de Evaluatie

Laat Rekenkamer eens kijken naar het nut en de kosten van die zwembaden vol evaluaties

Posted on 14 augustus 2025 by Gastauteur

*Door Frans L. Leeuw**

Evalueren van overheidsbeleid heeft in Nederland en andere OESO-landen een geschiedenis van zo'n 50-60 jaar. In de VS was Johnsons Great Society-politiek (1964) met talloze nieuwe wetten en interventies de aanjager van de eerste 'golden years' van dit onderzoeksterrein. Decennia lang was de VS ook nummer 1 in de toptien van de internationale evaluatie-atlas. Opvallend is dat het nu ook de VS is waar vanaf begin dit jaar zwaar bezuinigd wordt op evaluaties, maar ook op het toezicht/inspectie-wezen.

Van Willigenburg schreef in zijn Bericht uit de Samenleving dat 'we ons kapot evalueren' en maakt melding van het bestaan van een Evaluerende Klasse: zij die er dagelijks mee bezig zijn en er een goede boterham aan overhouden. Waar hij een heel breed evaluatie-begrip hanteert waaronder het beoordelen of de maat nemen

van ‘de kroeg, een film, een concert of populaire tv-programma’s’, gaat deze bijdrage over beleidsevaluaties: zijn er wel effecten van beleid in zijn diverse uitingen, welke dan, bij wie en voor hoe lang, zijn er ook (ongewenste) neveneffecten en is er sprake van doelmatigheid van de interventies, en van het evalueren?

Beleidsevaluaties zijn *booming business*

In een vergelijkend onderzoek (Sridharan & Leeuw, in uitvoering) gaan we ten eerste na of de wereld van de beleidsevaluaties (in Nederland en een aantal OESO-landen) als *booming business* te kenschetsen is. Dat geluid wordt vaak gehoord. Ons tel- en rekenwerk van aantallen publicaties, evaluatoren, instituten, centra, beroepsverenigingen en leden dat we voor de periode 1960-2024 uitvoerden, bevestigt dat geluid: in vergelijking met de jaren ’60-’70, toen er hooguit en met uitzondering van de VS, toen de beleidsevaluaties nog in de kinderschoenen stonden, is er nu sprake van een *booming business*.

De onlangs overleden Andries Hoogerwerf verweet de Nederlandse rijksoverheid 50 jaar geleden veel te weinig aandacht te hebben voor het evalueren van beleid, ‘een situatie die in een beschaafd land eigenlijk niet mag voorkomen’. De Algemene Rekenkamer begon in de jaren ’80 met rijksbrede onderzoekingen van diverse overheidsinterventies en was keer op keer kritisch over de te geringe aandacht ervoor.

Van dat beeld is niets meer over. Integendeel, meer en meer wordt over ‘evaluiitis’, ‘evaluation overload’ en de ‘evaluation society’ gesproken. Onderdeel daarvan is – sjiek uitgedrukt – de ‘institutionalisering’: talloze kaders, standaarden, criteria, richtlijnen, procedures, protocollen, evaluatiebeleid en regelgeving, afkomstig van overheden, VN-organisaties, ontwikkelingsbanken, OESO en NGO’s. Doorgaans gaat die ontwikkeling gepaard met hoera-geluiden: hoe meer institutionalisering, des te beter (voor ons allemaal).

Wat vergeten wordt is dat deze ontwikkeling een sterk (quasi)-gouvernementeel, ‘governance’-gericht karakter heeft, waar al in de jaren ’70 en ’80 voor gewaarschuwd is. Zonder veel effect dus. Evert Vedung uit Zweden is een van de weinigen die zich daar recent kritisch over uitliet. Is institutionalisering wel zo goed als gedacht wordt? Neen, zegt hij, er zijn neveneffecten zoals dat evaluaties routines worden, heel gemakkelijk door degenen die geëvalueerd worden bespeeld

kunnen worden en zelfs bedrog kunnen opleveren.

Donkere kanten

In onze vergelijkende studie analyseren we deze en een aantal andere donkere kanten. Eén eerste fenomeen zijn ‘evaluation machines’: periodiek draaiende metingen van diverse organisaties over van alles en nog wat. Vaak gaat het over beleid of over het functioneren van organisaties of hoe ‘de samenleving’ bepaalde ontwikkelingen beoordeelt. Ze komen voort uit wet- of regelgeving, internationale afspraken, ‘Agreements’, ‘Declarations’ en dergelijke.

Ooit was zo’n aanpak verstandig, maar het machinematige en soms bijna dwangmatige karakter ervan dreigt in het tegendeel daarvan te verkeren. Overheden maken er gebruik van om bij niet-naleving of te weinig ondersteuning van hun beleid door de bevolking, te pleiten voor nieuwe (en ‘betere’) interventies (of voor zogenaamde maatwerkbenaderingen), die vervolgens ook weer via deze ‘evaluatiemachines’ onderzocht worden. Het draagt allemaal bij aan routinisering, meer-van-het-zelfde en te weinig oog voor verschillende contexten waar beleid op probeert in te grijpen.

De onafhankelijkheidsparadox

Een ander zwarte rand-fenomeen is georkestreerde onafhankelijkheid: talloze zogenaamde waarborginstrumenten, zoals richtlijnen, gedragscodes, mandaten, klachtenregelingen, raden van advies/toezicht en ‘morele beraden’ creëren een beeld van onafhankelijkheid. Dat is belangrijk voor de reputatie, overlevings- en verdiencapaciteit van de professie, want wie daar niet aan voldoet, wordt snel als ‘gekke Henkie’ weggezet. Evaluaties en evaluatoren moeten immers (totaal) onafhankelijk zijn, maar omdat het moeilijk, zo niet onmogelijk is dat in de echte wereld met zijn diverse (semi-) publieke financieringsstromen en afhankelijkheden te realiseren, wordt aan beeldvorming gedaan via de weg van de orkestratie. Wim Derksen sprak in dit verband wel over het beeld dat onderzoekers (zoals evaluatoren) als ‘onafhankelijke helden’ worden gezien en beleidsambtenaren als ‘bedrieglijke paladijnen’.

Een ander probleem met orkestratie is dat zij die geëvalueerd worden, in staat zijn de onderzoekingen te bespelen en daarmee ook te anticiperen op wat (vermoedelijk) gewenste antwoorden zijn. Erger is dat de onafhankelijkheid van het

gedrag van evaluatoren nauwelijks verband houdt met de georkestreerde ‘onafhankelijkheid’ op papier. Niet voor niets bestaat er zoiets als de onafhankelijkheidsparadox.

Ten derde is er een verschijnsel dat zijn oorsprong vindt in wat verstrengeling heet van (gecoöpteerde) toezichthouders en de sector oftewel: *regulatory capture*. Daar hebben evaluaties ook mee te maken en dan heet het ‘evaluation capture’. Doorgaans zijn het evaluatoren, criminologen en andere onderzoekers die ‘regulatory capture’ ontmaskeren. Nu heeft de *booming business* van het evaluatiewerk er zélf mee te maken.

Het is het verschijnsel waarin evaluaties en evaluatoren omringd zijn door protocollen, richtlijnen, aanbestedingseisen, normen, criteria, sjablonen, toezicht en beoordelingen van nationale en supranationale organisaties, overheden, NGO’s en agentschappen die evaluaties aanvragen en (mede) financieren. Deze institutionalisering binnen bureaucratieën heeft de verkeerde prikkels gecreëerd voor evaluatoren, degenen die geëvalueerd worden en voor beleidsmakers: de prikkel is namelijk niet om primair de beste beleidsmaatregelen te ontwikkelen die het meeste bijdragen aan het verbeteren van de positie van burgers en hun organisaties, maar soms vooral om evaluatiesystemen te hebben die bruikbaar zijn om blunders, ‘foutje-bedankt’-gedrag, tegenvallers en vergelijkbare situaties te verhelpen en mede langs die weg ook reputatierisico’s voor bestuurders, politici, managers en sponsors te verminderen. Er is immers geëvalueerd. Dat het soms goed uitpakt en evaluaties wel bijdragen aan het realiseren van kansrijk beleid, is mooi meegenomen.

Vragen naar wat het kost is ‘kleinzielig’

Ten vierde en ten slotte: het verdwijnen van wat de ‘spulleboel’ aan evaluaties kost. Waar in de (internationale) overzichten van criteria en normen ‘transparantie’ en ‘publieksvriendelijkheid’ van evaluaties hoog in het vaandel staan, worden antwoorden op de vraag naar wat het kost, afgedaan óf als ‘gegevens-zijn-niet-beschikbaar, jammer maar helaas’, of als ‘zoek het zelf maar uit’ via de website van de zus-of-zo-organisatie (waarbij dan blijkt dat daar nauwelijks wat te vinden is en zeker niet over een langere periode).

Wat ik *evaluation hubris* noem, kan helpen dit fenomeen te verklaren: vragen naar geld is kleinzielig en misschien wel irrelevant. Immers, wie zou nu twijfelen aan de

doelmatigheid van het evaluatie-instrument of van toezicht en inspectie? Die begrijpt simpelweg niet dat het dáar dus niet over moet gaan. Furubo & Vestman vragen zich dat juist wél af: zijn evaluatoren wel altijd de 'good guys'? Anderen willen weten hoe het zit met de toegevoegde waarde van het huidige evaluatie-fenomeen en of het gebrek aan kennis daarover wellicht verband houdt met negatieve oordelen over de 'evaluatie-fabriek'. Ik stelde die vraag zelf voor het eerst in 2008.

Enfin, zo zijn er meer kritische kanttekeningen te maken. Het boek waar Sridharan en ik aan werken gaat op deze en andere zaken in, maar wil ook een bijdrage leveren aan de wijze waarop het serieuze en hoogwaardige evaluatiewerk te handhaven is maar dan zonder de donkere kanten.

Zwembaden vol routine-evaluaties

Er bestaan serieuze evaluaties van niveau, waar pittige vragen gesteld worden, daadwerkelijke onafhankelijkheid is, aandacht is voor theorieën die het geëvalueerde kunnen verklaren, transparantie bestaat over methodische aspecten en tegelijkertijd bescheidenheid in oordelen is. En er zijn zwembaden vol met routinematig werk die – gezien het volume ervan – ongetwijfeld ook nog veel geld kosten. Het zou mooi zijn als de Algemene Rekenkamer eens naar het Nederlandse beleid ten aanzien van evaluaties (en toezicht) zou kijken en ook de vraag naar het financiële belang ervan zou stellen.

Dat deed ze eerder in 1991 al een keer en het leverde een beperkte schatting, gebaseerd op weinig gegevens, op. Nut en impact van evalueren en niet alleen voor het overheidsapparaat zelf, zou ook bij zo'n exercitie horen maar ook het traceren van onbedoelde neveneffecten.

*****Frans L. Leeuw*** *is emeritus-hoogleraar aan de universiteiten van Maastricht en Utrecht en oud-directeur van de Algemene Rekenkamer en het WODC.*

Wynia's Week *brengt broodnodige, onafhankelijke berichtgeving: drie keer per week,* ***156 keer per jaar****, met artikelen en columns, video's en podcasts. Onze donateurs maken dat mogelijk. Doet u mee? Hartelijk dank!*