



- beleidsevaluatie
- niet-financiële beleidsinformatie
- evidence-based policy
- VBTB

# Ondoelmatig tot het tegendeel bewezen is?

## 'Evidence-based policy and practice': wat telt als bewijs?

*Evidence-based policy* is een mooi, aansprekend concept, dat zich wat lastig laat vertalen. 'Onderbouwd beleid' komt dicht in de buurt, maar misschien drukt 'verantwoord beleid' de bedoeling nog beter uit. Bij evidence-based policy draait het er namelijk om dat de uitkomsten van audits, inspecties en vooral beleidsonderzoek (het 'bewijs') daadwerkelijk worden gebruikt bij het verbeteren van beleid. In het Verenigd Koninkrijk is de afgelopen vijf jaar op uitgebreide schaal ervaring opgedaan met dit model.

Peter van der Knaap

Sinds 1999 heeft het Verenigd Koninkrijk een kleine 3 miljoen pond geïnvesteerd in het creëren van een zogenaamd Evidence Network. In totaal hebben zo'n 700 onderzoekers aan het project meegewerkt, gecoördineerd door het Centre for Evidence Based Policy and Practice. De overheid (op nationaal niveau in Londen, Edinburg en Cardiff, maar ook op lokaal niveau) onderstreept het belang ervan. Welke lessen kunnen wij uit de Britse praktijk trekken? En de hamvraag: wat telt als bewijs?

### Uit de *comfort zone* met kritische beleidsinformatie

Als het om beleid gaat, lijken politici en beleidsmakers het liefst in een soort

reservaat (de Engelsen spreken van een *comfort zone*) te blijven, waar globaal de volgende regels gelden:

1. We veronderstellen dat bestaand beleid werkt.
2. We nemen aan dat beleid dat bij probleem a lijkt te werken, ook bij probleem b zal werken.
3. We geloven dat maatregelen die eerder succesvol zijn bevonden, ook in de toekomst succesvol zullen zijn.
4. Horen we geen nieuws over tegenvallende resultaten, dan beschouwen we dat als goed nieuws.
5. Horen we onverhoopt toch geluiden over tegenvallende resultaten, dan zien we dat als een teken dat er meer geld moet worden uitgegeven (aan het opstellen van regels en het aanstellen van ambtenaren) om het

beleid alsnog tot een succes te maken.

Kort en goed: beleidsmaatregelen worden verondersteld te werken, tenzij het tegendeel bewezen is. Ze zijn doelmatig totdat anderszins is gebleken. Eigenlijk is dat raar.

### Nieuwe basisregel

We besteden veel belastinggeld aan maatregelen waarvan we niet weten of ze eigenlijk wel werken. De kans bestaat dat het beleid irrelevant is of zelfs contraproductief uitpakt. En daarmee zijn we beland bij de bedoeling van 'evidence-based policy and practice': ervoor zorgen dat het antwoord op de vraag of beleid werkt wordt gebruikt bij een nieuw besluit over dat beleid. Zodat slecht beleid wordt gestopt en succesvol beleid een duwtje in de rug krijgt. Vrij vertaald is het de bedoeling dat de regels van de *comfort zone* worden omgedraaid. De nieuwe basisregel luidt: we gaan ervan uit dat bestaand beleid niet werkt (ondoelmatig is) totdat het tegendeel is bewezen.

Dit uitgangspunt van evidence-based policy komt mooi tot uitdrukking in het White Paper on Modernising government, waarin de Britse regering haar plannen in 1999 als volgt verwoordde: "This Government expects more of policy makers. More new ideas, more willingness to question inherited ways of doing things, better use of evidence and research in policy making and better focus on policies that will deliver long term goals."<sup>1</sup> Al snel werd aan

### Samenvatting

In het Verenigd Koninkrijk is de afgelopen jaren op uitgebreide schaal ervaring opgedaan met *evidence-based policy and practice*. Het doel van dit project was om bestaande inzichten over de effectiviteit van beleid beter te benutten bij het herzien van beleidsmaatregelen. Uit de Britse ervaring blijkt dat het geen overbodige luxe is om te investeren in de kwaliteit van beleidsonderzoek. Aan de andere kant is het zaak om te waken voor een te perfecte bewijsvoering: perfectionisme of cijferfetisjisme kan doelmatigheid ook in de weg staan. Controllers moeten zich van dit dilemma bewust zijn.

het uitgangspunt dat beleid zich moest bewijzen de eis toegevoegd dat ook de uitvoering zo goed mogelijk 'geïnfomeerd' moest zijn over successen en mislukkingen.

## Eigenlijk is het raar: we besteden veel belastinggeld aan maatregelen waarvan we niet weten of ze eigenlijk wel werken

De afgelopen jaren is het ideaal van evidence-based policy door velen omarmd. Onderzoeksinstituten zagen er uiteraard een mooie markt in. Verrassender is de steun van bestuurders en politici die oprecht geïnteresseerd blijken te zijn in 'beleid dat werkt'. In het Verenigd Koninkrijk wordt deze hernieuwde aandacht voor *what works* toegeschreven aan het afnemende vertrouwen in politiek en bestuur: wie publiek geld besteedt aan beleid heeft er belang bij dat het succes van maatregelen ook echt aangetoond kan worden.

### Maar wat telt als bewijs?

Tot zover de veelbelovende perspectieven. Maar wat geldt eigenlijk als bewijs? Wanneer kunnen we zeggen dat beleid werkt? Aan welke kwaliteitscriteria moet bewijs voldoen voordat we het kunnen toelaten in de arena van de bestuurlijke en politieke besluitvorming?

Evidence, bewijs. De vergelijking met de rechtspraak dringt zich op. Een oud rechtsbeginsel luidt dat iemand onschuldig is totdat het tegendeel is bewezen. Een verdachte kan niet worden veroordeeld voordat overtuigend bewezen is dat hij of zij een bepaalde overtreding of misdaad heeft begaan. Dit principe is bedoeld om te voorkomen dat onschuldigen worden veroordeeld. Zoals we in Nederland hebben ervaren in de zaak van de Schiedamse parkmoord, vraagt een veroordeling om betrouwbaar bewijsmateriaal.

### Klassiek rijtje

Als het om wetenschappelijk onderzoek gaat, vormen betrouwbaarheid, repro-

duceerbaarheid en validiteit een klassiek rijtje. Ook aan gene zijde van de Noordzee geldt dat "the use of pre-set criteria is fairly widely accepted by researchers using quantitative methods."<sup>2</sup> Het is van

belang dat de gebruikers van onderzoeksresultaten ervan uit kunnen gaan dat de onderzoekers deze kwaliteitseisen in hun onderzoeksproces hebben geborgd. Door middel van standaarden en checklists, *peer reviews* en het instellen van begeleidingscommissies wordt hierop in het Evidence Network actief gestuurd. En hoewel ze nog niet overall even goed worden toegepast, 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. Daarnaast wordt er veel geïnvesteerd in het opleiden van onderzoekers. Maar ook beleidsmakers die onderzoek laten uitvoeren worden getraind in kennis en vaardigheden op onderzoeksgebied.

### Bewijshiërarchie

In het Verenigd Koninkrijk wordt verder – net als in de Verenigde Staten – steeds meer gewerkt volgens de 'bewijshiërarchie' uit het medisch onderzoek:

- systematische meta-analyses
- experimenten met een a-selecte experimentele groep en controlegroep
- quasi-experimenten (natuurlijke experimenten, tijdreeksonderzoeken)
- opinions van experts

Na een periode waarin met name kwalitatief onderzoek een opmars maakte, gaan steeds meer stemmen op voor een op de 'gouden standaard' van experimenteel onderzoek geënte benadering van beleidsonderzoek. Ook in ons land. Onlangs nog pleitten Frans Leeuw (directeur van het WODC van het ministerie van Justitie) en André de Jong (directeur-generaal van de Rijksbegroting) op een symposium over beleidsonderzoek gebroederlijk voor meer experimenteel onderzoek.<sup>3</sup> En gelijk hebben ze: zonder het determineren van de *counterfactual* (wat zou er zijn gebeurd zonder beleid?) mogen we geen harde causale uitspraken over de doeltreffendheid van beleid doen. Daarbij is er veel meer mogelijk dan onderzoeksbureaus vaak doen voorkomen. Een kanttekening daarbij is dat kwalitatief onderzoek, waarbij ook gekeken wordt naar verklaringen en beleidsalternatieven, zeker niet als inferieur beschouwd mag worden.

teur-generaal van de Rijksbegroting) op een symposium over beleidsonderzoek gebroederlijk voor meer experimenteel onderzoek.<sup>3</sup> En gelijk hebben ze: zonder het determineren van de *counterfactual* (wat zou er zijn gebeurd zonder beleid?) mogen we geen harde causale uitspraken over de doeltreffendheid van beleid doen. Daarbij is er veel meer mogelijk dan onderzoeksbureaus vaak doen voorkomen. Een kanttekening daarbij is dat kwalitatief onderzoek, waarbij ook gekeken wordt naar verklaringen en beleidsalternatieven, zeker niet als inferieur beschouwd mag worden.

### Toereikendheid

Naast relevantie en betrouwbaarheid geldt in het Verenigd Koninkrijk nog een derde eis: *sufficiency* oftewel 'toereikendheid'. Wanneer is bewijs toereikend? Wanneer is het *fit for purpose*? Deze vraag is vooral interessant omdat zij ons bewust maakt van de valkuil van perfectionisme of cijferfetisjisme: onderzoekers maar ook anderen hebben soms de neiging om te veel of te precies te willen meten. Een neiging die nog wordt 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 zoals 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."<sup>4</sup>

Relevant is ook dat men in het Verenigd Koninkrijk steeds nadrukkelijker een onderscheid maakt tussen de verschillende doelgroepen en doeleinden van de informatie. Waar beleidsinformatie wordt gebruikt in het kader van verantwoording, worden hogere eisen gesteld

dan wanneer 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 het beleid (zoals patiënten, onderwijzend personeel, ouders en leerlingen).

### Wat telt als bewijs? De Haagse discussie

De Britse ervaringen zijn rijk. Wie een vergelijking wil maken met Nederland, moet beseffen dat het woord *evidence* in Britse oren vooral een positieve klank lijkt te hebben, zeker in combinatie met beleid en beleidspraktijk. Ik ben er niet zeker van dat dit ook voor Nederlandse oren geldt. Als het streven naar resultaatgericht beleid (VBTB) in verband wordt gebracht met de control- en auditfunctie, is het wantrouwen over controletorens en bureaucratie niet van de lucht.

Wat telt in Nederland dan als bewijs? Over de mate waarin beleidsinformatie betrouwbaar kan en moet zijn, is de afgelopen vijf jaar op en rond het Haagse Binnenhof een verwoede discussie gevoerd. Als het om informatie gaat die gebruikt wordt bij begroting en verantwoording, houdt de Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid (RPE) uit 2002 het bij de klassieke kwaliteitseisen van wetenschappelijk onderzoek:

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

### Reconstructie-eisen

In de regeling wordt uitdrukkelijk aangegeven dat deze eisen onlosmakelijk zijn verbonden met het totstandkomingsproces. Naast deze inhoudelijke proceseisen worden ook 'reconstructie-eisen' vermeld. 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, eventueel door – bij wijze van uiterste controle – het onderzoek over te doen (reproduceerbaarheid).

In 2004 zijn VBTB en de RPE geëvalueerd. Uit die evaluatie komt naar voren dat het met de kwaliteit van evaluatieonderzoek bij de rijksoverheid nog niet goed gesteld is. Enigszins wollig geformuleerd: "In minder dan de helft van de evaluaties is de betrouwbaarheid minimaal van voldoende kwaliteit."<sup>5</sup>

## Wanneer is bewijs toereikend? Deze vraag is vooral interessant omdat zij ons bewust maakt van de valkuil van perfectionisme of cijferfetisjisme

Verder wordt er bedroevend weinig echt doeltreffendheidsonderzoek uitgevoerd. Naar aanleiding hiervan heeft het ministerie van Financiën voorgesteld om de eisen vooral toe te snijden op procesmatige zaken:

- Bij gepresenteerde beleidsinformatie wordt duidelijk de bron aangegeven.
- Tijdens de begrotingsbehandeling of bij aanvang van een 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 ordelijk, controleerbaar en deugdelijk totstandkomen van beleidsinformatie, indien de verantwoordelijkheden en bevoegdheden goed in het proces zijn belegd en het proces reconstrueerbaar is. Voorts dient de informatie die als uitkomst van het proces wordt opgele-

verd op volledige en juiste wijze in de verantwoording te worden opgenomen.<sup>6</sup>

Dat deze thematiek ook in de Tweede Kamer nauwlettend wordt gevolgd, bleek wel uit de motie-Mastwijk van 31 maart 2005. In reactie op de voorstellen werd de regering opgeroepen ervoor te zorgen dat de departementale auditdiensten per beleidsartikel een oordeel geven over de kwaliteit van het proces van totstandkoming van de niet-financiële en derdeninformatie. Om dat mogelijk te maken zou de minister van Financiën een normenkader moeten ontwikkelen "dat de departementale auditdiensten

kunnen gebruiken bij dat oordeel, welk normenkader gebaseerd is op de vraag of het proces dat tot die informatie heeft geleid van deugdelijke kwaliteit is, voldoende waarborgt dat die informatie juist en volledig is, en niet in strijd is met de financiële informatie."<sup>7</sup>

### Tot besluit: is evidence in control?

Voor ingewijden is het boeiend om de discussie tussen Tweede Kamer, Financiën en Algemene Rekenkamer te volgen. Binnenkort, bij het verschijnen van de nieuwe versie van de regeling voor evaluatieonderzoek, zal er duidelijkheid zijn over de uitkomst. Vooral de verwachtingen rond het nieuwe instrument 'beleidsdoorlichtingen' zijn hooggespannen.

Nog veel boeiender is de betekenis van betrouwbare beleidsinformatie voor een goed openbaar bestuur en effectief beleid. Want we moeten niet vergeten: het gaat steeds om de eenvoudige vraag of we meten wat we willen meten en of we dat op een geloofwaardige manier doen. 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 voldoende is gewaarborgd. Waarbij geldt dat echt goede onderzoekers dat natuurlijk al uit zichzelf doen.

#### **Wetenschappelijk verantwoord**

Het uitoefenen van 'control' over de mate waarin bewust gestuurd wordt op het verzamelen en benutten van goede beleidsinformatie, is naar mijn smaak in eerste instantie een verantwoordelijkheid van het management. Bij het sociale mandaat dat iedere 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 er vervolgens op 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, maar dat hij of zij erop let dat het management afdoende stuurt op de randvoorwaarden die gelden voor de totstandkoming van

betrouwbare en valide beleidsinformatie. Een kritische blik op de vraagstelling van het onderzoek, de omvang van de onderzoekspopulatie en de relatie tussen feitenmateriaal en conclusies, maar ook nagaan of er beredeneerd is gekozen voor bepaalde methoden en technieken – dat alles hoort er wat mij betreft zeker bij.

In de praktijk zie je her en der dat dit 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. Control op evidence houdt dus ook in: 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 haar prestaties – en daarmee de informatie over prestaties – niet beheerst, is dat in mijn ogen voldoende evidence voor de uitspraak: 'u bent niet in control!'

Dr. P. van der Knaap is directeur bij de Algemene Rekenkamer en redacteur van TPC. Hij woonde het afgelopen najaar in Cambridge een besloten werkconferentie van het UK Centre for Evidence Based Policy and Practice bij. Deze bijdrage schreef hij op persoonlijke titel.

#### **Noten**

1. William Solesbury, *Evidence Based Policy: Whence it Came and Where it's Going*, ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 1.
2. Annette Boaz & Deborah Ashby, *Fit for purpose? Assessing research quality for evidence based policy and practice*, ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 11.
3. Conferentie over *(Quasi-)Experimental Evaluations and Dutch Society: Trends and developments in criminal justice and prevention policies, social integration and education*, 4 november 2005, georganiseerd door het WODC van het ministerie van Justitie.
4. *Evidence for accountability. Background Paper for the Madingley Hall Conference*, Cambridge, september 2005.
5. Ministerie van Financiën, *Evaluatie ex post onderzoek*, bijlage bij het eindrapport VBTB-evaluatie, Den Haag, 2005.
6. Minister van Financiën, brief aan de Tweede Kamer van 18 juli 2005, TK 2004-2005, 29949 en 29950, nr. 34.
7. TK 2004-2005, 29949, nr. 11.