

ARTIKEL

Beleidsvaluatie in de praktijk: omgaan met uitdagingen en belangen*

Frank Bongers**

Inleiding

Beleidsvaluaties leunen bij voorkeur op economische en sociaalwetenschappelijke onderzoeksmethoden. Dat sluit aan bij het streven om beleid met wetenschappelijke kennis te onderbouwen (zie o.a. Derksen, 2011). Evaluatiestudies helpen te bepalen of beleid zal gaan werken, werkt of heeft gewerkt. Uiteraard moet daarbij worden voldaan aan bepaalde technieken om te leiden tot valide en betrouwbare uitspraken over beleidseffecten.

Het terechte streven om kennis voor overheidsbeleid te baseren op bepaalde methoden en technieken maakt het verhaal over evaluaties niet compleet. Evalueren vraagt in de praktijk om parate kennis en vaardigheden die niet vanzelfsprekend onderdeel uitmaken van de meer op onderzoeksmethoden gerichte handleidingen en opleidingen. Evaluaties gebeuren in een maatschappelijke en politieke omgeving met uiteenlopende en soms conflicterende belangen en verwachtingen. Een politieke omgeving maakt het voor beleidsonderzoekers een uitdaging (maar niet onmogelijk) om te komen tot een betrouwbaar resultaat.

In de afgelopen jaren heb ik als onderzoeker ervaring opgebouwd met praktische belemmeringen en belangen in beleidsvaluaties. Daar wil ik bij stilstaan in deze beschouwing.¹ Het eerste deel van deze beschouwing reflecteert (niet uitputtend) op enkele *praktische* uitdagingen en hoe een onderzoeker daarmee in de praktijk om kan gaan. Een eerste uitdaging is dat evaluaties wat betreft aantallen, scope en timing niet altijd aansluiten bij de informatiebehoeften van beleidsmakers. Evaluatoren zullen voorts bij het uitwerken van een onderzoeksopzet (niet geheel onverwacht) ervaren dat beleidsmakers doorgaans geen of weinig rekening houden met een (latere) evaluatie. De laatste uitdaging in dit deel betreft toenemende verwachtingen van beleidsmakers (als opdrachtgever van evaluaties) ten aanzien van kwantificering in en digitalisering van evaluaties.

* De auteur bedankt zijn collega's Pim den Hertog en Max Kemman voor hun suggesties bij een conceptversie van dit artikel.

** Frank Bongers is principal consultant bij Dialogic te Utrecht. Hij verricht beleidsvaluaties in onder meer de domeinen van economie, wetenschap, innovatie en cultuur. Hij is tevens lid van de rekenkamer Heusden.

1 Voor methodologische kennis over beleidsvaluaties wordt verwezen naar de literatuur ter zake (zie o.a. Van der Knaap et al., 2020; Van Hoesel & Herold, 2024; Van Thiel, 2021) of de Toolbox Beleidsvaluaties, zie <https://evaluaties.rijksfinancien.nl/toolbox>.

Frank Bongers

Het tweede deel van deze beschouwing staat stil bij het feit dat er natuurlijk meer politieke belangen spelen die de keuzeruimte voor een evaluator beperken. Het omgaan met politieke belangen in beleidsevaluaties vraagt om een vaardigheid om partijen te kunnen overtuigen van de juistheid van de keuzes die een evaluator in zijn (of haar) onderzoek maakt. Het negeren van deze belangen kan de doorwerking van een onderzoek beperken. Dit tweede deel staat daarom apart stil bij doorwerking.

Deel 1: Uitdagingen in de evaluatiepraktijk

Uitdaging 1: Meer evaluaties leiden niet vanzelf tot meer kwaliteit en samenhang

Het aantal evaluaties op rijksniveau neemt toe. De rijksoverheid houdt een overzicht bij van geplande en uitgevoerde evaluaties.² Een belangrijke oorzaak voor deze toename schuilt in wetgeving, strategische evaluatieagenda's of subsidiebeschikkingen, in het institutionaliseren van evaluatiefuncties binnen overheidsinstellingen in de vorm van afdelingen, programma's en evaluatiecoördinatoren, en in de professionalisering van het vak evalueren zelf.

In de Kaderwet zelfstandige bestuursorganen (zbo) staat bijvoorbeeld dat de minister elke vijf jaar een verslag aan beide kamers der Staten-Generaal zendt ten behoeve van de beoordeling van de doelmatigheid en doeltreffendheid van het functioneren van een zelfstandig bestuursorgaan (art. 39 lid 1). De Subsidieregeling instituten voor toegepast onderzoek bepaalt dat de subsidieontvanger medewerking verleent aan een evaluatie van de effecten van de door hem op grond van deze regeling uitgevoerde activiteiten (art. 10). Ook de Regeling Periodiek Evaluatieonderzoek (RPE) is een goede illustratie van een evaluatieverplichting. Dan zijn er nog allerlei (lokale en provinciale) rekenkamers, provincies en gemeenten, onderzoeksbureaus en kennisinstituten die op eigen initiatief of in opdracht evalueren.

Het gaat vaak om aparte evaluaties van instrumenten in een breder beleidsterrein. Beleidsdoorlichtingen – te beschouwen als meta-evaluaties van meerdere evaluaties binnen een artikel van de rijksbegroting in een specifieke periode – pogen het inzicht in de bijdrage van individuele instrumenten aan overkoepelend beleid te verbeteren. Deze doorlichtingen (veelal gebaseerd op de eerdergenoemde RPE) tonen dat evaluaties binnen een begrotingsartikel onderling vaak weinig samenhang vertonen. Hierdoor wordt de kans gemist om (resultaten van) evaluaties te koppelen aan de bredere doelstellingen in een bepaald domein (bijvoorbeeld via een overkoepelende beleidstheorie en/of een gezamenlijke evaluatieaanpak).

Een periodieke evaluatieverplichting (zoals een beleidsdoorlichting) heeft wel als voordeel dat beleid structureel doorgelicht en bijgestuurd kan worden en dat beleidsmakers en evaluatievoorwerpen (bijvoorbeeld een agentschap of instituut) anticiperen op een aanstaande evaluatie (door het formuleren van duidelijke doe-

2 Zie <https://evaluaties.rijksfinancien.nl>.

len, het samenstellen van beleidstheorieën, het bijhouden van kwantitatieve en kwalitatieve gegevens, het opstellen van zelfevaluaties, het ontwikkelen van evaluatie-expertise, etc.). Het voorkomt ook een stilzwijgende continuering van beleid zonder inachtneming van een veranderende omgeving.

Evaluatoren en overheden staan hier voor de uitdaging om (meer) samenhang te bevorderen in het groeiende aanbod van (verplichte) evaluaties. Overheden kunnen er op sturen dat aparte/kleinere instrumentevaluaties (beter) gekoppeld worden aan bredere doelstellingen in een beleidsterrein. Evaluatoren kunnen daaraan meewerken door bij instrumentevaluaties een breder perspectief te hanteren en een instrument niet alleen op doeltreffendheid en doelmatigheid te beoordelen. Onderzoekers kunnen tevens bepalen in welke mate een instrument bijdraagt aan bredere doelstellingen of wat de legitimiteit van een instrument is (welk probleem poot een instrument op te lossen?). Evaluaties zouden bijvoorbeeld beter kunnen aansluiten bij grotere maatschappelijke transitie en opgaven, bijvoorbeeld op het vlak van klimaat, energie en veiligheid. Deze koppeling kan bijvoorbeeld plaatsvinden door hier evaluatievragen over op te nemen of door de beleidstheorie van een individueel instrument te koppelen aan een meer overkoepelende beleidstheorie die visualiseert hoe de verschillende instrumenten bijdragen aan overkoepelende beleidsdoelstellingen en maatschappelijke impact.

Wat betreft de toename van het aantal evaluaties is het van belang dat opdrachtgevers rekening houden met de ‘absorptiecapaciteit’ van de gebruikers (bijvoorbeeld de volksvertegenwoordiging). Door de toename van het aantal evaluaties kan de (politieke) aandacht per rapport verslappen. Evaluatoren kunnen hierop anticiperen met korte samenvattingen die enkele centrale conclusies en aanbevelingen centraal stellen. Ook kunnen binnen een beleidsterrein meerdere evaluaties samengenomen worden om zodoende enkele (overkoepelende) conclusies en aanbevelingen te delen. Zeker wanneer hetzelfde type instrument (subsidie, wettelijke onderzoekstaak of een convenant) meerdere malen wordt onderzocht (binnen een beleidsterrein), kunnen de conclusies en aanbevelingen elkaar overlappen. Hier zou een onderzoeker op kunnen reflecteren in het eigen onderzoek (c.q. onderzoeksrapport).

Uitdaging 2: Scope en timing van evaluaties sluiten niet altijd aan bij beleid

Bij meerjarige beleidsprogramma’s treden effecten mogelijk pas na langere termijn op (bijvoorbeeld na vijf tot tien jaar). Beleid kan stabiel zijn, waardoor een nieuwe evaluatie weinig nieuws onder de zon brengt ten opzichte van een vorige evaluatie (maar dat kan nuttig zijn om vast te stellen). Een (verplichte) evaluatie sluit niet noodzakelijk goed aan bij de looptijd van beleid.

De politiek kan behoefte hebben aan een ‘early warning’-evaluatie (Korsten, 2023) om zich in een vroeg stadium (bijvoorbeeld vlak na een wetswijziging) een beeld te vormen van de eerste gevolgen in plaats van eerst vier jaren te wachten op een verplichte evaluatie (en het te laat is voor bijsturing). Verplichtingen zijn dus nuttig omdat zij evaluaties structureel verankeren, maar kunnen als nadeel hebben dat

Frank Bongers

de timing niet past bij de ‘incubatietijd’ van beleid of de informatiebehoefte van beleidsmakers.

Goed ingevoerde beleidsmedewerkers en doelgroepen van beleid weten vaak wel hoe de vlag erbij hangt. Een beknopte interviewronde of een gerichte zoektocht op internet kan bruikbare inzichten opleveren zonder dat er een uitgebreid beleidsonderzoek plaatsvindt (Van Hoesel, 2023). Evaluatiebepalingen in programma’s, convenanten of wetgeving bieden doorgaans voldoende ruimte om voor een kleinere evaluatie te kiezen.

Beleidsonderzoekers hebben – zeker bij een wettelijke verplichting – weinig invloed op de planning van een evaluatie. Zij zullen hier balanceren tussen enerzijds de verplichting (of wens) om te evalueren en anderzijds wat redelijkerwijs goed te evalueren is. Bij ex-post evaluaties zijn effectmetingen pas zinvol wanneer het aannemelijk is dat een beleidsinstrument voldoende lang bestaat om effect te sorteren bij een doelgroep. Bij een onderbenutting van een instrument (bijvoorbeeld een subsidieregeling) kan onderzoek naar oorzaken van deze onderbenutting relevanter zijn dan een effectmeting. Een evaluatie wint in deze situatie aan bruikbaarheid wanneer onderzoekers erin slagen om (eventueel in overleg met een opdrachtgever) de vraagstelling in een periodieke en verplichte evaluatie aan te laten sluiten bij de looptijd van beleid (en niet aan te sturen op een uitgebreide evaluatie van doeltreffendheid en doelmatigheid die zijn doel voorbijschiet). Dan zouden onderzoeksvragen zich eerder richten op de uitvoering, doelgroepenbereik, (tussentijdse) resultaten of noodzaak voor bijsturing.

Dit laatste vraagt van evaluatoren de vaardigheid om opdrachtgevers (vaak beleidsmakers) ervan te overtuigen dat bepaalde evaluatievragen te vroeg worden gesteld (en dus niet goed te beantwoorden zijn) of anders geformuleerd kunnen worden, of om tijdens een evaluatie in te zetten op het goed managen van verwachtingen. Bijvoorbeeld wanneer blijkt dat sommige evaluatievragen minder goed te beantwoorden zijn dan eerder werd aangenomen (door de evaluator en/of opdrachtgever). Ook kan een evaluator een vooronderzoek voorstellen om te bepalen of een evaluatie zinvol is, en zo ja, hoe deze dan aangepakt kan worden. De kunst is hier om als evaluator een (te) ambitieuze vraagstelling van een evaluatie om te zetten in een meer haalbaar en bruikbaar startpunt.

Uitdaging 3: Evaluaties zijn er voor beleid en niet andersom

Beleidsmakers richten zich niet automatisch op het scheppen van gunstige randvoorwaarden voor een evaluatie. Beleid bestaat eerst en vooral om beleidsproblemen op te lossen en (gelukkig) niet om het werk van evaluatoren eenvoudiger te maken. In sommige randvoorwaarden wordt overigens vaak wel voorzien. Denk aan het beschikbaar stellen van evaluatiemiddelen of politieke toezeggingen om te evalueren. Extra gunstige randvoorwaarden doen zich voor wanneer er een nulmeting of een ex-ante evaluatie beschikbaar is (en er latere metingen zullen volgen), wanneer er experimentele groepen en controlegroepen kunnen worden gevormd (Roper, 2020) of wanneer beleid gepaard gaat met monitoring (Hermans, 2017).

Het komt in de praktijk weinig voor dat deze gunstige randvoorwaarden allemaal worden vervuld.

Beleidsonderzoekers moeten hier dus omgaan met soms minder gunstige randvoorwaarden aan de ene kant en de wens een betrouwbaar en bruikbaar evaluatieresultaat te bereiken aan de andere kant. Hierbij speelt dat er spanning kan bestaan tussen een betrouwbare evaluatie en een bruikbare evaluatie. Een gebrek aan tijd, middelen of data kan de betrouwbaarheid van een evaluatie aantasten, maar dan is het nog steeds mogelijk dat een beleidsmaker verder kan met de resultaten van een kleinschalige, eenvoudige of minder optimale evaluatie.

Het is aan de evaluator om in deze context duidelijk te maken welke evaluatievragen, gegeven de mogelijkheden, geheel, gedeeltelijk of niet beantwoord kunnen worden. Vaak betekent dat dat onderzoekers een aanpak ontwikkelen die tegemoetkomt aan de mogelijkheden (budget, planning, databeschikbaarheid) en daarbij dus ook grenzen aangeven aan de mogelijke uitkomsten.

Evaluatoren kunnen – zeker wanneer zij institutioneel verbonden zijn aan overheidsinstellingen – stimuleren dat in vroege fasen van beleid rekening gehouden wordt met een evaluatie, bijvoorbeeld door een ex-ante evaluatie voor te stellen, een beleidstheorie op te stellen of aan te sturen op monitoring van beleid. Meer in het algemeen kunnen evaluatoren in elke evaluatie aanbevelingen opnemen die toekomstige evaluaties kunnen helpen. Dat kunnen hele praktische aanbevelingen zijn, bijvoorbeeld op het vlak van monitoring en dataverzameling, of meer strategische aanbevelingen, bijvoorbeeld over het ontwikkelen van een evaluatiefunctie in een overheidsinstelling.

Uitdaging 4: Evalueren is meer dan kwantificeren (en digitaliseren)

Steeds meer beleidsmakers verwachten dat evaluatoren gebruik maken van kwantitatieve en digitale gegevens. Kwantificering en digitalisering raken – los van deze verwachting – de evaluatieprofessie sowieso. Socioloog Mau (2019) wijst in dit verband op de volgens hem twijfelachtige trend om alle aspecten van ons sociaal leven (en dus ook beleid) te kwantificeren. In een dergelijke 'evaluatiesamenleving' wordt volgens hem alles en iedereen gereduceerd en beoordeeld op basis van kwantitatieve monitoring en evaluatie en ontstaan er nieuwe rangordes van wat als maatschappelijk waardevol wordt beschouwd.

De eerste uitdaging voor de evaluator bestaat eruit om aan beleidsmakers duidelijk te maken dat niet elk aspect van beleid in cijfers uit te drukken is. Beleidsmakers weten en begrijpen dat, maar de neiging bestaat het succes en falen van beleid uit te drukken in (kwantitatieve) indicatoren en streefwaarden. Het risico dreigt dan dat alleen gemeten wordt wat gemeten kan worden en er dus een (te) simplistisch beeld van de werkelijkheid ontstaat (zie ook Bongers & Verlet, 2013). Terugkomend op de beleidstheorie kan de (verwachte) relatie tussen middelen en doelen meer narratief worden uitgelegd, zodat in woorden duidelijker wordt hoe mechanismen in een beleid werken (of beogen te werken).

Frank Bongers

De sterke dynamiek op het gebied van kwantificering is volgens Mau (en vele anderen) onder meer te danken aan voortschrijdende digitalisering, waardoor het verzamelen, opslaan en analyseren van data aanzienlijk sneller en eenvoudiger worden. Sommige beleidswetenschappers (o.a. Valle-Cruz et al., 2020) wijzen erop dat digitalisering en artificiële intelligentie (AI) continu evalueren mogelijk maken en dat in elke stap in de beleidscyclus. Hierdoor vervagen grenzen tussen *ex-ante*, *ex-durante* en *ex-post* evaluaties.

Tyler et al. (2023) stellen dat AI-tools vooral openbare en doorzoekbare databestanden benutten; en die bestanden zijn niet altijd volledig of representatief. AI is dus geen hamer die voor beleidsmakers en -onderzoekers alles tot een spijker maakt. Een andere uitdaging voor de evaluator is dat hij zich (beter) moet scholen in het gebruik van digitale toepassingen zoals AI-tools in onderzoek. De uitdaging bestaat niet alleen uit het aanleren van nieuwe digitale vaardigheden, maar ook uit het correct gebruik van AI en het zich bewust zijn van de grenzen aan het gebruik van AI (bijvoorbeeld op het vlak van privacy en ethiek).

Deel 2: Belangen en doorwerking in de evaluatiepraktijk

Uitdaging 5: Evalueren vindt plaats tussen uiteenlopende en conflicterende belangen

Het is niet correct om te stellen dat beleidsbepaling alleen gestuurd wordt door de politiek en beleidsevaluatie alleen door onderzoek. Er bestaat wisselwerking. Evaluaties voeden politieke besluitvorming (naast andere factoren als machtsverhoudingen, ervaringen en politieke opvattingen). Voorts doen evaluaties zelf uitspraken over bijvoorbeeld de houdbaarheid of de legitimiteit van beleid (zie o.a. Weiss, 1973).

Evaluatoren moeten zich dus bewust zijn van de (politieke) gevolgen van evaluaties en moeten bevorderen dat de kwaliteit van evaluaties of zelfs evaluatoren niet ter discussie komen te staan. Dit betekent dat vanaf het begin van een evaluatie evaluatoren transparant moeten zijn over hun keuzes, deze keuzes afstemmen met opdrachtgevers en – *last but not least* – conclusies en aanbevelingen empirisch goed onderbouwen. Voorts moeten evaluatoren ervoor waken dat politieke wensen en beïnvloeding de onafhankelijkheid van een evaluatie niet aantasten en dat zij voldoende ruimte krijgen om goed te evalueren. Het voorbijgaan aan politieke belangen in een evaluatie kan de bruikbaarheid verminderen, maar zonder meer tegemoetkomen aan politieke wensen zet een rem op de betrouwbaarheid en geloofwaardigheid van de evaluatie (en de evaluator).

Voor beleidsmakers is het minder aangenaam om te horen dat beleid niet of onvoldoende werkt. Een evaluatie met een onwelkome boodschap over het gevoerde beleid kan leiden tot Kamervragen, berichten in de media of onvrede bij de doelgroep van het beleid. Desondanks is het belang van evaluaties dat zij bijdragen aan leren en verantwoorden. Zij voeden daarmee het politieke debat over het bereiken van doelen en de verantwoording over ingezette middelen. Zij ondersteunen dus de taken van onze volksvertegenwoordiging op het vlak van controle en kaderstelling.

Het helpt in dit opzicht om evaluaties vooral te positioneren als een kans om beleid te verbeteren en als een middel voor politieke verantwoording.

De beleidstheorie – het geheel van veronderstellingen waarop beleid berust – illustreert de verwevenheid van politieke factoren in beleid en evaluaties. De beleidstheorie is een hulpmiddel in een evaluatieonderzoek om zicht te krijgen op bijvoorbeeld causale relaties tussen middelen, doelen en effecten (en eventuele streefwaarden die daaraan gekoppeld kunnen worden) (zie o.a. Bongers, 2023). De beleidstheorie lijkt daarmee een rationeel en onafhankelijk hulpmiddel in onderzoek. Evaluatoren moeten zich echter realiseren dat elke beleidstheorie zich baseert op overheidsinterventies voor maatschappelijke problemen en voor beleidsdoelstellingen die de beleidsagenda hebben bereikt (en waarvan beleidsmakers veronderstellen dat ze effectief zijn). Beleid en (in het verlengde daarvan) beleidstheorieën en evaluaties bestendigen dan (onbedoeld) eventuele sociale of politieke ongelijkheid (oftewel laten potentieel ander beleid, andere instrumenten en andere doelgroepen buiten beschouwing). Evaluaties volgen wat betreft onderwerpkeuze dus de heersende politieke agenda.

Het negeren van politieke belangen hoort overigens niet de standaardreactie van een evaluator te zijn. Het is goed mogelijk om in een evaluatie een vraag of beleids optie mee te nemen waarvoor (extra) politieke belangstelling bestaat. Verder kunnen via democratische besluitvorming sommige belangen zwaarder wegen in beleid dan andere belangen (en dus ook in een evaluatie). Dat past in onze democratische rechtsstaat. Belangen hoeven dus een onafhankelijke evaluatie niet altijd in de weg te staan. Zij maken evaluatoren alert ten aanzien van de (politieke en maatschappelijke) omgeving waarbinnen zij onderzoek uitvoeren, voor wie zij dat doen en welke voor- en nadelen er kleven aan bepaalde keuzes binnen een evaluatie. Dat betekent onder meer dat evaluatoren onderwerpen, doelstellingen en vragen tegen het licht houden en deze waar nodig beargumenteerd aanpassen.

Ook kunnen ze goed letten op de ‘governance’ van een evaluatie: is een begeleidingscommissie zodanig samengesteld dat zij een onafhankelijk onderzoek mogelijk maakt, en dat zij betrokken blijft en vooral stuurt op voortgang en kwaliteit van het onderzoek (en geen platform wordt voor specifieke politieke en maatschappelijke belangenbehartiging)? Belangenbehartiging kan bij een evaluatie een plek krijgen in een klankbordgroep die door de evaluator tijdens de evaluatie inhoudelijk geïnformeerd wordt en waar leden suggesties kunnen doen voor de evaluatie. Kwaliteitsbewaking en voortgang blijven voorbehouden aan een begeleidingscommissie.

Bij voorkeur verleent een opdrachtgever aan evaluatoren keuzeruimte om in de evaluatieaanpak eigen inzichten te verwerken over de vraagstelling, methoden en doelgroepen van de evaluatie (en welke rol deze doelgroepen krijgen in een evaluatie). Dit kan bijvoorbeeld ruimte bieden om de formulering van een evaluatievraag breder te trekken en daarbij rekening te houden met inzichten van doelgroepen van beleid. Doelgroepen kunnen andere opvattingen hebben over het succes of falen van beleid dan beleidsmakers, uitvoerders en experts.

Frank Bongers

Evaluatoren moeten hier balanceren tussen de wensen van een opdrachtgever aan de ene kant en de belangen van degenen die beleid ondergaan aan de andere kant (en dus ook hun eigen wensen). Het gaat hier niet om tegenstrijdige rollen: een opdrachtgever kan er baat bij hebben dat een evaluatie vanuit andere en meerdere perspectieven wordt uitgevoerd. Het is dan aan onderzoekers om te sturen op dergelijke perspectiefwisselingen, bijvoorbeeld door in te zetten op responsieve evaluaties (Abma, 2024) en door transparant te zijn over gemaakte keuzes.

Uitdaging 6: Doorwerking van evaluaties gaat niet vanzelf

Een van de laatste uitdagingen (in dit artikel) betreft het bevorderen van de doorwerking van een evaluatie in het beleidsproces. In het rationele lineaire model van beleidsvorming vinden de resultaten – wanneer zij daartoe aanleiding geven – hun weg in een aanpassing van beleid (en de besluitvorming daarover). Daarmee sluit de cyclus zich. Dat gebeurt wanneer een minister of de volksvertegenwoordiging kennisneemt van de resultaten en waar nodig beleid aanpast. Daarnaast is doorwerking gediend met betrouwbaar en bruikbaar onderzoek dat bovendien zijn weg naar buiten vindt, bijvoorbeeld door (online) verspreiding en gebruik (ook buiten de kring van direct betrokkenen). Het is hier goed te verwijzen naar de vorige paragraaf, waarin gesteld wordt dat wanneer een evaluator volledig voorbijgaat aan de politieke kant van evalueren, de kans op doorwerking (gebruik) van een evaluatie kleiner wordt.

Een andere vorm van doorwerking betreft conceptuele doorwerking. Een evaluatie stimuleert dan vooral dat beleidsmakers een beter begrip krijgen van beleid. Bortz (2023) wijst in dit verband op de term ‘weerklink’ (*‘resonance’*). Wanneer resultaten beter aansluiten bij de mentale modellen van beleidsmakers, is de kans groter dat evaluatieresultaten resoneren. Andere factoren die volgens Bortz bijdragen aan doorwerking, zijn openheid van en vertrouwen in de beleidsonderzoeker.

Een betrouwbaar en bruikbaar evaluatieonderzoek is geen garantie voor doorwerking. Soms zijn aanbevelingen moeilijk vertaalbaar in nieuw beleid (of ontstaat daar onenigheid over in de politieke besluitvorming). Evaluaties kunnen – zoals eerder besproken – voelen als een verplichting, waardoor de timing van een evaluatie ongunstig is. Evaluaties moeten wat betreft doorwerking concurreren met andere factoren die van invloed zijn op politieke besluitvorming zoals ervaringen, politieke verhoudingen, aandacht en ideologieën van beleidsmakers. Het is wat dat betreft behulpzaam wanneer evaluatoren na afloop de ruimte geboden wordt om de evaluatieresultaten toe te lichten bij verschillende doelgroepen.

Voor evaluatoren blijft het continu een uitdaging om de doorwerking van een evaluatie te stimuleren. Wanneer een evaluatie onderdeel wordt van een publiek of politiek debat (wat op zichzelf al een voorbeeld van doorwerking is), blijft hun rol beperkt. Zij moeten balanceren tussen enerzijds het (technisch) toelichten van resultaten (deze aantrekkelijk presenteren in de vorm van visualisaties of effectieve samenvattingen) en het beantwoorden van vragen en anderzijds inhoudelijk meedoen aan het (politieke) debat over de evaluatie. Hier is een bijdrage duidelijk begrensd. Evaluatoren kunnen aangeven wanneer resultaten verdraaid of oneigenlijk

gebruikt worden, maar zullen zich verder inhoudelijk wat meer op de vlakke moeten houden.

Tot besluit

De gereedschapskist van beleidsonderzoekers behoort minimaal een vak met 'zachtere' kennis en vaardigheden te bevatten. Evalueren vraagt om meer dan alleen de toepassing van wetenschappelijke methoden (en de meeste evaluatoren zullen zich daar ongetwijfeld al van bewust zijn). Dat is geen excuus om niet te blijven streven naar methodologisch verantwoorde evaluaties die met feiten en kennis voeding geven aan beleid. Het is wel een oproep om te blijven reflecteren op andere en meer zachte aspecten van evalueren.

Evalueren gebeurt in een complexe en dynamische omgeving die flexibiliteit van onderzoekers vergt en hen uitdaagt keuzes te maken in een evaluatie die methodologisch soms moeilijk te verantwoorden zijn, maar die wel zorgen dat een evaluatie aan belang kan winnen en een grotere kans op doorwerking krijgt. De evaluator moet in staat zijn deze keuzes uit te leggen en te verdedigen. De beschikbaarheid van data is leidend voor het soort van methoden dat kan worden toegepast (en dat kan minder ideaal zijn dan methodologisch gewenst is). Tegelijkertijd hebben evaluaties te maken met groepsprocessen waarin uiteenlopende en zelfs conflicterende belangen kunnen spelen. Sommige actoren kunnen daarin meer invloed uitoefenen dan andere. Onderzoekers moeten dan schipperen tussen deze belangen met het oog op het bewaken van de onafhankelijkheid en kwaliteit van een onderzoek. Dat maakt dat evaluatoren soms hun onderzoekersrol moeten loslaten en in het domein van procesbegeleiding of zelfs pleitbezorging terechtkomen (zie ook Verlet et al., 2016).

Feitelijk komt het erop neer dat beleidsonderzoekers aan verschillende knoppen kunnen draaien om de kwaliteit en bruikbaarheid van evaluaties te borgen en te bevorderen. Deze knoppen kunnen in elke evaluatie variëren en zijn zeker niet alleen methodologisch van aard. Evaluatoren zullen de gemaakte keuzes wat betreft de evaluatieagenda moeten accepteren, maar daarbinnen kunnen zij ruimte pakken om bij een evaluatie maatwerk te leveren. Zolang de evaluator maar transparant is over de gemaakte keuzes binnen een evaluatie, ook wat betreft keuzes die voortkomen uit praktische uitdagingen en belangen (en deze keuzes kan verdedigen).

Het volledig voorbijgaan aan dit soort overwegingen kan een evaluatie onnodig vertragen en de impact ervan belemmeren. Een goed uitgedacht plan van aanpak voor een evaluatie is een prima startpunt, maar gaandeweg de route zullen evaluatoren ervaren dat er af en toe een zijpaadje ingeslagen moet worden en dat soms een stapje terug nodig is om toch tot dat bruikbare eindresultaat te komen. Bij voorkeur houdt de evaluator in het plan van aanpak daar al wat rekening mee, bijvoorbeeld door in overleg met betrokkenen te komen tot een realistische en haalbare set van evaluatievragen, door keuzemomenten in te bouwen en door duidelijk

Frank Bongers

aan te geven hoe verschillende belangen betrokken en gewogen worden in de evaluatie.

Literatuur

- Abma, T. (2024). Responsieve beleidsevaluatie: Theorie en praktijk van een participatieve methode met impact. *Beleidsonderzoek Online*, juni, DOI: 10.5553/BO/221335502024013.
- Bongers, F. (2023). De rol van de beleidstheorie in het ontwerpen en evalueren van beleid. *Beleidsonderzoek Online*, mei, DOI: 10.5553/BO/221335502023007.
- Bongers, F., & Verlet D. (2013). Indicatoren en beleid: Doe meer met ongeveer. *Bestuurswetenschappen*, 67(1), 36-55.
- Bortz, M. (2023). Advice that resonates: explaining the variability in consultants' policy influence. *Policy Sciences*, 56(2), april, DOI: 10.1007/s11077-023-09506-5.
- Derksen, W. (2011). *Kennis en beleid verbinden: Praktijkboek voor beleidsmakers*. Boom Lemma uitgevers.
- Hermans, L. (2017). Adaptief beleid en beleidsevaluaties. *Beleidsonderzoek Online*, november, DOI: 10.5553/BO/221335502017000009001.
- Korsten, A. F. A. (2023). Early warning evaluation en de invoeringstoets. *Beleidsonderzoek Online*, maart, DOI: 10.5553/BO/221335502023004.
- Mau, S. (2019). *The metric society: On the quantification of the social*. Polity Press.
- Roper, S. (2020). Using RCTs as a research method for SME policy research: The UK experience. In A. Saridakis & M. Cowling (Ed.), *Handbook of quantitative research methods in entrepreneurship* (pp. 116-135). Edward Elgar.
- Tyler, C., Akerlof, K. L., Alessandro, A., Zachary, A., Canino, H., Doornenbal, M. A., Goldstein, J. A., Budtz Pedersen, D., & Sutherland, W. J. (2023). AI tools as science policy advisers? The potential and the pitfalls. *Nature*, 622, 27-30.
- Valle-Cruz, D., Cradio, I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4). DOI: doi.org/10.1016/j.giq.2020.101509.
- Van der Knaap, P., Pattyn, V., & Hanemaayer, D. (2020). *Beleidsevaluatie in theorie en praktijk*. Boom bestuurskunde.
- Van Hoesel, P. (2023). Smpel beleidsonderzoek kan heel effectief zijn. *Beleidsonderzoek Online*, mei, DOI: 10.5553/BO/221335502023008.
- Van Hoesel, P., & Herold, M. (2024). *Beleidsonderzoek in de praktijk: Een routekaart voor onderzoekers en beleidsambtenaren*. Boom bestuurskunde.
- Van Thiel, S. (2021). *Bestuurskundig onderzoek: Een methodologische inleiding*. Coutinho.
- Verlet, D., Bongers F., & Devos, C. (2016). Beleidsevaluatie in woelige publieke wateren: Een vooruitblik. *Vlaams Tijdschrift voor Overheidsmanagement*, (2), 9-15.
- Weiss, C. H. (1973). Where politics and evaluation research meet. *Evaluation Practice*, 14(1), 93-106.