

7 Evaluatie van gezamenlijke kennisproductie

Leonie van Drooge, Stefan de Jong en Barend van der Meulen

7.1 Inleiding

In analyses van kenniscoproductie ligt de nadruk op de interactieprocessen tussen verschillende actoren met eigen doelen, opvattingen en belangen. De dynamiek van deze interactieprocessen is onzeker. Aan de ene kant kunnen betrokkenen afhaken om redenen die geheel buiten het project liggen, zoals verandering van wethouder of het implementeren van een nieuwe bedrijfsstrategie. Aan de andere kant kunnen nieuwe actoren zeer geïnteresseerd raken, waardoor er zich nieuwe mogelijkheden voordoen en belangen eventueel opnieuw moeten worden afgewogen. Dergelijke processen vragen een managementstijl die gericht is op interactie, leren, reflectie en acceptatie van onzekerheid en voorlopigheid van uitkomsten.

Projecten en programma's van kenniscoproductie hebben daarnaast te maken met het management binnen de eigen organisatie van deelnemers, met praktijken die vaak gedomineerd worden door vastgestelde budgetten, SMART-doelstellingen en vooraf gedefinieerde mijlpalen. Dit veronderstelt dat projecten geordend zijn en gemonitord kunnen worden aan de hand van tussentijdse *deliverables*. Voor nieuwe mogelijkheden en het verkennen van risicovolle trajecten is dan weinig ruimte.

De spanning tussen de twee managementstijlen wordt zowel op programma-niveau gevoeld, als in de projecten door betrokkenen. Consultants moeten de onzekere dynamiek inpassen in hun tijdschrijfsystematiek. Overheidsorganisaties moeten een lerende werkstijl afstemmen op politieke discussies. Onderzoekers moeten de door lokale omstandigheden gekleurde empirie vertalen naar algemene wetenschappelijke inzichten.

In de wereld van programma- en onderzoeksevaluatie hebben de twee managementstijlen tot verschillende manieren van evalueren geleid. Ruwweg kan een onderscheid gemaakt worden tussen *vormende evaluaties* en *verantwoordings-evaluaties*. Vormende evaluaties leren management en stakeholders over het verloop van het programma of het onderzoek en laten zien waar kansen zitten en waar knelpunten. Verantwoordings-evaluaties zijn bedoeld om management en opdrachtgevers inzicht te geven in de mate waarin het geld op de juiste wijze is gebruikt en of vooraf vastgestelde doelen zijn gerealiseerd.

In de praktijk van onderzoeksmanagement botsen die twee manieren steeds vaker. Voor programma's en projecten van kenniscoproductie is het van belang

dat er manieren gevonden worden om de twee benaderingen te verbinden. Dit hoofdstuk beschrijft hoe dat kan. Paragraaf 2 geeft enkele voorbeelden van de evaluatie van projecten van kenniscoproductie. Paragraaf 3 maakt duidelijk dat dergelijke evaluaties niet passen binnen de reguliere evaluaties van wetenschappelijk onderzoek. De problemen die zich specifiek voordoen bij het beoordelen van 'maatschappelijke kwaliteit' van wetenschappelijk onderzoek worden uiteengezet in paragraaf 4. Paragraaf 5 beschrijft de ERiC-methode die juist is ontwikkeld voor het beoordelen van maatschappelijke kwaliteit. De toepassing van de ERiC-methode wordt geïllustreerd met twee projecten. Hieruit worden in paragraaf 6 lessen getrokken. Als de nadruk gelegd wordt op de *productieve interacties* in het kennisproces, is het mogelijk om vormende evaluaties en verantwoordingsevaluaties met elkaar te verbinden. De nadruk op productieve interacties doet recht aan de dynamiek van het onderzoeksproces én maakt het mogelijk na te gaan hoe onderzoeksmiddelen geïnvesteerd zijn.

7.2 Evalueren en leren in kenniscoproductie

Nederland kent een groot aantal onderzoeksprogramma's waarin de afgelopen vijftien jaar is geprobeerd om innovatieprocessen in complexe omgevingen te stimuleren. Programma's als het Nationaal Initiatief Duurzame Ontwikkeling (NIDO), Transumo, Maatschappij Geaccepteerde Veehouderij, Kennis voor Klimaat (zie hoofdstuk 2), Kennisnetwerk Systeem Innovaties (KSI), Transform (zie hoofdstuk 6) initieerden en ondersteunden innovaties in landbouw, energie, water, voedsel, ruimtelijke ordening, en gezondheidszorg.

In deze programma's is veel aandacht besteed aan de organisatie van reflexief management en sociaal leren als voorwaarden voor de voortgang van complexe innovatieprocessen. Evaluaties zijn een belangrijk onderdeel van die leerprocessen. Drie voorbeelden laten dat zien.

Barbara van Mierlo en haar collega's analyseerden interactieprocessen in twee projecten van het NIDO-programma. Het ene project had tot doel om het watermanagement in de stad duurzamer te maken, waarbij drinkwaterbedrijf, waterschap en gemeentebestuur hun strategieën op elkaar moesten afstemmen. Het andere project was gericht op het verbeteren van de marktkansen voor duurzame producten en daarbij moesten innovaties aan het begin van de keten gekoppeld worden aan bedrijfsstrategieën van de keten. Door de interactieprocessen op systeemniveau in kaart te brengen en daarmee ook aan te tonen welke cruciale interacties zwak zijn, zien de betrokkenen hoe processen verbeterd kunnen worden (Van Mierlo et al., 2010).

Evaluaties in het kader van Health Technology Assessment richten zich voornamelijk op de efficiëntie en doelmatigheid van nieuwe zorgpraktijken. In een evaluatie van een initiatief in Nijmegen om de zorg voor patiënten na een zelfmoordpoging te verbeteren, gebruikte een interdisciplinair onderzoeksteam een andere aanpak. De verschillende specialisten, de verpleegkundigen en

de patiënten die meededen in het experiment, werden ondervraagd over hun perceptie van de problemen. Daarmee werd een interactieve aanpak geïntroduceerd die, beter dan de traditionele methoden, naar boven bracht waar de knelpunten zaten bij de invoering van het initiatief (Moret-Hartman et al., 2011).

Het programma Maatschappelijk Geaccepteerde Veehouderij ontwikkelde nieuwe vormen van veehouderij die diervriendelijk moesten zijn, maar ook duurzaam en economisch rendabel. In discussies over de toekomst van de veehouderij werden deze eisen vaak gezien als onverenigbaar. In het programma ontwikkelden ingenieurs, boeren, milieuorganisaties en bedrijven systemen waarin die eisen wel te verenigen waren. Zo ontstond bijvoorbeeld het Rondeel, een nieuwe legkippenhouderij als alternatief voor de legbatterij, waarvan het ontwerp later in het Transforum-programma werd gerealiseerd. In een evaluatie van het programma Maatschappelijk Geaccepteerde Veehouderij analyseerden Yvonne Cuijpers en Barend van der Meulen (2007) een aantal initiatieven via interviews, vragenlijsten en documentanalyse. Op basis daarvan kon gereconstrueerd worden hoe de kans op succes bij dergelijke initiatieven vergroot kan worden. De resultaten werden in de vervolgfase gebruikt om de aandacht te verschuiven van individuele boeren naar ketens en om de betrokkenheid van de overheid te vergroten.

Deze drie evaluaties hebben alle een interactief en reflexief karakter. Opvallend in het laatste voorbeeld is dat de evaluatie ook is gebruikt om indicatoren te ontwikkelen waarmee het programmamanagement verantwoording kon afleggen aan de formele opdrachtgever – in dit geval het toenmalige ministerie van Landbouw, Natuur en Voedselkwaliteit. Er werd gereconstrueerd welke interacties tussen de actoren en programma-activiteiten leidden tot nieuwe innovaties. Een dergelijke combinatie van verantwoorden en leren vormt tot op heden een uitzondering maar is veelbelovend, zoals in verderop in dit hoofdstuk wordt geïllustreerd.

7.3 Evalueren van wetenschappelijk onderzoek

In veel gevallen zijn de verschillende onderzoekers in kenniscoproductieprojecten in hun eigen organisatie ook onderworpen aan evaluaties. In de afgelopen decennia zijn deze onder druk van ideeën rond New Public Management (NPM) steeds meer gestructureerd rond SMART-doelen, outputparameters en prestatie-indicatoren. De New Public Management-trend is ook zichtbaar in de Nederlandse wetenschap. Vanaf eind jaren tachtig zien we een gestage toename van onderzoeksevaluaties. Maar de vraag is in hoeverre deze praktijk aansluit op de dynamiek van kenniscoproductie.

In de evaluatie van wetenschappelijk onderzoek zijn wetenschappelijke kwaliteit en wetenschappelijke impact vaste criteria. Er zijn twee methoden die veel worden gebruikt en naast elkaar voorkomen. Dit zijn bibliometrie als methode om de onderzoeksproductie te meten en *peer review* om tot een oordeel te

komen. Beide methoden kennen echter grenzen als het gaat om de beoordeling van kennisproductie. Dit geldt zowel voor evaluaties *ex ante*, bijvoorbeeld het aanvragen voor projectfinanciering, als voor evaluaties *ex post*, zoals bijvoorbeeld in Nederland bij de zesjaarlijkse onderzoeksvisitaties.¹²

Bibliometrie

Bibliometrie is veel gebruikt om inzichtelijk te maken welke resultaten in de vorm van teksten zijn geleverd, waar is gepubliceerd, hoeveel is gepubliceerd en hoe de waardering is. In een evaluatie vormen deze gegevens een belangrijk deel van het aangedragen bewijs voor wetenschappelijke productiviteit, kwaliteit en impact. Veelgebruikte bibliometrische indicatoren zijn het aantal artikelen in *peer reviewed journals* en het aantal citaties van die artikelen. Dat lijkt eenvoudig, maar in de loop der jaren is er een heel complex van maten en gewichten ontwikkeld om artikelen en citaties van de ene groep te vergelijken met die van een andere groep.

De nadruk op wetenschappelijke artikelen rijmt slecht met de praktijk van kennisproductie. Bij kennisproductie zien we dat de kennis op zeer verschillende manieren gecommuniceerd wordt: niet alleen via wetenschappelijke artikelen, maar ook via rapporten, artikelen in vaktijdschriften, workshops, ontwerpen, modellen of softwaretools. De overleggen met andere (maatschappelijke) partijen waar zij bijvoorbeeld leren van de manier van werken van wetenschappers en hun bredere kennis van het probleemveld, worden al helemaal niet meegenomen. Dit betekent dat een groot deel van de output die voortkomt uit kennisproductie niet zichtbaar is in bibliometrische analyses. Deze analyses leveren dan ook een beperkt beeld op. Deze spanning bleek bijvoorbeeld ook al uit het voorbeeld van het wijkverbeteringsproject (hoofdstuk 4).

Peer review

Peer review wordt in evaluaties veel gebruikt om tot een oordeel te komen. Wetenschappelijke vakgenoten – *peers* – worden voor de beoordeling gevraagd, omdat alleen zij in staat worden geacht om te oordelen of een onderzoeker, een voorstel of een resultaat vernieuwend is en aan kwaliteitsnormen voldoet. De veronderstelling is dat anderen dan wetenschappelijke vakgenoten kennis, inzicht en ervaring met het specifieke onderzoeksgebied missen; het ontbreekt hun aan visie op wat innovatie, kwaliteit en excellentie voor dat gebied betekent. Opvallend is daarbij dat *peers* in toenemende mate hun oordeel moeten samenvatten in een cijfer, dat zonder veel moeite geïnterpreteerd kan worden door managers en in financieringsmodellen.

¹² Voor universiteiten en onderzoeksinstituten: KNAW, VSNU, NWO (2009) Standard Evaluation Protocol (SEP) 2009-2015 en voor hogescholen: HBO-raad (2007) Brancheprotocol Kwaliteitszorg Onderzoek (BKO).

Ook voor peer review geldt dat de bestaande aanpak slecht past bij kennis-coproductie. Door de variatie in methoden en vormen van kennis die wordt gehanteerd in de projecten, is het uitermate moeilijk om peers te vinden die met voldoende afstand en deskundigheid een oordeel kunnen en willen vellen. Dergelijk onderzoek is discipline-overstijgend en zelfs wetenschapsoverstijgend gezien de nadruk op praktijk- en ervaringskennis. In sommige evaluaties worden peers wel gevraagd ook de maatschappelijke relevantie of maatschappelijke kwaliteit te beoordelen, maar in de praktijk blijken zij dit criterium moeilijk te kunnen hanteren. Er is wel gepleit om 'maatschappelijke peers' op te nemen in commissies. De vraag blijft echter wat de kwaliteitsnorm voor kennisco-productie is. Want anders dan in disciplinair onderzoek zijn dergelijke kwaliteitsnormen voor gebiedsoverschrijdend onderzoek veel minder duidelijk.

7.4 Evaluatie van maatschappelijke kwaliteit van onderzoek

Als evaluatiecriterium staat maatschappelijke kwaliteit van onderzoek in verschillende hoedanigheden (maatschappelijke relevantie, *societal impact*, kennisbenutting, valorisatie) al enkele decennia op de agenda; niet alleen in Nederland, ook in de ons omringende landen en bij de Europese Unie. Desondanks wordt het als een moeilijk te beoordelen criterium beschouwd en in veel gevallen genegeerd in evaluatieprocedures. Er zijn verschillende moeilijkheden bij het beoordelen van maatschappelijke kwaliteit en impact.

Ten eerste is er de vraag wat maatschappelijk is. Is het een verzamelterm voor 'anders dan academisch'? In dat geval is het een uitermate diffuus begrip. Of is het concreter en betekent het 'economisch'? In dat geval lijkt het veel duidelijker gedefinieerd. Het gaat dan om kennis die nuttig is voor de industrie, en waar de industrie ook voor wil betalen. Dergelijk nut kan bijvoorbeeld beoordeeld worden aan de hand van externe inkomsten van de kennisinstelling. Hierbij valt te denken aan de betalingen die een spin-off-bedrijf van het Leids Universitair Medisch Centrum ontvangt van een geneesmiddelenfabrikant in ruil voor een exclusieve licentie (hoofdstuk 3). Maar bij een economische opvatting van maatschappelijke relevantie wordt veel onderzoek buiten beschouwing gelaten. Want 'anders dan academisch' kan ook 'democratisch' betekenen (AWT, 2007), zoals bij bijdragen van politicologen en filosofen aan publieke debatten, en daarmee aan de mondigheid van burgers; het kan 'cultureel' betekenen, zoals in studies naar oude nederzettingen in Nederland die van invloed zijn op ons zelfbeeld; het kan 'sociaal' betekenen, wanneer resultaten van onderzoek bijdragen aan wijkverbetering (hoofdstuk 4).

Ten tweede is het publiek of de doelgroep, anders dan bij wetenschappelijke kwaliteit en impact waar het de wetenschappelijke peers betreft, vaak niet gedefinieerd. Zelfs als we uitgaan van de strikte definitie van economische

impact, is niet meteen duidelijk in welke sector van de maatschappij, bij welke stakeholder, we die impact kunnen vinden.¹³

Ten derde bestaan er voor het evalueren van maatschappelijke relevantie geen standaardmethoden die tegen een lage tijdsinvestering resultaat opleveren. Als methoden schieten klassieke bibliometrie en peer review als gezegd tekort. Vaak is diepgaand kwalitatief onderzoek nodig om maatschappelijke relevantie aan te tonen.

Ten slotte is het uitermate moeilijk om een gerealiseerde maatschappelijke impact te koppelen aan een bepaald onderzoeksproject of onderzoeker.¹⁴ Er is een tijdsprobleem: het duurt vaak jaren, zo niet decennia, voor de impact van onderzoek zichtbaar wordt, lang nadat een project is afgerond en geëvalueerd. Daarnaast is er het probleem van attributie: het is moeilijk om te identificeren wat het aandeel is geweest van een project of onderzoeker in het ontstaan van die impact; er zijn immers veel verschillende factoren die hierbij een rol spelen.

Deze bezwaren hebben gevolgen voor het funderen van uitspraken over kwaliteit in een evaluatie. Er is nog geen consensus over gegevens die maatschappelijke kwaliteit en impact onderbouwen. Wat voor soort gegevens zijn relevant? In welke sectoren van de maatschappij en bij welke doelgroep zijn die te vinden? Over welke termijn moeten gegevens worden verzameld? En als we gegevens vinden, bewijzen van maatschappelijke impact, in hoeverre zijn die dan aantoonbaar gerelateerd aan het onderzoek? En hoe belangrijk was dat onderzoek voor het realiseren van die impact?

7.5 Evalueren van onderzoek *in context*

In het Nederlandse project ERiC¹⁵ – *Evaluating Research in Context* – is een methode ontwikkeld voor de beoordeling van maatschappelijke kwaliteit en impact van onderzoek. Deze methode wordt hieronder uitgebreider besproken, omdat de aanpak een brug slaat tussen vormende evaluaties en verantwoordingsevaluaties. De kern van de methode zoals hieronder beschreven is van toepassing op ex post-evaluatie van onderzoeksgroepen (KNAW VSNU NWO, 2009) en bestaat uit vier stappen (ERiC, 2010). Twee studies naar de maatschappelijke impact van onderzoeksprogramma's laten vervolgens zien dat de aanpak ook voor andere evaluaties werkt.

¹³ Dit is anders bij de relatief nieuwe onderzoeksfunctie van de hogescholen. Hier is het onderzoek nadrukkelijk gekoppeld aan de beroepspraktijk. De nadruk ligt op het samen of in afstemming met de beroepspraktijk ontwerpen en ontwikkelen van producten, processen of diensten (<http://www.hbo-raad.nl/onderzoek?filter=2012>).

¹⁴ Dit geldt in theorie ook voor wetenschappelijke impact. In de praktijk is dat echter niet aan de orde, omdat er consensus lijkt te zijn over de oplossing. Als maat voor impact wordt veelal gebruik gemaakt van citaties over een relatief kort aantal jaren na verschijnen van een artikel.

¹⁵ ERiC (Evaluating Research in Context) was een samenwerkingsverband van HBO-raad, KNAW, NWO, VSNU en het Rathenau Instituut, dat liep van 2006-2011.

Beoordelingsmethode voor maatschappelijke kwaliteit en impact van onderzoek

De eerste stap van de genoemde methode vraagt om het formuleren of inventariseren van de missie en doelstellingen van het onderzoek. Wat wordt beoogd, hoe wordt dat bereikt en welke stakeholders of doelgroepen zijn betrokken bij het onderzoek? Met stakeholders worden de potentieel belanghebbenden van de onderzoeksresultaten bedoeld. Dat kunnen partijen in de bredere samenleving zijn, maar ook wetenschappers die niet tot het specifieke onderzoeksgebied behoren.

De tweede stap vraagt om een beschrijving van de maatschappelijke bijdrage van het onderzoek. Hierbij worden de volgende vragen gesteld:

- Welke inhoudelijke resultaten die maatschappelijk van belang zijn, heeft het onderzoek opgeleverd?
- Op welke manier is de kennis verspreid onder maatschappelijke stakeholders?
- Waaruit blijkt interesse en waardering van maatschappelijke stakeholders?
- Welke effecten hebben de onderzoeksresultaten gehad?

De antwoorden op deze vragen geven een kwalitatieve analyse van de maatschappelijke relevantie.

Voor deze tweede stap is het begrip *productieve interacties* van belang. Kennisontwikkeling is een sociaal proces, en onderzoekers hebben vele interacties. De tweede stap is erop gericht om interacties te identificeren die productief zijn, dat wil zeggen die ertoe leiden dat het denken en/of handelen van stakeholders beïnvloed wordt (Spaapen & Van Drooge, 2011).

In de derde stap wordt een selectie gemaakt uit mogelijke indicatoren voor maatschappelijke relevantie. Deze indicatoren sluiten aan op de tweede stap. Als bijvoorbeeld in de tweede stap geclaimd is dat de kennis verspreid is via professionele presentaties en publicaties, en dat de waardering blijkt uit contacten met professionele gebruikers, dan ligt het voor de hand om de publicaties en presentaties als indicatoren te gebruiken en onder de professionals te vragen via vragenlijsten of anderszins hoe zij het betreffende onderzoek waarderen. Deze stap draagt bij aan het inzicht in hoeverre gewenste prestaties ook gerealiseerd zijn. In verantwoordingsevaluaties kan dit van pas komen.

In de vierde stap wordt door de evaluatiecommissie een oordeel gevormd over de maatschappelijke doelstellingen en de bereikte resultaten. Voor een goede beoordeling hiervan zijn stakeholders nodig die inzicht hebben in de maatschappelijke context. Uit pilotstudies is gebleken dat dit het beste kan door ontmoetingen met stakeholders te arrangeren, of door stakeholders op te nemen in de evaluatiecommissie.

De ERiC-aanpak biedt goede aanknopingspunten voor de evaluatie van kenniscoproductie. Het uitgangspunt van een dergelijke evaluatie vormt de

missie en doelstelling van het onderzoek en de beoogde doelgroep (stap 1). In het geval van kennisproductie verschillen de doelstelling en doelgroep in vergelijking tot die van de meer traditionele academische projecten; hiermee wordt rekening gehouden. Voor de belangrijke rol van maatschappelijke stakeholders is ook ruimte in stap 1, als doelgroep van het onderzoek.

Maatschappelijke bijdragen en effecten in gebruikerscontexten komen naar voren in de antwoorden in stap 2, en de indicatoren en data in stap 3. Door het betrekken van maatschappelijke stakeholders bij de beoordeling in stap 4, wordt het belang van deze groep actoren als doelgroep nogmaals onderstreept. Deze aanpak lost het eerder genoemde definitieprobleem op. Maatschappelijke kwaliteit en impact worden namelijk gedefinieerd in termen van de doelstelling van het onderzoek. Ook de stakeholders van het onderzoek worden door de doelstelling en missie gedefinieerd. De implicatie is dat de definities van 'maatschappelijk' en 'stakeholder' per onderzoek zullen verschillen, want ze zijn gebonden aan de context van het onderzoek.

Evalueren van een faculteit en een programma in context

Een van de ERiC-pilotstudies is uitgevoerd bij de faculteit Bouwkunde van de TU Delft. Het onderzoek in deze faculteit varieert van constructietechnologie, bouwmanagement en ruimtelijke ordening tot de (kunsthistorische) studie van gebouwen en architecten. Het onderzoek heeft kenmerken van kennisproductie, omdat het onderzoek en de maatschappelijke praktijk nauw met elkaar verweven zijn.

Op basis van interviews met onderzoekers en stakeholders zijn de productieve interacties gereconstrueerd. De academici bleken veel en vaak samen te werken met maatschappelijke stakeholders, zoals architecten, stedenbouwkundigen en bouwbedrijven. Dit werd ook zichtbaar in de vele dubbelaanstellingen, in de onderzoeksagenda, en in de gezamenlijke projecten met een breed scala aan maatschappelijke stakeholders, van nationale en regionale overheden tot bouwondernemingen en musea. De in de projecten ontwikkelde kennis werd op veel manieren overgedragen, waarbij vakpublicaties en professionele bijeenkomsten van groot belang waren.

Op basis van deze analyse hebben we een aantal nieuwe indicatoren geformuleerd die recht doen aan de productieve interacties in het wetenschappelijke en professionele veld van Bouwkunde. Voorbeelden van deze indicatoren zijn: tentoonstellingen, lidmaatschappen van overheidsadviescommissies, jurylidmaatschappen en redacteurschappen van vaktijdschriften (De Jong, Arensbergen, et al., 2011). De faculteit Bouwkunde heeft de ERiC-aanpak gebruikt bij het schrijven van de zelfevaluatie en was zo in staat om maatschappelijke kwaliteit en relevantie te presenteren op een manier die in evenwicht was met de wetenschappelijke kwaliteit en relevantie en recht deed aan de aard van het onderzoek (Van der Hoeven, 2010).

Een vergelijkbare aanpak is ook toegepast in studies naar de maatschappelijke impact van de BSIK-onderzoeksprogramma's Klimaat voor Ruimte, Kennis voor Klimaat (Merkx et al., 2011a) en Next Generation Infrastructures (NG Infra) (De Jong, Hessels, et al., 2011). Ook hier was het van belang om te begrijpen hoe het onderzoek in de specifieke disciplinaire, politieke en maatschappelijke context impact had. Ter voorbereiding op de evaluatie van deze programma's reconstrueerden we aan de hand van interviews en documentstudies de programmalogica. Zo werd zichtbaar wat volgens de betrokkenen de productieve interacties waren voor kenniscoproductie en het realiseren van kennisimpact. Vervolgens werd via vragenlijsten de dynamiek van de projecten en hun mogelijke impact onderzocht. Daarmee kon niet alleen de mate van impact vastgesteld worden, maar konden ook uitspraken worden gedaan over wat voor soort projecten welke impact hadden.

Het interessante van deze studies is dat ze een uitweg bieden uit het in het begin van het artikel geschetste dilemma tussen vormende en verantwoordingsevaluaties. De studies brengen enerzijds de (beoogde) dynamiek van de onderzoeksprojecten en -programma's in kaart en laten zien hoe verschillende stakeholders de projecten ervaren. Anderzijds identificeren ze indicatoren die specifiek zijn voor het onderzoeksgebied, het onderzoeksinstituut of het onderzoeksprogramma.

7.6 Lessen uit de praktijk: evaluatie van kenniscoproductie

Op basis van deze ervaringen kunnen we een aantal lessen trekken die ook voor de evaluatie van projecten en programma's voor kenniscoproductie van belang zijn. Deze lessen moeten bijdragen aan het inrichten van evaluaties die een scala van betrokkenen helpt om te leren van en te reflecteren op de productieve interacties tijdens het onderzoek, en die bijdragen aan het afleggen van goede verantwoording aan financiers.

1 Gebruik extended peer review om een compleet beeld te krijgen

Betrokkenen hebben vaak verschillende visies op de waarde van samenwerking. Dit bleek in verschillende pilotstudies, waaronder de hiervoor beschreven studies bij de faculteit Bouwkunde en de BSIK-onderzoeksprogramma's. Verschillende malen gaven onderzoekers aan dat een bepaalde stakeholder een belangrijke rol had gespeeld in onderzoek of dat het onderzoek belangrijk was voor een met name genoemde stakeholder. In interviews met de stakeholder zelf waardeerde deze zijn bijdrage veel lager en was de stakeholder soms niet eens goed bekend met het onderzoek. De vraag is of er sprake is van kenniscoproductie als er zo'n groot verschil in waardering is van de interacties tijdens het onderzoek. Het laat ook zien dat er groot verschil bestaat in de waardering van onderzoek en onderzoeksprocessen door onderzoekers en stakeholders. Dit pleit ervoor om stakeholders op te nemen in een extended *peer review*-commis-

sie. Juist omdat de ervaring en perceptie van partijen kunnen verschillen, is het verstandig om alle relevante betrokkenen aan het woord te laten.

Extended peer review kan ook het antwoord zijn op een ander probleem van kenniscoproductie. Daar waar werkelijk sprake is van kenniscoproductie is het eigenaarschap diffuus. Problemen, kennis en argumenten zijn niet meer te onderscheiden naar eigenaar, discipline of gebruikscontext, maar vormen een geheel met de betrokkenen. Ook de verantwoordelijkheid om tot een oplossing te komen wordt gezamenlijk gedragen en gedeeld; er is geen duidelijk onderscheid tussen verantwoordelijkheden. De enige manier om in dit geval een zinvolle uitspraak te doen over de waarde van de bijdrage van de wetenschap of wetenschappers, is om in gesprek te gaan met de verschillende deelnemers aan het project, in de vorm van *extended peer review*.

2 Focus op proces in plaats van op product

Wanneer wordt gefocust op het proces van kenniscoproductie, in plaats van op de producten die eruit voortkomen, biedt dat twee belangrijke voordelen. Ten eerste sluit dit beter aan op de beleving van de betrokken onderzoekers. Dat is vooral van waarde bij evaluaties waarin leren het belangrijkste doel is. Het begrip productieve interacties speelt dan een centrale rol om het proces te beschrijven. Een directeur van een Engels onderzoeksinstituut waarvan in een Europese studie de maatschappelijke impact onderzocht werd, verwoordde de waarde van het begrip productieve interacties als volgt:¹⁶

Het helpt om activiteiten te benoemen, te duiden en te legitimeren die waardevol zijn voor wetenschappers en stakeholders, maar die in een traditionele evaluatie ondergewaardeerd blijven. Wetenschappers herkennen in het begrip productieve interacties hun eigen werk. Dat de interacties 'productief' worden genoemd, stimuleert ze om relaties met stakeholders aan te gaan en te onderhouden. Het legitimeert ook de strategie om stakeholders te ontmoeten, te netwerken, bijeenkomsten bij te wonen en daar te presenteren. Feit is dat dit belangrijk is om mogelijkheden voor onderzoek te creëren, maar het voelt niet altijd als 'echt' wetenschappelijk werk.

Productieve interacties vormen dus een integraal onderdeel van wetenschappelijk werk en zouden ook als zodanig moeten worden gemonitord.

¹⁶ Report of SIAMPI Workshop 10, geraadpleegd 8 december 2011.
Zie: <http://siampi.eu/Content/SIAMPI/Report%20SIAMPI%20workshop.pdf>.

3 Gebruik nieuwe indicatoren

In de Nederlandse universiteiten is stilzwijgend overeenstemming ontstaan over relevante outputindicatoren. Deze worden verzameld en bijgehouden in centrale databestanden. Nieuwe indicatoren voor productieve interacties zijn in eerste instantie vaak slecht bruikbaar, omdat de hiervoor benodigde gegevens (nog) niet worden geregistreerd. Naar gegevens over gebruik en impact van onderzoek wordt gewoonlijk niet gevraagd en onderzoekers houden deze gegevens dus niet bij. Bovendien zijn voor het monitoren en evalueren van kenniscoproductieprogramma's de gebruikelijke indicatoren onvoldoende.

Met de systematische aanpak die in dit hoofdstuk is gepresenteerd, is het voor onderzoekers relatief eenvoudig te identificeren wie gebruik zou kunnen of moeten maken van hun onderzoeksresultaten, waar de impact zich zou moeten manifesteren, welke productieve interacties hierbij een rol spelen en dus welke gegevens onderzoekers zouden moeten bijhouden (Van Drooge & Vandeberg, 2011).

4 Van ex post-evaluatie van wetenschappelijke en maatschappelijke kwaliteit naar integrale evaluatie

Onderzoeksevaluaties worden meestal achteraf uitgevoerd, ter verantwoording. Evaluatie van onderzoek kan echter ook een andere rol spelen en betekenis krijgen tijdens projecten voor kenniscoproductie. Gezamenlijke evaluatie tijdens dergelijke projecten kan een middel zijn om kennis te verankeren. Hierbij wordt evaluatie dus een managementinstrument, dat er mede toe dient projecten te verankeren in de participerende organisaties.

Samenwerking vindt plaats tussen personen, niet tussen organisaties. Een stakeholder (in dit geval: een specifieke medewerker van een betrokken maatschappelijke partij) kan vaak moeilijk inschatten wat het onderzoek van een bepaalde universiteit betekent voor zijn of haar werk. Diezelfde stakeholder blijkt vaak prima in staat om aan te geven wat het onderzoek van of de samenwerking met een bepaalde persoon heeft opgeleverd. Dit geldt ook omgekeerd zoals bleek uit de pilotstudies: als diezelfde medewerker niet meer werkzaam is voor de betrokken maatschappelijke partij, dan blijkt het vaak uitermate moeilijk een collega te vinden die het project kent en er een goed oordeel over kan geven.

Dat roept de vraag op in hoeverre de bestudeerde projecten gedragen worden door de betrokken organisaties. Bij kenniscoproductie is deze vraag uitermate relevant, aangezien de projecten en programma's het belang van individuele personen overstijgen.

Een optie om een project meer te verankeren in de betrokken organisaties is om al tijdens de uitvoering te evalueren. De rollen en verantwoordelijkheden worden in zo'n geval gearticuleerd en daarmee ook de verantwoordelijkheid van de betrokken individuen als vertegenwoordigers van hun eigen organisatie. Hierdoor kan worden voorkomen dat het onduidelijk is dat wordt samengewerkt, met wie wordt samengewerkt, waarom wordt samengewerkt en wat het resultaat is. De vormende evaluatie wordt op deze manier een middel om de voortgang en het bereiken van doelstellingen te monitoren. Daarmee krijgt evaluatie een centrale rol in het management van het proces van kenniscoproductie.