Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

BeleidsonderzoekOnline

DOI: 10.5553/BO/221335502020000010001

De waarde van lerend evalueren in de praktijk: ervaringen van onderzoekers en beleidsbetrokkenen van de lerende evaluatie van het Natuurpact (2014-2017)

Lisa Verwoerd, *Pim Klaassen* en *Barbara J. Regeer*

Aanbevolen citeerwijze bij dit artikel
Lisa Verwoerd, Pim Klaassen en Barbara J. Regeer, 'De waarde van lerend evalueren in de praktijk: ervaringen van onderzoekers en beleidsbetrokkenen van de lerende evaluatie van het Natuurpact (2014-2017)', *Beleidsonderzoek Online* november 2020, DOI: 10.5553/BO/221335502020000010001

1 Introductie

Evaluatie is onmisbaar in een goed functionerend beleidsproces: het helpt beleidsbetrokkenen om grip te krijgen op de werking en de toegevoegde waarde van het beleid. Echter, wanneer sprake is van beleid voor complexe maatschappelijke opgaven, dat gekenmerkt wordt door verwevenheid en onzekerheid, een meerlagig, multi-actor karakter en een grote mate van persistentie, kunnen reguliere evaluatiemethoden soms onvoldoende een betekenisvolle rol spelen om het beleid vooruit te brengen. Reguliere evaluaties vinden doorgaans achteraf plaats (*ex post*) en onderzoeken meestal de doelmatigheid en doeltreffendheid van beleid. Daarbij wordt verondersteld dat er concrete beleidsdoelen zijn geformuleerd waarnaar via lineaire en rationele beleidsprocessen kan worden toegewerkt. Beleid voor complexe beleidsdossiers – het natuurdossier, de vluchtelingencrisis, jeugdzorg, de klimaatopgave en meer recent de coronacrisis – kennen echter een grilliger en reflexief karakter, waarin talloze interacties tussen actoren en beleidselementen leiden tot onvoorspelbare en onvoorziene ontwikkelingen (Van der Steen et al., 2018). Men ruilt de taal van *doelen* in voor die van *ambities* of *opgaven*: stippen op de horizon waar we naartoe reizen langs tal van deels parallelle, deels elkaar kruisende en beïnvloedende paden. Naarmate het beleid vordert, worden (tussen)doelen geconcretiseerd, maar het beleid blijft een doelzoekend, deliberatief en experimenteel karakter houden. Passende evaluatiemethoden zijn methoden die aansluiten bij dit karakter (Van der Steen et al., 2018). Eén methode die de laatste jaren in zwang lijkt te zijn, is die van lerend evalueren. Tijdens lerend evalueren werken beleidsonderzoekers samen met beleidsbetrokkenen om gaandeweg het beleid te verbeteren door gelijktijdig beleid te verantwoorden en te leren van de ervaringen met het beleid. Op deze manier, zo is de veronderstelling, vergroot lerend evalueren de kwaliteit en bruikbaarheid van de geproduceerde kennis, evenals de impact op beleid.

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

Het Planbureau voor de Leefomgeving (PBL; met Wageningen Research & University en Vrije Universiteit Amsterdam) heeft ervaring opgedaan met lerend evalueren in het natuurdossier (Folkert et al., 2018). De methode is relatief nieuw (ook voor de organisatie) en de toepassing op natuurbeleid vindt voor het eerst plaats op nationale schaal en in een interbestuurlijke (gedecentraliseerde) context.[1] Vanzelfsprekend roept dat vragen op. Kloppen de veronderstellingen: vergroot lerend evalueren inderdaad de kwaliteit, bruikbaarheid en impact van de geproduceerde kennis? En tegen welke kosten, wat is de doelmatigheid van zo'n evaluatiemethode? Voorliggend artikel presenteert een verkennend onderzoek dat zich richt op de eerste vraag. Onderzoekers van de VU hebben in opdracht van het PBL de ervaringen van beleidsbetrokkenen en de evaluatoren van de lerende evaluatie van het Natuurpact (2014-2017) opgetekend om zicht te krijgen op de waarde die deze methode heeft voor hun respectievelijke praktijken (Verwoerd et al., 2017; 2019). Het doel was om de doorwerking van de methode in beeld te brengen en lessen te trekken met betrekking tot de kwaliteit van de uitvoering voor de volgende evaluatieperiode (2018-2020). Dit artikel vat de uitkomsten van het VU-onderzoek samen en biedt een eerste inzage in de praktijk en manieren waarop lerend evalueren van waarde kan zijn voor zowel beleid als onderzoek. Ook worden spanningen waarmee de methode gepaard gaat en risico's van de methode – en hoe hier mee kan worden omgegaan – uitgelicht. Het artikel dient ter ondersteuning van de beleidsonderzoekers (en hun kennisorganisaties) en beleidsbetrokkenen die eenzelfde aanpak overwegen.

2 Achtergrond

2.1 Lerend evalueren: een ander perspectief op beleidsevaluatie

Lerend evalueren combineert elementen van verschillende 'evaluatiescholen', zoals tabel 1 uiteenzet. Het streven in lerende evaluaties is om zowel beleidsresultaten te verantwoorden (voor het gemak gevat onder *reguliere evaluaties*) als om aan te sluiten bij de leerbehoeften van de betrokkenen bij een beleidsdossier (hier aangemerkt als *responsieve evaluaties*). Tijdens een lerende evaluatie leveren evaluatoren maatwerk. Zij werken daarbij samen met de beleidsbetrokkenen om het evaluatieonderzoek, dat parallel loopt aan het beleidsproces (ex durante), goed aan te laten sluiten op de ontwikkelingen die daar spelen. Evaluatoren en beleidsbetrokkenen ontwikkelen samen een evaluatiekader, delen (voorlopige) analyses tijdens het onderzoek, en trachten in gezamenlijkheid betekenis te geven aan de conclusies die evaluatoren trekken, om in reactie daarop handelingsperspectieven te formuleren. Zo'n samenwerking vraagt om zorgvuldige borging van de onafhankelijke positie van de evaluatoren, die zowel onafhankelijk expert als facilitator van leerprocessen zijn.

Tabel 1 Kenmerken van reguliere, lerende en responsieve evaluaties (Van Veen et al., 2016)

	Reguliere evaluatie	**Lerende evaluatie**	**Responsieve evaluatie**
Doel van de evaluatie	Beleidsverantwoording gericht op doelrealisatie	*Beleidsverantwoording én leren van ervaringen om beleid en beleidsuitvoering te verbeteren*	Leren van ervaringen, beleid en beleidsuitvoering verbeteren

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

Type geproduceerde kennis	Expertkennis	*Co-creatie tussen expert- en praktijkkennis die leidt tot sociaal-robuuste kennis*	Praktijkkennis
Relatie met de beleidspraktijk	Gescheiden: evaluatie en beleidspraktijk zijn strikt gescheiden om onafhankelijkheid te garanderen	*Geïntegreerd: evaluatie is afgestemd op beleidspraktijk op een wijze die onafhankelijkheid borgt*	Gefuseerd: evaluatie is onderdeel van beleidspraktijk, onafhankelijkheid is geen criterium
Doelen en kaders van de evaluatie	Vooraf vastgestelde beleidsdoelen en kaders	*Vooraf vastgestelde beleidsdoelen en kaders en verwachtingen van beleidsbetrokkenen*	Verwachtingen van beleidsbetrokkenen zijn leidend voor doelen en kaders
Relevante stakeholders en hun rol	Opdrachtgever van de evaluatie, beleidsmaker wiens beleid wordt geëvalueerd. Zij verstrekken passief data ten behoeve van het onderzoek.	*Beleidsbetrokkenen die met handelingsperspectieven uit het evaluatieonderzoek aan de slag gaan. Zij hebben een actieve rol in het onderzoek.*	Alle relevante stakeholders: beleidsbetrokkenen, begunstigden en 'slachtoffers' van het beleid. Zij werken actief mee aan alle facetten van het evaluatieonderzoek.
Rol van de evaluator	Onafhankelijk expert	*Onafhankelijk expert en facilitator*	Facilitator

Het evaluatieonderzoek en de beleidspraktijk, zo is de aanname, verrijken elkaar middels deze samenwerking. Evaluatoren kunnen de beleidsrealiteit beter benaderen tijdens hun onderzoek en deze beter duiden door de kennis en perspectieven van beleidsbetrokkenen zoveel mogelijk te integreren met inzichten uit de studie (Shulha et al., 2016). Hierdoor ontstaat een groter inzicht in de effectiviteit en de werking van het beleid, en wordt de kwaliteit van het evaluatieonderzoek verbeterd. Beleidsbetrokkenen krijgen door hun actieve betrokkenheid meer gevoel voor hoe de kennis tot stand komt, hoe deze te interpreteren en hoe deze toe te passen in de praktijk (Scarinci et al., 2017). Omdat de kennis en perspectieven van beleidsbetrokkenen worden geïntegreerd in het onderzoek, is de kennis die ontstaat niet alleen wetenschappelijk, maar ook sociaal-robuust (Nowotny, 2000; Polk, 2014) – dat wil zeggen: legitiem, herkenbaar en toepasbaar voor beleidsbetrokkenen. Lerend evalueren vergroot zo de kwaliteit en bruikbaarheid van de geproduceerde kennis en draagt bij aan een kennis-verrijkte beleidspraktijk (Van Veen et al., 2016). Uiteindelijk, zo is de hypothese, heeft lerend evalueren de potentie om een grotere impact te hebben op de beleidspraktijk en de samenleving.

De beloften van lerend evalueren zijn behoorlijk. Daarbij vraagt het uitvoeren van een lerende evaluatie om een investering van budget en capaciteit van zowel evaluatoren en beoogde deelnemers uit de beleidspraktijk (Edelenbos & Van Buuren, 2005). Logischerwijs is er behoefte aan inzicht in wat de opbrengsten zijn van lerend evalueren, en of de beloften waar worden gemaakt. Voorliggend

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

artikel is een eerste verkenning van de waarde van lerend evalueren in de praktijk van het natuurbeleid. Een impact of kosten-batenanalyse van de methode valt buiten de scope van dit artikel. De onderzoekers hebben zich gericht op het in beeld brengen van de ervaringen van onderzoekers zelf en de beleidsbetrokkenen, en wat in hun optiek de waarde is van lerend evalueren in het licht van complexe maatschappelijke vraagstukken, zoals natuurbeleid.

2.2 Het Natuurpact

De centrale casus betreft de eerste evaluatieperiode van de lerende evaluatie van het Natuurpact (2014-2017). Het Natuurpact (2013) vormt het bestuurlijke sluitstuk van de onderhandelingen tussen het Rijk en provincies over de decentralisatie van het natuurbeleid. De provincies zijn hiermee verantwoordelijk geworden voor het formuleren en implementeren van natuurbeleid om zo de ambities uit het pact te realiseren in 2027. Het Rijk blijft verantwoordelijk voor de Europese biodiversiteitsverplichtingen uit de Vogel- en Habitatrichtlijn (VHR) en de Kaderrichtlijn Water (KRW).

De natuurambities zijn in het Natuurpact verbreed. Naast internationaal overeengekomen biodiversiteitsdoelen – een weerbarstige opgave an sich – staan ook het versterken van maatschappelijke betrokkenheid en het verbinden van economie en natuur centraal. Om de ambities te realiseren zijn de provincies volop aan het experimenteren en werken zij in toenemende mate samen met maatschappelijke partijen.

Op verzoek van het ministerie van Landbouw, Natuur en Voedselkwaliteit (LNV, destijds Economische Zaken) en het Interprovinciaal Overleg (IPO; vertegenwoordiger van de provincies) evalueren het PBL en partners eens in de drie jaar de voortgang van het Natuurpact. De nieuwe bestuurlijke context en de verbrede ambities maakten dat er veel valt te leren door Rijk en provincies. Daarnaast stonden de relaties tussen Rijk en provincies bij aanvang van de evaluatie onder spanning, een gevolg van een stroef decentralisatieproces. Daarom is gekozen voor een lerende evaluatieaanpak die tussentijds inzichten geeft in de werking van beleid om tijdig bij te kunnen sturen, die uitwisseling van ervaringen centraal stelt, en die het mogelijk maakt om Rijk en provincies op constructieve wijze te betrekken bij het evaluatieonderzoek.

2.3 De lerende evaluatie van het Natuurpact

Inhoudelijk

Het Rijk en provincies hebben het PBL verzocht te onderzoeken of het gevoerde natuurbeleid van de provincies effectief en efficiënt is, met aandacht voor de rol van de geboden rijkskaders. Daarnaast is verzocht om beleidsopties en handelingsperspectieven voor te stellen waarmee de ambities uit het Natuurpact effectiever en efficiënter kunnen worden gerealiseerd (PBL & WUR, 2017). Deze eerste evaluatieperiode had het karakter van een ex ante; er is hoofdzakelijk gekeken naar de potentiële effecten van provinciaal natuurbeleid op de ambities, omdat provincies nog aan het begin van hun uitvoeringsperiode stonden.[2] Gezamenlijk met Rijk, provincies en maatschappelijke organisaties heeft het PBL het evaluatiekader vastgesteld. De drie ambities uit het Natuurpact zijn hierin opgenomen. Voor de biodiversiteitsambitie zijn de doelen van de Europese Vogel- en Habitatrichtlijn (VHR) en de Kaderrichtlijn Water (KRW) gebruikt als

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

concrete uitwerking en als toetsingskader voor effectiviteit. Voor de verbrede ambities ontbreken zulke concrete doelstellingen. Het onderzoek had voor deze ambities een meer verkennend karakter. Het provinciaal beleid werd dan ook verantwoord op het vlak van voortgang met de biodiversiteitsambitie. Voor beleidsleren lag de focus op beleidsvernieuwingen voor natuurontwikkeling en -beheer, en beleidsstrategieën voor maatschappelijke betrokkenheid en verdienmodellen voor natuur.

Procesmatig

Een interdisciplinair kernteam van zes onderzoekers van het PBL en de WUR (inclusief projectleiders) heeft de lerende evaluatie uitgevoerd. Hun ervaring met lerende evaluatiemethoden was relatief beperkt. Daarom heeft het PBL het Athena Instituut (VU Amsterdam) verzocht het team vanuit hun expertise op het gebied van lerende onderzoeksmethoden te adviseren. Vanuit de beleidspraktijk namen hoofdzakelijk provinciale beleidsbetrokkenen en vertegenwoordigers van het Rijk die betrokken waren bij het Natuurpact, deel aan de evaluatie. Maatschappelijke partijen zijn ook betrokken, maar meer als consultant dan als actief deelnemende actoren.[3]

De evaluatie bestond uit drie fasen: (1) ontwikkeling van een gedeeld evaluatiekader; (2) analyse en gedeelde interpretatie van de bevindingen; en (3) trekken van conclusies en verspreiding daarvan (figuur 1). De lerende evaluatie is opgezet als een interactief proces tussen beleid en onderzoek. Deze interacties bestonden ten eerste uit vier overkoepelende leer- of werksessies waar vertegenwoordigers van Rijk, provincies en maatschappelijke partijen aan deelnamen. Het doel verschilde per evaluatiefase; van het verzamelen van evaluatiebehoeften, naar het gezamenlijk formuleren van handelingsperspectieven. Ten tweede vond structureel interactie plaats met de IPO-werkgroep Natuurbeleid om de kwaliteit en bruikbaarheid van de evaluatie af te stemmen. In deze werkgroep is elke provincie vertegenwoordigd door een beleidsmedewerker. Samen vertegenwoordigen zij de evaluatiebehoeften van de provincies. Ten derde vond er interactie plaats tijdens reguliere onderzoeksactiviteiten, zoals interviews met beleidsbetrokkenen en groepsgesprekken. Ten slotte was er een aantal keren per jaar een ambtelijk en bestuurlijk overleg met LNV en IPO om de voortgang en landing van het onderzoek te waarborgen.

Figuur 1 Schematische weergave van de lerende evaluatie van het Natuurpact (2014-2017) met daarin de drie fases van het onderzoek inclusief de centrale activiteiten

3 Methodologie

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

Voor deze studie is een volledig kwalitatieve onderzoeksbenadering toegepast om een rijkheid aan ervaringen met de lerende evaluatie van het Natuurpact te vergaren. De vraagstelling richting beleidsbetrokkenen en onderzoekers was daarbij verschillend. Bij beleidsbetrokkenen lag de nadruk op het in beeld brengen van de waarde van de lerende evaluatie voor hun beleidspraktijk via interviews, en het zoeken naar aanwijzingen in beleidsplannen en documenten die deze waarde bevestigen om zo de bevindingen te trianguleren (Denzin, 2012). Bij de onderzoekers stonden hun ervaringen en de uitdagingen van het uitvoeren van de methode centraal, en de betekenis daarvan voor hun onderzoekspraktijk. De weergave van hun perspectieven in de resultaten heeft dan ook een wat ander accent. In de volgende paragrafen lichten we de aanpak nader toe.

3.1 Dataverzameling

Interviews

Er zijn elf semi-open diepte-interviews uitgevoerd met beleidsbetrokkenen (vijf provinciale beleidsbetrokkenen van wie drie uit de werkgroep natuurbeleid, drie vertegenwoordigers van LNV, twee vertegenwoordigers van het IPO). Semi-open diepte-interviews laten ruimte aan de respondent voor het aansnijden van onderwerpen die volgens hem of haar relevant zijn en voor het uitdiepen en duiden van ervaringen en opvattingen (Creswell, 2009). De respondenten zijn geselecteerd op diversiteit in de mate van betrokkenheid (van intensief tot matig) en thuisorganisaties om zo diversiteit aan ervaringen en perspectieven te kunnen vangen. De respondenten is gevraagd naar de wijzen waarop de lerende evaluatie (zowel de deelname aan het proces als het eindresultaat) van waarde is geweest voor hun beleidspraktijk. Deze vraagstelling was bewust open en verkennend. Door geen voorbedacht ‘impactkader’ aan te reiken is gepoogd de respondenten zo min mogelijk te sturen en zo ook mogelijk onverwachte vormen van impact te verzamelen.

Voor de perspectieven van de betrokken onderzoekers zijn zeven semi-open diepte-interviews gehouden. Hiervoor is geselecteerd op diversiteit in thuisorganisaties (PBL/WUR), disciplinaire achtergrond (ecologie, bestuurskunde, economie) en functie (plaatsvervangend projectleider, deelprojectleider, onderzoeker). Tijdens de interviews is de onderzoekers gevraagd wat in hun beleving de waarde is geweest van de lerende evaluatiemethode voor kwaliteit, bruikbaarheid en impact en wat daarvoor succes- en risicofactoren zijn geweest.

Deskresearch

De interviews met beleidsbetrokkenen hebben plaatsgevonden tijdens de afronding van de eindrapportage. Mogelijk meer structurele aanpassingen in beleid konden destijds nog beperkt worden aangetoond. Een latere analyse van de provinciale natuurplannen (de Natuurvisies) is gedaan om hier alsnog zicht op te krijgen. Daarbij is gekeken naar verwijzingen naar de eindrapportage in deze documenten en andere aanwijzingen dat de aanbevelingen zijn opgepakt.

3.2 Data-analyse

Van alle interviews zijn audio-opnamen gemaakt, die vervolgens zijn getranscribeerd. De transcripten zijn thematisch gecodeerd (Hsieh & Shannon, 2005) met gebruik van MAXQDA-software om perspectieven op de waarde voor

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

beleids- en onderzoekspraktijk te identificeren.

De VU-onderzoekers waren behalve via deze studies ook op andere manieren betrokken bij de lerende evaluatie. Zij hadden een ondersteunende rol in het evaluatieontwerp en het proces; er is actie-onderzoek uitgevoerd. Dit biedt een waardevol 'insider'-perspectief op het evaluatieonderzoek en op de ervaringen van beleidsbetrokkenen en onderzoekers (Verwoerd et al., 2020). Om de validiteit en betrouwbaarheid van het onderzoek te bewaken zijn er verschillende strategieën toegepast om 'observational biases' (Robson, 2002) te beperken en de intersubjectiviteit (Creswell, 2009) te borgen. Onderzoekers die onbekend waren met het Natuurpact, zijn betrokken bij de analyses en interpretaties voor de onafhankelijkheid in oordeelsvorming. Daarnaast zijn om de validiteit en betrouwbaarheid te toetsen en de data te verrijken voor beide studies de conceptresultaten voorgelegd aan zowel de respondenten als andere relevante actoren uit de beleids- en onderzoekspraktijk. Ook hebben er wetenschappelijke reviews plaatsgevonden door onafhankelijke experts van de Universiteit Utrecht en Erasmus Universiteit Rotterdam op het gebied van lerend onderzoek voor leefomgevingsvraagstukken.

4 Resultaten

4.1 Kwaliteit

Een van de veronderstellingen van lerend evalueren is dat de samenwerking tussen onderzoekers en beleidsmakers ertoe leidt dat het onderzoek de beleidsrealiteit beter kan benaderen, waardoor de validiteit en betrouwbaarheid toenemen. De ervaringen van verschillende onderzoekers bevestigen dit voor de Natuurpact-evaluatie. Tijdens zogenaamde 'bilaterale consultaties' gingen zij langs bij provincies om samen hun natuurkaarten en beleidsplannen te bespreken. Hierdoor konden onderzoekers de plannen beter interpreteren en gebruiken als data voor hun modelanalyses. De samenwerking leidde tot een meer accurate interpretatie en soms ook tot meer data, omdat provincies visa versa beter begrepen waar onderzoekers naar op zoek waren.

> *'De bilaterale consultaties maakten dat we hun beleidsplannen veel beter snapten. Vaak waren die of niet compleet, of begrepen wij gewoon niet wat er werd bedoeld. Dan moesten we maar wat proberen af te leiden. Zonder hun uitleg hadden we dat nooit even goed kunnen doen.'* – Onderzoeker 1

Tegelijkertijd klinkt er ook een kritischer geluid onder onderzoekers over de waarde van de samenwerking. Hier worden objectiviteit en onafhankelijkheid in de oordeelsvorming als maat genomen voor de kwaliteit van kennis, en sommige onderzoekers stellen dat het bewaken daarvan extra aandacht behoefde door de vele interacties die er waren. Door de korte lijnen tussen onderzoek en beleid konden provincies onderzoekers bijvoorbeeld makkelijker benaderen om bevindingen ter discussie te stellen:

> *'Wat betreft wetenschappelijke kwaliteit, sommige provincies stuurden* [aan] *op andere framing van bepaalde bevindingen. De druk was wel hoog, want we wilden tegelijkertijd natuurlijk niet dat ze zouden zeggen: "We herkennen dit niet, dus we nemen er afstand van." Dat is wel een inherente spanning van lerend evalueren.'* – Onderzoeker 2

Volgens sommige onderzoekers stonden door de interacties de objectiviteit en

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

onafhankelijkheid onder spanning met het realiseren van sociale robuustheid – beide graadmeters voor de kwaliteit van de geproduceerde kennis. Om hiermee om te gaan heeft het Natuurpact-team extra acties uitgevoerd om hun onafhankelijke oordeelsvorming te bewaken zonder de samenwerking teniet te doen. Zo zijn onder andere de rollen in het team verdeeld tussen zij die interacteerden met beleid en zij die conclusies trekken. Ook is de Natuurpactevaluatie, om de wetenschappelijke en sociale robuustheid van het onderzoek te waarborgen, tijdens het traject met regelmaat intern en extern kritisch tegen het licht gehouden door zowel collega-onderzoekers als vertegenwoordigers uit de beleidspraktijk (zie ook Folkert et al., 2018). De bevindingen laten zien dat de onderzoekers, ondanks de verschillende invullingen die zij geven aan kwaliteit, en ondanks de spanning die deze invullingen soms teweegbrachten, over het algemeen inschatten dat de lerende evaluatiemethode een positief effect heeft gehad op de kwaliteit van het onderzoek.

Vanuit het perspectief van de beleidsbetrokkenen ontstaat een vergelijkbaar beeld: ook zij staan in het algemeen positief tegenover de kwaliteit van de evaluatie. Opvallend is echter dat zij kwaliteit minder benaderen vanuit een wetenschappelijk oogpunt (in termen zoals wetenschappelijke robuustheid of validiteit), maar meer vanuit hun eigen praktijk. Voor hen waren de herkenbaarheid en bruikbaarheid van de evaluatiebevindingen de voornaamste graadmeter voor de kwaliteit van de evaluatie. In de volgende paragraaf werken we dit verder uit.

4.2 Bruikbaarheid

Volgens beleidsbetrokkenen

Volgens beleidsbetrokkenen is de lerende evaluatie op diverse manieren bruikbaar geweest. ‘Bruikbaar’ wordt hier breed geïnterpreteerd, namelijk als alle manieren waarop de evaluatie van waarde is geweest voor hun beleidspraktijk. We identificeren vijf verschillende effecten: (i) *cognitief-inhoudelijk,* (ii) *politiek-strategisch,* (iii) *netwerk,* (iv) *affectief* en een (v) *instrumenteel effect*. In de volgende tekst lichten we deze toe, tabel 2 laat diverse illustratieve quotes zien.

(i) Cognitief-inhoudelijk effect: Het proces en de uitkomsten van het onderzoek droegen bij aan inhoudelijke verdieping van de kennis van beleidsbetrokkenen over het natuurbeleid, vooral bij de provincies. De kennis over de (potentiële) doelmatigheid en effectiviteit van huidige beleidsstrategieën voor biodiversiteit is vergroot. Ook zijn provincies bekend(er) geraakt met de werking van alternatieve beleidsstrategieën, zoals nieuwe uitvoeringsarrangementen en omgaan met groene burgerinitiatieven. Voor de verbrede ambities – maatschappelijke betrokkenheid, en verbinden economie en natuur – is geconstateerd dat er nog meer aandacht moet worden besteed aan de uitwerking hiervan in beleidstheorieën. De respondenten beschrijven daarbij dat ze door de lerende evaluatie beter bekend zijn geraakt met elkaars perspectieven op natuurbeleid. Er is een meer gemeenschappelijke taal ontstaan en meer congruentie in de verschillende benaderingen.

Tabel 2 Quotes van respondenten om de diverse wijzen waarop lerend evalueren van waarde was voor de natuurbeleidspraktijk te illustreren

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

Type effect voor beleidspraktijk	Quotes van respondenten
(i) Cognitief-inhoudelijk	*'In onze provincie benaderen we het Natuurnetwerk best wel op een behouden, zeg maar ouderwetse wijze – daar werd ik me ineens bewust van. Er zijn veel meer maatregelen mogelijk dan aankoop en herbestemming voor het Natuurnetwerk.'* – Provincievertegenwoordiger 1 *'We hadden het over dingen als: Waarom hebben we dit eigenlijk zo gedaan? En werkt het eigenlijk? Daardoor gingen we veel dieper dan dat we zelf ooit zouden doen.'* – Provincievertegenwoordiger 2
(ii) Instrumenteel	*'We gebruiken de drie ambities en de aanbevelingen uit het rapport om vorm te geven aan onze provinciale Natuurvisie.'* – Provincievertegenwoordiger 1 *'De inhoud van de evaluatie structureert straks het gesprek tussen Rijk en provincies, wanneer we met elkaar afstemmen hoe we vervolg gaan geven aan de aanbevelingen. Daar is de evaluatie belangrijke input voor.'* – EZ-vertegenwoordiger 1
(iii) Politiek-strategisch	*'Het helpt enorm om richting maatschappelijke organisaties te kunnen zeggen: "PBL zegt zus en zo, dus we gaan ons óók richten op maatschappelijke betrokkenheid bij natuur."'* – Provincievertegenwoordiger 3
(iv) Netwerk-	*'Je ziet elkaar ineens wat meer, en je hebt de tijd om er wat meer "boven" te gaan hangen. En ik raakte bekend met nieuwe partijen die ik niet eerder heb gesproken.'* – Provincievertegenwoordiger 4
(v) Affectief	*'Het was eigenlijk ook pure opluchting: we zijn niet de enige die zoekend zijn! En een gevoel van: we staan er samen voor.'* – Provincievertegenwoordiger 2

(ii) Instrumenteel effect: Het cognitief-inhoudelijk effect was op verschillende vlakken instrumenteel en droeg bij aan concrete beleidsacties van provincies. De evaluatie bood bijvoorbeeld direct input voor de Natuurvisies waar diverse provincies op dat moment mee bezig waren. Ook op interbestuurlijk vlak was er sprake van instrumentele waarde: de aanbevelingen boden houvast bij de gesprekken tussen Rijk en provincies over de voortgang en richting van het natuurbeleid, en droeg zo bij aan beleidscoördinatie.

(iii) Politiek-strategisch effect: Het evaluatieonderzoek stelde provincies in staat beleidskeuzes richting collega's, andere overheden en maatschappelijke organisaties beter te legitimeren, met name wat betreft beleidsinzet op het vergroten van maatschappelijke betrokkenheid. De evaluatie droeg bij aan het gevoel van urgentie voor natuurbeleid door de concrete boodschap dat met de huidige plannen in 2027 maximaal 65 procent van de VHR-doelstellingen gehaald zullen worden. Ook droeg de evaluatie bij aan het op de agenda zetten van de verbreding van de ambities van het natuurbeleid in de verschillende

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

provincies, aldus betrokkenen. Belangrijk voor dit politiek-strategisch effect was de status van onafhankelijkheid van het eindrapport, wat bijdroeg aan de geloofwaardigheid van het onderzoek.

(iv) Netwerkeffect: Deelname aan de lerende evaluatie heeft betekenis gehad voor het netwerk van deelnemers. Tijdens verschillende workshops zijn nieuwe contacten gelegd tussen vertegenwoordigers van provincies, Rijk, terreinbeheerders en maatschappelijke organisaties, en zijn bestaande contacten versterkt. Provincies geven aan beter zicht te hebben gekregen op 'wie waarmee bezig is' en dat de lijntjes korter zijn geworden.

(v) Affectief effect: Verrassend was de constatering van het *affectief effect*: als gevolg van deelname is een gevoel van saamhorigheid ontstaan onder provincies. Constateren dat ze worstelen met vergelijkbare uitdagingen was een belangrijke eyeopener en gaf een soort opluchting. Het affectief effect werd heel belangrijk gevonden voor de gemeenschapszin onder provincies en werkte versterkend voor de kennisuitwisseling onderling. Ook spraken respondenten over een voorzichtige verbetering van de verstandhouding tussen EZ en de provincies dankzij regelmatig contact in het kader van de evaluatie. Het netwerkeffect lijkt daarmee een randvoorwaarde voor het realiseren van een affectief effect.

Niet alle geluiden waren lovend. Een aantal beleidsbetrokkenen is kritisch op de cognitief-inhoudelijke en instrumentele waarde. De beleidsstrategieën die zijn onderzocht, zoals nieuwe uitvoeringsarrangementen, procesbeheer en natuur-inclusieve landbouw, sloten niet altijd aan op hun specifieke provinciale beleidscontext. Het evaluatieonderzoek was daardoor niet voor elke provincie even relevant. Sommigen hadden überhaupt minder behoefte aan uitwisseling met andere provincies of het Rijk:

> *'Het is wel leuk om collega's uit andere provincies te zien, maar dat zoeken we ook zelf wel op als dat nodig is.'* – Provincievertegenwoordiger 5

Daarnaast werd de bruikbaarheid voor provincies beïnvloed door het schaalniveau waarop de bevindingen betrekking hadden. Als nationaal instituut doet het PBL uitspraken op nationaal niveau, en bood daardoor in eerste instantie verminderd inzichten in provincie-specifiek beleid. We komen hierop terug in de volgende paragraaf, waar we bruikbaarheid bespreken vanuit het perspectief van onderzoekers.

Volgens onderzoekers

Onderzoekers schatten dat de lerende aanpak de bruikbaarheid van de evaluatie over het algemeen heeft vergroot, vooral voor provincies. Door het onderzoek en de workshops af te stemmen op het gros van de evaluatiebehoeften is er gestuurd op relevantie. De samenwerking heeft daarnaast bijgedragen aan een gemeenschappelijke taal, waardoor het onderzoek beter aansloot bij de beleefwereld van beleidsbetrokkenen, bijvoorbeeld de termen 'plan- en uitvoeringspotentieel' om het potentiële effect van beleidsplannen op biodiversiteit en succes- en risicofactoren voor de uitvoering daarvan te bespreken:

> *'Je kon echt zien dat we aansloten bij hun referentiekaders. Ze konden de bevindingen snel vertalen naar hun eigen praktijken en deze vervolgens met ons bediscussiëren. Dat onderstreept wel de toepasbaarheid.'* – Onderzoeker 3

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

Zoals de ervaringen van de beleidsbetrokkenen ook laten zien, was het borgen van bruikbaarheid tegelijk ook een zoektocht voor onderzoekers. De evaluatiebehoeften liepen uiteen, en voor Rijk en provincies had 'bruikbaar onderzoek' een andere betekenis. De nationale schaal waarop het PBL normaliter uitspraken doet, diende vooral een verantwoordingsfunctie, en bood voor provincies beperkt inzage in hun eigen natuurbeleid en was onvoldoende bruikbaar om van te leren. Voor het evaluatieteam was dit schakelen. Om teleurstelling te voorkomen hebben zij extra activiteiten ondernomen om deze bruikbaarheid toch aan te kunnen bieden, zoals het maken van factsheets per provincie en een aanvullende achtergrondrapportage over potentiële effecten van provinciale beleidsstrategieën op biodiversiteit per provincie.

> *'Dat we geen bevindingen zouden produceren op het provinciale schaalniveau was best een verrassing voor de provincies. Voor ons was het gewoon logisch: dat model geeft gewoon geen betrouwbare output op regionale schaal. Om hen toch tegemoet te komen hebben we voor elke provincie een factsheet gemaakt met toch zoveel mogelijk informatie op hun schaalniveau. Dat besloeg het grootste aandeel van mijn werk uiteindelijk … maar het kwam de bruikbaarheid enorm ten goede.'* – Onderzoeker 4

Daarnaast speelden de perspectieven van onderzoekers zelf op wat relevant en bruikbaar onderzoek is ook een rol en moesten zij een balans vinden tussen vraag- en aanbodgestuurd werken. Een aantal onderzoekers had een eigen standpunt over wat relevante onderwerpen waren voor het evaluatieonderzoek. Door samen met beleidsbetrokkenen de scope te bepalen zijn niet al deze onderwerpen volledig geland en uitgevoerd. Bijvoorbeeld een deelstudie naar de kosteneffectiviteit van natuurbeleid: gegeven de recente decentralisatie vonden provincies dit te prematuur en zij stelden dat uitspraken hierover mogelijk nieuw beleid in de kiem zouden smoren. Hun angst voor afrekening door het Rijk speelde hier een rol. Onderzoekers vonden dat een dergelijke deelstudie juist de kennis over effectieve beleidsstrategieën ten goede was gekomen en daardoor van betekenis zou zijn geweest voor de bruikbaarheid.

4.3 Impact

Onder impact wordt verstaan structurele beleidsmatige aanpassingen waaraan het evaluatieonderzoek heeft bijgedragen. Dergelijke bijdragen zijn zichtbaar in bijvoorbeeld de natuurvisies die provincies sinds 2017 verplicht zijn om op te stellen. De helft[4] van de provincies refereert daarin naar de lerende evaluatie van het Natuurpact en licht toe hoe wordt ingezet op de natuurambities – met in de regel veel aandacht voor maatschappelijke betrokkenheid – en de aanbevelingen uit het rapport. Een ander vlak waarop de aanbevelingen doorwerking hebben gehad, betreft de oprichting van het Leernetwerk Vernieuwing Natuurbeleid. In dit formele leernetwerk pakken provincies leervragen op die buiten de scope van de evaluatie vallen, met thema's zoals klimaatinclusief natuurbeleid en natuurinclusief bouwen. Ook zijn de aanbevelingen uit het eindrapport opgepakt door het Rijk en zijn er werkgroepen aangewezen om de aanbeveling voor het vergroten van de maatschappelijke betrokkenheid op te pakken en is in het licht hiervan opdracht gegeven tot de uitwerking ervan. Al met al: vanuit het perspectief van de beleidspraktijk zijn verschillende soorten doorwerking op verschillende bestuurslagen zichtbaar.

Ook het evaluatieteam beschouwt het feit dat de evaluatie heeft bijgedragen aan de natuurvisies als concreet bewijs hiervoor. Maar, net als voor kwaliteit en

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

bruikbaarheid, is een aantal onderzoekers ook kritisch op de lerende evaluatiemethode voor wat betreft impact. Al eerder kwam ter sprake dat sommigen de samenwerking als een risico hebben ervaren voor de objectiviteit en onafhankelijkheid in de oordeelsvorming, en dat beeld wordt hier bevestigd:

> *'Dat interactieve proces tijdens de rapportage, daardoor zijn de scherpe randjes van sommige formuleringen er wel af gegaan. Inhoudelijk staat er niets anders, het was meer een glas-halfvol-halfleeg-discussie. Het was navigeren tussen impact en wetenschappelijke kwaliteit.'* – Onderzoeker 2

Bovenstaand citaat illustreert de al eerder besproken spanning, namelijk tussen enerzijds het aanmoedigen van eigenaarschap van beleidsbetrokkenen over de bevindingen door zorg te dragen voor de sociale robuustheid van het onderzoek, en anderzijds het waarborgen van geloofwaardig onderzoek door de onafhankelijke oordeelsvorming en autonomie te bewaken. Door transparant en openbaar te zijn in het onderzoek, door bijvoorbeeld de aanpassingen die zijn gemaakt naar aanleiding van input van beleidsbetrokkenen en de onderbouwing daarvoor goed te documenteren en beschikbaar te stellen, heeft het Natuurpact-team gepoogd beide perspectieven op impact samen te brengen en de impact te vergroten.

5 Conclusie en discussie

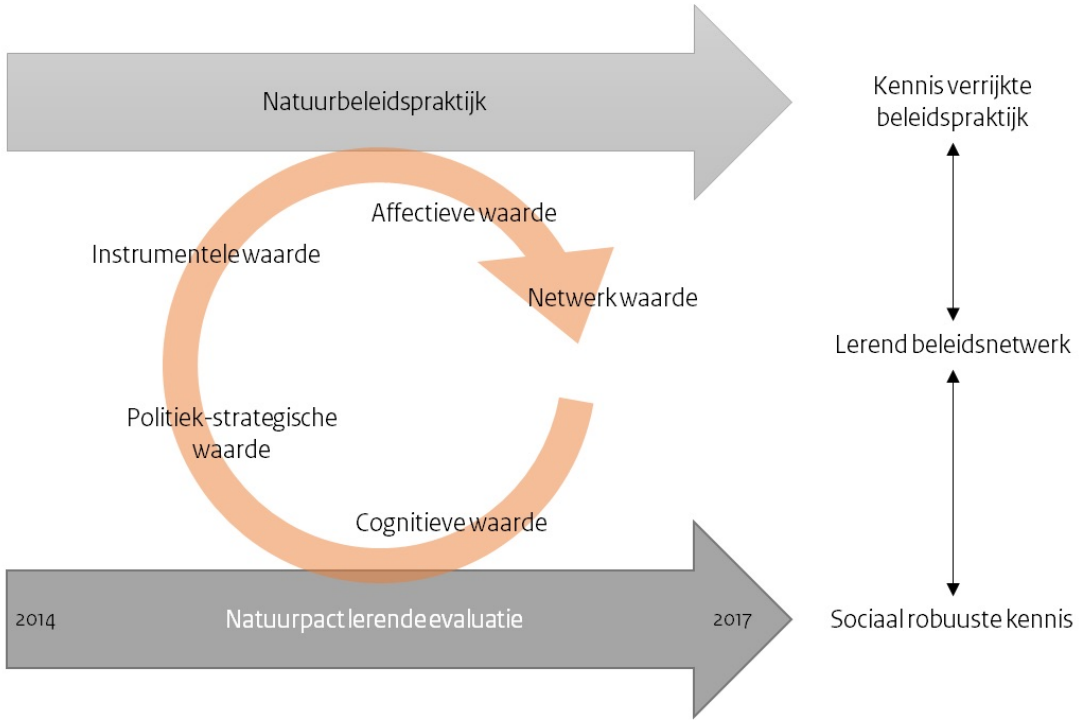
Dit artikel nam als vertrekpunt de veronderstelling dat lerend evalueren middels een interactief proces en het combineren van leren en verantwoorden van beleid, de kwaliteit, bruikbaarheid en impact van de geproduceerde kennis vergroot. Deze veronderstelling is in beschouwing genomen door de ervaringen van beleidsbetrokkenen en betrokken onderzoekers aan de lerende evaluatie van het Natuurpact (2014-2017) in beeld te brengen en hun kijk op de waarde van lerend evalueren op te tekenen. In deze slotsectie relateren we de bevindingen aan wat er in vakliteratuur wordt geschreven over de doorwerking van reguliere en lerende (en participatieve) evaluaties. Ook staan we stil bij risico's en uitdagingen die lerend evalueren met zich meebrengt en hoe daarmee kan worden omgegaan.

De belangstelling voor de doorwerking van evaluatiestudies kent een lange geschiedenis. Waar in de twintigste eeuw doorwerking hoofdzakelijk als het instrumenteel gebruik van resultaten in beleidsprocessen werd begrepen, is deze opvatting met de jaren verbreed. Naast een instrumenteel effect dat direct leidt tot beleidsactie (beleid wordt bijgestuurd of stopgezet), wordt ook de cognitieve betekenis van evaluatie onderschreven (kennis over beleidsdoelen en beleidswerking is vergroot) en het symbolisch of strategisch gebruik van evaluatie (bijvoorbeeld om urgentie te creëren voor een politieke agenda) (Kirkhart, 2000). De ervaringen van de deelnemers aan de lerende evaluatie van het Natuurpact wijzen erop dat doorwerking inderdaad langs deze assen heeft plaatsgevonden en dat er sprake is van een met kennis verrijkte beleidspraktijk. Was het lerende karakter van de evaluatie hier dan doorslaggevend voor? Op basis van deze studie lijkt het antwoord hierop positief: door de interactie tussen beleidspraktijk en onderzoek sloot het kennisproductieproces aan bij het gros van de belevingswerelden van de deelnemers, met meer diepgaand begrip en acceptatie van de resultaten als uitkomst. In de literatuur wordt met betrekking tot interactieve of participatieve evaluaties geschreven over sociaal-robuuste kennis als uitkomst: geloofwaardige en bruikbare kennis om van te leren en beleid te verbeteren (Chouinard, 2013; Nowotny, 2000). De overwegend positieve geluiden van de deelnemers onderschrijven deze uitkomst. Aanvullend

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

zijn een netwerk- en een affectief effect van de evaluatie geïdentificeerd, effecten die kunnen worden toegeschreven aan het interactieve evaluatieproces. In studies naar de doorwerking van meer interactieve en participatieve (transdisciplinaire) onderzoeksmethoden zijn dergelijke effecten al langer bekend (Walter et al., 2007;Wiek et al., 2014). Ontwikkelingen op het vlak van gevoelens van saamhorigheid (het 'community-gevoel') en een toenemende mate van beleidscoördinatie worden dan als voorbeelden genoemd. De lerende evaluatie heeft op dit vlak betekenis gehad voor de ontwikkeling van een lerend beleidsnetwerk onder de deelnemende partijen, onder provincies onderling, en (in mindere mate) tussen Rijk en provincies (zie figuur 2). De bevindingen bevestigen de veronderstelling dat een lerende evaluatieaanpak bijdraagt aan de kwaliteit, bruikbaarheid en (hoewel minder aantoonbaar) de impact van de geproduceerde kennis.

Figuur 2 De ontstane tussentijdse effecten die bijdragen aan sociaal-robuuste kennis, een lerend beleidsnetwerk en een kennis-verrijkte beleidspraktijk



Desalniettemin is er ook kritiek op de methode. De evaluatiebehoeften van Rijk en provincies onderling lopen uiteen en niet voor alle betrokkenen was de bruikbaarheid daardoor even groot. Provincies onderling verschillen van mening over de vraag of de evaluatie door het lerende karakter niet te vriendelijk is geweest, waardoor de harde boodschap dat het provinciaal natuurbeleid onvoldoende effectief is om de natuurdoelen te halen (te) beperkt is geland. De ervaringen van onderzoekers laten zien dat het een zoektocht was naar het bewaken van de balans tussen het borgen van betrokkenheid en draagvlak onder provincies voor de bevindingen, en hun onafhankelijke oordeelsvorming – een balans tussen leren en verantwoorden. Het combineren van leren en verantwoorden in één evaluatiestudie brengt uitdagingen met zich mee. Zogenaamde dichotomieën – kennis coproduceren en wetenschappelijke autonomie; vraag- en aanbodgestuurd onderzoek doen; participatief onderzoek en onafhankelijkheid – manifesteren zich als spanningen op het vlak van conflicterende rollen en kwaliteitsstandaarden van evaluatieonderzoek (Chouinard, 2013), ook zichtbaar in de Natuurpact-casus. In de vakliteratuur wordt geschreven dat het samenbrengen van leren en verantwoorden vraagt om het overstijgen van dergelijke ogenschijnlijke dichotomieën (Guijt, 2010; Regeer

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

et al., 2016). Acties van het Natuurpact-team illustreren hoe dit mogelijk wordt door leren en verantwoorden als wederzijds versterkend in plaats van tegenstrijdig te benaderen. Het feit dat het provinciaal beleid werd verantwoord, gaf urgentie aan betrokkenheid bij de evaluatie, en maakte zo leerprocessen mogelijk. De risico's van afhankelijk onderzoek en subjectiviteit zijn daarbij niet opgelost. Tijdens een lerende evaluatie vragen deze van de onderzoekers continu aandacht en zorg voor een transparant, openbaar en navolgbaar onderzoeksproces.

Een ander kritiek punt is dat de diversiteit onder de deelnemende partijen relatief laag was: het evaluatieonderzoek was sterk gericht op de provincies en de rol van het Rijk als systeemverantwoordelijke is beperkt aan bod gekomen. Maatschappelijke organisaties en het bedrijfsleven waren beperkt vertegenwoordigd, en inwonersorganisaties waren niet betrokken. Voor een legitiem en geloofwaardig kennisproces wordt een diversiteit aan betrokken stakeholders juist aanbevolen (Shulha et al., 2016). De breedte aan perspectieven maakt de kennis meer sociaal-robuust en relevant om complexe opgaven het hoofd te bieden. Omdat de lerende methode nieuw was voor onderzoekers en beleidsbetrokkenen, en vanwege de spanningen die er bij aanvang waren tussen Rijk en provincies, is de diversiteit van deelnemers bewust laag gehouden om de bereidwilligheid van de provincies voor actieve deelname te borgen. Dit heeft consequenties gehad voor de bredere legitimiteit van het onderzoek en het is aannemelijk dat dit eveneens effect gehad heeft op de mate van doorwerking. Om deze reden is voor de volgende evaluatieperiode (2018-2020, reeds afgerond) door het PBL ingezet op een bredere kring van maatschappelijke betrokkenheid tijdens het evaluatieonderzoek.

Mede door de vele interacties en sessies die plaatsvonden tijdens de evaluatie, waren de kosten hoog. Het onderzoek vroeg van zowel beleid als onderzoek intensieve betrokkenheid en capaciteit. Respondenten van beide werelden onderschrijven dat deze intensiviteit van belang was voor de diepgang van de politiek-strategische, inhoudelijke en instrumentele waarde, en de netwerk- en affectieve effecten die zijn ontstaan. Een analyse van de doelmatigheid van de methode ontbreekt echter nog. Opvallend is dat beleidsbetrokkenen voor de tweede periode vroegen om een 'leaner en meaner' lerende evaluatie – ook daar is dus behoefte aan een minder kostbaar evaluatieproces. Voor de tweede lerende evaluatieperiode zou het relevant zijn voor het begrip van de methode om navolging te geven aan deze vraag over doelmatigheid en om te bezien of een lichtere invulling van het lerende karakter afdoet aan het ontstaan van deze effecten. Ook gegeven andere aanpassingen die zijn gemaakt in de uitvoering door het PBL (o.a. een bredere kring aan partijen betrekken), is het interessant om de uitwerking daarvan te bezien op de kwaliteit en doorwerking van lerend evalueren.

Het voorliggende artikel vatte twee studies naar de ervaringen van beleidsbetrokkenen en onderzoekers aan de lerende evaluatie van het Natuurpact samen om hun perspectief op de waarde van deze methode in beeld te brengen. Het had daarbij een verkennend karakter, waarbij de eerste indicaties van de impact van de evaluatie op het natuurbeleid zichtbaar werden. Door een breedheid aan betrokkenen te interviewen is gepoogd een genuanceerd beeld te geven van wat lerend evalueren teweeg kan brengen. Huidig onderzoek laat zien dat het combineren van leren en verantwoorden voor onderzoekers geen sinecure is, en dat verschillende opvattingen over onder andere de rol van onderzoekers, en de functie en bijbehorende kwaliteitsstandaarden voor beleidsevaluatie onder

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

spanning kunnen komen te staan. Om risico's ten aanzien van geloofwaardigheid, onafhankelijkheid en objectiviteit het hoofd te bieden hebben onderzoekers specifieke vaardigheden nodig om leren en verantwoorden succesvol te navigeren. Hierbij is een leergierige en 'reflexieve' houding vereist (Kunseler & Verwoerd, 2019; Popa et al., 2015): naast hun inhoudelijke kennis hebben onderzoekers dan voelsprieten voor de wisselwerking tussen hoe kennis tot stand komt en de politiek-bestuurlijke context waarin deze wordt gebruikt, en weten zij hun rol en het evaluatieonderzoek in te richten op een manier die past bij de voorliggende opgave en deze context. Door middel van een verbreed perspectief op de functie van beleidsevaluatie kan zo opgavegericht en meer betekenisvol worden geëvalueerd (Van der Steen et al., 2018). Belangrijk daarbij is dat niet alleen bij onderzoekers zelf, maar ook bij hun kennisorganisaties en de beleidsbetrokkenen waarmee zij samenwerken dit verbrede perspectief voet aan de grond vindt en normaliseert. We hopen met dit artikel te inspireren tot toepassing van lerend evalueren in andere dossiers en contexten om zo de lerende evaluatiebenadering verder te professionaliseren.

Literatuur

Chouinard, J. A. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation, 34*(2), 237-253. https://doi.org/10.1177/1098214013478142

Creswell, J. W. (2009). *Research design: Qualitative, quantitative and mixed methods approaches* (3rd ed.). SAGE Publications.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*(2), 80-88. https://doi.org/10.1177/1558689812437186

Edelenbos, J., & Van Buuren, A. (2005). The learning evaluation: A theoretical and empirical exploration. *Evaluation Review, 29*(6), 591-612. https://doi.org/10.1177/0193841X05276126

Folkert, R., Verwoerd, L., & Verwest, F. (2018). Lerend evalueren: navigeren tussen verantwoorden en leren. *Beleidsonderzoek Online*. https://doi.org/10.5553/BO/221335502018000007001

Guijt, I. (2010). Accountability and Learning: Exploding the myth of incompatibility between accountability and learning. In J. Ubels, N.-A. Acquaye-Baddoo, & A. Folwer (Eds.), *Capacity development in practice* (pp. 277-291). Earthscan.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation, 88*, 5.

Kunseler, E., & Verwoerd, L. (2019). *Kennis met impact - Reflexief werken bij publieke kennisorganisaties*. PBL, Den Haag

Nowotny, H. (2000). Re-thinking science: From reliable knowledge to socially robust knowledge. In *Jahrbuch 2000 des Collegium Helveticum* (pp. 1-19).

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

PBL & WUR. (2017). *Lerende evaluatie van het Natuurpact: Naar nieuwe verbindingen tussen natuur, beleid en samenleving*. PBL, Den Haag

Polk, M. (2014). *Achieving the promise of transdisciplinarity: A critical exploration of the relationship between transdisciplinary research and societal problem solving*. https://doi.org/10.1007/s11625-014-0247-7

Popa, F., Guillermin, M., & Dedeurwaerdere, T. (2015). A pragmatist approach to transdisciplinarity in sustainability research: From complex systems theory to reflexive science. *Futures, 65*, 45-56. https://doi.org/10.1016/j.futures.2014.02.002

Regeer, B. J., De Wildt-Liesveld, R., Van Mierlo, B., & Bunders, J. F. G. (2016). Exploring ways to reconcile accountability and learning in the evaluation of niche experiments. *Evaluation, 22*(1), 6-28. https://doi.org/10.1177/1356389015623659

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers*. Blackwell Publishing

Scarinci, I. C., Moore, A., Benjamin, R., Vickers, S., Shikany, J., & Fouad, M. (2017). A participatory evaluation framework in the establishment and implementation of transdisciplinary collaborative centers for health disparities research. *Evaluation and Program Planning, 60*, 37-45. https://doi.org/10.1016/j.evalprogplan.2016.08.020

Shulha, L. M., Whitmore, E., Cousins, J. B., Gilbert, N., & Al Hudib, H. (2016). Introducing evidence-based principles to guide collaborative approaches to evaluation. *American Journal of Evaluation*. https://doi.org/10.1177/1098214015615230

Van der Steen, M., Faber, A., Frankowski, A., & Norbruis, F. (2018). *Opgavegericht evalueren: Beleidsevaluatie voor systeemverandering*. Den Haag: NSOB.

Van Veen, S. C., Verwoerd, L., & Regeer, B. J. (2016). *Characteristics of reflexive evaluation - a literature review conducted in the context of the Natuurpact (2014-2027) evaluation*. Amsterdam: Athena Instituut.

Verwoerd, L., De Wildt-Liesveld, R., & Regeer, B. J. (2017). *The value of reflexive evaluation in the eyes of participants - a review of the Natuurpact reflexive evaluation (2014-2017)*. Amsterdam: Athena Instituut.

Verwoerd, L., Klaassen, P., & Regeer, B. J. (2019). *The value of a reflexive evaluation approach in the eyes of researchers*. Amsterdam: Athena Instituut.

Verwoerd, L., Klaassen, P., van Veen, S. C., De Wildt-Liesveld, R., & Regeer, B. J. (2020). Combining the roles of evaluator and facilitator: Assessing societal impacts of transdisciplinary research while building capacities to improve its quality. *Environmental Science and Policy, 103*, 32-40. https://doi.org/10.1016/j.envsci.2019.10.011

Walter, A. I., Helgenberger, S., Wiek, A., & Scholz, R. W. (2007). Measuring societal effects of transdisciplinary research projects: Design and application of

Dit artikel uit Beleidsonderzoek Online is gepubliceerd door Boom bestuurskunde en is bestemd voor anonieme bezoeker

an evaluation method. *Evaluation and Program Planning, 30*(4), 325-338. https://doi.org/10.1016/j.evalprogplan.2007.08.002

Wiek, A., Talwar, S., O'Shea, M., & Robinson, J. (2014). Toward a methodological scheme for capturing societal effects of participatory sustainability research. *Research Evaluation, 23*(2), 117-132. https://doi.org/10.1093/reseval/rvt031

Noten

* De auteurs bedanken de respondenten voor hun deelname aan de studie, en interne en externe reviewers voor hun waardevolle feedback op het onderzoek.

1 Momenteel wordt lerend evalueren door het PBL ook toegepast bij de evaluatie van het Interbestuurlijk Programma Vitaal Platteland en het PBL-onderzoeksprogramma Regio Deals voor brede welvaart.

2 De vraag naar efficiëntie is deze periode mede daarom niet aan bod gekomen; daarvoor ontbrak inzicht in de kosten van verschillende provinciale beleidsstrategieën.

3 De gehele rapportage van de lerende evaluatie van het Natuurpact (2014-2017) is te downloaden via: http://themasites.pbl.nl/evaluatie-natuurpact/.

4 Zes provincies verwijzen expliciet naar de lerende evaluatie van het Natuurpact in hun natuurvisie, in drie is geen verwijzing gevonden, en drie andere natuurvisies zijn gemaakt vóórdat de lerende evaluatie van start ging.

De auteurs bedanken de respondenten voor hun deelname aan de studie, en interne en externe reviewers voor hun waardevolle feedback op het onderzoek.

© Boom bestuurskunde