoeringspraktijk. Welke lessen kunnen uit de Britse praktijk worden geleerd?

'BEWIJS' EN BELEID

Allereerst is het goed om stil te staan bij de betekenis van het woord 'evidence' of 'bewijs'. Een van de oudste principes van een correcte rechtspraak is 'onschuldig tot bewezen'. Iemand kan niet veroordeeld worden totdat 'bewezen' is dat hij of zij een bepaalde overtreding of misdaad heeft begaan. Het doel van dit uitgangspunt is om te voorkomen dat mensen ten onrechte tot het betalen van boeten worden veroordeeld of, erger, ten onrechte worden opgesloten. Zoals we in ons eigen land recentelijk op pijnlijke wijze hebben ervaren met de Schiedamse parkmoord, vraagt dat om betrouwbaar bewijsmateriaal.

Als we kijken naar beleid, lijkt het wel of een omgekeerd principe van toepassing is. Beleidsmaatregelen worden vaak verondersteld te werken, tenzij bewezen is dat het tegendeel waar is. Doeltreffend en doelmatig dus, tot anderszins bewezen. Eigenlijk is dat vreemd: we besteden veel geld aan maatregelen waarvan we niet weten of ze eigenlijk wel werken. Daarmee bestaat de kans dat het beleid irrelevant is of zelfs contraproductief uitpakt.

In wezen hebben we daarmee de bedoeling van 'evidence based policy and practice' te pakken: ervoor zorgen dat het bewijs dat beleid werkt (of niet werkt) wordt benut bij de beslissing of het beleid moet worden voortgezet, veranderd of beëindigd (oud-Financiën verloochent zichzelf nooit), zodat slecht beleid wordt gestopt en succesvol beleid een duwtje in de rug krijgt. Een goede bepaling die

in de Britse praktijk hier en daar al lijkt te werken, is het hanteren van evaluatietermijnen: als na een aantal jaar het bewijs dat beleid werkt niet geleverd is, wordt het stopgezet.¹

Overigens is het goed om te beseffen dat het woord 'evidence' aan de andere kant van de Noordzee een louter positieve connotatie heeft. Zeker wanneer het in combinatie met beleid en beleidspraktijk wordt gebruikt. Ik ben er niet zeker van dat dat ook voor Nederland geldt. Zeker wanneer het streven naar resultaatgericht beleid (VBTB) in verband wordt gebracht met de control- en auditfunctie is het wantrouwen niet van de lucht. Zoals De Jong laatst al in ditzelfde blad opmerkte, zijn beleidsprofessionals meer geneigd hun eigen overtuigingen te volgen als het gaat om 'goed beleid', dan de beheersstandaarden van de controllers en auditors.²

“EVIDENCE' HEEFT
AAN DE ANDERE KANT
VAN DE NOORDZEE
EEN LOUTER POSITIEVE
CONNOTATIE”

KWALITEIT VAN BEWIJS: CRITERIA EN DISCUSSIE

En dan: wat geldt als een bewijs? Wanneer is 'onomstotelijk aangetoond' dat beleid al dan niet werkt? Aan welke kwaliteitscriteria moet bewijs voldoen vooraleer we het kunnen toelaten in de arena van de bestuurlijke en politieke besluitvorming? En hoe zorgen we er vervolgens voor dat dat bewijs er komt en op de goede manier gebruikt wordt? Hoe zorgen we dat de bewijslast geen bewijsverlast wordt? En, ten slotte, welke rol is weggelegd voor de doelgroepen, de ontvangers van beleid?

1 Vergelijk Nutley, S., H. Davies & I. Walter: *Evidence based policy and practice. Cross sector lessons from the UK*. Working Paper 9, ESRC UK Centre for Evidence Based Policy and Practice.

2 F. de Jong (2005): VBTB: een proces van lange adem. MCA 2/2005.

Ik zal in het bestek van deze bijdrage niet iedere vraag langslopen, maar om met de eerste vraag te beginnen: wat telt als bewijs? Wat is bruikbare beleidsinformatie? In het rapport *Staat van de beleidsinformatie 2005* geeft de Algemene Rekenkamer aan dat bruikbare beleidsinformatie aan drie eisen voldoet: het is concreet geformuleerd, het geeft een relevant en valide beeld van de geleverde prestaties en de mate waarin de effecten zijn bereikt en het is betrouwbaar. Over met name dat laatste aspect is de afgelopen vijf jaar een verwoede discussie gevoerd over de eisen die je aan beleidsmatige verantwoordingsinformatie mag – of moet – stellen.

De Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid (RPE) uit 2002 hield het bij de klassieke kwaliteitseisen van wetenschappelijk onderzoek. De drie basis-eisen zijn:

- validiteit (van opzet van informatiesysteem of onderzoek; van gevolgtrekkingen);
- betrouwbaarheid en nauwkeurigheid (van de in informatiesysteem of onderzoek gebruikte meetmethoden);
- bruikbaarheid (van de uitkomsten van systeem of onderzoek).

Naast deze formele inhoudsnormen worden in de RPE ook procesnormen onderscheiden. De essentie daarvan is dat de uitkomsten van evaluatieonderzoek ordelijk en controleerbaar tot stand komen. Het moet mogelijk zijn om de kwaliteit en de onafhankelijkheid van evaluatieonderzoek aan de hand van een onderzoeksdossier vast te stellen.

Uit de VBTB-evaluatie komt naar voren dat het met de kwaliteit van evaluatieonderzoek bij de rijksoverheid niet goed gesteld is: “in minder dan de helft van de evaluaties is de betrouwbaarheid minimaal van voldoende kwaliteit”.³ Naar aanleiding hiervan werd door het ministerie van Financiën voorgesteld de eisen vooral te beperken tot dergelijke procesmatige eisen (een ordelijk en controleerbaar proces met onafhankelijke onderzoekers en/of begeleidingscommissie). De redenering is daarbij dat niet-financiële informatie niet vergelijkbaar is met financiële informatie: de mate van exactheid die bij financiële informatie wel mogelijk is, kan niet worden gerealiseerd bij het meten van maatschappelijke effecten en gerealiseerde beleidsprestaties. Financiën stelt daarom de volgende waarborgen voor:

- bij gepresenteerde beleidsinformatie wordt duidelijk de bron aangeven;
- tijdens de begrotingsbehandeling of een start Algemeen Overleg kunnen afspraken met de Tweede Kamer worden gemaakt over de kwaliteitseisen;
- in specifieke gevallen kan de Tweede Kamer de minister vragen om een onderzoek naar de kwaliteit van de beleidsinformatie, waarbij externe deskundigen worden betrokken;
- bij beleidsevaluaties en beleidsdoorlichtingen zullen externe onafhankelijke deskundigen worden betrokken;
- het proces van totstandkoming moet ordelijk, controleerbaar en deugdelijk zijn.

Er is volgens Financiën sprake van ordelijke, controleerbare en deugdelijke totstandkoming van beleidsinformatie, indien de verantwoordelijkheden en bevoegdheden goed in het proces zijn belegd en het proces achteraf reconstrueerbaar is. Voorts dient de informatie die als uitkomst van het proces wordt opgeleverd op volledige en juiste wijze in de verantwoording te worden opgenomen.⁴ De discussie hierover is nog niet afgerond.

DE BRITSE AANPAK

Hoe zit dat in het VK? Naar mijn bescheiden mening is het niet verassend dat de Britten wél vasthouden aan relevantie en betrouwbaarheid. Door middel van standaarden en checklists, peer reviews en het instellen van begeleidingscommissies wordt hier actief op gestuurd. Daarbij worden inhoudelijke vragen als ‘hoe valide is het onderzoeksontwerp’, ‘wat is de kwaliteit van de onderliggende gegevens’ of ‘mogen deze conclusies op basis van het materiaal wel worden getrokken’ niet geschuwd. Aangezien er veel op het spel staat wanneer beleidsinformatie (waaronder de uitkomsten van onderzoek) een rol gaat spelen in een echt verantwoordingsdebat, wordt er veel geïnvesteerd in het opleiden van onderzoekers. Maar ook beleidsmakers die onderzoek laten uitvoeren worden getraind in kennis en vaardigheden op onderzoeksgebied.

“HET IS NIET VERASSEND DAT DE BRITTEN WÉL VASTHOUDEN AAN RELEVANTIE EN BETROUWBAARHEID”

In het VK – maar ook in de Verenigde Staten trouwens – wordt daarnaast steeds meer gewerkt volgens de ‘bewijs-hiërarchie’ uit medisch onderzoek:

systematische meta-analyses
 experimenten met aselect getrokken experimentele groep en controlegroep
 quasi-experimenten (natuurlijke experimenten, tijdreeksonderzoeken)
 opinies van experts

3 Ministerie van Financiën (2005): Evaluatie ex post onderzoek (bijlage bij het eindrapport VBTB-evaluatie), Den Haag.

4 Minister van Financiën (2005): Brief aan de Tweede Kamer van 18 juli 2005, *Kamerstukken II* 2004/05, 29 949 en 29 950, nr. 34.

Ook in Nederland gaan steeds meer stemmen op voor een dergelijke, op de 'gouden standaard' van experimenteel onderzoek geënte, benadering van beleidsonderzoek.⁵ En daar valt veel voor te zeggen: zonder een dergelijke aanpak blijft het lastig doeltreffendheid en doelmatigheid *echt* aan te tonen. Daarbij is er vaak veel meer mogelijk dan onderzoeksbureaus vaak doen voorkomen. Een kanttekening daarbij is dat kwalitatief onderzoek, waarbij ook op meer explorerende wijze gekeken wordt naar verklaringen en beleidsalternatieven, niet als inferieur versleten wordt: beide hebben hun functie.

ENOUGH IS ENOUGH: DE EIS VAN 'SUFFICIENCY'

Naast relevantie en betrouwbaarheid geldt er in het VK nog een derde, boeiende eis: 'sufficiency' oftewel 'toereikendheid'. Wanneer is bewijs toereikend? Deze vraag is vooral boeiend omdat het ons bewustmaakt van de valkuil van perfectionisme. Zoals bekend is het betere vaak de vijand van het goede. Er bestaat nogal eens de neiging om te veel of te precies te willen meten. Dit wordt nog versterkt door de mogelijkheden van de toepassing van ICT.

Het laveren tussen 'te veel' en 'te weinig' komt mooi tot uitdrukking in het streven van de Britse Healthcare Commission. Men houdt bij het verzamelen van informatie expliciet rekening met de bevragsingslast van organisaties als ziekenhuizen: "The Healthcare Commission has to deal with the tension between ensuring consistent and robust judgements, and not burdening the inspected organisations with a heavier than necessary regime which distracts from, rather than enhancing, service improvement."⁶

Interessant is ook om te zien dat men in het VK steeds nadrukkelijker een onderscheid maakt tussen de verschillende doelgroepen, maar ook tussen verschillende doeleinden van de informatie. Waar beleidsinformatie wordt gebruikt in het kader van verantwoording, worden hogere eisen gesteld dan waar het eerste doel is om te komen tot meer inzicht voor beleidsmakers of een beroepsgroep zelf. Bij dat laatste wordt tegelijkertijd meer en meer het nut ingezien van het betrekken van de doelgroepen of klanten van beleid (zoals patiënten, onderwijzend personeel maar ook ouders en leerlingen).

BESLUIT: EVIDENCE IN CONTROL?

Ik meen dat validiteit en betrouwbaarheid een centrale rol verdienen in beleidsonderzoek. Maar ook in systemen die leiden tot beleidsinformatie. Het gaat om de eenvoudige vraag: meten we wat we willen meten en doen we dit op een betrouwbare en dus geloofwaardige manier. Beide eisen komen erop neer dat de gegevens waarop de niet-financiële informatie is gebaseerd, op adequate wijze moeten worden verzameld en bewerkt. Een goede toepassing van deze eisen leidt ertoe dat de kwaliteit (juistheid en volledigheid) van de informatie in het jaarverslag vol-

doende is gewaarborgd. Het niet naleven van deze eisen vergroot de kans op onjuiste en onvolledige informatie. De eisen voor ordelijke en controleerbare totstandkoming, zoals verwoord in de huidige regeling (de RPE), zijn nog steeds toepasbaar. Deze eisen kunnen gehandhaafd worden.

Het uitoefenen van 'control' over de mate waarin dit gebeurt, is in eerste instantie een management verantwoordelijkheid. Bij het sociale mandaat dat een overheidsorganisatie nu eenmaal heeft, hoort het streven naar goed geïnformeerd beleid. Je zou dat – om een vergelijking met maatschappelijk verantwoord ondernemen te maken – wetenschappelijk verantwoord beleid kunnen noemen. De controlfunctie zou erop moeten toezien dat het management die verantwoordelijkheid ook werkelijk neemt. Daarbij gaat het er niet om dat de controller elke stap in het onderzoeksproces als een methodologisch expert nagaat. Wel moet hij erop letten dat door het management afdoende wordt gestuurd op de randvoorwaarden die gelden voor de totstandkoming van betrouwbare en valide beleidsinformatie. Een kritische blik op de vraagstelling van onderzoek, de selectie en omvang van de onderzoekspopulatie en de relatie tussen feitenmateriaal en conclusies, maar ook het nagaan of er beredeneerd is gekozen voor bepaalde methoden en technieken, horen daar zeker bij.

CONTROL OP EVIDENCE BETEKENT HET SCHERP HOUDEN VAN BELEIDSDIRECTIES

In de praktijk zie je dat dit zo kan werken. Zo heeft de accountantsdienst van het ministerie van Economische Zaken (EZ) de betrouwbaarheid vastgesteld van de totstandkoming van de bij EZ gebruikte prestatie-indicatoren. Het werk van de controllers heeft er in de eerste jaren van VBTB toe bijgedragen dat het percentage van de prestatie-indicatoren die betrouwbaar tot stand komen, is toegenomen van 36% in 2001 naar 94% in de begroting 2003.⁷ Control op evidence betekent in mijn ogen dan ook geen extra bureaucratie, maar het scherp houden van beleidsdirecties. En dat heeft weer alles te maken met het serieus werk maken van sturen op resultaten. Waarbij ten slotte geldt dat er een enorm verschil bestaat tussen informatie over prestaties en informatie over effecten van overheidsbeleid. Als een organisatie zijn prestaties – en daarmee de informatie over prestaties – niet beheerst, is dat in mijn ogen voldoende evidence voor de veroordeling: u bent niet in control.

MCA

- 5 Zie bijvoorbeeld Leuven, E., M. Lindahl, H. Oosterbeek & D. Web-bink: *New evidence on the effect of time in school on early achievement*. Örebro University, www.econpapers.repec.org.
- 6 *Evidence for accountability. Background Paper for the Madingley Hall Conference*, September 2005.
- 7 *Kamerstukken II 2004/05*, 29 949, nr. 4.