



Bespreking van: S.B. Nielsen, R. Turksema & P. van der Knaap (Eds.), Success in evaluation: Focusing on the positives, New Brunswick/London: Transaction Publishers 2015

Carolien Klein Haarhuis *

Aanbevolen citeerwijze bij dit artikel

Carolien Klein Haarhuis, 'Bespreking van: S.B. Nielsen, R. Turksema & P. van der Knaap (Eds.), Success in evaluation: Focusing on the positives, New Brunswick/London: Transaction Publishers 2015', *Beleidsonderzoek Online* februari 2017, DOI: 10.5553/BO/221335502017000002001

De positieve benadering van evaluatie in relatie tot het bereiken van succes is een frisse wind. Het wordt inderdaad tijd dat evaluatoren gebruikers meer helpen en stimuleren om van (beleids)ervaringen te leren. En: dat geëvalueerden zich daarvoor vanaf het begin openstellen. Aandachtspunt is wel dat de evaluator daarbij zijn of haar waardevolle kritische (kennis)blik niet verliest.

Het boek in het kort

De boodschap van de bundel *Success in Evaluation* is helder: evaluaties vervullen in de praktijk niet de gedroomde rol in het 'leren' door organisaties – waaronder beleidsorganisaties. Dat kan anders.

Evaluaties worden volgens de redacteuren steeds vaker ter discussie gesteld: zijn ze relevant? Dragen ze echt bij aan betere besluitvorming? Zo dreigt evalueren steeds meer in de negatieve sfeer, bijvoorbeeld van kritiek en afrekenen, te worden getrokken.¹ Drempels voor het leren van evaluatie-inzichten zijn nog steeds, vrij naar het boek:

- dominantie van politieke of strategische argumenten;
- nadruk op achteraf verantwoording afleggen;
- geëvalueerden – de beoogde gebruikers – staan niet open

voor feedback en ontberen de *skills*; anderzijds leven evaluatoren zich te weinig in;

- het ‘mosterd-na-de-maaltijd-effect’: evalueren vergt tijd en die is er niet altijd (meer);
- evaluaties en ook *evidence-based* ‘kookboeken’ sluiten niet aan op de beleidspraktijk.

Redactie en auteurs Bojsen & Dybdal wijzen op de volgende realiteit: beleidsmakers zijn geen rationeel-objectieve wezens die evidentie opslokken en deze vervolgens één-op-één omzetten in top-down beleid. Ze zijn verwikkeld in een complex interactieproces met vele, waaronder politieke, actoren met *stakes* in een onderwerp, waarin uiteindelijk wordt toegewerkt naar een breed gedragen ‘besluit’ en de verdere ‘praktijkvorming’. De evaluator moet samen met deze actoren overeenstemming zien te bereiken over de focus, reikwijdte, afbakening, producten et cetera, en ook over de onderlinge rollen (Van der Knaap, Nielsen & Turksema, p. 232). Niemand in deze bundel zegt dat dit een simpele taak is.

Hoe we evalueren en *wat*, is volgens de redactie cruciaal. Er zitten drie kernelementen in het door hen opgestelde conceptuele kader, die samen een ‘positieve cyclus’ vormen: evaluatieonderzoek, leren en succes.

Dit toegankelijke boek biedt twee centrale aanknopingspunten voor de positieve insteek, die aansluiten op de delen 1 en 2, waarin de bijdragen van de auteurs zijn onderverdeeld.

(1) Steek *de evaluatie zelf* positief in, via het design en via de te evalueren cases. Het eerste deel van het boek biedt verschillende mogelijke evaluatie-aanpakken en technieken passend bij een positieve insteek. Van al te objectief-kritische – mogelijk negatieve – feedback leer je als *geëvalueerde* niet hoe het wél zou moeten, en krijg je ook niet bepaald frisse moed om opnieuw aan de slag te gaan. Inzichten uit de neuropsychologie ondersteunen dit. Kracht is dus gelegen in het uitbuiten van positieve leerervaringen – temeer als daar concrete handvatten uit voortvloeien, en condities waaronder die kunnen werken; geheel in de *realist evaluation*-traditie.

Enkele voorbeelden van de positieve aanpakken in het eerste deel van het boek:

- Evaluatiefocus en selectie van gevallen: een beleid is nooit een volslagen succes maar ook nooit een volslagen mislukking. Er zijn altijd gevallen waarin het heeft gewerkt. Borduur daarop voort met behulp van verschillende onderzoeksmethoden; zeg

de positieve verhalen voort en bespreek ze vervolgens actief met alle belanghebbenden (in plaats van een evaluatierapport op te leveren). Dit wordt bijvoorbeeld geïllustreerd door de door Stame & Lo Presti uitgelegde *Succes Case Method*.

- Het blootleggen van handvatten en condities waaronder succes optreedt, is als het om leren gaat belangrijker dan een evenwichtig plaatje krijgen van positief en negatief, aldus Stame & Lo Presti. Een focus op het gemiddelde leidt maar tot een hang naar het middelmatige (in verwijzing naar Perrin). Omdat een programma verstrengeld raakt met uiteenlopende contextfactoren, gaat de vlieger van het lineaire oorzaak-gevolg-denken niet op. Om te kunnen leren gaat het erom de condities te benoemen waaronder successen worden of zijn behaald.
- Welke meet- en leermethoden dan te gebruiken? Verschillende bijdragen bespreken methoden inclusief big of open data-georiënteerde strategieën om gefundeerde feedback een voet tussen de deur te geven. Deze methoden en technieken hebben de toekomst. De vraag is mijns inziens wel steeds, in hoeverre we ze kunnen inpassen in evaluatiedesigns die op valide wijze iets zeggen over de werkzaamheid van (beleids)initiatieven.

(2) Ten tweede is nodig dat er een *leereffect* ontstaat bij de afnemer, de producteigenaar. Het tweede deel van het boek geeft een overzicht van positieve ervaringen hiermee. Met het conceptueel kader laten die zien dat aan succesgericht evalueren (te omschrijven als: evalueren om te leren teneinde beter te worden) succesgericht *opereren* voorafgaat. Idealiter is de start-attitude van de beleidsmaker of -organisatie al op leren gericht, vergelijk Mayne's *evaluative thinking*. Evaluatie is dan een onmisbaar leerinstrument. De evaluator versterkt het gevoel van relevantie bij de geëvalueerde door die te betrekken bij de operationalisering van het doel, de opzet en de indicatoren, de gegevensverzameling, et cetera. Dat maakt de evaluator een metgezel van de beleidsmaker en anderen, een kritische vriend; geen onafhankelijke die in zijn hok blijft. En uiteraard is daarbij de manier waarop de evaluator bevindingen brengt van belang: 'Strike a balance between criticism, constructive suggestions, and even praise' (Turksema, Van der Knaap & Nielsen, p. 10).

Een paar concrete handvatten in dit verband:

- Empathie van de kant van de evaluator: weet bijvoorbeeld wat de (economische c.q. strategische) doelen van de beleidsorganisatie zijn (Bojzen & Dybdal).

- Betrek *geëvalueerden* vanaf het begin bij het evalueren en het evaluatiedesign, inclusief performance-indicatoren, (evaluatie)data en -rapportage (Nielsen). Adagium: *No stories without numbers & no numbers without stories*.

Een paar overdenkingen

De boodschap van dit boek heeft nogal wat voeten in de aarde. Om te kunnen leren, moet evalueren positiever en met meer empathie. Evaluatoren moeten dus meer affiniteit krijgen met de lerende beleidsmaker en met de context waarin die opereert. Op hun beurt zouden beleidsmakers zich (meer) eigenaar moeten gaan voelen van de feedback op hun producten en meer de regie moeten nemen over wat zij uit evalueren willen leren – liefst tot in de kernfilosofie achter hun beleid.

Het boek heeft mij als onderzoeker in elk geval aardig aan het denken gezet. Hoe positief opereer ik eigenlijk? Hoe responsief ben ik naar de *geëvalueerde(n)* toe, ben ik een kritische vriend? Alleen dat al maakt lezing van dit boek dik de moeite waard. Tegelijk dienen zich ook meer fundamentele vragen aan bij het lezen van dit boek, namelijk: past dit bij mij en mijn rol als evaluator? Wil en kan ik die kritische vriend zijn? Ook lezende ‘beleidsmakers’ (c.q. leden van beleidsorganisaties of -netwerken) zullen zich wellicht afvragen: wat betekent succes voor mij – is dat bijvoorbeeld maatschappelijk, politiek of organisatiesucces? Wil en kan ik vanaf het begin openstaan voor leerervaringen uit evaluatie en deze opvolgen?

- Hoe realistisch is het?

Hoe laten we de positieve vonk overslaan om de waargenomen tekorten aan openheid bij de beleidsmaker en inlevingsvermogen bij de evaluator op te heffen? Hoe zorgen we dat, naast ideologie, politiek-strategische belangen en ad hoc inzichten, ook leerzame feedback uit evaluaties nadrukkelijker deel gaat uitmaken van beleids-‘r&d’?

Als hij er nog niet is, ontstaat de eerste vonk niet vanzelf. Ik zal de lezer een betoog over evalueren in de politieke werkelijkheid besparen. Voor attitudeverandering is overtuiging cruciaal. Een eerste tip: *geëvalueerden* en evaluatoren, en zeker ook Kamerleden: lees dit boek!

Een andere tip. Met meerdere auteurs en de redactie (p. 231) ben ik een voorstander van voortbouwen aan *evaluatievermogen* aan de vraagzijde – de kant van de *geëvalueerde(n)* dus. Globaal betekent dat: voortborduren op positieve ervaringen met (a) processen rond evaluaties en (b) de inhoud – bijvoorbeeld designs van evaluaties, die voor de organisatie in kwestie eerder goed hebben gewerkt. Uiteraard

rekening houdend met verschillende praktijksituaties: wanneer werkte dit goed, onder welke condities? De betrekkelijk jonge aanpak van evaluaties binnen het ministerie van EZ is exemplarisch: evalueren – zowel vroeg experimenteren als het ex post-werk – wordt beschouwd als middel tot het uiteindelijke doel, namelijk dat van beter en effectiever beleid maken. Daartoe moet, en dit is vastgelegd in een handleiding, vanaf de conceptie van beleidsinitiatieven al over het evalueren worden nagedacht.

In dit verband is het trouwens een beetje spijtig dat de bijdrage met de casus over de Britse *National Audit Office* (NAO) (Lonsdale & Stewart) vooral ingaat op de rol van positieve feedback voor de NAO als organisatie. Een vraag die de lezer blijft plagen, is: hoe kunnen de Rekenkamer en ook het ministerie van Financiën (verder) stimuleren dat beleidsorganisaties hun evaluatiecapaciteit op een open, transparante manier en voor maatschappelijk gewin inzetten? Hoe geven we ‘afrekenen op leren’ handen en voeten?

- Is het positief-lerende altijd toepasbaar?

Je hebt praktische evaluaties van projectjes van kleinere organisatieonderdelen, waaruit de feedback te behappen is en ook de terugkoppeling relatief overzichtelijk is. Veel bijdragen lijken geschreven met dit type situatie in het achterhoofd. Maar je hebt vaak ook te maken met grootschalige en complexe evaluatiestudies, bijvoorbeeld naar de werking van de ‘3D’s’. In die situaties kunnen geëvalueerden op een ingeslagen weg niet zomaar even omdraaien of een andere koers gaan varen. De vraag is of en hoe we het positief-lerende kunnen extrapoleren naar zulke complexe situaties met een wijdvertakt beleidsnetwerk. Je hebt strikt genomen dat hele netwerk nodig om gefocust en relevant evaluatieonderzoek te doen, om iedereen aan boord te houden én: om bevindingen te laten landen (gebruik). Dit vergt een duizendpoot-evaluator, met: analytisch vermogen voor de inhoud, contextsensitiviteit en flexibiliteit voor het vele afstemmen, doorhebben welk ‘spel’ er wordt gespeeld en navenant kunnen manoeuvreren (vgl. Mark, 2012). Dat vergt een sterk team. Het boekomslog met het zegevierende elftal is hier extra toepasselijk. Ook vooraf bepaalde spelregels helpen, over hoe en met wie interactie plaatsvindt over het evalueren.

Klein beginnen kan de boel vereenvoudigen: koppel evaluaties bijvoorbeeld aan fases of organisatieonderdelen. Leer van het kleine en koppel die lessen zodra beschikbaar terug naar het grotere geheel; vooral daar waar inzicht ontstaat in het *waarom* van succes of gebrek daaraan. Dat gaat extra makkelijk als óók de beleidsmaker niet te groot start, bijvoorbeeld met een of meer experimenten of pilots. De evaluator volgt deze initiatieven op de voet met monitoring en evaluatieproducten – in plaats van ze achteraf in één klap te

beoordelen (vgl. Patton, 2010). De evaluator biedt dan dus *real time*-producten over topics die potentiële gebruikers op dat moment bezighouden, en dus: relevantie hebben. Goede data zijn overigens cruciaal daarvoor; gaat het bijvoorbeeld om registraties, dan is niet de evaluator maar de geëvalueerde bepalend voor de datakwaliteit.

Ik zou wel graag zien – met Mayne in het boek – dat de opeenvolgende evaluatieproducten duidelijk te onderscheiden zijn van de doorgaande beleidsontwikkeling. Dat voorkomt dat evaluaties verworden tot deeltjes in een maalstroom van input, waarbij vervaagt wat de meerwaarde van het evalueren ook alweer was. En het nageslacht heeft er ook nog wat aan.

- Is het verenigbaar met de verschillende soorten inbreng die evaluatoren kunnen hebben?

Ik zie een verschil tussen een lange termijn ‘kennisperspectief’ op evaluatie, en een perspectief op evaluatie als instrument om de in situatie X concreet gestelde doelen te bereiken. Het tweede perspectief lijkt te domineren in dit boek. Met een lange-termijn-kennisperspectief bedoel ik het gebruiken, in nieuwe evaluaties, van kennis die door de tijd heen is aangegroeid. In mijn optiek is leer- en doelgericht werken geen reden om dit lange-termijn-kennisperspectief te laten vallen en vice versa: de perspectieven versterken elkaar.

Gebruik wat er is! Ook al houd je dat klein

Aan de voorkant van een beleidstraject – liefst bij het afwegen van beleidsinstrumenten – kun je, op een doel- of succesgerichte manier, de al beschikbare evaluatiekennis en -inzichten in kaart brengen. Daartoe reken ik ook wetenschappelijke inzichten met betrekking tot interventies of interventiefamilies. Het is dan zaak om de concreetheid van je eigen project te overstijgen en te zoeken naar ervaringen met vergelijkbare interventies of beleidsfundamenten eerder en elders, juist ook internationaal. De evaluator heeft daarvoor de skills, en de ervaren *geëvalueerden* de concreet-inhoudelijke aanknopingspunten. Ga bijvoorbeeld na: wat werkte eerder en elders goed, en onder welke condities was dat? Deze aanpak kreeg een naam: *Realist Synthesis* (Pawson, 2002) en leent zich mijns inziens uitstekend voor een lerend perspectief. *Realist Synthesis* biedt een alternatief voor ‘jarenlang’ wachten op een finaal oordeel uit experimenten of metastudies. Het zou mooi zijn als beleidsmedewerkers en evaluatoren meer initiatieven in die sfeer op zich zouden nemen. Evalueren aan de voorkant (*ex ante*) krijgt zo meer gewicht en dat prikkelt geëvalueerden misschien ook om het lerende pad te blijven volgen.

Tot robuuste lange-termijnkennis reken ik toch ook de in een reeks van jaren opgestapelde *evidence* uit (internationale) experimenten en

metastudies. Starre kookboeken, aldus bijvoorbeeld Deense gemeenten in het boek ... Maar ik acht die overstijgende kennis toch een eerste ankerpunt op de weg naar succes – zeker daar waar sprake is van duidelijk hoge – of lage – algemene successcores van de voorlopers van een mogelijke interventie.

Blijf kritisch

Ook succes heeft zowel een korte- als een lange-termijnkant. In het evaluatiedesign is het goed om korte- en lange-termijnsuccessen van elkaar te onderscheiden. Het proces van ontwikkeling en evaluatie als zodanig kan als heel positief en stimulerend worden ervaren – maar hoe staat het met de uitkristallisering van de harde succesindicatoren na bijvoorbeeld twee jaar? Het mooiste is als niet alleen de operatie slaagt, maar de patiënt er ook nog beter van wordt.

Bij bepaalde bijdragen, zoals die van Stame & Lo Presti, borrelt toch de vraag op: waarom minder succesvolle aanpakken en condities via selectie terzijde schuiven? Hoe motiverend ook, positiviteit is en blijft maar één kant van de medaille. Juist de evaluator kan ook lessen uit een wat mindere aanpak betrekken in zijn of haar feedback, die daardoor alleen maar completer en geloofwaardiger wordt. Mayne spreekt terecht van een ‘tricky balance’ tussen *helping and fostering learning and being critical and ‘objective’* (p. 64). Voor een leereffect is het natuurlijk wel slim om die feedback opbouwend te brengen. Benadruk bijvoorbeeld de positieve werkingscondities. Dus: evalueer gebalanceerd, communiceer positief.

Kortom, niet compleet met het verleden breken, maar positief-opbouwend aan de slag. En, in lijn met de boodschap van dit boek: begin klein. Ik denk in elk geval aan het volgende:

- *Evaluator*: betrek de geëvalueerde(n) bij het evaluatieproces om relevantie en eigenaarschap te stimuleren. Het lukt je niet alleen en bovendien hebben *geëvalueerden* kennis en informatie te bieden.
- *Evaluator & geëvalueerde*: blijf zo mogelijk experimenteren en evalueren op kleine schaal, liefst in een pril stadium (r&d-achtig).
- *Geëvalueerde*: denk na over de kwaliteit van je registraties en betekenis van die data voor het leerproces.
- *Evaluator & geëvalueerde*: breng – hoe bescheiden of pragmatisch ook – al aan het begin van beleidsprojecten betrouwbare kennis en inzichten bijeen over wat werkt, hoe en onder welke condities.

- *Evaluator & geëvalueerde*: hak al te grote evaluatieprojecten in slimme mootjes en leer van de delen – leer niet alleen van het geëvalueerde beleid maar ook van het evalueren zelf (evaluatievermogen).
- *Evaluator*: produceer in plaats van één dik evaluatierapport liever meerdere dunne, over verschillende fasen en/of onderdelen van een project. Vergeet de waarde van *real-time* monitoring(-data) niet.

Laten we over een paar jaar een tweede editie van *Success in Evaluation* uitbrengen met nieuwe leerervaringen!

Literatuur

Mark, M.M. (2012). Influences on evaluation quality: Reflections and elaborations. *American Journal of Evaluation*, 33(1), 85-87.
Doi:10.1177/1098214011426470.

Patton, M.Q. (2010). Utilization-focused evaluation: Real-time and prospective aspects. In Wereldbank (red.), *Exploring the potential of real-time and prospective evaluations, summary of a workshop* (pp. 6-11). Washington DC: World Bank Independent Evaluation Group.

Pawson, R. (2002). Evidence-based policy: The promise of 'Realist Synthesis'. *Evaluation*, 8(3), 340-358.

Noten

* In de tekst van deze review wordt op enkele punten verwezen naar individuele bijdragen in dit boek (alleen auteursnamen).

1 De auteurs signaleren niettemin ook aanzienlijke verbeteringen in en rond evalueren, bijvoorbeeld in termen van inhoudelijke kwaliteit.