



Hogeschool
van Amsterdam

LECTORALE REDE

De Voorbeschouwing

Over impact, valorisatie
en doorwerking

Harry van Vliet

Lector Doorwerking
Praktijkgericht Onderzoek

Creating Tomorrow

De voorbeschouwing

De voorbeschouwing

Over impact, valorisatie en doorwerking

Lectorale rede

In verkorte vorm uitgesproken

op dinsdag 20 september 2022 door

Harry van Vliet

Lector Doorwerking Praktijkgericht Onderzoek



**Hogeschool
van Amsterdam**

ISBN 978-94-6301-424-3

Academische Uitgeverij Eburon, Utrecht
www.eburon.nl

Beeld omslag: Lady Kirschen / Shutterstock.com
Omslagontwerp: Textcetera, Den Haag
Figuren: Armchair Adventure

© 2022 H. van Vliet / Hogeschool van Amsterdam. Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of op enig andere manier, zonder voorafgaande schriftelijke toestemming van de rechthebbende.

Inhoud

Proloog	9
Lineland	13
Eendimensionaal	15
SLM en kritiek	19
Impact	22
Valorisatie	26
En doorwerking?	27
Flatland	29
Doorwerking	31
Een logische analyse	31
Een narratief	35
Een tool	38
Exploratief onderzoek #1: inspanning versus effect	40
Doorwerking redux	42
Intermezzo	47
Indicatoren	49
Context, context, context	50
Criteria en indicatoren	51
Exploratief onderzoek #2: criteria praktijkgericht onderzoek	52
Verscheidenheid en vergelijkbaarheid	54
Exploratief onderzoek #3: indicatoren voor onderzoek	55
Kwaliteitscriteria voor indicatoren	59
Indicatoren redux	61
Spaceland	65
Ecosysteem	67
Productieve interacties	68

Elementen van het ecosysteem	70
Exploratief onderzoek #4: het ecosysteem van krachtige SPRONG-groepen	76
Zwitsers zakmes	78
Pointland	79
5A-model	81
Analysekader	82
Attitude	87
Attributie	88
Aantoonbaar	92
Aggregatie	95
Epiloog	99
Bronnen	103

"Well, that is my fate: and it is as natural for us Flatlanders to lock up a Square for preaching the Third Dimension, as it is for you Spacelanders to lock up a Cube for preaching the Fourth. Alas, how strong a family likeness runs through blind and persecuting humanity in all Dimensions! Points, Lines, Squares, Cubes, Extra-Cubes – we are all liable to the same errors, all alike the Slaves of our respective Dimensional prejudices"

Edwin A. Abbott – *Flatland. A Romance of Many dimensions* (1884)

*Another dimension, another dimension
Another dimension, another dimension*

Beastie Boys – *Intergalactic* (1998)

Proloog

Een vingertop raakt een wateroppervlakte aan, de ontstane rimpeling vormt cirkels in het water. Een bijna poëtisch beeld van een eerste beweging die iets in gang zet en alles beroert in haar omgeving en verder weg, in ruimte en tijd. Een beweging onmiskenbaar herleidbaar tot de aanraking van die vinger, hoe voorzichtig deze het water ook streelt. Een aanraking die zich verspreidt, gevolgen heeft, ertoe doet. Het is een beeld van kracht en onontkoombaarheid, iets is teweeg gebracht. Het is een beeld dat staat in de kunsthistorische traditie van de aanraking die verbindt, die doorgeeft, zelfs het leven. Een begin, een geboorte...

Al jaren sinds dit beeld voorbijkomt bij de lectorale redes van lectoren aan de Hogeschool van Amsterdam vraag ik me af: van wie is die vinger eigenlijk? En wat doet die vinger daar? Zijn er nog meer vingers in het spel? En als iedereen met zijn vinger in het water zit te roeren, doven de rimpels elkaar dan ook niet uit? En waar is die vijver? Of is het een poel in een stiltegebied? Is het een laagje water op de plaatselijke schaatsbaan dat ligt te wachten op vorst, waarbij het laatste wat we moeten doen is er steeds met onze vinger in prikken? Komt er een afvoerbuus op uit van de plaatselijke fabriek? Wordt er afval van een drugslab in geloosd, zodat het meer een stroperige drab is dan helder kabbelend water?

Met het lectoraat 'Doorwerking Praktijkgericht Onderzoek' heb ik een plek kunnen geven aan deze vragen, want daarin blijken ze helemaal zo gek nog niet: wat zet onderzoek nu in beweging? Hoe moeilijk of makkelijk is dat? Wat is het effect van onderzoek? En wie speelt daar een rol in en waarom? Geen makkelijk te beantwoorden vragen, maar wel relevante en urgente vragen in het licht van de toenemende nadruk dat onderzoek een bijdrage aan grote maatschappelijke vraagstukken moet leveren. Een nadruk die vooral ook het onderzoek aan hogescholen aanspreekt door hun praktijkgerichtheid.

In deze voorbeschuiving zal ik enkele antwoorden verkennen die gegeven zijn op deze vraag naar impact bewerkstelligen. Hierbij zal ik vooral aan het opruimen zijn. Daarmee bedoel ik dat ik erachter probeer te komen wat nou de essentiële vraagstukken zijn, door voorbij dwalingen, details en randzaken te kijken. Dat zal niet direct leiden tot een parade van originele inzichten, maar ach originaliteit wordt nogal overschat, en ik ben niet de eerste die dat zegt... Tijdens dat opruimen dringen zich bepaalde ideeën naar voren wat die vragen, en zelfs al mogelijke antwoorden, zouden kunnen zijn. Dat wordt snel ingewikkeld en levert nog niet een compleet antwoord op. Met de titel 'De voorbeschuiving' heb ik me daarvoor ingedeekt, het spel is nog niet gespeeld.

Tijdens het schrijven zocht ik naar een eenvoudige maar bruikbare rode draad om de soms rommelige en ingewikkelde discussies het hoofd te bieden. Het boek *Flatland* van Edwin A. Abbott uit 1884 lag al wat langer te wachten om gelezen te worden en bleek, hoewel geen meesterwerk, een amusante vertelling. Een kubus die een vierkant probeert ervan te overtuigen dat er een derde dimensie is, waarop het vierkant tegenwerpt dat er dan mogelijk ook een vierde dimensie moet zijn die de kubus niet 'ziet'. De kubus ontkent dit furieus... Om recht te doen aan hoe mij dit geholpen heeft, nodig ik u uit mij te volgen naar *Lineland*, *Flatland*, *Spaceland* en *Pointland* middels dit reisverslag vanachter de schrijftafel.

Lineland

Eendimensionaal

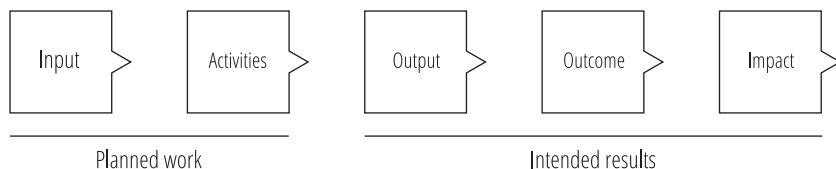
Drie specialisten wordt gevraagd een veehouder te adviseren om de tegenvallende melkproductie weer op te schroeven. De hoogleraar dierenwelzijn oppert 'Mozart in de stal', de veearts stelt voor meer lijnzaad in het voer, waarna de fysicus zijn advies begint met: 'stel de koe is bolvormig...'.¹ Waar deze grap de draak mee steekt, is de veelgebruikte methode binnen onderzoek om van een complex probleem een model te maken. Het model versimpelt noodzakelijkerwijs de werkelijkheid, maar geeft wel betere mogelijkheden exacte onderzoeksvragen te stellen en mogelijke verklarende relaties te leggen.

Het complexe vraagstuk van de vaststelling van de impact van onderzoek is ook aangepakt door modellen te maken. Een van de eerste modellen, en het meestgebruikte, is het zogeheten logische model. Een logisch model heeft tot doel een inzichtelijk stappenplan te bieden voor hoe tot bepaalde resultaten te komen. Dat stappenplan bestaat uit een beschrijving van een keten van activiteiten en de verwachte resultaten. Die keten loopt van input, oftewel middelen (mensen, geld, tijd, faciliteiten et cetera), naar activiteiten en resultaten. Die resultaten worden veelal onderverdeeld in output, outcome en impact, veelal getypeerd langs een tijdsdimensie: korte termijn, middellange termijn respectievelijk lange termijn (zie bijvoorbeeld W.K. Kellogg, 2004; Van de Sanden et al., 2018). De simpelste grafische vorm van een logisch model is een keten van input – activiteiten – output – outcome – impact (zie figuur 1). Ik zal dit voor het gemak verder aanduiden als het Standaard Logisch Model (SLM).²

In het SLM is een onderscheid te maken tussen geplande inspanningen (input en activiteiten) en verwachte resultaten (output, outcome en impact). Hieraan kunnen we ook verschillende vormen van evaluatie koppelen, zoals procesevaluatie van de activiteiten, outputevaluatie van de opgeleverde output, en

1 Voorbeeld ontleend aan Berends & Van Delft (2019).

2 Dit is een typering van het slm op een microniveau, namelijk de uitvoering van een onderzoeksproject. Er is een parallelle discussie over het lineaire/logische model op het macroniveau, daar waar het gaat om de volgorde van fundamenteel onderzoek – toegepast onderzoek (*applied research*) – ontwikkeling en productie (zie onder andere Godin, 2006; Balconi, Brusoni & Orsenigo, 2010). Deze twee niveaus zijn met elkaar verbonden, zie bijvoorbeeld Smith & Hessels (2021) die een direct verband leggen: 'The linear model adds that knowledge exchange runs only from fundamental research to society, eventually causing technological change and economic growth' (p. 324). Dat ligt mijns inziens iets subtieler, maar ik zal hier niet ingaan op de precieze verwevenheid van deze niveaus.

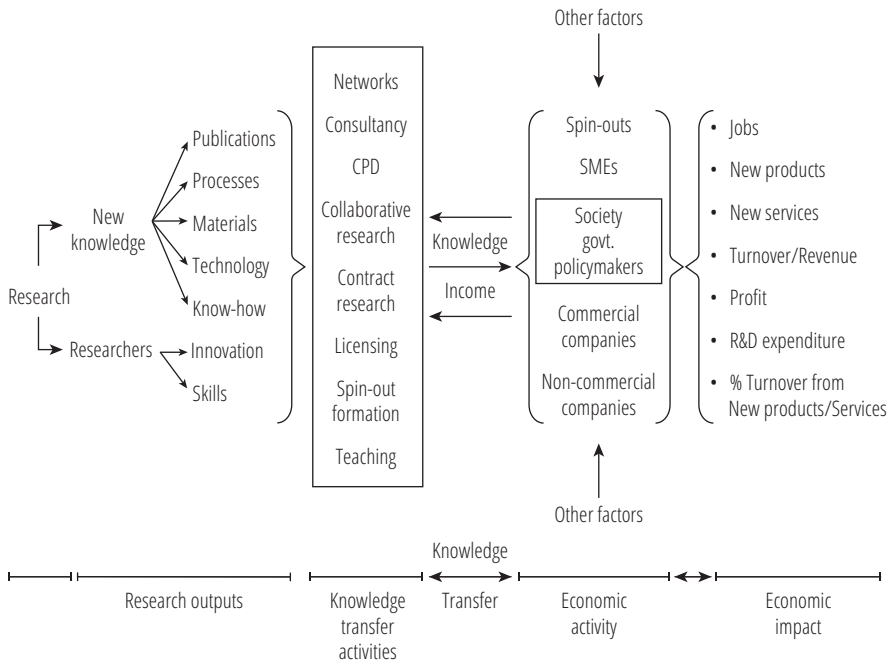


Figuur 1. Het Standaard Logisch Model (SLM)

impactevaluatie van de bereikte impact. Bij de verschillende vormen van evaluatie kunnen verschillende soorten indicatoren worden ontwikkeld, zoals tijdigheid bij procesevaluatie en kwaliteit bij outputevaluatie. Een belangrijke bijdrage van logische modellen is de aandacht die zo ontstaat voor uitkomsten en de verdere invloed van projectresultaten, in plaats van projecten en programma's alleen af te meten aan input (hoeveelheid budget) en activiteiten (verrichte inspanning).

Hoewel het SLM een model is, een abstractie, zien we het model als sjabloon terug in de beschrijving van daadwerkelijke processen van bijvoorbeeld kennistransfer. Een goed voorbeeld hiervan is de studie van Holi et al. (2008) waarin met focusgroepen van stakeholders (onderzoeksfinciers, universiteiten, bedrijven) gezocht is naar mechanismen van kennistransfer en indicatoren om dit te meten. Het resultaat is een raamwerk (figuur 2) waarin de stappen van output (research output), outcome (*economic activity*) en impact (economic impact) duidelijk zijn te herkennen. De studie leidt uit dit proces ook kwantitatieve en kwalitatieve indicatoren af, verdeeld over acht mechanismen van kennistransfer: netwerken, consultancy, leven lang ontwikkelen (CPD: *Continuing Professional Development*), gezamenlijk onderzoek, contractonderzoek, licenties, spin-outs en educatie. Een groot aantal van de indicatoren uit de studie van Holi et al., evenals het model van kennistransfer, hebben hun weg gevonden in de studie van Finne et al. (2011), waaruit de toenmalige VSNU ruimhartig heeft geput voor haar keuzemenu aan indicatoren (VSNU, 2013a/b).

Recent zien we hoe het SLM ook doorklinkt in hoe NWO haar verschillende impactbenaderingen beschrijft. De benadering Impact Outlook betreft onderzoek dat gericht is op een wetenschappelijke vraag waar geen maatschappelijke vraag aan vooraf hoeft te gaan. De benadering Impact Plan gaat over onderzoek dat gericht is op een bijdrage leveren aan maatschappelijke vraagstukken. En de benadering Impact Focus legt de focus op de toepassing van reeds ontwikkelde kennis en inzichten (NWO, 2019). Deze benaderingen zijn eenvoudig te plotten op het SLM (figuur 3). Impact Outlook is dan te typeren als het wetenschappelijk



Figuur 2. Model van kennistransfer in het innovatie-ecosysteem volgens Holi et al. (2008)

nieuwsgierheidsgedreven onderzoek zonder rekening te houden met impact en dat 'ophoudt' bij het publiceren van de output. Impact Plan is de benadering die van het begin af aan gericht is op maatschappelijke impact maar wel het startpunt heeft in onderzoek. En de benadering Impact Focus is te zien als voortzetting van het 'oude' valorisatieprogramma, gericht op het doorvertalen van kennis naar toepassing (kennisbenutting). Voor aanhangers van Donald Stokes' categorisering van onderzoek in vier kwadranten zou je hierin Bohr (*outlook*), Pasteur (*plan*) en Edison (*focus*) kunnen herkennen.³

3 Zie ook Borgdorff, van Staa & Van der Vos (2007) die praktijkgericht onderzoek aan de hand hiervan typeren door te zeggen dat het een Pasteur-/Edison-karakter heeft, en daartoe een lijn trekken van begrip – gebruik. Ik vind dit weinig toevoegen en ben het eens met de constatering van De Jong et al. (2011): 'Classifications of research into two 'modes' (Gibbons et al, 1994) or into four 'quadrants' (Stokes, 1997) seems too general for analysing the dynamics of the science system and particularly for research evaluation. The frequently discussed change of "the relations between science and society" has resulted in a large variety of types and contexts of scholarly research and appropriate evaluation approaches are needed that reflect that heterogeneity.' (p. 71).

	Input	Activities	Output	Outcome	Impact
Impact outlook benadering Onderzoek gericht op een wetenschappelijke vraag, waaraan een maatschappelijke vraag aan vooraf hoeft te gaan			→		
Impact plan benadering Onderzoek dat erop is gericht een bijdrage te leveren aan maatschappelijke vraagstukken			→		
Impact focus benadering Onderzoek met focus op de toepassing van reeds ontwikkelde kennis en inzichten					→

Figuur 3. De drie impactbenaderingen van NWO (2019) geplott op het SLM

Het SLM, als model, is een versimpelde weergave van de werkelijkheid en maakt volgens McCawley (2001) communiceerbaar wat het doel is van een project en wat de volgorde van handelingen is hoe daar te komen. De simpelste vorm volstaat voor kleinschalige projectmonitoring met een duidelijk doel en zichtbare uitkomsten (producten, diensten) – ook omdat sommige processen op deze manier werken, dat wil zeggen erg gericht zijn op het uitvoeren van activiteiten, zogenoemde *activities approach models* (bijvoorbeeld het uitrollen van een nieuwe update van Windows binnen een organisatie). Indien de problematiek ingewikkelder wordt (multidisciplinair, multicausaal, multistakeholder) krijgen logische modellen vaak een andere invulling: of we richten ons meer op de planning van de uitkomsten, zogenoemde *outcome approach models* (zoals bijvoorbeeld de politieke duurzaamheidsagenda met streefwaarden voor de komende jaren) of we richten ons op een uitgebreide (theoretische) analyse van de probleemsituatie en het onderkennen van meerdere *impact pathways*. Dit laatste is de kern geworden van methoden zoals *intervention mapping* en *Theory of Change*. Daarbij is het goed om te beseffen dat logische modellen en bijvoorbeeld een *Theory of Change* dezelfde roots hebben, namelijk in de zogeheten *programme theory* (Rogers, 2008). Ergens zijn ze elkaar echter kwijtgeraakt.

SLM en kritiek

De laatste decennia is er veelvuldig en harde kritiek geformuleerd op het SLM, onder andere door Carol Weiss zo'n veertig jaar geleden, en recent nog door onder andere Raftery et al. (2016) en Pedersen, Grønvad & Hvidtfeldt (2020). Ik noem hier drie onderling samenhangende kritiekpunten die voor het praktijkgericht onderzoek relevant zijn.

1. Het onderscheid tussen een onderzoeksfase en daarna een fase van kennisoverdracht en kennis toevoeging is niet houdbaar voor onderzoek in al zijn verscheidenheid. Er waren al effectieve vaccins voor gele koorts en polio voordat er 'wetenschappelijk' begrip was hoe ze werkten, de tweede wet van de thermodynamica is 'ontdekt' door de stoommachine efficiënter te maken – niet door fundamenteel onderzoek (Balconi, Brusoni & Orsenigo, 2008). Er zijn nog vele andere voorbeelden (Kline & Rosenberg, 1986). Je kunt zelfs twifelen of onderzoek altijd de eerste stap moet zijn voor innovaties: 'Contrary to much common wisdom, the initiating step in most innovations is not research, but rather a design' (Kline and Rosenberg, 1986, p. 302).

Praktijkgericht onderzoek werkt samen met en in de praktijk, waarbij sprake is van kennisontwikkeling en -delen tijdens de onderzoeksactiviteiten en zelfs al in de voorbereiding van het onderzoek bij de zogeheten vraagarticulatie. Interventies in en met de praktijk leveren al veranderingen op in kennis, gedrag en houding, niet alleen bij de onderzoekers maar ook bij de stakeholders. Pedersen, Grønvad & Hvidtfeldt (2020) verwoorden dit als volgt: 'The evolution of impact studies has shown that public research organizations do not just release their benefits to society following a linear model of growth and application. Instead, real-world effects of research occur at different stages in the research process, extending from knowledge dissemination and knowledge mobilization to long term applications and dynamic effects' (p. 1).

2. Het lineaire van het SLM doet geen recht aan de gecompliceerdheid van hoe onderzoek plaatsvindt. In veel evaluatieraamwerken is er de onderkenning dat onderzoek complex, interactieel en dynamisch is, waarbij vele processen simultaan verlopen (onder andere Kline & Rosenberg, 1986; Keustermans et al., 2018). De argumenten en bewijzen hiervoor zijn eindeloos. Alleen al vanuit het perspectief van de 'Social Construction of Technology' (SCOT) zijn er vele analyses voorhanden van het complexe spel hoe tot nieuwe producten en diensten wordt gekomen waarvoor kennis nodig is. Of het nou gaat om bakeliet of de fiets (Bijker, 1997). Over de fiets gesproken, Kline & Rosenberg (1986) merken

hier al over op: 'Had the idea been true that science is the initiating step in innovation, we would never have invented the bicycle' (p. 288). Maar ook concepten als *Mode 2* van Gibbons, Triple Helix en de aandacht voor innovatiesystemen benadrukken de verstrengeling van allerlei activiteiten, processen, instituten en randvoorwaarden in onderzoek doen. Met andere woorden: 'The production and use of knowledge is understood as a process of interaction and co-creation rather than as a linear process that eventually leads to an effect or 'impact' outside of research' (Sivertsen & Meijer, 2020, p. 68). Raftery et al. (2016) zijn nog uitgesprokener: 'Logic models reflect a deterministic, research-into-practice mind set that is incapable of fully capturing the messiness and non-linearity of the relationship between research, practice and policy' (p. 50).

3. Binnen logische modellen is er wel aandacht voor stakeholders en het in gezamenlijkheid formuleren van een bepaalde problematiek, maar dit wordt uitsluitend gezien als randvoorwaardelijk om tot gezamenlijk (eind)resultaat te komen. Net zoals het SLM de inherente onzekerheid in onderzoek probeert weg te gummen door een strak causaal verband te propaganderen, is er ook geen ruimte voor de rol van verschillende, vaak tegenstrijdige, belangen van stakeholders. Onder andere Arkesteijn, Mierlo & Leeuwis (2015) maken dit punt: 'the logical framework renders disagreement and conflicts invisible as the focus is on what is 'agreed' during project formulation' (p. 105). Ze zetten daar aanpakken tegenover die 'start from the recognition that people have different values and perceptions of reality and hence of the interrelations in a system. The generation of knowledge is not seen as a role of experts only, but as a collective process including different stakeholders' (p. 105). Met de toevoeging dat in dit proces 'power relationships' een essentiële rol spelen. Ook Termeer & Dewulf (2019) wijzen op dit punt in hun aanpak met *small wins*: 'a small win for one person could be a small loss for someone else' (p. 304).

Tot slot is nog op te merken dat een SLM vaak samengaat met een positivistisch kader en een summatieve opvatting van evaluatie (Greenhalgh et al., 2016; Coombs & Meijer, 2020), hoewel dit geen wetmatigheid is: zo hanteert een 'realistic' perspectief ook een (lineair) model van *context-mechanism-output-impact*.

Deze kritiekpunten op het SLM zijn al decennia bekend en stevig onderbouwd, en het model werd in de jaren negentig zelfs al dood verklaard. Ondanks allerlei nuancerings van het SLM (zie bijvoorbeeld Rogers, 2008; Balconi, Brusoni & Orsenigo, 2010) blijft het standaardmodel hardnekkig zijn kop opsteken en is het in essentie nog steeds de ruggengraat van menig evaluatiemethodiek (bijvoorbeeld het *Payback Framework*), evaluatiekader (NWO impactkader, BKO)

en beleidskader (bijvoorbeeld Van de Sanden, 2018; Vankan et al., 2020), zelfs als er eerst een opmerking is dat het zo (lineaire model) niet werkt en het achterhaald is.⁴ Je zou kunnen zeggen dat we ondertussen het 'domheidspunt' wel meer dan bereikt hebben.⁵

Dat hardnekkige heeft te maken met het feit dat het SLM het eerste model was van de relatie tussen wetenschap en de samenleving c.q. economie (Godin, 2006), en dat frame is bijzonder persistent gebleken (Schot & Steinmueller, 2016), hoewel andere frames zich hebben aangediend (zie 'Pointland'). Die persistentie ziet Godin vooral gelegen in *statistics*, het model is onderdeel geworden in de standaardisering van het meten van en methodisch rapporteren over bestedingen aan wetenschap en de bijdrage van wetenschap aan economische groei, zoals in de Frascati-handleiding van de OECD in 1963. Daardoor werd het volgens Godin een *rhetorical entity* en een *social fact*.

Het SLM is inderdaad ingedaald in vele evaluatieprotocollen, subsidieregelingen en ook de onderzoekspraktijk. Een onderzoekspraktijk van subsidie verwerven, activiteiten uitvoeren en disseminatieactiviteiten als afronding, en door naar het volgende project. De simpelheid van het model heeft een grote aantrekkingskracht. Rogers (2008) stelt dan ook dat een SLM een zekere *reassurance* geeft van houvast in een complexe omgeving 'even when this is not appropriate' (p. 45). Dit laatste is een serieus probleem: om oneigenlijke redenen wordt een model gebruikt. Het SLM is daarmee een molensteen om de nek geworden, een ingesleten werkwijze die geen recht doet aan de werkelijkheid in al haar verscheidenheid. Het heeft geleid tot tunnelvisie en tot een comfortzone die het lastig maakt om het model te verlaten. De pogingen om het SLM te 'verrijken' met feedbackloops, iteraties et cetera, zijn lapmiddelen op een niet-passend fundament, en leiden tot een lelijke lappendeken, zeker als meerdere concepten uit verschillende perspectieven en met verschillende uitgangspunten worden gebruikt als lijm (Van Vliet, 2021). Het missiegedreven innovatiebeleid legt de beperking van het SLM nog verder bloot door

4 Een typisch voorbeeld hiervan is het rapport *Maatschappelijke impact in kaart* uit 2018, waarin nadrukkelijk wordt gesteld: 'Het wetenschappelijke onderzoeksproces verloopt al lang niet meer lineair van onderzoek naar toepassing' (p. 5) en 'Dit zogenaamde 'pijplijnmodel' [fundamenteel – toegepast – uiteindelijke toepassing] is echter achterhaald' (p. 12). Om vervolgens de relatie tussen wetenschap en maatschappelijke impact te beschrijven aan de hand van de 'keten': output – outcome- impact.

5 'Domheid betreft niet zozeer de dwaling, maar veeleer het volharden in de dwaling, vaak tegen beter weten in. Zelfs denkpatronen die de mens in eerste instantie verder helpen, zullen zich op den duur tegen ons keren, een omslag die plaatsvindt op het domheidspunt' (Van Boxsel, 2016, p. 82). En ook: 'One must simplify the world to discover something new about it. The problem comes when, long after the discovery has been made, people continue to simplify' (Graeber & Wengrow, 2021, p. 21).

complexe vraagstukken centraal te stellen. Het SLM is hier te beperkt voor. Het praktijkgericht onderzoek met zijn praktijkverbondenheid verdient een beter model. Kortom, de invloed van onderzoek is niet te vangen in een lijn.

Voordat we afscheid nemen van het SLM is het simpele karakter ervan wel bruikbaar om een aantal begrippen te omschrijven en te positioneren. Begrippen die we tegen zullen blijven komen en die deels ook geworteld zijn in het SLM. En begrippen zijn er genoeg bij de evaluatie van onderzoek. In het Engels zijn dit termen zoals *utilization*, *payback* en *diffusion*, in het Nederlandse taalgebruik zijn dit begrippen zoals 'impact', 'valorisatie' en sinds enkele jaren ook 'doorwerking'. Ik zal me hier beperken tot een korte bespreking van de begrippen 'impact' en 'valorisatie', en neem een voorzet op de discussie rond 'doorwerking' dat in 'Flatland' uitgebreider aan bod komt.

Impact

In het SLM is impact de laatste stap in het (lineaire) proces. De typering van deze laatste stap is echter niet altijd eenduidig. In de literatuur zijn minimaal vier soorten typering van impact aan te treffen (zie tabel 1). Deze typering baseren zich op: óf een tijdsdimensie (korte termijn, middellange termijn, lange termijn); óf op bereik/schaal (projectdeelnemers, netwerk van de deelnemende partijen, iedereen of in ieder geval iedereen binnen het toepassingsdomein); óf op het object van verandering (individu, organisatie, systeem), óf op het soort 'product' dat wordt opgeleverd (projectdeliverables, kennisname van deelnemers, gedragsverandering bij deelnemers). Omdat de typering elkaar niet uitsluiten, komen ook allerlei combinaties voor zoals: impact als systeemveranderingen (objecttypering) door gedragsverandering (producttypering) in een tijdspanne van 5 à 10 jaar (tijdstypering). Dat lijkt logisch, veranderingen vragen nu eenmaal tijd is dan het argument, maar dit is niet vanzelfsprekend. De COVID-19 pandemie heeft duidelijk gemaakt dat je bijvoorbeeld gedragsveranderingen ook relatief snel kunt 'doorvoeren'.

Dat de verschillende typering soms onbewust door elkaar worden gebruikt is goed te zien in het door NWO gepresenteerde impactmodel. Impact staat in dit model aanvankelijk voor veranderingen die 'pas gerealiseerd [worden] vele jaren nádat een onderzoeksproject is afgesloten' (NWO, 2019, p. 2). Dit is een typering op basis van een tijdsaspect. Op een ander moment wordt gesteld: 'Onder de uiteindelijk beoogde maatschappelijke impact verstaan we veranderingen in de maatschappij' (p. 3) Dit is een typering op basis van het

object (maatschappij/systeem). Ook wordt er nog een relatie gelegd tussen het verbinden van het onderzoek met veel kennisgebruikers voor impact. Dat is een typering op basis van bereik/schaal. Het blijft onduidelijk of het noemen van deze verschillende types van impact een bewuste keuze is of een slordigheid. Een argument voor het laatste is dat het standpunt uiteindelijk is dat onderzoek moet rapporteren over output en outcomes, omdat de impact vaak (ver) na het onderzoeksprogramma zal plaatsvinden (tijdstypering), zie ook Van de Sanden et al. (2018).

Tabel 1. Verschillende typering van het impactbegrip

	Output	Outcome	Impact
Tijd	Korte termijn	Middenlange termijn	Lange termijn
Bereik	Projectdeelnemers	Netwerk	Iedereen in domein
Object	Persoon	Organisatie	Systeem
Product	Deliverables	Kennistoename	Gedragsverandering

Niet alleen wordt het begrip impact op verschillende manier gebruikt, het wordt ook steeds verder opgerekt. Sivertsen & Meijer (2020) stellen een onderscheid voor tussen *normal* versus *extraordinary* maatschappelijke impact. Normal impact wordt gedefinieerd ‘as the results of active, productive, and responsible interactions between (units of) research organizations and other organizations according to their purposes and aims in society. Within the research organizations, such interactions will often occur informally at the individual researcher or research group level, but they may also follow formalized agreements or well established traditions for collaboration’ (p. 67). Hier klinkt het concept van productieve interacties in door. Productieve interacties gaan over contacten tussen onderzoekers en andere stakeholders, waarin kennis wordt gegenereerd en gebruikt (zie ‘Spaceland’). Sivertsen & Meijer sluiten hierbij aan maar beogen een bredere typering van de interacties: die hoeven niet uitsluitend productief te zijn maar kunnen ook *active* of *responsible* zijn. Dit is in lijn met kritiek van Van Vliet (2021) op de typering ‘productief’ in het concept productieve interacties, waarbij hij voorstelt om over ‘waardevolle’ interacties te spreken. Daarbij gaat hij ook nog verder dan Sivertsen & Meijer door erop te wijzen dat dergelijke interacties niet alleen voorbehouden zijn aan onderzoekers onderling en in relatie tot stakeholders, maar dat die ook gelden voor stakeholders onderling.

Tegenover deze normal impact plaatsen ze extraordinary impact: 'more rare incidences where traditional and typical or new and untypical interactions between science and society have unexpected widespread positive or negative implications for society' (p. 67). Het zijn de aansprekende voorbeelden van (individuele) wetenschappers die zorgen voor onverwachtse doorbraken. Die voorbeelden worden vaak aangehaald voor hoe wetenschap bijdraagt aan de maatschappij. Dit overschaduwet het dagelijkse handwerk in onderzoeken dat op kleinere schaal veel effect sorteert. Hierin klinkt een echo door van de vaak gemaakte observatie dat incrementele veranderingen een belangrijker motor zijn van technologische vooruitgang dan revolutionaire innovaties (Balconi, Brusoni & Orsenigo, 2006). En ook het onderscheid dat Kline & Rosenberg (1986) maken tussen *highly visible* innovaties en de *almost invisible sort*: 'A large part of the technological innovation that is carried out in industrial societies takes the form of very small changes, such as minor modifications in the design of a machine' (p. 282).

Een voorbeeld dat aangehaald wordt in de studie van Sivertsen & Meijer (2020) is de werkwijze van de Hartstichting. Om interacties tussen onderzoek en potentiële gebruikers van de opbrengsten van het onderzoek te bevorderen, hanteert de Hartstichting twee strategieën: 1) gezamenlijke agendabepaling: zo hebben 11.000 burgers geholpen bij de prioritering van de onderzoeksagenda, resulterend in vijf hoofdonderwerpen; 2) evaluatie van onderzoek: een eindgebruikerscommissie evalueert de onderzoeksvoorstellen op criteria zoals relevantie, participatie, activiteiten en interacties met gebruikers, parallel aan een wetenschappelijke adviesraad. Deze interacties zorgen voor een gezamenlijke verantwoordelijkheid voor de beoogde impact.

De consequentie van normal impact is de focus op alledaagse interacties tussen onderzoek en praktijk, het gesprek over wat (gedeelde) belangen, verantwoordelijkheden en doelen zijn en hoe deze goed onderling af te stemmen om te komen tot veranderingen. Zo opgeschreven lijkt 'normal' impact verdacht veel op doorwerking (zie 'Flatland'), en de werkwijze in het voorbeeld van de Hartstichting is in veel praktijkgericht onderzoek terug te vinden.

Een ander voorbeeld is de aanpak van small wins van Katrien Termeer (2019; Termeer & DeWulf, 2019). Deze aanpak is niet zozeer ontwikkeld vanuit een impactdiscussie, maar vanuit de discussie rond (maatschappelijke) transitie en hoe hierin vorderingen te maken. Small wins zijn kleine diepgaande veranderingen met tastbare resultaten voor direct betrokkenen, in de vorm van een beleidsinstrument, een technologie, een verdienmodel of een ketensamenwerking. Het gaat om (radicale) veranderingen die barrières proberen te slechten, en die ook schuren. Het gaat niet om laaghangend fruit van makkelijk te scoren resultaten binnen bestaande kaders. Small wins worden gepresenteerd als

krachtige bouwstenen voor een transformatie: 'Door de focus op het kleine voorkomt een small win dat mensen overweldigd raken door de complexiteit van een vraagstuk waardoor ze minder vrij en precies kunnen denken zich laten verleiden tot abstracties. Bovendien helpt een focus op het kleine om uitstel te voorkomen. Omdat het toch maar klein is hoeven partijen niet te wachten op alle informatie en kunnen ze snel beginnen en kijken wat het oplevert. Het bevordert daarmee dat mensen niet te lang blijven steken in praten over maar overgaan tot handelen. Kleine stapjes zijn ook sneller uit te voeren omdat ze minder weerstand oproepen en er minder competitie voor het claimen van successen is. Small wins vormen daarmee de cruciale zaadjes voor maatschappelijke transitie.' (2019, p. 5).

Om deze small wins verder omhoog te stuwen, zijn er verschillende aanjaagmechanismen zoals: zichtbaar maken van het resultaat zodat mensen er energie van krijgen en geïnspireerd worden (*energizing*) en het ook meer mensen en middelen aantrekt (*logic of attraction*); verbinding zoeken van de small wins met daar waar het andere elementen in het systeem raakt (*verbinden*); en ook verbinden met andere smalls wins zodat er massa ontstaat en de voordelen al geïnternaliseerd worden (robuustheid). Dit soort mechanismen kunnen helpen om te komen tot een verdere verspreiding van de small win (opschalen, uitrollen), het verbreden van de small win door deze toe te passen op een ander terrein of te verbinden met andere vraagstukken, en een verdieping door een verdere radicalisering van de small win. Zeker met deze verspreidingsmechanismen van opschalen, verbreden en verdiepen, klinkt hierin ook de aanpak door van transitie-experimenten (zie Van den Bosch & Rotmans, 2008).

Het is onmiskenbaar dat small wins effect sorteren, impact maken, maar zich niet zo zeer kenmerken door een lange termijn en een extraordinary karakter. Het zijn meer kleinschalige interventies die hanteerbaar en concreet zijn met aantoonbare resultaten, maar tegelijkertijd wel ingrijpend zijn en verreikende consequenties kunnen hebben, zeg maar extraordinary zijn in hun *smallness*.

Deze andere perspectieven op impact kun je verbinden aan specifieke methoden. Sivertsen & Meijer (2020) noemen zelf al methodieken zoals SIAMPI, ASIRPA en *contribution mapping*, waarin de aandacht voor allerlei interacties in het onderzoeksproces centraal staan als belangrijk mechanisme voor impact. De aanpak van small wins lijkt in te passen in de methode *Theory of Change*. In die methode staat een (logische) reeks van interventies centraal die concrete resultaten opleveren (*outcomes*) waarop gereflecteerd dient te worden in het licht van de eigen aannames over de beoogde verandering en de vervolgstappen. Deze 'tussentijdse' *outcomes* zou je kunnen zien als small wins waarop je verder kunt voortbouwen.

Valorisatie

De term valorisatie is geworteld in het wetenschaps- en innovatiebeleid van de rijksoverheid sinds het begin van deze eeuw. Valorisatie legt de nadruk op de benutting van ontwikkelde kennis voor welzijn en welvaart. Een veel gehanteerde definitie van valorisatie is de volgende: 'Kennisvalorisatie is het proces van waardecreatie uit kennis, door kennis geschikt en/of beschikbaar te maken voor economische en/of maatschappelijke benutting en te vertalen in concurrerende producten, diensten, processen en nieuwe bedrijvigheid' (Nederland Ondernemend Innovatieland, 2009, p. 8). Op zich een definitie waarmee ook het onderzoek aan de hogescholen uit de voeten kan, zoals gebeurd is door het ontwikkelen van hbo valorisatie-indicatoren (Van der Zwan & Van der Vos, 2015), die uiteindelijk terug te vinden zijn in het *Brancheprotocol Kwaliteitszorg Onderzoek 2016–2022* (Vereniging Hogescholen, 2015). Valorisatie als 'overdracht van kennis' is overigens de derde taakstelling van hogescholen, en universiteiten, naast onderwijs en onderzoek (Rosenthal et al., 2019).

In de definitie en zeker in de manier waarop valorisatie(beleid) verder is vormgegeven zijn minstens drie aspecten aan te wijzen die het gebruik van de term in de context van praktijkgericht onderzoek ongemakkelijk maken.

1. Het valorisatiebeleid sprak in eerste instantie nadrukkelijk het academische onderzoek aan, niet in de laatste plaats omdat het onderzoek aan hogescholen nog in ontwikkeling was. Dit is duidelijk zichtbaar in het ontwikkelde beleid. Zo was de Subsidieregeling Kennis Exploitatie (SKE) bedoeld om (techno)starters wetenschappelijke kennis sneller te laten benutten. De oprichting van zogenoemde Technologische Topinstituten (TTI's) waren bedoeld als brug tussen wetenschappelijke kennis van universiteiten en het (grote) bedrijfsleven in het ontwikkelen van nieuwe toepassingen. Denk aan het Telematica Instituut in Enschede waar begin deze eeuw al werd gewerkt aan bijvoorbeeld mobiel betalen, video-over-IP en elektronische programmagidsen. Latere beleidsinitiatieven waarbij mbo en hbo betrokken waren, gingen vooral over het verhogen van ondernemerschap, inzet van studenten, disseminatie van kennis et cetera, zoals bijvoorbeeld in de *Centres of Entrepreneurship*.

2. Het valorisatiebeleid speelde zich ook af in de context van de zogeheten kennisparadox, de constatering dat het Nederlandse (academische) onderzoek goed is in het ontwikkelen van nieuwe kennis, maar dat die kennis onvoldoende zijn weg vindt naar toepassing in de samenleving. Het is het beeld van kennis die

op de plank ligt en niet gebruikt wordt. Kennis belandt zo in de *valley of death*. Hierin zit een duidelijke (causale) tweedeling: eerst is er kennis en vervolgens is er het probleem hoe deze te benutten. Dit zit ook in de gegeven definitie van valorisatie: kennisvalorisatie als een proces dat vanuit kennis waarde creëert. Door zo nadrukkelijk een koppelpunt te definiëren van eerst kennis ontwikkelen en dan benutten, wordt in feite al de mogelijkheid van een kloof gecreëerd. Termen als 'vertalen', 'overbruggen', 'schakels' en 'transfer' veronderstellen deze (potentiële) kloof.

Dat ook de Vereniging Hogescholen niet ongevoelig was voor dit taalspel blijkt uit de volgende opmerking in de strategische agenda *Wendbaar & Weerbaar* uit 2015 'Hogescholen zijn belangrijke schakels in het doorbreken van de kennisparadox' (2015a, p. 13). In enkele recente SPRONG-aanvragen van hogescholen klinkt dit nog door in beschrijvingen van het onderzoek aan de hogeschool als een 'brug', 'toepasbaar maken', 'kloof overbruggen'. Dit gaat vaak samen met een typering van het onderzoek aan de hogeschool als 'toegepast onderzoek' en niet als 'praktijkgericht onderzoek' (van Vliet, 2022). De notitie *De HvA van waarde* (Smeulders, 2019) typeert het onderzoek aan de hogeschool als een circulair valorisatieproces, waarmee het zich onderscheidt van het lineaire valorisatieproces van de universiteit. De discussie over doorwerking als alternatieve typering van het onderzoek aan hogescholen wordt nergens in de notitie genoemd.

3. Hoewel er in de definitie van valorisatie 'maatschappelijke benutting' wordt genoemd, is de benutting van de kennis, de waardecreatie, vooral een economische benutting geweest. Valorisatiebeleid, zeker internationaal, kent een stevige nadruk op spin-offs, starters, ondernemerschap. Indicatoren van valorisatie zijn bijvoorbeeld octrooien en patenten (Van Drooge et al., 2011). Valorisatie is gericht op het verbeteren van de kennistransitie 'naar het private systeem, het bedrijfsleven' (Wakkee, Van der Sijde & Nuijens, 2013, p. 3). Kennis wordt hierbij gezien als een middel, net als arbeid en kapitaal, om tot productiviteit en *return on investment* te komen (Joly et al., 2015). Dit aspect van economische benutting, samen met dat van een lineair proces, komt naadloos samen in het toentertijd (kabinet-Rutte 1, 2010–2012) door de minister van Economische Zaken, Maxime Verhagen, gehanteerde adagium 'van kennis naar kunde naar kassa'.

En doorwerking?

De term doorwerking is niet nieuw in de context van praktijkgericht onderzoek. Zo treffen we de term al aan in het allereerste beoordelingskader van

de RAAK-regeling van Regieorgaan SIA waar wordt gesproken over 'duurzame doorwerking' als een van de vijf pijlers (Van Vliet & Slotman, 2006).

In het adviesrapport van de commissie-Pijlman (Pijlman et al., 2017) over de kwaliteit van het praktijkgericht onderzoek gebruikt deze commissie de term doorwerking voor het eerst structureel, en wel voor het effect van het praktijkgericht onderzoek op de beroepspraktijk, samenleving, onderwijs en wetenschap. Het gebruik van de term doorwerking wordt ook nadrukkelijk gedaan om recht te doen aan de eigen taal en eigen benadering van het onderzoek aan hogescholen. In een vervolgrapport van de commissie-Franken (Franken et al., 2018) definieert deze commissie, onder verwijzing naar de commissie-Pijlman, doorwerking als 'de invloed van zowel het proces van onderzoek als van de onderzoeksresultaten op het onderwijs, de praktijk en de samenleving'; waarbij de doorwerking naar het onderzoeksveld c.q. de wetenschap hier lijkt te zijn vergeten. Doorwerking wordt getypeerd als 'breder' dan het valorisatiebegrip. Dit 'bredere' betreft de onderkenning dat gedurende het onderzoeksproces al kennisontwikkeling en -uitwisseling plaatsvindt tussen deelnemers, en dat er actief interventies plaatsvinden om het object van studie te veranderen c.q. te verbeteren.

Opvallend in deze rapporten is dat bij de commissie-Pijlman doorwerking nog een laatste stap is in een keten van activiteiten (2017, p. 16). De commissie-Franken 'citeert' het figuur uit het rapport van Pijlman, met een belangrijke wijziging: de doorwerking is nu een pijl geworden die alle voorafgaande onderzoeksfases omvat (2018, p. 10). Dit is een logische consequentie van de gegeven definitie in het rapport van Franken, maar de wijziging in het figuur wordt nergens expliciet benoemd. De pijl blijft nog wel één kant op wijzen in een keten van activiteiten. Beide rapporten bevatten geen analyse van de relatie tussen doorwerking en impact.

Beide commissies verwoorden een behoefte het praktijkgericht onderzoek een eigen profiel te geven in waar en hoe het effect van het onderzoek aan te tonen. Enerzijds door een duidelijke gerichtheid te geven aan het praktijkgericht onderzoek (niet alleen onderzoek maar ook onderwijs en beroepspraktijk/samenleving); en anderzijds door ook aandacht te geven aan de effecten die ontstaan tijdens het onderzoeksproces, dus al tijdens de fase van 'activiteiten' in het SLM. Desalniettemin blijven deze typeringen nog erg verankerd in de context van het SLM en in bijbehorende begrippen als 'valorisatie': doorwerking als 'breder' dan valorisatie. In het onderdeel 'Flatland' presenteer ik een analyse van doorwerking die niet a priori uitgaat van een model zoals het SLM.

Flatland

Doorwerking

Doorwerking is in relatie tot het standaard logische model geformuleerd als: staand voor de effecten die óók al plaatsvinden tijdens het onderzoeksproces ('ook tijdens, niet alleen na'). Hier kunnen we de effecten van de voorbereiding op het onderzoek nog aan toevoegen. De vaak gebruikte term 'vraagarticulatie' verwijst hiernaar. De formulering van de onderzoeksvraag, verwachte oplossingen en gewenste opbrengsten samen met stakeholders kunnen zeker beschouwd worden als een proces van kennisuitwisseling en kennisontwikkeling. Doorwerking beslaat hiermee de gehele duur van het proces van het onderzoek: voorbereiding, uitvoering en disseminatie, en het effect van de onderzoeksopbrengsten na afsluiting van het onderzoeksproject. Het beslaat de gehele 'lengte' van de lijn die SLM trekt.

Daarnaast is de gerichtheid van doorwerking niet globaal ingevuld ('impact op samenleving') maar drievoudig geformuleerd. Het gaat niet alleen om de effecten op de samenleving, maar ook om die op het onderzoeksveld én het onderwijs. Bovendien krijgt 'de' samenleving ook meer gerichtheid door te spreken over het specifieke effect op de beroepspraktijk. Zo ontstaan als groepen waar doorwerking op gericht is: onderzoek – onderwijs – beroepspraktijk, de befaamde driehoek (Ministerie van OCW, Vereniging Hogescholen & Regieorgaan SIA, 2019)

Naast de 'lengte' en de 'gerichtheid' van doorwerking, ontbreekt echter nog een verdieping van het concept: wat is nu precies de *aard* van doorwerking? Ik zal dit op drie manieren uitwerken, middels een logische (declaratieve) analyse, een narratief en een aanzet tot een tool.

Een logische analyse

Uitgangspunt voor de logische analyse is dat we het hebben over doorwerking in de context van onderzoeksactiviteiten waarbij er een bepaald doel wordt nagestreefd. Of dit nu inzicht verkrijgen is, een probleem oplossen of het opleveren van een specifieke dienst of product.

1. Doorwerking is de bewuste inspanning om tot een beïnvloeding te komen van iets, zodanig dat het iets in een andere toestand geraakt, verandert.
 - 1.1 De bewuste inspanning is een formuleerbare intentie.
 - 1.1.1 Deze intentie heeft een actor.
 - 1.1.2 Deze intentie kent een zekere gerichtheid.
 - 1.1.3 Die gerichtheid heeft een beoogd begunstigde (het 'iets').
 - 1.2 Actor en begunstigde staan door de intentie in een relationele verhouding tot elkaar.
 - 1.2.1 Deze verhouding is niet per definitie gelijkwaardig of wederkerig en hoeft dat ook niet te zijn.
 - 1.2.2 Deze verhouding laat toe dat er sprake kan zijn van 'rolwisseling' waarbij actor begunstigde wordt en omgekeerd.
 - 1.3 Er is een rationale (het 'waarom') voor de formuleerbare intentie.
 - 1.3.1 Die rationale is de ruimte waarin de relationele verhouding betekenis krijgt.
 - 1.3.2 Die rationale vereist een mate van gedeelde achtergrondkennis tussen actor en begunstigde om de relationele verhouding betekenis te *kunnen* geven.

2. De inspanning kent een bepaalde organisatiegraad.
 - 2.1 Een organisatiegraad is de mate waarin de intentie planmatig wordt aangepakt en tot uitvoering kan worden gebracht.
 - 2.2 Actor en begunstigde zijn aanwijsbare entiteiten of instanties binnen de organisatiegraad.
 - 2.2.1 Een instantie is een specifiek voorkomen van een actor of begunstigde binnen de relationele verhouding.
 - 2.2.2 Een entiteit is een (formele) vertegenwoordiging van een verzameling van gelijksoortige instanties.
 - 2.2.3 Afhankelijk van de organisatiegraad kunnen er ook meer verzamelingen van instanties zijn waaraan een formuleerbare intentie is toe te schrijven; en die verzamelingen zijn als zodanig als actor te zien (granulariteitsprobleem).

- 2.3 Actor en begunstigde kunnen afspraken maken over de organisatiegraad van de inspanningen (verantwoordelijkheden, rollen, monitoring, et cetera).

- 3. De inspanning kan leiden tot opbrengsten.
 - 3.1 Opbrengsten hebben dragers.
 - 3.1.1 Een drager is een instantie die kennis, inzichten, ervaringen en dergelijke representeert.
 - 3.1.2 Een drager kan een instantie zijn die los van de actor is over te dragen, zoals een product of publicatie.
 - 3.1.3 De actor kan ook zelf de drager zijn, waarbij kennis, inzichten en ervaringen in direct contact met de begunstigde wordt gearticuleerd, zoals bij een gesprek, training of workshop.

- 4. De inspanning is aantoonbaar te maken.
 - 4.1 Er zijn minimaal drie typeringen die de inspanning kunnen aantonen.
 - 4.1.1 Een typering van de inspanning is de **hoeveelheid** inspanning die verricht is.
 - 4.1.2 Een typering van de inspanning is de **aard** van de inspanning die verricht is.
 - 4.1.3 Een typering van de inspanning is de **kwaliteit** van de inspanning die verricht is.
 - 4.2 Een typering van de inspanning krijgt concrete invulling in relatie tot de intentie van de inspanning (1.2) en haar rationale (1.3).
 - 4.3 Een typering van de inspanning staat in directe relatie tot de organisatiegraad van de inspanning (2).

- 5. De begunstigde heeft een bepaalde mate van ontvankelijkheid.
 - 5.1 Ontvankelijkheid is de mogelijkheid, gelegenheid en bereidheid zich te laten beïnvloeden.
 - 5.1.1 Bij een totaal gebrek aan ontvankelijkheid verliest de begunstigde haar rol.

- 5.2 Ontvankelijkheid laat zich op minimaal vier manieren typeren
- 5.2.1 Ontvankelijkheid is te **activeren**.
- 5.2.1.1 Activering kan plaatsvinden door dragers (4.1) zoals een actor, of door afspraken (2.3). Dit laatste betreft dan in wezen een intentie tot ontvankelijkheid van de begunstigde.
- 5.2.2 Ontvankelijkheid is gerelateerd aan de mate van (veronderstelde) **relevantie** van de beïnvloeding c.q. de opbrengsten.
- 5.2.3 Ontvankelijkheid is gerelateerd aan de mate van (veronderstelde) **bruikbaarheid** van de beïnvloeding c.q. de opbrengsten.
- 5.2.4 Ontvankelijkheid is gerelateerd aan de mate van (veronderstelde) **verandering** die plaats gaat vinden door de beïnvloeding c.q. met de opbrengsten.
- 5.3 Een typering van de ontvankelijkheid krijgt concrete invulling in relatie tot de intentie van de inspanning (1.2) en haar rationale (1.3). En spiegelt daarmee de typering van de inspanning (4.2), wat terug te voeren is op de rationale en gedeelde achtergrondkennis (1.3).

- 6. De inspanning is eindig, de doorwerking niet.
- 6.1 De bewuste inspanning (tot onderzoek) en de organisatiegraad (bijvoorbeeld een projectplan) kunnen veranderen of stoppen, echter de dragers kunnen los daarvan blijven bestaan.
- 6.1.1 Dragers hebben een zekere autonomie.
- 6.1.2 Door de autonomie van de drager kunnen ook niet-beoogd begunstigen zich ertoe verhouden.
- 6.1.3 Door de autonomie van de drager kan beïnvloeding ook in niet geïntendeerde beïnvloeding resulteren.
- 6.2 Door de oneindigheid van de doorwerking ontvouwt zich een tijdsdimensie voor de doorwerking.
- 6.2.1 Dit is goed voorstelbaar voor dragers zoals producten en publicaties die blijven bestaan, en minder van toepassing op de eindigheid van de actor als drager, hoewel gedacht kan worden aan een drager die 'school' maakt en middels opeenvolgende actoren voortleeft.

- 6.3 Vraagstukken die vele verschillende ‘begunstigde – actor’-relaties activeren over een bepaalde tijdsspanne, met ook rolverwisselingen, vergroten de moeilijkheid om de doorwerking toe te kennen aan een enkele drager (attributieprobleem).

- 7. Doorwerking is gestructureerd in kaart te brengen
 - 7.1 De evaluatie van doorwerking moet in ieder geval uitspraken doen over *waar* je naar kijkt en *hoe* je ernaar kijkt.
 - 7.1.1 De typering van de twee soorten dragers (3.1) is een antwoord op de vraag ‘waar’ je naar kijkt bij doorwerking.
 - 7.1.2 De typering van de inspanning (volume, aard en kwaliteit) alsook de typering van de ontvankelijkheid (activatie, relevantie, bruikbaarheid, veranderingen) zijn een antwoord op de vraag ‘hoe’ je naar doorwerking kunt kijken.
 - 7.2 De rationale (1.3), relationele verhouding (1.2), de organisatiegraad (2.) en de tijdspanne (6.2) vormen de context voor betekenisgeving in de evaluatie.
 - 7.3 Voor het ‘laden’ van de typering (inspanning en ontvankelijkheid) kun je indicatoren formuleren.
 - 7.3.1 Een typering heeft altijd een of meerdere indicatoren.
 - 7.3.2 Een indicator is een (kwantitatieve of kwalitatieve) uitspraak die aantoont wat er bereikt is in relatie tot wat belangrijk wordt gevonden.
 - 7.3.3 Een indicator krijgt betekenis vanuit de rationale en relationele verhouding (1.3). Dat komt doordat de indicator de eigenschappen overerft van de inspanning (4) en de ontvankelijkheid (5).
 - 7.3.4 Aan een indicator kun je een streefwaarde/norm koppelen om vast te stellen of een bepaalde grenswaarde is gehaald of niet. Daar aan kun je bepaalde kwalificaties koppelen, zoals (on)voldoende en goed.

Een narratief

De logische analyse van het begrip doorwerking zal ik nu uitschrijven in een narratief over doorwerking van het praktijkgericht onderzoek aan een hogeschool. Dit narratief gaat als volgt.

De hogeschool heeft bepaalde taakstellingen en doelstellingen (1.1 formuleerbare intentie), waarbij de hogeschool eindverantwoordelijk is voor het bereiken van deze doelstelling (1.1.1 actor). De doelstellingen zijn gericht op het verbeteren van de (toekomstige) professionals in een veranderende wereld (1.1.2 gerichtheid; 1.1.3 begunstigde). Hiertoe onderhoudt de hogeschool (formele) relaties met studenten, docenten, alumni, beroepenveld, et cetera (1.2 relationele verhouding), die participeren in onderwijs en onderzoek (1.2.1 en 1.2.2 verhouding). Dit doet de hogeschool vanuit een bepaalde analyse van de rol van de hogeschool in de samenleving (1.3 rationale). Die rol is ook het kader waarbinnen ambities, activiteiten en opbrengsten worden geëvalueerd (1.3.1 betekenisruimte), samen met betrokken partijen (1.3.2 gedeelde kennis).

De inspanningen die de hogeschool pleegt ten aanzien van de taak- en doelstellingen zijn georganiseerd in een organisatiestructuur, in governance, in een kwaliteitszorgsysteem, et cetera (2 en 2.1 organisatiegraad). Daarbinnen zijn mensen en afdelingen werkzaam om uitvoering te geven aan de taak- en doelstellingen (2.2 aanwijsbaarheid). Zo geeft een docent les aan een student, doet een onderzoeker een experiment bij de lokale boekhandel (2.2.1 instantie), gaat een Centre of Expertise een samenwerking aan met een externe partij of brengt een kenniscentrum een publicatie uit (2.2.2 entiteit), gebaseerd op afspraken (2.3 afspraken). Zowel de docent, de onderzoeker, het kenniscentrum als het CoE geven ieder invulling aan de doelstelling van de hogeschool ('ze dragen hun steentje bij') en kunnen daarover bevraagd en beoordeeld worden (2.2.3 granulariteit).

De inspanningen van de hogeschool slaan neer in allerhande producten, diensten en persoonlijke ontwikkeling; als opbrengsten van het handelen (3 opbrengsten), waarin kennis, ervaring en inzichten zijn gepresenteerd (3.1.1 drager). Dit zijn voor de hand liggende opbrengsten als publicaties en prototypes (3.1.2 product als drager). Maar die kennis en ervaring zit ook in de docenten en onderzoekers zelf die dat kunnen overdragen in trainingen, presentaties, workshops en door samen te werken (3.1.3 actor als drager). In deze situaties vindt de beïnvloeding plaats in de directe (kennis)interactie met bijvoorbeeld een student, professional of collega-onderzoeker. Hoewel hierbij veelal ook een product aanwezig is (presentatie, materiaal, een video van de presentatie; 3.1.1 drager) is dit als een specifieke situatie te onderscheiden.

Over de opbrengsten aan de hogeschool kunnen we uitspraken doen over hoeveel er is opgeleverd (4.1.1 hoeveelheid). Vaak gaat het om een grote verscheidenheid aan opbrengsten, van publicaties tot prototypes, trainingen en gevraagd worden voor een adviescommissie (4.1.2 aard). De meeste opbrengsten zijn niet vrijblijvend maar onderdeel van een kwaliteitszorgcyclus en

geborgd in onderzoeksmethoden, en hebben daardoor een zekere kwaliteit (4.1.3 kwaliteit). Door een beschrijving en inventarisatie van al deze opbrengsten kan de inspanning van de hogeschool inzichtelijk worden. Door dit te relateren aan profilering en doelstelling van de hogeschool (1.3 rationale) kunnen we betekenis geven aan die beschrijving en inventarisatie, namelijk wat de waarde is (4.2 betekenis). Door het te relateren aan de organisatiegraad kan een cyclus van evalueren en kunnen we verbetering in gang zetten (4.3 organisatiegraad). Vandaar dat vaak ook de organisatiegraad onderdeel is van visitaties en audits, omdat die als randvoorwaarde een belangrijke rol speelt.

De hogeschool werkt met vele partijen samen, zoals bedrijven, publieke organisaties, overheid, andere onderwijsinstellingen, et cetera; niet in de laatste plaats om af te stemmen hoe de inspanningen zo effectief mogelijk kunnen zijn. Dit leidt tot (formele) afspraken over samenwerking met daarin ook uitspraken over doel, inspanning en opbrengsten (2.3 vastlegging). In andere gevallen is de samenwerking lossier, zijn het partijen die aanschuiven in klankbordgroepen, beroepenveldcommissies, stuurgroepen, expertiseteams, of is er geen direct samenwerkingsverband maar zitten de partijen 'verder weg' van de hogeschool. In alle gevallen is belangrijk in hoeverre een partij is geactiveerd, in de zin van op de hoogte van de opbrengsten van het onderzoek aan de hogeschool en het hebben van zijn/haar aandacht (5.2.1 activatie). In feite is dat een eerste stap in het tonen van ontvankelijkheid⁶ (5 ontvankelijkheid), dat wil zeggen: het kunnen en willen inzetten van de opbrengsten van het onderzoek. Een andere afweging die een partij zal maken, is in hoeverre de opbrengsten relevant zijn (5.2.2) en hoe bruikbaar ze zijn (5.2.3). Een winkelier die het moeilijk heeft om te overleven in een winkelstraat, moet weet hebben dat er onderzoek is naar het inzetten van VR in winkels om de bezoekduur te verlengen (activatie), vervolgens zal de winkelier een afweging moeten maken of het relevant is (aan welke KPI draagt het bij en heb ik daar een probleem mee voor wat betreft mijn assortiment en klanten?) en indien het relevant is, is er de afweging of het bruikbaar is (kosten-batenanalyse, risico's, inpassing in *customer journey*, et cetera). De uiteindelijke 'afrekening' is natuurlijk of het tot een daadwerkelijke verandering heeft geleid: meer en langer bezoek in de winkel in dit geval, met meer omzet tot gevolg (5.2.4).

Docenten en lectoren komen en gaan; projecten beginnen en worden afgerond; speerpunten worden Centres of Expertise; kenniscentra krijgen

6 Zie ook het rapport van de wrp (2013) waar dit het absorptievermogen wordt genoemd, hoewel 'ontvankelijkheid' niet alleen gaat om het niet kunnen maar ook om het (niet) willen ('willingness') om wat voor reden dan ook.

gezelschap van expertisecentra; kortom, een hogeschool heeft een constante dynamiek in haar organisatie en de mensen die de organisatie dragen (6.1 verandering). Vele opbrengsten (een MOOC, een publicatie, een website, een prototype, et cetera) 'leven voort' (6 eindigheid) en vinden zelf hun weg in allerhande situaties van kennisdeling en kennistoepassing (6.1.1 autonomie). Ze kunnen daarbij mensen en organisaties bereiken die niet voorzien waren of niet de beoogde 'afnemers' waren (6.1.2 niet-beoogd), en kunnen zelfs tot effecten leiden die niet tot de originele intentie te herleiden zijn (6.1.3). Dit speelt zich af in de tijd, snel of langzaam worden de effecten zichtbaar (6.2 oneindigheid): direct bij bijvoorbeeld projectdeelnemers die gezamenlijk aan een product werken en daardoor hun kennis delen en ontwikkelen; bij een bezoeker aan een conferentie die geïnspireerd raakt door opbrengsten van een onderzoek en dat meeneemt naar zijn eigen werkomgeving; of bij opbrengsten, zoals bijvoorbeeld zorgprotocollen, die doorwerken in hoe een ecosysteem opereert – en veranderingen in bijvoorbeeld armoedebestrijding in grote steden op basis van inzichten, ervaringen met interventies en onderzoeksdata. Bij bijdragen van de hogeschool aan complexe maatschappelijke vraagstukken wordt het wel moeilijker de precieze bijdrage van de hogeschool c.q. onderzoeker aan de doorwerking te isoleren, gegeven allerlei onderlinge uitwisselingen van kennis en inzichten (6.3).

De hogeschool stimuleert een omgeving waarin vele onderwerpen, aanpakken en opbrengsten kunnen floreren en effect sorteren. Die verscheidenheid brengt wel de discussie met zich mee op welk niveau van een hogeschool nog de evaluatie van het onderzoek en zijn doorwerking zinvol is te doen. Hoe 'tel' je diverse opbrengsten bij elkaar op? Wat is belangrijker om te laten zien? Waar aan besteed je aandacht en wat zet je in de etalage als een typisch voorbeeld van het praktijkgericht onderzoek? Zonder inhoudelijke discussie te hoeven slechten over methodologieën, kennistheoretische uitgangspunten en paradigma's kun je altijd de vraag stellen waar je uitspraak over wilt doen (7.1.1) en wat voor soort uitspraak je daarover wilt doen (7.1.2). Dit geeft niet alleen al richting aan de indicatoren die hierbij passen (7.3), maar vindt zijn noodzakelijke invulling ook in wat belangrijk wordt gevonden (1.3 rationale) en hoe alle deelnemers in het onderzoek zich hiertoe verhouden (1.3.1 relationele verhouding).

Een tool

De bovenstaande ideeën over doorwerking heb ik uitgewerkt in de zogeheten doorwerkingsmatrix. De doorwerkingsmatrix heb ik in eerste instantie

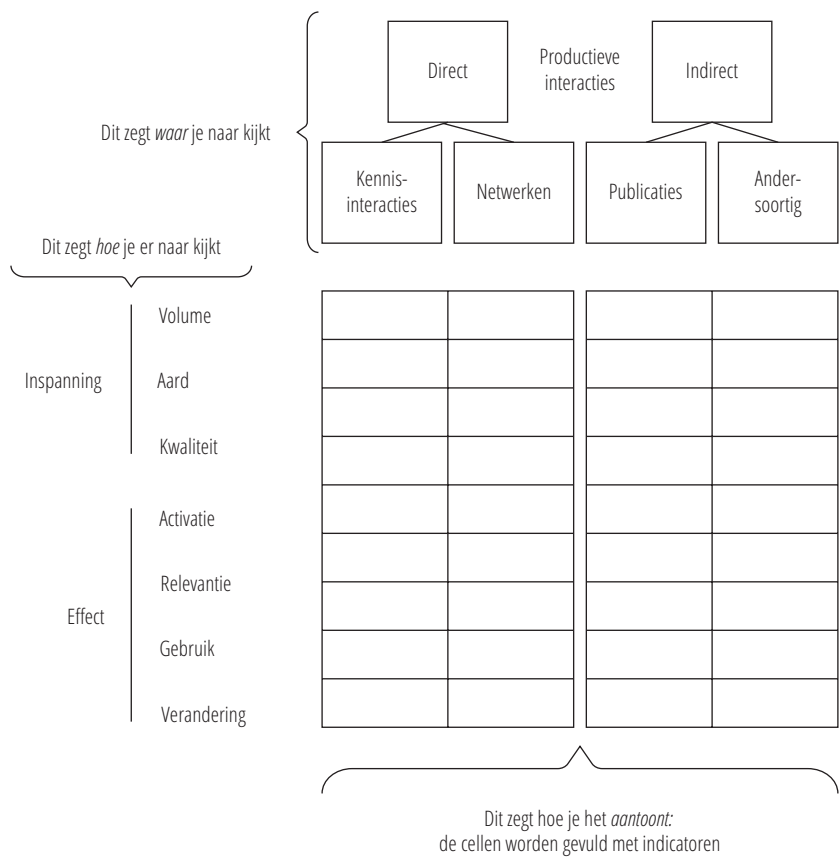
ontwikkeld als een hulpmiddel om de doorwerking van onderzoek in kaart te brengen en hierover te rapporteren. Het is een conceptueel model dat geen voorwaarden stelt over het gevolgde onderzoeksproces of een onderliggende structuur (bijvoorbeeld een logisch model), noch over wanneer (voor, tijdens, na) je naar de doorwerking moet kijken. De enige procesaanname is dat er inspanningen zijn die effecten sorteren.

Centraal in de doorwerkingsmatrix staan twee vragen: waarover wil je een uitspraak doen, en: wat voor soort uitspraak wil je daar dan over doen? De eerste vraag gaat over de 'objecten' waarnaar je wilt kijken. De grote diversiteit aan opbrengsten van onderzoek maakt het noodzakelijk hier een bepaalde categorisering in aan te brengen. In een eerste versie van de doorwerkingsmatrix werden drie bottom-up vastgestelde categorieën gebruikt: producten, netwerken en kennisinteracties (Van Vliet et al., 2020). In een tweede versie is een onderverdeling gebruikt gebaseerd op het onderscheid in directe en indirecte productieve interacties (Van Vliet, 2021; zie ook 'Spaceland'). In principe kun je iedere andere indeling gebruiken, afhankelijk van onder andere context en doel.

De tweede vraag wordt beantwoord door zeven aspecten van doorwerking te onderscheiden, bij doorwerking kun je vragen naar: volume (hoeveel is er gedaan?), aard (wat is er gedaan?), kwaliteit (wat is de kwaliteit?), activatie (wie is er bereikt en geactiveerd?), relevantie (welke waarde wordt er toegekend?), bruikbaarheid (wordt er gebruik van gemaakt?) en verandering (wat is er daadwerkelijk veranderd?).

Door deze twee vragen als twee dimensies te zien van een matrix, ontstaat een systematisch overzicht waarbij in de cellen van de matrix vragen (en antwoorden) terechtkomen (figuur 4), zoals: in welke netwerken nemen we deel (netwerken – aard)? Worden onze prototypes gebruikt (producten – gebruik)? Hoeveel workshops hebben we gegeven (kennisinteracties – volume)? Hoe relevant worden onze websites gevonden (producten – relevantie)? Et cetera.

De matrix kun je gebruiken als hulpmiddel om een narratief te schrijven door te bepalen aan welke aspecten het narratief met name aandacht wil besteden, bijvoorbeeld de verschillende netwerken waarin de onderzoekers betrokken zijn (netwerken – aard) en/of het gebruik van bepaalde producten als resultaat van het onderzoek (producten – gebruik). Een tweede manier om de matrix te gebruiken is als systematisering van indicatoren. Iedere cel in de matrix kun je vullen met bestaande indicatoren of met nieuw te ontwikkelen indicatoren. Aan de hand van de indicatoren kun je voorbeelden en data verzamelen. Op deze manier is er een systematisch overzicht van datgene wat gerapporteerd wordt over het praktijkgericht onderzoek, en kunnen we eenvoudig en snel eventuele lacunes opsporen of 'afwijkingen' van gestelde doelen.



Figuur 4. Doorwerkingsmatrix

Exploratief onderzoek #1: inspanning versus effect

In de voorbereiding op de externe visitatie van de kenniscentra aan de Hogeschool van Amsterdam in het najaar van 2021 is een analyse uitgevoerd op de midterm zelfevaluaties uit 2018 van de zes betreffende kenniscentra (Van Vliet, 2020). Deze analyse bestond uit een tekstanalyse aan de hand van een vooraf opgesteld scoreformulier. Iedere zin in de rapporten is bekeken met de vraag of het een uitspraak bevatte over opbrengsten van het praktijkgericht

onderzoek en wat de aard was van die uitspraak. Voor het vaststellen van die aard is het onderscheid gebruikt tussen inspanning en effect. Hierbij is inspanning onderverdeeld in uitspraken over hoeveelheid (volume), de verschijningsvormen van de opbrengsten (aard) en de kwaliteit van de inspanning (kwaliteit). Het effect is onderverdeeld in uitspraken over bereik en activatie van doelgroepen (activatie), de relevantie van de opbrengsten (relevantie) en uitspraken over het gebruik van de opbrengsten (gebruik). Deze indeling volgt de logische analyse van doorwerking (zie 7.1.2). Het aspect van 'verandering' is pas later toegevoegd aan deze logische analyse en werd in dit eerste onderzoek nog niet meegenomen.

De uitkomst van dit onderzoek was dat over alle 6 de zelfevaluaties heen 66% van de uitspraken over opbrengsten gingen over inspanningen, en 34% over het effect van de inspanningen. Bij inspanningen droeg met name het aspect van 'aard' hieraan bij: uitspraken over de grote verscheidenheid aan opbrengsten werden in alle rapportages het meest aangetroffen. Dit komt ook wel tegemoet aan het beeld dat praktijkgericht onderzoek niet alleen is af te meten aan publicaties, maar op zoveel andere manieren kennis deelt (workshops, prototypes, kennisclips, dashboards, hackatons, living labs, et cetera). Wel moet ik opmerken dat de verschillende kenniscentra onderling verschillen in de verhouding inspanning – effect. Deze waarde varieert tussen de 76% en 54% procent, waarbij met name de relatieve aandacht voor het effect bepalend is. In sommige zelfrapportages wordt hier weinig over gezegd, terwijl andere zelfrapportages hier redelijk uitgesproken over zijn.

Er is zeker iets voor te zeggen dat de uitspraken ook uitgelokt worden. De zelfrapportages conformeren zich aan de BKO-standaarden en de vraagstellingen daarin hebben een bias voor (kwantitatieve) inspanningsantwoorden. Voor de rapportage over de input en de producten worden uniforme kwantitatieve indicatoren aangedragen in het BKO, voor de impact (gebruik/erkenning) wordt dit overgelaten aan de hogeschoolen, als het er maar minimaal zes zijn. Gebruik en erkenning overlappen maar zeer ten dele met de in de logische analyse onderscheiden aspecten van activatie, relevantie en gebruik. Het kan dus zijn dat wat niet gevraagd wordt, ook niet in de rapportages wordt geschreven. Dit is misschien ook een verklaring voor de onvrede die soms doorklinkt over de eigen rapportages doordat mensen zich strikt houden aan de BKO-richtlijnen.

Een ander onderdeel in de analyse betrof de vraag *waar* naar gekeken werd als er gevraagd werd om over de opbrengsten van het praktijkgericht onderzoek te rapporteren. Voorafgaand aan de analyse was hier al aandacht aan besteed in een sessie met de lectoren van de Hogeschool van Amsterdam.

Aan de hand van een kaartspel met zo'n vijftig indicatoren kregen lectoren in groepjes de vraag te komen tot een ranking van welke indicatoren zij het relevantst vonden om de meerwaarde van hun onderzoek te benoemen (Van Vliet et al., 2020). In de analyse van de resultaten kwamen drie groepen van opbrengsten naar boven: producten (onder andere publicaties, websites, apps, *demonstrators*, dashboards, et cetera), netwerken (lid van adviesraden, denktanks, redactieraden, EU-panels, et cetera) en kennisinteracties (presentaties, workshops, consultancy, expertsessies, trainingen, et cetera). Deze driedeling is dus bottom-up ontstaan, maar goed te verbinden met het onderscheid in de logische analyse van indirecte kennisdeling – via een drager/product – en een directe kennisdeling middels een interactie in netwerken of kennisinteracties (zie lemma's 3.1.2, 3.1.3 en 7.1.1).

De uitkomst van deze analyse was dat de 3 categorieën nagenoeg in gelijke verhouding voorkomen in de rapportages: producten (36%) noemden respondenten iets vaker dan netwerken (32%) en kennisinteracties (31%). De verschillen tussen de kenniscentra zijn echter opmerkelijk. Zo liggen de scores op kennisinteracties tussen de 5% en 44%, en bij netwerken is dit 20% en 48%. Bij producten is het verschil het kleinst, scores liggen tussen de 29% en 48%. Ook hier kun je het argument van uitlokking gebruiken: van de 24 indicatoren die als voorbeeld dienen voor gebruik/erkenning in het BKO-protocol gaan er maar 5 over kennisinteracties (expertmeetings, onderwijs geven) en netwerken (lid wetenschappelijke redactie, deelname adviesraden, participatie in netwerken). Het vergt dus enige inspanning, creativiteit en lef om het belang hiervan zelf voor het voetlicht te brengen.

Doorwerking redux

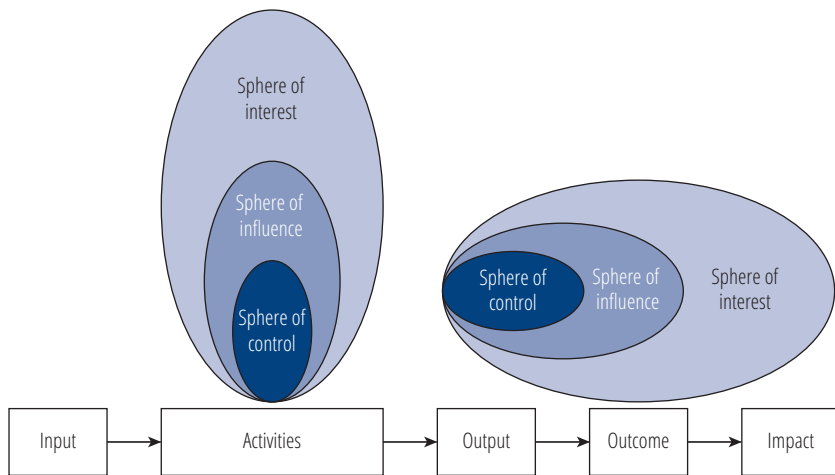
De bovenstaande analyse van de aard van doorwerking maakt duidelijk dat de typering daarvan in ieder geval berust op drie pijlers: communicatie (een actor en een begunstigde die in een relationele verhouding tot elkaar staan), sociale werkelijkheid (actoren en begunstigden opereren in een geheel van gedeelde kennis, rationale, afspraken en entiteiten en instanties die daar invulling aan geven) en een handelingsperspectief (het interveniëren om te komen tot een verandering). In feite een iets gedetailleerde parafrasering van de vele opmerkingen dat onderzoek doen 'vernetwerkt is', interactief, dynamisch... kortom 'messy'.

Dit vraagt om een verdere theoretische uitwerking. Hier zijn genoeg kandidaten voor, van grote theorieën zoals over communicatief handelen (Habermas)

of over actornetwerken (Latour) tot meer specifieke theorieën zoals die van *speech acts* (Searle) en misschien zelfs 'lichte' verkenningen van wetmatigheden in de opbrengsten van handelingen tussen twee actoren, zoals Cipolla nastreeft in zijn zoektocht naar de wetmatigheden van menselijke stupiditeit (Cipolla, 2016). Deze uitwerking blijft in deze voorbeschouwing achterwege. Termen die in ieder geval lijken te passen bij de analyse zijn termen als 'performatief', 'realistisch' en 'constructivistisch' (Coombs & Meijer, 2021). Ik kom hier nog op terug (zie 'Pointland').

Ook zonder een theoretisch fundament kan ik hier al twee consequenties noemen van de gevolgde gedachtegang. Een eerste consequentie is dat doorwerking zich niet alleen in het verloop van de tijd kan voordoen, maar op ieder moment. Het SLM is in feite een 'diachronische' typering: kennis bouwt zich op, cumuleert, in de lengte van de tijd tot er een overdracht kan plaatsvinden naar een begunstigde. Daartegenover staat een 'synchronische' typering: in ieder tijdsmoment is er een handelingsperspectief om tot effecten te komen. Dit is al naar voren gekomen in de typering van doorwerking middels: 'ook tijdens, niet alleen na', maar krijgt hier door de beschrijving van de aard van doorwerking een krachtiger inhoud: (communicatief) handelen in een sociale context met een keuze hoe ieder moment te benutten voor doorwerking. Als literaire beschrijving kunnen we – analoog aan het 'leven in de breedte' van schrijver A.F.Th. van der Heijden – spreken over 'doorwerking in de breedte': ieder moment, iedere handeling verbreden om het waardevoller te maken. Het 'leven in de lengte' is niet te stoppen en heeft een onvermijdelijkheid die zich zal laten gelden. Zo cumuleert onderzoek wel of niet in overgedragen kennis, in impact, maar in de breedte ligt een aanvullend handelingsperspectief. Met het opgeroepen beeld van lengte en breedte, kunnen we ook stellen dat doorwerking meer is dan een lijn, in ieder geval een vierkant.

Dit argument is vrij eenvoudig in beeld te brengen. Hiervoor kun je een veelgebruikte visuele voorstelling gebruiken van de impact van onderzoek, zoals het SLM dat zich voorstelt. Na alle activiteiten worden de fases van output – outcome – impact getypeerd aan de hand van drie sferen: de *sphere of control*, de *sphere of influence* en de *sphere of interest* (figuur 5). Maar waarom zou je dit niet kunnen kantelen? Een kwartslag draaien en deze 'sferen' ook een rol geven op eerdere en eigenlijk alle momenten in het onderzoeksproces? Komt dit niet veel meer tegemoet aan de al eerder aangehaalde 'dynamiek' van het onderzoeksproces? Belangrijker nog: geeft dit niet veel meer mogelijkheden om doorwerking als handelingsperspectief te integreren in het gehele onderzoeksproces?



Figuur 5. De tweede dimensie van doorwerking

Een tweede consequentie is dat de vraag naar boven komt of er überhaupt een procesmodel van onderzoek, zoals het SLM, nodig is om over doorwerking te kunnen spreken. Een procesmodel krijgt al gauw een voorschrijvend, normatief karakter, hoe dynamisch ook weergegeven. De onvrede met het SLM heeft al geleid tot voorstellen met aanpassingen op het SLM, met allerlei pijlen die heen en weer wijzen waardoor het overzicht snel verdwijnt. Deze alternatieven zijn dan ook weggezet als *a plate of spaghetti with meatballs*. Maar wat als we een conceptueel model voor doorwerking ontwikkelen dat geen principiële uitspraak doet over het onderzoeksproces zelf? De gepresenteerde doorwerkingsmatrix kun je zien als zo'n conceptueel model, immers de enige 'proces'-aanname is dat er inspanningen zijn met bepaalde effecten. Die inspanningen zijn getypeerd als communicatief, als sociale handelingen en niet gekoppeld aan 'processtappen'.

De doorwerkingsmatrix is ontwikkeld als een evaluatie-instrument (Van Vliet et al., 2020) maar kun je eenvoudig omkatten in een planningsmatrix, omdat dezelfde vragen relevant blijven. De cellen in de matrix vul je dan niet met indicatoren, maar met (voorgenomen) activiteiten. Vragen over inspanningen: wat gaan we doen (aard) en hoeveel (volume) en hoe zorgen we voor kwaliteit van wat we doen (kwaliteit) zijn typisch vragen die we in projectplannen beantwoorden: werkpakketten (aard), budget en mensen (volume) en bijvoorbeeld

Research Data Management (kwaliteit) zijn herkenbare uitingen hiervan. Vragen over het effect kunnen we ook in de planning stellen: hoe en welke stakeholders activeren we? Wat is de relevantie van het onderzoek en hoe houden we dit scherp? Hoe kunnen we ervoor zorgen dat we inzichten en andere opbrengsten gebruiken? En op welke verandering zijn we uit? Antwoorden kunnen gestructureerd invulling krijgen door te kijken naar de gerichtheid van de activiteiten (onderwijs, onderzoek, beroepspraktijk) en de productieve interacties: welke invulling geef je aan je handelen?

Door de doorwerkingsmatrix te beschrijven als planningstool, naast evaluatietool, lijkt ik een knieval te maken voor toch een procesperspectief. Nee, het argument is juist dat de doorwerkingsmatrix een overkoepelend perspectief biedt, of het nu voor of na of tijdens is. Voor dit laatste, 'tijdens', zal ik in 'Pointland' bestaande methoden beschrijven die hier invulling aan geven, maar eerst een intermezzo dat de deur opent naar 'Spaceland'.

Intermezzo

Indicatoren

Ik neem hier de ruimte voor een intermezzo. De discussie over impact van onderzoek kunnen we namelijk niet alleen voeren aan de hand van een onderliggend model, maar ook middels de zoektocht naar geschikte indicatoren voor de impact van onderzoek. Dit model en deze indicatoren hangen nauw samen, maar dat staat niet in de weg om apart te reflecteren op deze zoektocht naar indicatoren. Niet in de laatste plaats omdat het de deur zal openen naar 'Spaceland'.

Er is geen gebrek aan indicatoren als we de impact van onderzoek willen vaststellen. De standaardprotocollen voor evaluatie van onderzoek in Nederland hanteren een relatief beperkte set van indicatoren. Voor universitaire onderzoeksevaluatie is dit het SEP (*Standard Evaluation Protocol*). Het vastgestelde SEP voor de periode 2015–2021 (KNAW, VSNU & NWO, 2016) bevat dertig indicatoren. Voor onderzoek aan de hogescholen is dit het Brancheprotocol Kwaliteitszorg Onderzoek (BKO) van de Vereniging Hogescholen. De meest recente versie van de BKO betreft de periode 2016–2022 en bevat veertig indicatoren (Vereniging Hogescholen, 2015). Een alleszins overzichtelijk aantal.

Op beide protocollen zijn echter diverse aanvullingen gekomen, zoals vanuit de geesteswetenschappen voor het SEP (QRiH, 2017). Binnen het hoger onderwijs heeft een tweetal commissies de BKO verder aangevuld (Pijlman et al., 2017; Franken et al., 2018). Maar dit is slechts het topje van de ijsberg. Zo heeft het Rathenau Instituut een protocol ontwikkeld voor de evaluatie van de zes zogenaamde TO2 instellingen (TNO, MARIN, NLR, Deltares, ECN en de WUR) (Deuten et al., 2015). En onderzoeksbureaus zoals Biggar Economics (2014) en Birch (Cloosterman et al., 2019) hanteren een eigen set van indicatoren. Voorts zijn er initiatieven zoals U-multirank (www.umultirank.org), waarop je universiteiten kunt vergelijken, onder andere op basis van een reeks indicatoren over onderzoek, en ontwikkelingen zoals altmetrics (www.altmetric.com), waarbij de online conversatie wordt gemonitord over gepubliceerd onderzoek zoals op sociale media, Wikipedia en blogs. In de sectie 'Exploratief onderzoek #3' breng ik meer dan zeshonderd indicatoren in kaart vanuit twintig bronnen.

De hoeveelheid indicatoren is geen garantie voor verscheidenheid. Het rapport van STW, Rathenau Instituut en Technopolis (Van Drooge et al., 2011) over valorisatie-indicatoren is hier heel stellig in: 'De meeste indicatoren betreffen economische benutting, er zijn weinig indicatoren voor maatschappelijke

benutting; de meeste indicatoren zijn van toepassing op medisch, technisch of natuurwetenschappelijk onderzoek, indicatoren voor andere onderzoeksgebieden zijn er nauwelijks; de meeste indicatoren betreffen output, er zijn weinig indicatoren voor impact, interactie of voor andere fasen van het onderzoeksproces' (p. 17). Ook Godin benadrukt de dominantie van de economische dimensie in indicatoren (Godin & Doré, 2005). Dit zal ongetwijfeld (deels) de onvrede verklaren die er bestaat over de inzet van indicatoren. Die onvrede zit in het bijvoorbeeld niet inzichtelijk kunnen maken van maatschappelijke impact, of in het feit dat de indicatoren niet aansluiten bij de onderzoekspraktijk. Ik zal aan de hand van vier thema's reflecteren op deze spanning tussen de hoeveelheid beschikbare indicatoren en de voortdurende zoektocht naar 'juiste' indicatoren.

Context, context, context

Een eerste reflectie heeft als startpunt het gebruik van indicatoren, dit is namelijk vaak kwantitatief: hoeveelheid publicaties, percentage deelname studenten aan onderzoek, omvang van totale tweede geldstroom, et cetera. Cijfers hebben de neiging een eigen leven te leiden als ze los komen van de context waarin ze zijn ontstaan. Een getal geeft weinig inzicht in voor wie die impact was, wat het effect is geweest van die impact, wat heeft bijgedragen aan het bereiken van de impact, et cetera. Zeker in een evaluatieve context kan een getal een (nieuwe) betekenis krijgen van 'te weinig' of 'onder de norm'. Dit 'decontextualiseren' van informatie in indicatoren en de indicator vervolgens weer contextualiseren voor het 'afrekenen' van programma's is een realistisch gevaar (Coombs, 2019). Met andere woorden: indicatoren moeten we altijd voor een context selecteren en in een context presenteren. 'Indicators used in impact assessment cannot be universal. Instead, they need to be developed for given contexts' (Pedersen, Grønvad & Hvidtfeldt, 2020, p.1). Als leidraad kunnen we hierbij de volgende uitspraak hanteren: 'Een werkelijkheid die gereduceerd wordt tot indicatoren is armoedig. Een werkelijkheid die wordt aangevuld met indicatoren is rijk en geeft heldere communicatie' (Van Drooge et al., 2011, p. 7).

Contextualiseren betekent dat we goed moeten beschrijven in welke context we de indicatoren gebruiken. Hiervoor kunnen we in ieder geval de volgende vier aspecten onderscheiden (zie ook De Jong et al., 2011; Van Drooge et al., 2011):

1. De scope of het aggregatieniveau van de onderzoeksevaluatie. Veelal gebeurt dit binnen hogescholen op basis van een organisatorische eenheid

(kenniscentrum, lectoraat) of programmatische lijn (Centre of Expertise, groot onderzoeksprogramma). In principe kan de scope ook een geografische eenheid zijn: wijk, stad of regio. De keuze van de scope heeft consequenties voor welke indicatoren van belang zijn en hoe we die wegen. Bovendien wordt hier een probleem zichtbaar: hoe individuele onderzoeksprojecten optellen tot een groter geheel, zoals een programma of een kenniscentrum (zie 'Pointland')?

2. Het domein, de sector of de discipline waarover gerapporteerd wordt. Creatieve industrie, zorg, landbouw en transport zitten anders in elkaar, wat betekent dat de activiteiten en de (aard van) opbrengsten van het onderzoek verschillen. Daar moeten we recht aan doen. Huidige indicatoren zijn vaak 'domeinneutral', wat een sterkte en een zwakte is. Het voordeel is dat we de indicatoren breed kunnen inzetten en dat ze een gedeeld kader bieden, hierdoor wordt het makkelijker onderzoek te vergelijken en er onderling over te communiceren. Het nadeel is dat de eigenheid van een domein niet doorklinkt. Wel recht doen aan het domein betekent vragen stellen zoals: wat is er relevant in dit domein? Hoe zit het domein 'in elkaar'? Hoe wordt kennis ontwikkeld, gedeeld en gebruikt? Wie zijn belangrijke stakeholders en waarom? Dit zal per domein leiden tot de keuze van indicatoren, een bepaalde weging van deze indicatoren, en de noodzakelijke contextualisering c.q. interpretatie van de indicatoren.
3. De partij of doelgroepen waarover de indicator iets willen zeggen. Binnen bijvoorbeeld het BKO is er een onderscheid tussen beroepspraktijk/samenleving, onderwijs (studenten en docenten) en het onderzoeksveld c.q. de wetenschap. Een indicator voor de bijdrage aan de beroepspraktijk is een andere dan een indicator voor de bijdrage aan het onderwijs.
4. De fase waarin het onderzoek zich bevindt. Een beginnende onderzoeksgroep of een sector/discipline in opkomst, bijvoorbeeld AI, vergt deels andere indicatoren dan een onderzoeksgroep die al meer volwassen is en al een gevormde traditie heeft van onderzoeksmethoden, netwerken en gedeelde onderzoeksagenda's.

Criteria en indicatoren

Een tweede reflectie is het verschil tussen een criterium en een indicator. Een criterium is een stelregel waaraan je wilt voldoen; een indicator is de maat die aantoont hoe je je verhoudt tot die stelregel, het criterium. Zo geldt bijvoorbeeld voor olympische deelname van sporters de stelregel dat ze op niveau

moeten presteren; de afstand of hoogte die wordt gesprongen is de indicator, en vervolgens is de olympische limiet de norm. Voor praktijkgericht onderzoek is bijvoorbeeld een stelregel dat resultaten bruikbaar moeten zijn voor de praktijk, vervolgens is een indicator nodig die aantoont dat dit het geval is, bijvoorbeeld door uitspraken van stakeholders over die bruikbaarheid. Een ander voorbeeld is een criterium als 'resultaten zijn toegankelijk', een indicator zou dan kunnen zijn het relatieve aandeel van publicaties dat via open access is gepubliceerd, en een norm dat dit voor 50% van de publicaties geldt. Een criterium als 'onderzoekers zijn integer' is al een stuk lastiger aan te tonen, een indicator zou kunnen zijn het aantal onderzoekers dat een cursus integriteit heeft gevolgd of geslaagd is voor een 'integriteitstoets'. Indicatoren zijn dus afgesproken manieren om aan te tonen wat er bereikt is in relatie tot een bepaalde stelregel; met een eventuele norm wordt bepaald of het wel of niet gehaald is.

Dit onderscheid tussen criteria en indicatoren wordt niet altijd gemaakt en leidt tot verwarring (voor voorbeelden zie Van Vliet et al., 2020). Het is soms ook lastig een onderscheid te maken. 'Tevredenheid van stakeholders' kun je bijvoorbeeld als een criterium lezen met dan de vervolgvraag 'hoe toon ik dit aan' oftewel welke indicator kies ik? Maar 'tevredenheid van stakeholders' kun je ook lezen als een indicator, met dan de vervolgvraag: 'Wat toon ik hiermee aan?' Oftewel, aan welk criterium wordt voldaan?

Het verschil tussen criterium en indicator is iets wat we goed in de gaten moeten houden. Het is het verschil tussen waar ik mezelf aan wil houden en de keuze hoe te bewijzen of je je eraan gehouden hebt (indicator) met mogelijk een streefwaarde (norm). Tegelijkertijd is het zaak dat je criterium en indicator als tweespan inzet bij een evaluatie: zowel formuleren waaraan het onderzoek gehouden moet worden als de keuze hoe hier bewijs voor aan te dragen. Het voorstel is expliciet te zijn in welke criteria je onderscheidt en welke indicatoren je ziet om aan te tonen wat er bereikt is op het criterium.

Exploratief onderzoek #2: criteria praktijkgericht onderzoek

In hetzelfde onderzoek waaraan ik eerder refereerde ('exploratief onderzoek #1') is ook gekeken naar de genoemde criteria die voorkwamen in de kritische zelfreflecties, uit 2018, van de zes kenniscentra aan de Hogeschool van Amsterdam. Hiertoe is allereerst een lijst gemaakt van criteria die genoemd worden voor praktijkgericht onderzoek. Deze lijst is tot stand gekomen door bronnen zoals de rapporten van de commissie-Pijlman (2017), de commissie-Franken (2018) en de

commissie-Reiner (2019) te scannen op criteria. Uiteindelijk resulteerde dit in een lijst van achttien criteria, zoals: 'de beroepspraktijk is betrokken bij het onderzoek', 'onderzoekers nemen deel aan (regionale) netwerken', 'in het onderzoek is sprake van *co-makership*' en 'het onderzoek is relevant voor de praktijk' (zie tabel 2).

Uit de analyse komt naar voren dat nagenoeg alle criteria in de kritische zelfreflecties worden genoemd, behalve het criterium van 'onderzoekers zijn lid van (regionale) adviescommissies' dat in geen van de zes rapporten werd genoemd. Tegelijkertijd werd wel een aantal malen het criterium van 'internationalisering' genoemd, dat zouden we dus nog aan de lijst van criteria kunnen toevoegen. De conclusie was, op basis van de analyse van deze zes kritische zelfreflecties, dat de lijst een goede afspiegeling vormt van wat kenniscentra belangrijk vinden aan praktijkgericht onderzoek.

Daarmee is niet gezegd dat alle criteria evenveel aandacht kregen. Criteria zoals 'er is sprake van een diversiteit aan onderzoeksresultaten' en 'onderzoeksresultaten worden gebruikt', werden maar in drie van de zes rapportages van kenniscentra genoemd. Criteria als 'er zijn strategische (meerjarige) partnerships met diverse instellingen, bedrijven en organisaties' en 'onderzoeksresultaten zijn toegankelijk' werden door vier van de zes kenniscentra genoemd. De vijf meest genoemde criteria in de kritische zelfreflecties waren, in volgorde: 'het onderwijs is betrokken bij het onderzoek', 'het onderzoek is relevant voor de praktijk', 'de beroepspraktijk is betrokken bij het onderzoek', 'het onderzoek heeft kwaliteit' en 'het onderzoek levert een bijdrage aan het onderwijs'.

Belangrijk is ook op te merken dat de kenniscentra onderling verschilden in hun profiel van scores op de criteria. Bij het ene kenniscentrum speelt het criterium 'het onderzoek is multidisciplinair/interdisciplinair van aard' een prominente rol, terwijl dit bij andere kenniscentra nauwelijks wordt genoemd. Ook bijvoorbeeld een criterium als 'onderzoekers nemen deel aan (regionale) netwerken' krijgt bij het ene kenniscentrum meer aandacht dan bij een ander kenniscentrum. Dit mag eigenlijk niet meer verrassend zijn, gegeven de eerdere opmerkingen over de context van het onderzoek dat bepalend is voor de evaluatie van dat onderzoek: de eigen strategie en het profiel van het kenniscentrum, het onderzoeksdomein van het kenniscentrum, en de ontwikkelingsfase van een kenniscentrum zijn relevante factoren in de keuze van wat mensen als belangrijk zien.

Die keuze zou dan ook moeten doorklinken in de gebruikte indicatoren. Het is echter geen uitzondering om in de analyse van de kritische zelfreflecties criteria aan te treffen zonder bewijslast c.q. indicator, en omgekeerd bewijs gepresenteerd te krijgen waarvan niet duidelijk is waarvoor het bewijs is. Dit typeren we als *orphans* (Van Vliet et al., 2020). Een gestructureerde aanpak

van wat als belangrijk wordt gezien en hoe dit aan te tonen, zou dit kunnen voorkomen.

Tabel 2. Criteria voor praktijkgericht onderzoek (Van Vliet et al., 2020)

	 criterium
1.	Er zijn strategische (meerjarige) partnerships met diverse instellingen, bedrijven en organisaties.
2.	Er is een gezamenlijke visie van de samenwerkingspartners.
3.	Het onderzoek sluit aan bij (regionale/nationale/internationale) agenda's.
4.	Onderzoekers nemen deel aan (regionale) netwerken.
5.	Onderzoekers zijn lid van (regionale) adviescommissies.
6.	De beroepspraktijk is betrokken bij het onderzoek.
7.	Het onderwijs is betrokken bij het onderzoek.
8.	De onderzoekswereld is betrokken bij het onderzoek.
9.	Het onderzoek is relevant voor de praktijk.
10.	Het onderzoek is multidisciplinair/interdisciplinair van aard.
11.	Het onderzoek draagt bij aan kennisverhoging.
12.	Het onderzoek heeft kwaliteit.
13.	Er is sprake van een diversiteit aan onderzoeksresultaten.
14.	In het onderzoek is sprake van co-maatschap.
15.	Het onderzoek levert een bijdrage aan het onderwijs.
16.	Het onderzoek wordt (internationaal) erkend.
17.	Onderzoeksresultaten worden gebruikt.
18.	Onderzoeksresultaten zijn toegankelijk.

Verscheidenheid en vergelijkbaarheid

Een derde reflectie heeft te maken met de vraag of er een gedeelde set van indicatoren is aan te reiken als er sprake is van een grote verscheidenheid in doelen, uitvoering en resultaten bij onderzoek. Een negatief antwoord hierop klinkt door in het rapport van de commissie-Franken: 'Op basis van ervaringen in binnen- en buitenland en de wetenschappelijke literatuur mag de conclusie

getrokken worden dat het nagenoeg onmogelijk is om een gemeenschappelijke set aan indicatoren voor alle hoger onderwijsinstellingen te vinden, te meer omdat de doelstellingen van het onderzoek en de praktijk in de verschillende onderzoeksgebieden vaak erg verschillen' (Franken et al., 2018, p. 28). Een ander geluid laten Pedersen, Grønvad & Hvidtfeldt (2020) horen: 'Drawing on the strong methodological pluralism emerging in the literature, we conclude that there is considerable room for researchers, universities, and funding agencies to establish impact assessment tools directed towards specific missions while avoiding catch-all indicators and universal metrics' (p. 1). We vinden eenzelfde standpunt terug bij een EU expertgroep: "An EU expert group that was to develop indicators for the evaluation of RRI [Responsible Research and Innovation], concluded that RRI, being a dynamic and multifaceted concept, would not benefit from a fixed set of indicators. It was rather in need of a toolbox of quantitative and qualitative indicators. The expert group concluded that the assessment of RRI required both indicators of the process and the outcome and impact of research and innovation. The indicators should support the learning process of the actors and organizations involved ((Expert Group on Policy. Indicators for RRI 2015, in: Van Drooge & Spaapen, 2017, p. 6).

Er zit nog ruimte tussen een aantal indicatoren voor iedereen verplicht voorschrijven en de vrijblijvendheid zoals bijvoorbeeld in het Brancheprotocol Kwaliteitszorg Onderzoek ten aanzien van 'gebruik' ('iets over gebruik melden'). Dit is in feite de roep om een toolbox. De League of European Research Universities (LERU) verwoordt het als volgt: 'LERU is in favour of the creation of a "toolbox" of indicators, some short, some medium and some longer term, from which, per call or programme, those indicators can be chosen that are most relevant' (Keustermans et al., 2018, p. 10). De geesteswetenschappen hanteren eenzelfde concept van de toolbox: 'De indicatoren vormen als het ware een toolbox of gereedschapskist waaruit een keuze kan worden gemaakt voor de zelfevaluatierapporten' (QriH, 2017, p. 11). Binnen de hogescholen is een roep om een dergelijke toolbox nog niet concreet verwoord, maar kan deze wel als invulling worden gezien van een gezamenlijk statement van ministerie, nationaal regieorgaan SIA en de Vereniging Hogescholen om te 'werken aan een systeem voor impactmeting' (2019, p. 9).

Exploratief onderzoek #3: indicatoren voor onderzoek

Een onderzoek is opgezet naar gebruikte indicatoren voor de evaluatie van onderzoek. Hiertoe zijn tien relevante bronnen geanalyseerd die uitspraken

doen over indicatoren voor universitair onderzoek, en er zijn elf relevante bronnen geanalyseerd die uitspraken doen over indicatoren voor onderzoek aan hogescholen.⁷ Dit leverde 640 genoemde indicatoren op. Er is veel overlap tussen de indicatoren omdat verschillende bronnen veel van elkaar lenen (Holi, Finne, VSNU enerzijds, Pijlman, Franken, BKO anderzijds) en dezelfde auteurs regelmatig terugkeren (Spaapen, Van Drooge; en Andriessen, Van der Zwan). Tegelijkertijd verschillen ze veelal in de mate van gedetailleerdheid. Algemene aanduidingen van indicatoren zoals licenties (Spaapen & Van Drooge (2011b) en *publications* (Holi et al., 2008) kennen ook preciezere uitwerkingen zoals 'International licensing trade from HEIs and PROs' (Finne et al., 2011) respectievelijk 'aandeel open access publicaties' (VSNU, 2013b). Die gedetailleerd zit soms ook in het uitsplitsen naar doelgroepen, zoals boeken voor een algemeen publiek of voor een professioneel publiek (QRiH) of aantal contracten met bedrijfsleven of publieke organisaties (VSNU, 2013b).

De hoeveelheid indicatoren, de overlap en ook de verschillen in detaillering maakten het lastig tot een gerichte analyse te komen, daarom zijn de indicatoren ingedeeld in categorieën. Uitgangspunt hierbij is geweest om een hanteerbare set van categorieën te hebben die dicht bij de bronnen blijft in de typering en clustering van indicatoren, en een goede balans weet te vinden tussen algemeen en specifiek. Verder is besloten de categorieën niet te baseren op doelgroepen. In navolging van Van Drooge et al. (2011) en Van Vliet et al. (2020) worden doelgroepen gezien als een aanvullend aspect, net zoals aggregatieniveau en het onderzoeksdomein, die de keuze van indicatoren bepaalt en de precieze interpretatie c.q. toepassing. Ook is besloten om meerdere categorieën voor output te hanteren, meer dan alleen publicaties en 'anders', vooral ook omdat diverse bronnen dit zelf al doen (bijvoorbeeld QRiH), maar ook om de variëteit aan opbrengsten goed tot haar recht te laten komen.

De indeling was soms lastig omdat bronnen verschillen in de presentatie van de indicatoren, sommige bronnen volstaan met een lijstje woorden (onder andere ERiC, 2011; Spaapen & Van Drooge, 2011b; VSNU, 2013b), andere bronnen geven definities en uitgebreide toelichting (onder andere Finne et al., 2011; Van Drooge et al, 2011; QRiH, 2017). Daarnaast komt het voor dat een indicator samengesteld is, in die zin dat het in feite over meer dan één ding

7 Bronnen voor indicatoren universitair onderzoek: Van Drooge et al. (2011), ERiC (2011), Finne et al. (2011), Holi et al. (2008), QRiH (2017), knaw, vsnu & nwo (2016), Spaapen & Van Drooge (2011b), vsnu (2013b), vsnu (2017). Bronnen voor indicatorenonderzoek aan hogescholen: Andriessen & Franken (2015), Van der Zwan & Van der Vos (2015), De Poot (2019), Van Drooge et al. (2011), ERiC (2011), Pijlman et al. (2017), Franken et al. (2018), Van Vliet et al. (2020), Moerman et al. (2016), Scholtes et al. (2011), Vereniging Hogescholen (2015b).

gaat, zoals 'Abstracts, editorships, inaugural lectures, designs and prototypes, media appearances, protocol, boekbespreking, congresbijdrage (abstract, poster), ontwerp, database, digitale of visuele producten, software, webpublicatie, overige' (VSNU, 2013b). Of het gaat om een term die staat voor meerdere dingen, zoals 'artefacten' (Van Drooge et al., 2011). Dit soort indicatoren zijn terechtgekomen in meerdere categorieën. Hierdoor werd het aantal indelingen hoger dan het aantal aangetroffen indicatoren.

De geformuleerde categorieën zijn bottom-up ontstaan door dicht bij de originele bronnen te blijven, dit betekent onder andere dat ze niet orthogonaal zijn. Categorieën zoals 'samenwerking', 'consortia' en 'gezamenlijke afstemming van agenda's/roadmaps/programma's' hangen samen en vallen soms ook samen doordat in een consortium, als samenwerkingsverband, een gezamenlijke roadmap wordt gemaakt. Maar er zijn ook situaties waarin we ze kunnen onderscheiden, er kan prima samengewerkt worden zonder een onderliggende roadmap of zonder een formele typering als consortium; en niet alle partijen die een roadmap opstellen zullen noodzakelijkerwijs in een consortium terechtkomen dat de roadmap uitvoert, zoals een gemeente of ministerie. Een ander voorbeeld is de categorie 'Uitnodigingen': de indicatoren hieronder zouden ook ondergebracht kunnen worden bij bijvoorbeeld presentaties (uitnodiging voor een keynote), of lidmaatschappen redactie (uitnodiging gastredacteur), maar in de onderliggende bronnen wordt het belang benadrukt dat het gaat om het 'gevraagd worden' als teken van erkende expertise of reputatie. Om daar recht aan te doen, is 'uitnodigingen' een aparte categorie geworden. Met andere woorden, er is zeker wat op te merken over de categorieën. Een reshuffle is mogelijk, maar het doel hier is geen formele analyse als wel een pragmatische categorisering van de hoeveelheid aangetroffen indicatoren om tot een overzichtelijke analyse te kunnen komen.

De 10 bronnen over indicatoren van universitair onderzoek leverde 352 indicatoren op, de 10 bronnen van indicatoren voor onderzoek aan hogescholen leverde 288 indicatoren op. Uiteindelijk zijn 39 subcategorieën ontstaan waarop de 640 indicatoren zijn ingedeeld, waarbij sommige indicatoren dus in meerdere categorieën zijn opgenomen. Doordat bijgehouden is welke indicator uit welke bron afkomstig is, kan een vergelijking worden gemaakt van indicatoren (subcategorieën) die door de universiteit en door de hogescholen worden genoemd (tabel 3).

Tabel 3. Top 10 van indicatorcategoriegèen met de meeste indicatoren

	Universitair	HBO
1	Publicaties	Onderwijs
2	Financiering 3de geldstroom	Samenwerking en netwerken
3	Octrooien/patenten/licenties	Publicaties
4	Valorisatie	Kwaliteit
5	Fincanciering 2de geldstroom	Onderzoeksteam
6	Samenwerking en netwerken	Producten als demonstrators, prototypes
7	Advisering en consultancy	Financiering 3de geldstroom
8	Producten als digitale media, software, apps	Fincanciering 2de geldstroom
9	Arbeidsmarkt	Presentaties
10	Presentaties	Masterclasses, workshops, co-creatie

Er is veel overlap in de indicatoren categorieën, wat niet al te verwonderlijk is omdat vooral in de beginfase van het opstellen van indicatoren voor het onderzoek aan hogescholen gekeken is naar systematiek en inhoud van het universitaire kwaliteitssysteem (bijvoorbeeld Van der Zwan & Van der Vos, 2015).

Toch zijn er een aantal verschillen. Ten eerste zijn voor het hbo geen indicatoren gevonden die vallen in de categorie 'nevenfuncties', voor het universitair onderzoek zijn hier wel (vijf) expliciete indicatoren gevonden. Dit is wel in lijn met het feit dat bij de universiteiten meer aandacht is voor inbedding van onderzoekers in andere geledingen, zoals ook blijkt uit de aandacht voor indicatoren voor 'dubbele aanstellingen' en 'lidmaatschappen van adviesraden en commissies'.

Ten tweede worden voor het universitair onderzoek geen expliciete indicatoren aangetroffen voor de categorieën '1ste geldstroom' en 'kwaliteit', deze komen wel voor bij het onderzoek aan hogescholen: vier respectievelijk twintig keer. Het is niet geheel duidelijk waarom de indicator van de '1ste geldstroom' voorkomt, mogelijk heeft dit te maken met ontwikkeling van de bekostiging van het onderzoek aan de hogescholen, van in eerste instantie lopend via SKO en later via een lumpsum in de begroting van de hogescholen zelf. De categorie

'kwaliteit' ontbreekt in de universitaire bronnen, er is binnen het SEP wel een assessmentcriterium 'Research quality', maar deze is in de indicatoren geoperationaliseerd in indicatoren zoals 'publicaties in *peer reviewed* tijdschriften'. In de bronnen van het onderzoek aan de hogescholen zijn de kwaliteitseisen veel explicieter verwoord, zoals rondom integriteit, validiteit en ook bijvoorbeeld de inzet van *research datamanagement* (RDM).

Vijf indicatorencategorieën komen in beide top 10's voor: publicaties, financiering 2de geldstroom, financiering 3de geldstroom, samenwerking en netwerken, presentaties. Ranking voor de beide top 10's verschilt echter significant van elkaar, met andere woorden het hbo heeft een eigen profiel wat betreft indicatoren.

Verdere analyse van het eigen profiel van het hbo laat zien dat de hbo-bronnen de volgende indicatoren meer noemen dan universitaire bronnen: onderwijs (1), kwaliteit (4), onderzoeksteam (5) en masterclasses, workshops, cocreatie (10). Daarentegen noemen de hbo-bronnen octrooien, patenten, licenties (13), valorisatie (19), advisering en consultancy (20) en arbeidsmarkt (26) minder.

Kwaliteitscriteria voor indicatoren

Een vierde reflectie is in te leiden met de volgende kritische opmerking over indicatoren: 'de meeste indicatoren zijn niet getest en zijn voor zover bekend niet gebruikt na publicatie van de desbetreffende studie' (Van Drooge et al., 2011, p. 7). Los van de inhoudelijke vraag welke indicator te kiezen, is er dus ook de vraag te stellen of de indicator goed is. De ene indicator is namelijk de ander niet. Indicatoren kunnen we toetsten op hoe goed ze zijn, met andere woorden: er zijn kwaliteitscriteria op te stellen om een keuze te maken uit mogelijk te gebruiken indicatoren. Een voorbeeld is de betrouwbaarheid van de gebruikte data voor een indicator. De literatuur verschilt echter over welke kwaliteitscriteria er voor indicatoren zijn. Tabel 4 geeft een overzicht van een aantal kwaliteitscriteria voor indicatoren vanuit vier verschillende bronnen.

Er zijn overeenkomsten tussen de criteria die de diverse bronnen noemen, zoals de betrouwbaarheid en validiteit van de data en de relevantie en communicatiekracht van de indicator; maar er zijn ook verschillen. Wat de verschillende bronnen niet duidelijk maken is dat de genoemde criteria over verschillende 'dingen' gaan. Criteria zoals beschikbaarheid en continuïteit gaan over de databron waarop een indicator zich beroept. Zo'n databron kan een administratief

Tabel 4. Kwaliteitscriteria voor indicatoren

Bron	RBA-methodiek (Friedman, 2005)	Waardevol (Van Drooge et al., 2011)	PIPA-methodiek (Van Drooge & Spaapen, 2017)	Scholtes et al. (2011)
Kwaliteitscriteria	<ul style="list-style-type: none"> - Communication power: Is de indicator goed communiceerbaar naar alle betrokkenen? Is makkelijk te begrijpen wat de indicator betekent? Ervaart de community de indicator hetzelfde? - Proxy power: Is de indicator een goede vertegenwoordiger (proxy) van andere indicatoren zodat kan worden volstaan met in totaal maximaal 2 of 3 indicatoren? - Data power: Is er voldoende data beschikbaar die valide en betrouwbaar is en ook regelmatig beschikbaar om de voortgang te monitoren? 	<ul style="list-style-type: none"> - Meetbaar: is de indicator meetbaar en is de interpretatie eenduidig? - Beschikbaar en betrouwbaar: zijn benodigde gegevens beschikbaar en betrouwbaar? - Manipuleerbaar: hoe eenvoudig zijn scores te manipuleren? - Valide: meet de indicator dat wat we willen meten? 	<ul style="list-style-type: none"> - function to monitor whether the project will reach the goals set; - need to enable learning during the project, so as to stimulate changes and improvements; - should reflect characteristics of the specific project; - should be realistic to use, and this includes that it should be financially justified to collect evidence; - need to be endorsed by policy and politics. 	<ul style="list-style-type: none"> - Relevance: is the indicator acceptable to the salient stakeholders as relevant indicator? - Validity: does the indicator measure what it is supposed to measure? - Robustness: is the indicator reliable? - Feasibility: is data collection for the indicator possible?

systeem zijn (financiële gegevens), een bibliografisch systeem (publicaties, citaties) of een stakeholderpanel als focusgroep die bevraged wordt over bijvoorbeeld de tevredenheid over en de relevantie van het onderzoek. Criteria zoals betrouwbaarheid, manipuleerbaarheid en validiteit gaan over de data zelf. Een financieel systeem kan continu beschikbaar zijn om informatie te leveren, maar de data zelf kunnen, om wat voor reden dan ook, niet betrouwbaar zijn of gemanipuleerd worden om gunstiger voor de dag te komen.

Op deze manier zijn er dus in ieder geval al drie aspecten of lagen te onderscheiden waarover kwaliteitscriteria zijn op te stellen: indicator, data en databron. De gelaagdheid van de verschillende kwaliteitscriteria voor indicatoren geeft al meer houvast waar op te letten en welke afwegingen te maken (zie tabel 5). Het is verstandig een beperkt aantal kwaliteitscriteria op te stellen om indicatoren aan te toetsen. Waarbij je waakzaam moet zijn om niet juist die indicatoren te verwijderen die erg relevant maar niet volledig robuust zijn of waarvoor de data lastig te verzamelen zijn, want dan blijven al snel alleen de *usual suspects* over van 'geld, mensen, publicaties en patenten' (Brouns, 2016).

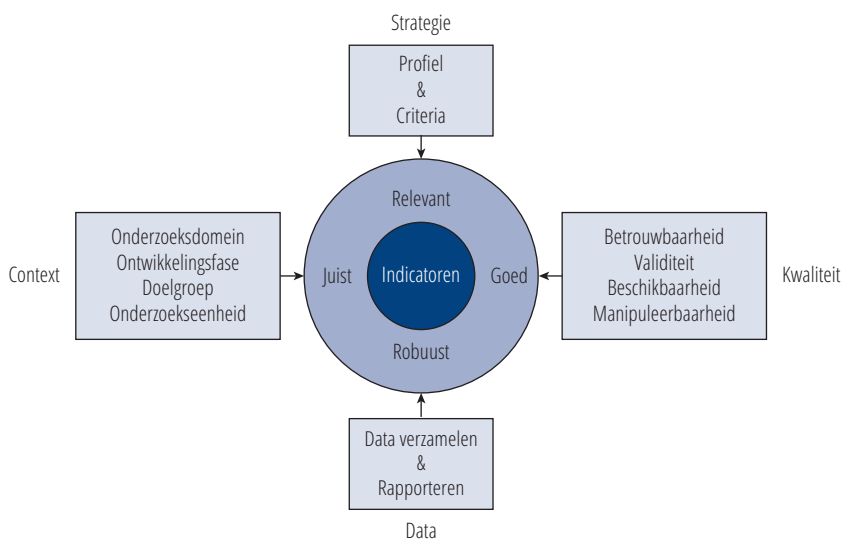
Tabel 5. Kwaliteitscriteria ingedeeld naar verschillende aspecten

Aspect	Kwaliteitscriteria
Indicator	<i>proxy power, communication power, draagvlak</i>
Data	betrouwbaarheid, manipuleerbaarheid, validiteit
Databronnen	beschikbaarheid, continuïteit

Indicatoren redux

Indicatoren krijgen veel aandacht. Enerzijds is dit terecht omdat een indicator gaat over het aantonen wat bereikt is, en omdat indicatoren van invloed zijn op de onderzoekspraktijk – hoewel die relatie complex is (zie De Rijcke et al., 2015). Anderzijds is dit onterecht, omdat er veel afwegingen vooraf (moeten) gaan alvorens een indicator te kiezen, afwegingen van inhoudelijke en praktische aard. Dat verdwijnt snel uit het zicht als het zoveelste lijstje met indicatoren prominent wordt gepresenteerd. Met de bovenstaande reflecties kunnen we die afwegingen systematisch in kaart brengen, waardoor indicatoren meer ingebed worden in een (keuze)proces (zie figuur 6):

1. Indicatoren volgen logisch uit een inhoudelijke context van wat het profiel is van de onderzoekseenheid en dat wat belangrijk wordt gevonden aan het onderzoek, de criteria. Dit maakt dat de gekozen indicatoren *relevante* indicatoren zijn.
2. Indicatoren houden rekening met de context waarin het onderzoek plaatsvindt. Dit maakt dat de *juiste* indicatoren worden gekozen.
3. Indicatoren voldoen aan bepaalde kwaliteitseisen. Dit maakt dat *goede* indicatoren worden gekozen.
4. Indicatoren zijn afhankelijk van praktisch haalbare datavergaring. Dit maakt dat *robuuste* indicatoren worden gekozen.



Figuur 6. Keuzeproces indicatoren

Een tweede herpositionering van indicatoren heeft van doen met de constatering van Godin (2006) dat de rol van *statistics* essentieel is geweest in het borgen van SLM in veel evaluatiekaders. Godin ziet dat nieuwe interactieve modellen om het SLM te vervangen niet succesvol zijn (geweest) doordat ze hun *statistics* nog niet op orde hebben: 'The succes of the linear model suggests how *statistics* often are required to give (long) life to concepts but also how their

absence is a limitation in changing analytical models and frameworks' (2006, p. 660). Een 'reframing' van onderzoeksevaluatie zal dus met een eigen set van statistics moeten komen. Godin & Doré (2005) doen hier al een voorzet voor. Wat ook helpt om te bepalen welke kant het alternatief op moet gaan, is het plotten van bestaande indicatoren op alternatieve modellen. Zo maakt het plotten van bestaande indicatoren in de doorwerkingsmatrix (zie 'Flatland') duidelijk dat de huidige indicatorensets inspanningen meer benadrukken dan effecten.

Spaceland

Ecosysteem

Tot nu toe heb ik voornamelijk gekeken naar de doorwerking van onderzoeksprojecten en -programma's, en heb ik aan het eendimensionale beeld ('Lineland') een tweede dimensie toegevoegd ('Flatland') met het begrip 'doorwerking'. Maar daarmee zijn we er nog niet. Het is evident dat deze projecten en programma's in een context zijn ingebed. Bij de analyse van indicatoren zijn we die context al tegengekomen, bijvoorbeeld in de vorm van het profiel van de hogeschool en de ontwikkelingsfase waarin de onderzoekseenheid zich bevindt. De inbedding van het onderzoek in de hogeschool wordt ook nadrukkelijk bevraagd door het BKO middels randvoorwaarden zoals de organisatie van het onderzoek en de kwaliteitszorg. Inbedding betreft ook een bredere scope, namelijk in de (inter)nationale kennisinfrastructuur en uiteindelijk de samenleving. Die context zal invloed hebben op de doorwerking, in de vorm van stimulansen, barrières of onverschilligheid.

Een aanwijzing dat die inbedding relevant is, is af te lezen aan het programmaportfolio van Regieorgaan SIA. De tweedegeldstroomfinanciering vanuit Regieorgaan SIA is een belangrijke stimulans voor het praktijkgericht onderzoek. In het geheel van regelingen is een tweedeling te maken. Enerzijds zijn er regelingen die zich richten op de uitvoering van het praktijkgericht onderzoek zoals de diverse RAAK-regelingen, themaprogramma's en 'valorisatie'-regelingen zoals Take off en Top-up. Daarnaast is er een tweede categorie van regelingen die de onderzoeksinfrastructuur zelf adresseren, voorbeelden hiervan zijn de hbo-postdoc en L.INT regelingen die gaan over professionalisering, en de SPRONG-regeling, die gaat over het versterken van onderzoeksgroepen. Deze regelingen zijn meer randvoorwaardelijk om onder andere de kwaliteit, slagkracht en continuïteit van het praktijkgericht onderzoek te versterken. Ook in de analyse van indicatoren zien we categorieën van indicatoren die meer gaan over de randvoorwaarden van het onderzoek dan over de feitelijke uitvoering en het effect van het onderzoek. Zo zijn er categorieën van indicatoren over financiering, professionalisering en over de samenstelling van het onderzoeksteam (zie Intermezzo). De vraag die zich opdringt is of er een analyse te maken is van factoren die een relevante rol spelen in de context waarin het praktijkgericht onderzoek is ingebed. Een bruikbare introductie op dit vraagstuk is het concept van productieve interacties.

Productieve interacties

Het concept van productieve interacties is afkomstig uit het Europese project SIAMPI (Social Impact Assessment Methods through Productive Interactions), waarin het concept is geïntroduceerd (Spaapen & Van Drooge, 2011a). Het heeft navolging gehad in verschillende reflecties op het evalueren van onderzoek zoals onder andere in LERU (Van den Akker & Spaapen, 2017; Keustermans et al., 2018) en ERIC (2011), en heeft inmiddels een prominente plek verworven in de impactstrategie van NWO (2019, 2020) en ZonMw. Hoewel het concept niet zonder problemen is (zie Van Vliet, 2021), is het een bruikbaar concept voor de verdieping van de discussie rondom doorwerking.

Productieve interacties gaan over contacten tussen onderzoekers en andere stakeholders waarin kennis wordt gegenereerd en gebruikt. Een belangrijke typering van deze productieve interacties is dat ze de vorm kunnen aannemen van directe/persoonlijke, indirecte interacties of financiële interacties (Spaapen & Van Drooge, 2011a/b). Dit is op twee manieren interessant:

- 1) de indeling direct/indirect gaat over het 'waar' je naar kijkt, en kunnen we vergelijken met de eerder gemaakte bottom-up ontstane indeling van producten, netwerken en kennisinteracties (zie 'Flatland');
- 2) de stelling dat er een vorm van productieve interacties is te typeren als 'financieel', legt precies het punt bloot dat ik hier probeer te maken over contextfactoren.

De directe/persoonlijke interacties betreffen zowel bijeenkomsten als complexe samenwerkingsverbanden en alles wat daartussen ligt. Indirecte interacties zijn interacties die verlopen via een medium c.q. een artefact zoals een publicatie, zorgprotocol of adviesrapport. De categorieën 'direct' en 'indirect' kunnen we zien als een typering van interactiesoorten, maar zijn wel wat grofmazig. Voor de categorie 'direct' kunnen we een verdere onderverdeling maken die eerder is aangedragen (zie *Exploratief onderzoek #1*) namelijk die tussen kennisinteracties en netwerken. Kennisinteracties zijn die interacties waar er direct contact is tussen bijvoorbeeld een onderzoeker en een doelgroep, zoals bij een training of in een cocreatiesessie. Andere voorbeelden zijn workshops, consultancygesprekken, (gast)colleges, hackatons, seminars, expertsessies, et cetera. Bij kennisinteracties gaat het om kennis ontwikkelen, delen en toepassen in direct contact tussen betrokkenen. Waar kennisinteracties een eenmalige of ad-hockarakter kunnen hebben, daar legt de aandacht voor netwerken de nadruk op de deelname in formele en ook informele netwerken.

Er zijn drie redenen om dit als een aparte (sub)categorie te zien: ten eerste komt het tegemoet aan de verknoping van het praktijkgerichte onderzoek met

het werkveld en de samenleving, oftewel de ‘praktijkverbondenheid’ (Brouns, 2016) en het is goed om dit dan zichtbaar te maken.

Ten tweede zal door het inter- en transdisciplinair werken en het missiegedreven innovatiebeleid een onderzoeker zich steeds meer moeten verbinden met netwerken om onderzoek uit te voeren en een vervolg te geven.

Ten derde maakt het expliciet dat onderzoek gezien wordt als ‘to take place in a network in which different partners with diverse expertise and knowledge collaborate on the basis of a joint agenda. Paramount are successful interactions between stakeholders (each with their own expertise, knowledge and insight) and the formulation of common goals and joint achievement of results. It requires insight into the relevant networks of stakeholders and the varied interactions and communications that are part of those networks, and this insight should as much as possible start right from the design of research programmes’ (Keustermans et al., 2018, p. 11).

De categorie ‘indirect’ gaat over interacties middels een specifiek medium, waardoor het niet noodzakelijk is dat de afzender zelf aanwezig is. Een boek kun je lezen zonder dat de auteur aanwezig is. Voor de categorie ‘indirect’ kun je de onderverdeling volgen die in ERiC (2011) wordt voorgesteld, namelijk tussen teksten en artefacten. Om beter aan te sluiten bij de staande praktijk lijkt het verstandiger te spreken over publicaties en andersoortige producten. Publicaties in de vorm van wetenschappelijke publicaties, vakpublicaties en populaire publicaties zijn een gebruikelijke manier om te rapporteren over de opbrengsten van onderzoek, zoals ook vastgelegd in evaluatieprotocollen (SEP en BKO).

Het is onmiskenbaar dat er naast publicaties nog vele andersoortige dragers zijn die de interactie zoeken met doelgroepen en gebruikers. Van Vliet et al. (2020) treffen in evaluatierapporten van praktijkgericht onderzoek voorbeelden aan als: websites, apps, (simulatie)modellen, tentoonstellingen, ontwerpen, software/algorithmes, toolkits, methodieken, spellen, interactieve dashboards, kennisclips, et cetera. Door hier oog voor te hebben kunnen we tegemoetkomen aan de vaak aangehaalde typering van praktijkgericht onderzoek dat het een grote diversiteit aan opbrengsten heeft.

Aldus zijn er twee hoofdcategorieën van interacties, direct en indirect, die vóórkomen in een ecosysteem waarbinnen bewust en onbewust invloed wordt uitgeoefend op het wel of niet optreden van deze interacties. Deze hoofd categorieën kunnen we verder onderverdelen in kennisinteracties en netwerken voor de directe interacties, en publicaties en andersoortige producten voor de indirecte interacties.

De derde categorie ‘financieel’ gaat over financiële investering en andere manieren van materiële ondersteuning. Hierover wordt gezegd dat: ‘This type

of interactions cannot take place in isolation: it will be accompanied by direct or indirect interactions, or both´ (Spaapen & Van Drooge, 2011a, p. 14). De typering ‘cannot take place in isolation’ geeft al aan dat dit een apart soort categorie is, want deze uitspraak gaat verder dan de realistische opmerking dat de drie soorten interacties vaak gezamenlijk voorkomen. Bij deze derde categorie gaat het namelijk om *support mechanisms* (ibid, p.4). Daarvan zijn er nog wel meer te formuleren dan alleen financiële ondersteuning, bijvoorbeeld de uitwisseling van mensen tussen deelnemende organisaties in een onderzoek, of de agenderende rol van instituten in de kennisinfrastructuur die innovatie en onderzoeken aanjagen, of in ieder geval selecteren.

Wat hieraan ten grondslag ligt is de noodzaak een onderscheid te maken tussen de interactiesoorten zelf en de context, het ecosysteem, waarin ze (kunnen) optreden. Die context bepaalt de parameters waarin interacties worden gestimuleerd, afgeremd, verboden of beloond. Soms direct door financiering van specifieke samenwerking of bijvoorbeeld bekostiging van uitwisselingsprogramma’s van onderzoekers. Soms indirect doordat technologie het mogelijk maakt op afstand samen te werken, of doordat wetgeving het moeilijker maakt data uit te wisselen in verband met privacy, of doordat het beleid van kennisinstellingen het mogelijk maakt meer resultaten te delen (open access).

Elementen van het ecosysteem

In het gedachtegoed van de productieve interacties is ‘financieel’ in feite een randvoorwaardelijk mechanisme voor onderzoek. In de SPRONG-regeling van Regieorgaan SIA gaat het om de versterking van onderzoeksgroepen als borging van het praktijkgericht onderzoek, sterke onderzoeksgroepen als randvoorwaarde voor de continuïteit en kracht van het onderzoek. Daarmee zijn al twee variabelen genoemd die een rol spelen in de randvoorwaarden van de doorwerking in onderzoeksprojecten. Het is niet moeilijk hierop verder te brainstormen en lijstjes op te stellen van wat nodig is voor praktijkgericht onderzoek, zoals: geld, tijd en ruimte (Van Staa & Mobach, 2020).

Er is ook systematischer onderzoek gedaan naar randvoorwaardelijke mechanismen (actoren, factoren, regels) voor onderzoek en innovatie. Veelal gaat het dan over ecosystemen of innovatiesystemen (Godin, 2010), zoals missiegedreven innovatiesystemen (MIS), losjes te typeren als een, vaak geografisch of thematisch, af te bakenen community van onderling afhankelijke partijen in relatie tot situationele factoren, die gezamenlijk een complexe maar relatief coherente en verbonden eenheid vormen. Deze ecosystemen bestaan

‘bij de gratie van interacties en netwerken waarin complementaire kennis, vaardigheden en resources zodanig worden gecombineerd dat ze leiden tot nieuwe toepassingen’ (Vankan et al., 2020, p. 23). Deze literatuur is uitgebreid en breed, zeker daar waar het erom gaat determinanten te vinden voor bijvoorbeeld groei in ecosystemen of de voorwaarden voor een *minimum viable ecosystem* (Van Bree et al., 2020). Ook doemen fundamentele vragen op zoals de afgrenzing van systemen. Ik kies er hier voor om twee studies toe te lichten die voor nu voldoende inzicht geven in waar we aan moeten denken bij randvoorwaardelijke mechanismen in de context van onderzoek.

De eerste studie is het rapport van Dialogic (Vankan et al., 2020) dat in opdracht van de overheid is geschreven als achtergrondstudie bij de kabinetsstrategie van ‘Versterken van onderzoeks- en innovatie-ecosystemen’ (Kabinet, 2019). In het rapport beschrijft Vankan met collega’s acht uitdagingen die betrekking hebben op specifieke elementen in onderzoeks- en innovatie-ecosystemen: 1) investeringen in onderzoeks- en testfaciliteiten, zodat de benodigde infrastructures voor onderzoek voor langere tijd behouden blijven; 2) toptalent, ofwel kwaliteit en kwantiteit van geschikt *human capital*; 3) financiering voor start-ups en scale-ups om kansrijke innovatieve jonge bedrijven te behouden; 4) marktcreatie en het betrekken van gebruikers bij onderzoek en innovatie; 5) het faciliteren van verbindingen tussen ecosystemen; 6) vaardigheden en absorptiecapaciteit in het mkb zodat hoogwaardige kennis beter wordt toegepast; 7) versterken kennisoverdracht en valorisatieproces voor meer impact; 8) eerder in het proces aandacht besteden aan wet- en regelgeving.

Hierbij moet ik opmerken dat de referentielijst van 56 onderzoeks- en innovatiesystemen een behoorlijke oververtegenwoordiging laten zien van technologische en industriële clusters (biotech, nieuwe materialen, chemie, energie, water, AI, logistiek, voedseltechnologie, et cetera), met slechts enkele ecosystemen voor zorg, een enkele rond onderwijs en als laatste voor design, gaming en cultureel erfgoed. Daarmee lijken de geconstateerde uitdagingen een bias te hebben richting technologie/industrie en economie. Dat laatste blijkt ook uit bijvoorbeeld het wel noemen van mkb als ‘afnemer’ maar niet publieke organisaties en hun absorptiecapaciteit, en het nadrukkelijk zien van bedrijven c.q. start-ups als (economische) motor. Systeemuitdagingen in relatie tot sociaal-maatschappelijke transitie lopen hiermee de kans ondervertegenwoordigd te zijn, een herkenbaar patroon (zie ‘Lineland’)

Behalve deze uitdagingen presenteert het rapport ook een conceptueel kader innovatie- en onderzoeksecosystemen met een set van elementen voor een beter begrip. Dit is gedaan langs de lijn van een *logic* framework (input – activiteiten – output – outcome – impact). Het leidt tot, eerlijk gezegd, een

ratjetoe aan concepten waarin juist de systeemrelaties naar de achtergrond verdwijnen. Zo worden sommige systeemvariabelen vanuit een transitieperspectief wel opgenomen (richting geven, *directionality failure*) en andere niet (*reflexivity failure*, onvermogen te reflecteren op waarom een transitie uitblijft); worden wel sociale normen als randvoorwaarden genoemd maar niet onderzoekscultuur; is human capital geen input maar wel een activiteit, et cetera.

De tweede studie betreft een analyse van zogeheten *entrepreneurial ecosystems* door Stam & Van de Ven (2018). Deze studie is mede gebruikt in een notitie die is geschreven voor Regieorgaan SIA over een nieuw impactkader (Van den Toren & Van Vliet, 2021). Bij deze *entrepreneurial ecosystems* ligt de nadruk op het komen tot coördinatie van actoren en factoren om productief ondernemerschap mogelijk te maken. Dit vanuit de stelling dat het succes van een individuele ondernemer afhankelijk is van de context waarin hij of zij opereert: sociaal, economisch, geografisch, et cetera. In de studie worden eerdere voorstellen in academische studies en allerlei *long laundry lists of relevant factors* (p. 3) aangaande bepalende factoren in *entrepreneurial ecosystems* samengebracht in tien elementen: 1) formele instituties als vertegenwoordigers van regels die gelden binnen het ecosysteem (wet- en regelgeving) en daarmee allerlei handelingen mogelijk maken (en andere niet), 2) cultuur, zijnde normen en waarden, van wat gewaardeerd wordt en wat niet, 3) netwerken als de verbandingen tussen actoren en de sociale context waarin actoren zich bewegen en kennis, arbeid en kapitaal worden uitgewisseld, 4) fysieke infrastructuur als bepalend voor de ontmoeting van actoren (*proximity*), 5) financiering als investeringen in (nieuwe) activiteiten, 6) leiderschap, de aanwezigheid van zichtbare leiders en mentoren die richting geven en als rolmodel dienen, 7) talent, het human capital als vaardigheden, kennis en ervaring die individuen bezitten, 8) kennis, investeringen in kenniscreatie, 9) vraag, de vraag die er is naar goederen en diensten, 10) intermediaire diensten die zorgen voor aanvullende en gespecialiseerde diensten die processen vergemakkelijken, bijvoorbeeld *time-to-market* bekorten.

Stam & Van de Ven wisten deze tien elementen te operationaliseren en vervolgens ook de relatie met productief ondernemerschap te onderzoeken, als output van het ecosysteem. Als eenheid van ecosystemen namen ze hierbij de provincies. Ze troffen grote variatie aan op de waarden van de tien elementen per provincie. Ook zagen ze sterke correlaties tussen elementen, zoals tussen 'kennis' en 'leiderschap', en tussen 'fysieke infrastructuur' en 'vraag'. Uiteindelijk poneerden ze een *entrepreneurial ecosystem index*, een enkel getal op basis van scores op de tien elementen, waarmee ze regio's met elkaar kunnen vergelijken. Opmerkelijk is dat we in het rapport van Dialogic juist argumenten

vinden voor zeer grote terughoudendheid aangaande dergelijke vergelijkingen van contextueel heel verschillende ecosystemen (Vankan et al., 2020, p. 9-10).

Beide studies tonen flinke onderlinge overeenkomsten in de elementen die ze noemen, en overigens ook met de studie van Van Bree et al. (2020) die ik hier verder niet bespreek. Er zijn wel verschillen in interpretatie, gegeven de context (innovatie bij Vankan et al., ondernemerschap bij Stam & Van de Ven). Zo is 'fysieke infrastructuur' bij Stam & Van de Ven geoperationaliseerd in onder andere het aantal wegen en treinverbindingen, terwijl het in het rapport van Dialogic gaat om onderzoeksfaciliteiten zoals labs en experimenteeruimtes. In tabel 6 verwoord ik de ecosysteemelementen die Vankan et al. (2020) en Stam & Van de Ven (2018) noemen in de context van praktijkgericht onderzoek en benoem ik enkele voorbeelden per element.

Een dergelijke systematische opsomming geeft meteen handelingsperspectieven voor het praktijkgericht onderzoek c.q. een lectoraat. Van den Toren & Van Vliet (2021) noemen er een aantal, zoals: het onderzoek zo organiseren dat netwerken zich ontwikkelen en uitbreiden, met speciale aandacht voor intermediaire organisaties als brancheverenigingen en beroepsverenigingen die toegang tot stakeholders kunnen versnellen; onderzoek programmeren in afstemming met het leiderschap in regio's en sectoren; studenten onderzoekend vermogen meegeven als onderdeel van hun talent, zorgen voor een infrastructuur waarin je samenwerking stimuleert zoals in labs en hybride leeromgevingen.

Niet alleen handelingsperspectieven worden evident, maar ook sturingsmogelijkheden. Binnen bijvoorbeeld de KIA Maatschappelijk Verdienvermogen wordt nadrukkelijk gekeken naar interventies in (regionale) ecosystemen om te komen tot meer impact. Zo staat in de tweede call in 2022 multi-stakeholder inter- en transdisciplinaire samenwerking in ecosystemen centraal. Ook Regieorgaan SIA heeft zich met het nieuwe SPRONG-programmakader gericht op de randvoorwaarden voor onderzoek. Bij de beschrijving van het programmakader (Regieorgaan SIA, 2020) komen nagenoeg alle onderscheiden elementen van het ecosysteem aan bod, behalve 'financiering' maar dat kan komen doordat SPRONG zelf een financieringsinstrument is. De elementen worden alleen in wisselende abstractie/concreetheid beschreven en niet gepresenteerd als elementen van een samenhangend ecosysteem, wat – gegeven de bovenstaande analyse – een gemiste kans is. Het gaat bovendien erg om de eigen positionering van het praktijkgericht onderzoek in het ecosysteem, vandaar dat zaken als profilering, kwaliteit en het onderzoeksteam centraal staan in wat als 'kracht' wordt gezien. Criteria zoals netwerkvorming en professionalisering (talent/human capital) verwijzen dan weer meer naar aansluiting bij ecosysteemelementen.

Tabel 6. Voorstel ecosysteem elementen voor praktijkgericht onderzoek

Element	Definitie	Voorbeelden in relatie tot <i>reo</i>	Vankan et al. (2020)	Stam & Van de Ven (2018)
Netwerken	Formele en informele verbanden tussen (groepen van) actoren	Leergemeenschappen, lecturenplatforms, economic boards, consortia	Faciliteren van verbindingen (netwerkvorming)	Netwerken
Leiderschap	Actoren die (zichtbaar) richting geven	'Ambassadeurs' van KIAs, leden van CvB, boegbeelden	(Leiderschap)	Leiderschap
Financiering	Investerings in onderzoek en de randvoorwaarden voor onderzoek	Regioorgaan SIA subsidieregelingen, ZonMw	Investerings, financiering (kapitaal)	Financiering
Talent/Human Capital	Vaardigheden, kennis en ervaring die individuen bezitten	Opleidingen, Leven Lang Ontwikkelen, Professional Doctorate ontwikkeling	Toptalent (arbeid en talent)	Talent
Nieuwe kennis	Ontwikkelen, delen en implementeren van kennis	Praktijkgericht onderzoek	Absorptiecapaciteit, kennisoverdracht, valorisatie (kennis)	Kennis
Intermediaire diensten	Diensten die bemiddelen in het bereiken van een doel	Branche-organisaties, PublinoVA		Intermediaire diensten

Element	Definitie	Voorbeelden in relatie tot <i>peo</i>	Vankan et al. (2020)	Stam & Van de Ven (2018)
Formele instituten / wet- en regelgeving	Wet- en regelgeving en haar institutionalisering	De hogeschool en haar (wettelijke) taakstelling, waaronder onderzoek	Wet- en regelgeving	Formele instituten
Cultuur	Normen en waarden, wat wordt gewaardeerd en wat niet	Gedragscode onderzoek, prijzen voor onderzoek		Cultuur
Fysieke/digitale infrastructuur	Voorzieningen (fysiek en digitaal) die nodig zijn om onderzoek (beter/snel)er uit te kunnen voeren	Labs, Virtual Research Environments	Onderzoeks- en test-faciliteiten (locaties en gebouwen)	Fysieke infrastructuur (wegen/verbindingen)
Vraag (articulatie)	Gearticuleerde vragen om verandering	(Regionale) agenda's, klankbordgroepen, quadrupel helix	Marktcreatie en betrekken van gebruikers bij onderzoek en innovatie	Vraag om producten en diensten

Met de ecosysteemelementen wordt een aanvullend perspectief geboden op het perspectief van het individuele onderzoeksproject en -programma wanneer gesproken wordt over doorwerking. Met de ecosysteemelementen expliciet gemaakt zien we ook dat tot nu toe gepubliceerde evaluatierapporten over praktijkgericht onderzoek veelal een nogal 'smalle' insteek hebben: omvang (geld, mensen, netwerk) en profilering (thema's, aansluiting bij agenda's) overheersen in analyses van het praktijkgericht onderzoek (De Jonge, 2016; Cloosterman et al., 2019). Meer oog voor de genoemde randvoorwaarden van het praktijkgericht onderzoek zou een sterkere analyse opleveren en ook meer handelingsperspectieven openen.

Tot slot is nog de opmerking noodzakelijk dat het benoemen van elementen in een ecosysteem niet het eindantwoord is. Het gaat er ook om vanuit welk perspectief je naar het systeem kijkt. Ik licht dit verder toe in de volgende sectie door verschillende manieren te beschrijven waarop naar het 'falen' van systemen is gekeken (zie 'Pointland'). Voorts kun je systemen niet alleen als object van analyse zien, maar kun je ze ook 'reactant' zien door bijvoorbeeld te vragen hoe een systeem omgaat met complexe problemen en hoe de stabiliteit van een systeem hier een rol in speelt (Arkesteijn, Van Mierlo & Leeuwis, 2015), wat ook een specifieke methodiek meebrengt (reflexieve monitoring; Van Mierlo et al., 2020). Een aanpak die overigens in de uitwerking heel herkenbaar is in relatie tot andere methodieken die oog hebben voor ecosysteemelementen zoals netwerken en de machtsrelaties daarbinnen, en het constant ter discussie stellen van aannames c.q. ingesleten patronen (bijvoorbeeld *Theory of Change*).

Exploratief onderzoek #4: het ecosysteem van krachtige SPRONG-groepen

Het SPRONG-programma van Regieorgaan SIA gaat over het opbouwen en (kwalitatief) versterken van onderzoeksgroepen, aangeduid als krachtige SPRONG-groepen. Het regieorgaan ziet praktijkgericht onderzoek als antwoord op de groeiende behoefte aan praktijkkennis en innovaties voor actuele maatschappelijke uitdagingen. Door samenwerking in meerjarige, strategische samenwerkingsverbanden met relevante partners én hoogwaardig onderzoek van onderzoeksgroepen vindt er een verdere uitbouw plaats van de positie van praktijkgericht onderzoek, professionalisering van het praktijkgericht onderzoek én meer doorwerking naar onderwijs en praktijk.

Hierbij richt het programma zich niet, of niet direct, op de uitvoering van praktijkgericht onderzoek an sich, maar op de randvoorwaarden om tot goed

praktijkgericht onderzoek te komen. En hoewel het programmakader zelf (Regieorgaan SIA, 2020) niet spreekt over ecosystemen, zijn de randvoorwaarden waarop de regeling zich richt te lezen als zodanig (Van Vliet, 2022).

Een analyse is uitgevoerd op de zeventien goedgekeurde SPRONG-aanvragen van ronde 1 (2021). Hierbij is gekeken naar welke elementen van een ecosysteem worden genoemd in de analyse van het domein waar de aanvraag zich op richt én welke ecosysteemelementen onderdeel zijn van de uitvoering van de aanvraag. Voor deze analyse zijn de tien elementen gebruikt zoals afgeleid door Van den Toren & Van Vliet (2021; zoals ook genoemd in tabel 6), met toevoeging van 'onderzoeksteam' als element, gegeven het primaire belang hiervan binnen de SPRONG-regeling.

In de aanvragen wordt bij de analyse van het betreffende domein maar een zeer beperkt aantal elementen van het ecosysteem genoemd. Bijna de helft (48%, acht aanvragen) noemt maar één of twee van de elf ecosysteemelementen bij de analyse van het domein waarop de aanvraag zich richt. Er is één aanvraag die zeven elementen noemt, de overige aanvragen noemen er tussen de drie en zes. Het element 'nieuwe kennis' wordt door alle aanvragen genoemd, vier aanvragen noemen zelfs alléén dit element. Dit vormt een contrast met de activiteiten die de aanvragen ontplooiën, deze bewegen zich op veel meer elementen (gemiddeld zo'n zeven elementen, binnen een beperkte range van vijf tot acht). Alle aanvragen gaan activiteiten ontplooiën op vier elementen van het ecosysteem: netwerken, talent/human capital, nieuwe kennis en vraag(articulatie). Er zijn vijf elementen die een groot verschil laten zien tussen waar wel activiteiten op worden ontplooid maar die niet genoemd worden in de analyse van het ecosysteem: fysieke infrastructuur, talent/human capital, netwerken, vraag(articulatie) en onderzoeksteam.

Een conclusie is dat 'nieuwe kennis' de lens is waardoor onderzoekers naar het domein kijken. Dat kan tot tunnelvisie leiden waardoor ze andere uitdagingen niet zien en ze er dan geen rekening mee houden. Verder komen veel activiteiten in de aanvragen logisch en sympathiek over maar wordt er geen rationale gegeven vanuit het perspectief van het ecosysteem waarom dit van belang is. Dat zou er wel moeten zijn: interventies plegen zonder een onderbouwde rationale van 'waarom' lijkt niet wenselijk, zeker niet daar waar de doelstelling is het ecosysteem c.q. de kennisinfrastructuur te verbeteren.

Zwitsers zakmes

Het benoemen van randvoorwaarden voor onderzoek brengt een extra laag, een extra dimensie, aan in de discussie over doorwerking van onderzoek. Niet alleen is doorwerking iets van onderzoeksprojecten, het is ook afhankelijk van de constellatie van randvoorwaarden die dit wel of niet faciliteren. Het wordt nog extra ingewikkeld als we die randvoorwaarden ook in een ontwikkelingsperspectief plaatsen. De SPRONG-regeling bijvoorbeeld vraagt ook een ontwikkeling op het gebied van impact, dus het nadenken over en het ontwikkelen van een impactkader dat bijdraagt aan sterke onderzoeksgroepen. Hier gaat het dus niet direct om de impact van het onderzoeksproject, maar om dit op metaniveau beter mogelijk te maken, zichtbaar te maken, et cetera. Net zoals je een kwaliteitskader ontwikkelt waartoe ieder individueel onderzoeksproject zich moet verhouden, zijn voordeel ermee kan doen, zo kun je ook een impactkader ontwikkelen waartoe ieder individueel onderzoeksproject zich moet verhouden, zijn voordeel mee kan doen. Niet iedere SPRONG-aanvrager heeft deze dubbele laag gezien (Van Vliet, 2022).

Dit heeft ook zijn weerslag op de eerder ontwikkelde ideeën. Voor de indicatoren heb ik al opgemerkt dat je ze moet uitsplitsten naar het perspectief waarvoor ze meer geschikt zijn, of in ieder geval dat je ze anders moet verwoorden voor onderzoeksprojecten dan wel infrastructuur c.q. ecosystemen. Voor beide kun je financiële indicatoren hanteren: in het eerste geval kijk je naar omvang (volume) en soort financiële bron (aard) die gebruikt is in het onderzoeksproject, in het tweede geval kijk je naar de omvang en de financiële instrumenten die voorhanden zijn in het ecosysteem en hoe ze op elkaar inwerken.

De doorwerkingsmatrix is ontwikkeld voor onderzoeksprojecten. De onderliggende twee vragen waarop de matrix gebaseerd is, zijn ook van toepassing op het niveau van het ecosysteem: waar kijk je naar? En hoe kijk je ernaar? Voor het 'waar' zou je de eerdere invulling van productieve interacties kunnen vervangen door de tien onderscheiden ecosysteemelementen. Voor het 'hoe' blijven dezelfde vragen relevant op het niveau van het ecosysteem: welke inspanningen worden er gepleegd om het ecosysteem aan te passen en wat is het effect ervan. Of daarmee de doorwerkingsmatrix een multidimensionaal Zwitsers zakmes is geworden zal nog verder uitgewerkt en onderbouwd moeten worden.

Pointland

5A-model

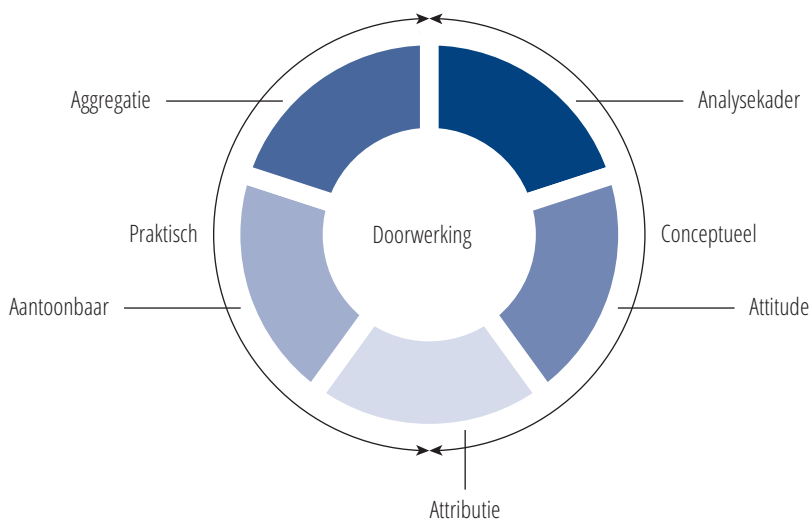
Wat is het punt dat ik wil maken? Het eendimensionale model van het SLM is een beperkend model doordat het de nadruk legt op het eindproduct als uitkomst van een stapsgewijs proces met impact als eindstation. Onderzoek als een estafette met een stokje (kennis) dat je doorgeeft en dat bij een slechte wissel valt. In het kielzog hiervan zijn instrumenten ontstaan (kennistransfer, valorisatie) en problemen (attributie) die discussies over impact domineren. Door de borging van het SLM in allerlei evaluatiemodellen, indicatorensets en subsidieprogramma's is het hardnekkig aanwezig en ook sturend in de praktijk van onderzoek doen. Aan de hand van het begrip 'doorwerking' is een ander perspectief geopend dat vooral zichtbaar wordt bij een focus op de effecten voor en tijdens het onderzoeksproces. Het praktijkgerichte onderzoek is een praktijk waarin dit met name van belang is: dat komt door de verbondenheid met de beroepspraktijk, de gerichtheid op interventies in die praktijk en het leerproces dat hierbij inherent is. Onderzoek als contactsport⁸, een worsteling om gezamenlijk tot veranderingen te komen. Doorwerking is op verschillende manieren verkend, bijvoorbeeld in de discussie over indicatoren. Er zijn ook mogelijke (deel)oplossingen geschetst zoals de doorwerkingsmatrix.

Hoe nu verder? In ieder geval wordt de opdracht nog eens krachtig neergezet in een gezamenlijk statement van het Ministerie van OCW, Regieorgaan SIA en de Vereniging Hogescholen, met als stip op de horizon: 'Hogescholen zijn in staat om structureel de impact van hun onderzoek kwalitatief en kwantitatief te duiden' (ibid., p. 4). Dat kunnen we nog niet. De conclusie dat er nog geen raamwerk is om (maatschappelijke) impact aan te tonen, is op diverse plekken te vinden (onder andere Godin & Doré, 2005; Bornmann, 2013; Joly et al., 2015). Ook in een specifieke studie over het onderzoek aan hogescholen is de conclusie 'Given the short history of research conducted by UASs, it is not surprising that there is no recognised approach for evaluating the impact created by the research of these institutions' (Coombs & Meijer, 2021, p. 7). Daarbij worden wel kenmerken genoemd die dat raamwerk zou moeten hebben zoals performatief, coproductie en realistisch. Die staan zeker niet haaks op wat hier betoogd is, integendeel. Complementair aan deze bottom-upaanpak (welke elementen uit bestaande raamwerken zijn bruikbaar?), wil ik afsluiten met het

8 Ontleend aan Kogan & Henkel's studie uit 1983, zoals aangehaald in Raftery et al. (2016).

uitgangspunt dat een antwoord begint met het stellen van de juiste vragen. Ook als agenda voor de komende jaren van het onderzoek in het lectoraat.

Om tot een verdere uitwerking en integratie te komen van de verschillende gegeven aanzetten is het nodig de onderliggende vraagstukken scherp te formuleren. Terugkijkend op de gevoerde discussies zie ik er vijf. Als ezelsbruggetje heb ik ze allemaal met een 'A' laten beginnen: analysekader, attitude, attributie, aantoonbaar en aggregatie. In figuur 7 zijn ze weergegeven, niet om een proces te suggereren dat je met de klok mee zou moeten doorlopen, als wel een losse ordening. 'Analysekader' en 'attitude' betreffen meer conceptuele beslissingen en 'aantoonbaar' en 'aggregatie' meer praktische beslissingen. 'Attributie' lijkt een soort scharnierfunctie te hebben tussen het conceptuele en het praktische. Ik zal deze vijf A's kort toelichten.



Figuur 7. Het 5A-model

Analysekader

'Analysekader' betreft het vraagstuk vanuit welk interpretatiekader je naar doorwerking kijkt, wat de consequenties hiervan zijn voor de selectie van een methodiek waarmee de doorwerking wordt aangetoond, en welke inzichten het gekozen analyseperspectief insluit én uitsluit.

Uitspraken doen over de relatie tussen onderzoek en doorwerking steunt op aannames over onder andere wat kennis is; het doel van onderzoek; de manier waarop kennis tot stand komt en je kwaliteit kunt bereiken; de mechanismen die leiden tot doorwerking en wat als bewijs dient dat die doorwerking heeft plaatsgevonden (Greenhalgh et al., 2016). Die aannames sturen ook de methodieken die je inzet, de indicatoren die je kiest en hoe je bewijs verzamelt. Dit is zichtbaar te maken door een aantal bestaande perspectieven te schetsen die, vaak impliciet, aangetroffen kunnen worden in methodieken. Ik zal er hier ter illustratie vier noemen (zie Greenhalgh et al., 2016; Raftery et al., 2016).

1. Een positivistisch perspectief ziet kennis bestaan uit feiten die onafhankelijk van de onderzoeker bestaan. Onderzoek is erop uit uitspraken te doen over wetmatigheden van deze feiten door toetsing en modelvorming. Door specifieke onderzoeksmethoden en -designs te gebruiken kan de kwaliteit van het onderzoek worden gewaarborgd. Feiten, wetmatigheden en modellen vinden direct hun weg naar de praktijk en naar beleid. Een logisch model kan in kaart brengen hoe de onderzoeksresultaten, als overdraagbare feiten over wat werkt, hun weg hebben gevonden in de maatschappij.
2. Een constructivistisch perspectief ziet kennis als een interpretatie van een bepaald fenomeen in de historische, culturele en sociale context. Onderzoek is duiden en betekenisgeving, niet in algemeenheden maar in voorkomens, door in de betreffende context het onderzoek te doen (sociale interventie). Kwaliteit bereik je door te reflecteren over meerdere interpretaties, en in dialoog en debat. Uitkomsten van onderzoek bereiken indirect de praktijk en beleidsmakers, via interacties en bewustwording. Impact is onvoorspelbaar, beter is om zich te richten op activiteiten en interacties met diegenen die ertoe doen.
3. Een performatief perspectief stelt dat kennis ontstaat in de praktijk van actornetwerken van mensen en technologie. Het is de dynamiek van deze actornetwerken die onderzoek in kaart moet brengen door naar de veranderingen in het netwerk en veranderingen daarvan in de tijd te kijken. Die veranderingen komen door formele en informele interacties tot stand. Kwaliteit zit in de rijkheid van de beschrijving van actornetwerken. Impact wordt bereikt door actoren die andere actoren tot nieuwe configuraties van actornetwerken bewegen. De evaluatie van impact moet zich richten op hoe deze nieuwe configuraties tot stand zijn gekomen (translations) en stabiel worden in een netwerk.
4. Een realistisch perspectief stelt dat de waarde van onderzoek ligt in zoverre zij passende ideeën en kansen (mechanismen) introduceren in een specifieke context. Realistische benaderingen gaan ervan uit dat niets overal voor

iedereen werkt: context maakt een groot verschil. Een realistische evaluatie vraagt niet: 'Wat werkt?' Zij vraagt: 'Hoe of waarom werkt dit, voor wie, in welke omstandigheden?' Realistische evaluatie richt zich daarom op hoe mechanismen werken in een bepaalde context en vraagt de mensen die dat zouden kunnen weten om bewijs daarvoor aan te dragen. Impact is de interactie tussen betrokken stakeholders en de middelen die beschikbaar zijn voor de toepassing van de onderzoeksopbrengsten in een bepaalde context, zogeheten 'contextmechanismen – outcome – impact'-configuraties.

Deze perspectieven zijn terug te vinden in verschillende methodieken. Zo typeren Coombs & Meijer (2021) methoden binnen het *Payback framework* als positivistisch, de methode van SIAMPI als (deels) constructivistisch, de methode van *contribution mapping* als performatief en de methode van ASI-RPA als realistisch. Kortom, een methodiek is niet 'vrij' maar geworteld in een bepaald filosofisch perspectief over kennis, kennisproductie, de rol van kennis c.q. onderzoek in de samenleving, de functie van onderzoek, et cetera. En een methodiek brengt op haar beurt weer bepaalde voorkeuren met zich hoe bewijs te verzamelen over impact. Zo geven Pedersen, Grønvald & Hvidtfeldt (2020) en Smith & Hessels (2021) overzichten van hoe evaluatiemethodieken voor impact samenhangen met manieren van data verzamelen over impact. Zo wordt duidelijk dat (impliciete) aannames over onderzoek en impact doorwerken in de gekozen methodieken tot en met de manieren van data verzamelen.

Waar deze perspectieven nog op het onderzoek(sproces) zelf gericht zijn, heb ik ook aandacht gevraagd voor het ecosysteem waarin het onderzoek opereert. Hier speelt ook de vraag met welk (analytisch) perspectief naar verandering überhaupt wordt gekeken en naar systeemontwikkeling in het bijzonder.

Voor wat betreft verandering kun je de vraag stellen welke soorten veranderingen er zijn. Zo onderscheidt Reeler (2007) drie soorten veranderingen: emergent (dagelijkse onbewuste en bewuste leerprocessen die leiden tot veranderingen), transformatief (verandering door crisissituaties) en voorspelbaar (vanuit visie/strategie veranderingen plannen). In het zogeheten *Payback Framework* worden de *paybacks* c.q. veranderingen, langs vijf dimensies ingedeeld: kennis, toekomstig onderzoek, beleid, praktijk en bredere maatschappelijke veranderingen (Greenhalgh et al., 2016). In transitieliteratuur wordt gesproken over veranderingen in cultuur, *practices* en structuur, om tot een transitie te komen (Van den Bosch & Rotmans, 2008). In het PRIME-model (Greven & Andriessen, 2019) wordt impact (*impact pathways*) gekoppeld aan veranderingen in persoonlijke ontwikkeling (leren), productontwikkeling (ontwerpen),

kennisontwikkeling (onderzoeken) en systeemontwikkeling (veranderen). Een organisatie als Keystone gebruikt een tool om vier dimensies van veranderingen te analyseren: individueel, relationeel, institutioneel en cultureel (James, 2011). Voor alles lijkt iets te zeggen, maar de verschillen maken ook duidelijk dat een gedegen analyse nodig is van soorten van verandering.

Een aanvullende vraag is hoe veranderingsprocessen in zijn werk gaan. Een van de bekendste theorieën hierover is die van Kurt Lewin. Onderliggend gaat het voor Lewin om een equilibrium van tegengestelde krachten (*driving forces* en *restraining forces*), waarbij verandering het bereiken van een nieuw equilibrium is. Een veranderingsproces verloopt in drie fases: *unfreezing*, *change*, *refreezing*. In de fase van *unfreezing* gaat het erom in bijvoorbeeld gedrag patronen te doorbreken die niet meer gewenst zijn. Dat kan door te 'duwen' in de richting van het nieuwe gedrag (*driving forces*) en/of het verminderen van krachten die oude patronen in stand houden (*restraining forces*). In de fase van *change* gaat het om de beweging naar een nieuw equilibrium toe, een beweging in denken, voelen en/of gedrag. In de fase van *refreezing* gaat om het verankeren van de verandering in een nieuw patroon, een nieuwe gewoonte waardoor je het gevaar van afglijden naar oude gewoonten voorkomt. Er zijn door de jaren heen vele uitbreidingen en variaties ontstaan op dit basispatroon van verandering zoals beschreven door Lewin.

Ook ten aanzien van ecosystemen kun je de vraag stellen vanuit welk perspectief je ernaar kijkt. Inzichtelijk hierbij is bijvoorbeeld de driedeling gebaseerd op interventiegronden, namelijk die van marktfalen, systeemfalen en transitiefalen (Schot & Steinmueller, 2016; Vankan et al., 2020).

Marktfalen treedt op als de mechanismen voor marktwerking inefficiëntie vertonen, bijvoorbeeld door gebrekkige coördinatie waardoor actoren elkaar niet weten te vinden of zogeheten informatieasymmetrie waarbij actoren de waarde van een product niet goed beoordelen.

Systeemfalen richt zich op tekortkomingen in op elkaar afgestemde organisaties, vaardigheden en instituties. Dit falen kun je recht trekken door in het systeem bijvoorbeeld essentiële infrastructuur in te brengen, of door wet- en regelgeving te maken, of actoren vaardigheden aan te leren (ondernemerschap, absorptievermogen van bedrijven). Er is een punt te maken dat het valorisatiebeleid een interventie is op marktfalen/systeemfalen: de vele kennis die op de plank ligt en zijn weg niet weet te vinden naar bedrijven. Ook adviezen om te zorgen dat innovaties zich beter verspreiden door het ecosysteem passen hierin (Rosenthal & Bovens, 2018) of analyses over de *stickiness* van informatie waardoor die zich niet makkelijk verspreidt (Von Hippel, 1994).

Ondernemerschap als thema binnen het hoger onderwijs of zelfs als profilering (de 'ondernemende' hogeschool) en het actieve beleid rond spin-offs, start-ups en scale-ups zijn hier een andere zichtbare uiting van. Bekende begrippen zoals triple helix en Gibbons' mode 2 passen in dit systeemdenken door hun nadruk op samenwerking en interacties (Schot & Steinmueller, 2016).

Een derde perspectief is dat van het transformatieperspectief dat door de lens kijkt van het in gang zetten van grootschalige veranderingen, bijvoorbeeld richting een duurzame samenleving. Centraal staan transities die nodig zijn als antwoord op weerbarstige of *wicked* maatschappelijke vraagstukken en die vragen om complexe, integrale en completerende veranderingen (sociaal, technologisch, bestuurlijk, gedrag, cultuur, et cetera). Ook hier kan sprake zijn van falen. Anders dan bij marktfalen en systeemfalen gaat het hier om zaken als het ontbreken van een duidelijke richting/visie op de verandering (*directionality failure*), afwezigheid van de vraag naar opbrengsten van de transitie (*demand articulation failure*) en falen in het bepalen van waarom een transitie uitblijft of hoever die al is (*reflexivity failure*).

Deze vormen van falen vanuit het transformatieperspectief zijn ook inzichtelijk bij het praktijkgericht onderzoek. De inspanningen en nadruk op regionale/landelijke samenwerking en het vormen van massa op onderzoeksthema's stuurt op het voorkomen van de *directionality failure*, en is af te lezen aan (gestimuleerde) initiatieven zoals lectorenplatforms, thematafels en krachtige onderzoeksgroepen (SPRONG-regeling). De vraagarticulatie als onderdeel van de praktijkverbondenheid van het onderzoek is al sinds jaar en dag onderdeel van de werkwijze van lectoraten. Zij is een mechanisme om de *demand articulation failure* te voorkomen, ook aangaande transitievraagstukken. In de analyse van SPRONG-aanvragen (zie 'exploratief onderzoek #4') is ook geconstateerd dat vraag(articulatie) een ecosysteemelement is dat in alle aanvragen een vanzelfsprekende rol speelt (Van Vliet, 2022). Dit is anders als het gaat om de *reflexivity failure*. In bijna de helft van de SPRONG-aanvragen is geen wezenlijk ontwikkelingsperspectief terug te vinden wat betreft impact/doorwerking, terwijl het zeer onwaarschijnlijk is dat de huidige systematiek van impact/doorwerking geheel toereikend is, volledig toekomstbestendig en optimaal past bij maatschappelijke transities. Dit is zorgwekkend, zeker omdat er methoden zijn waarin dit 'reflexieve' op veranderingen naar aanleiding van interventies ingebakken zit, zoals in *Theory of Change* en *Result-based accountability*, en er methoden zijn waarin maatschappelijke impact een integraal onderdeel is van het onderzoeksproces zoals SIAMPI, *Contribution mapping* en ASIRPA (Smit & Hessels, 2021). Zo wordt nogmaals duidelijk dat een bepaald perspectief voorsorteert op de inzet van bepaalde methodieken en de praktische uitvoering en evaluatie van doorwerking.

Attitude

Attitude gaat over het onderliggende doel c.q. de onderliggende motivatie om naar doorwerking te kijken, en daarmee over de houding ten opzichte van deze vraagstukken.

Bij de beschrijving van het analyseperspectief is een relatie gelegd tussen bepaalde onderliggende aannames en ideeën over kennis(productie), en methodieken. Er zijn nog meer dimensies waarlangs je methodieken met elkaar kunt vergelijken, zoals of het gaat om een ex post, ex ante of in media res type van evalueren, of het wel of niet uitgaat van een logisch/lineair model, et cetera (zie Coombs & Meijer, 2021; Smit & Hessels, 2021). Een dimensie die ook vaak terugkeert is of de methode een summatieve of een formatieve insteek heeft. Hoewel die overigens niet altijd leidt tot eenzelfde indeling, zo typeren Coombs & Meijer (2021) de methode ASIRPA als formatief, terwijl Smit & Hessels (2021) deze methode als summatief classificeren.

Het verschil tussen summatief en formatief is kort gezegd de nadruk leggen op verantwoorden versus de nadruk leggen op het leerproces. Bij summatief is de insteek: vaststellen of een bepaald doel is gehaald of in welke mate het is gehaald, om zo te komen tot kwalificaties als onvoldoende, voldoende of goed. Bepaalde scores worden dus beloond of juist bestraft, wat ook leidt tot ander gedrag om de scores te halen. Veelal zijn er afspraken gemaakt over hoe tot een specifieke (kwantitatieve) score te komen, waardoor er ook een vergelijking mogelijk is met anderen, bijvoorbeeld andere onderzoekseenheden. Bij formatief is de insteek om vast te stellen *waarom* bepaalde uitkomsten zijn bereikt, met daarbij de reflectie op wat goed gaat, wat beter kan en hoe dit bereikt kan worden gegeven de context. Die context maakt het vaak lastiger onderling te vergelijken. Formatief is dus een standpunt waarin de vraag centraal staat hoe doorwerking is bereikt en hoe dat nog verder is te verbeteren.

Methodieken nemen hier een verschillend standpunt in. De methode SIAMPI is bijvoorbeeld heel uitgesproken over haar formatieve standpunt: 'where learning is the prime concern instead of judging' (Spaapen & Van Drooge, 2011b, p. 211), en 'The essence of this approach [productive interactions] is that it is not competitive but collaborative by nature. Its premise is that to achieve societal impact it is most fruitful if various relevant stakeholders work together, combining different kinds of knowledge and expertise, designing a joint research and innovation agenda, and being open and inclusive because one never knows in advance where things are going. In this approach there also *is a joint responsibility for every actor to evaluate the project and learn from it*' (Van den Akker &

Spaapen, 2017, p. 17, mijn cursivering). Ook andere methoden leggen meer de nadruk op hoe je impact kunt bereiken en hoe dat beter kan, of dat nou is door *outcome pathways* te analyseren (*Theory of Change*) of door nauwgezet bijdragen van verschillende stakeholders in kaart te brengen (*contribution mapping*). Hierbij is veel meer aandacht voor een verdere optimalisering van het proces, in de veronderstelling dat je daarmee ook de impact vergroot.

Ook hier geldt dat het standpunt formatief of summatief veelal samengaat met andere standpunten: 'When comparing the purpose of the methods with their conceptual principles, we see that most summative methods hold a linear model of knowledge exchange, and formative methods take either a cyclical or co-production model (Smith & Hessels, 2021, p. 332). Gegeven de bedenkingen bij het lineaire model (zie 'Lineland') zou dat zomaar tot de conclusie kunnen leiden van 'formative good, summative bad' (Lau, 2016). Maar zoals Lau duidelijk maakt, is het maar de vraag in hoeverre beide perspectieven strijdig met elkaar zijn omdat ze aanvankelijk ook gezamenlijk zijn geformuleerd. Dit maakt Lau duidelijk met een smakelijk voorbeeld 'when the cook tastes the soup, that's formative evaluation; when the guests taste the soup, that's summative evaluation' (ibid., p. 511). Hoewel het bij onderzoeksevaluatie iets ingewikkelder ligt dan het bereiden van een gerecht (zie verder onder attributie), maakt het wel duidelijk dat methoden die meer naar 'voren kruipen' in het onderzoeksproces bijna noodzakelijkerwijs bij een formatief standpunt uitkomen en methoden die meer 'op het eind' kijken bij een meer summatief standpunt uitkomen. Het is dan ook eerder de vraag hoe de balans hiertussen te vinden dan een harde keuze te maken. Wel is vanuit de gepresenteerde exploratieve onderzoeken te concluderen dat het (kwantitatief) summatieve vaak overheerst, vanuit een verantwoordingsdoctrine, en het formatieve perspectief te weinig wordt uitgedragen of in ieder geval zichtbaar gemaakt en gewaardeerd.

Attributie

Attributie gaat over de vraag hoe doorwerking is te herleiden tot een aanwijsbare partij of persoon die de veroorzaker is van die doorwerking.

Attributie wordt in discussies zonder meer gekenschetst als een uitdagend probleem (zie onder andere Greenhalgh et al., 2016; Raftery et al., 2016; Pedersen, Grønvad & Hvidtfeldt, 2020). Hoe kun je nu een uiteindelijk bereikt effect onomstotelijk herleiden tot een actor? Zeker als de consensus is dat onderzoek doen voortbouwt op kennis van anderen, niet kan zonder ideeënuitswisseling, berust op peer review, op interacties, netwerken, et cetera. Onderzoek is

kortom een contactsport. Het feit dat de vraag naar attributie toch gesteld blijft worden is ten eerste omdat het in sommige gevallen wel lukt. De *lone genius* die voor een omwenteling zorgt in het denken of tot een nieuw product komt dat revolutionair is, hoewel uit historische analyses vaak blijkt dat gelijksoortige ideeën en producten ook bij anderen en elders zich begonnen te vormen – het ‘hing in de lucht’. Inzoomen op grote ontdekkingen laten van dichtbij zien dat het veelal een aaneenschakeling is van kleine stapjes.

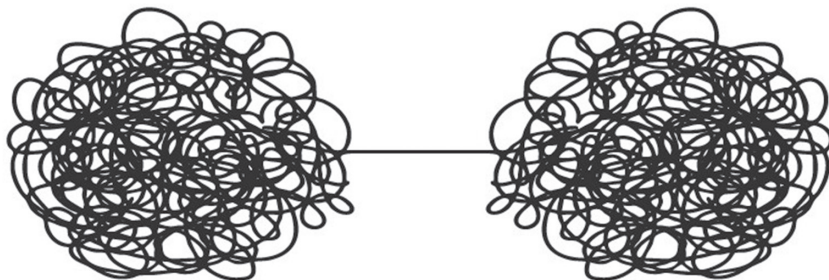
Ten tweede, laten we niet vergeten dat door de ‘scheiding’ van fundamenteel onderzoek en toegepast onderzoek er de noodzaak ontstond om de overdracht van kennis te regelen via bijvoorbeeld licenties en intellectueel eigendom, waarmee een administratief spoor werd gecreëerd om te achterhalen van wie het oorspronkelijk idee was.⁹ Met overigens een nadruk op economische impact (Godin & Doré, 2005). Eigenaarschap claimen is niet ongebruikelijk in de wetenschap. Zo noteerde een van Nederlands grootste wetenschappers Christiaan Huygens (1629–1695) sommige ideeën in anagrammen, waarna hij verder werkte aan de onderbouwing ervan. Met die anagrammen kon hij later wel aan concurrenten laten zien dat hij het idee al eerder had, eerder dan zij (Aldersey-Williams, 2020).

Ten derde blijft het attributievraagstuk terugkomen omdat het nadrukkelijk verbonden is aan het SLM met zijn lineaire push van toewerken naar het eindstation van impact. Het mooie van onderzoek is dat je ook onderzoek kunt doen naar hoe onderzoek werkt. Raftery et al. (2016) halen diverse studies aan die onderzochten of er een directe lijn te trekken was tussen onderzoeksprojectuitkomsten en impact. De conclusie is ontvankelijk: in maar zeer weinig gevallen is dat te doen. Veeleer gaat het om een *body of research*, een *research cloud*, dat in zijn totaliteit zorgt voor een conceptuele verschuiving in beleidsvorming en praktijken. Kanttekening is dat deze studies vaak vanuit een constructivistisch perspectief zijn uitgevoerd en gericht op onderzoek in de sociale wetenschappen. Maar het is niet moeilijk gelijksoortige conclusies te vinden vanuit andere hoeken, zoals bij Kline and Rosenberg (1986, p. 302): ‘Any model that describes innovation as a single process, or attributes its sources to a single cause, or gives a truly simple picture will therefore distort the reality and thereby impair our thinking and decision making’ (p. 302) en: ‘The notion that innovation is initiated by research is wrong most of the time’ (p. 288). En: ‘In reality, attribution is often an inexact science in which accuracy attenuates with each stage in the logic model’ (Raftery et al., 2016, p. 50). Zeker een definitie van

9 Met de kanttekening dat dit niet altijd een 1-op-1-relatie oplevert: zo maakt de cd-rom gebruik van meer dan 20.000 gepatenteerde technische vindingen (Rosenthal & Bovens, 2018).

impact als verder weg in de tijd helpt niet. 'The more impact is defined at a large distance in time and in causal relations from the research activities, the less appropriate it is for evaluation purposes' (Spaapen & Van Drooge, 2011a, p. 9).

Met ontwikkelingen als open innovation, open science, de gerichtheid op complexe maatschappelijke problemen met een onderkenning van de noodzaak tot interdisciplinair en transdisciplinair werken, en de helices die over elkaar heen struikelen – van triple helix naar quadrupel helix en intussen ook quintupel helix – als opeenstapelingen van belangen die een rol spelen, maken we het onszelf ook niet makkelijker om de lijn te trekken tussen een resultaat en impact. Dit komt uiteindelijk neer op de onderkenning dat voor het onderzoek zelf een complexe samenwerking vereist is én dat het resultaat landt in een complexe omgeving namelijk de samenleving. Daartussen willen we dan een lijn trekken tussen een punt in de kluwen aan de ene kant en een punt in de kluwen aan de andere kant (figuur 8). Succes! Het attributievraagstuk lijkt daarmee steeds meer op een zelf gecreëerd probleem met bijbehorende deskundigen in de zin van: 'Een deskundige is iemand die zich specialiseert in zelf opgeworpen problemen' (Van Boxsel, 2016, p. 206)



Figuur 8. Het wankele koord waarop de attributieacrobaat zich begeeft

Hoe gaan verschillende methoden om met de onderkenning dat het attributieprobleem een uitdagend probleem is? Ik zie in ieder geval drie soorten aanpakken.

De eerste aanpak is de stapsgewijze aanpak waarbij je in stapjes toewerkt naar de finale impact. Die stapjes zijn goed herkenbaar en kun je ook beoordelen op hun potentiële bijdrage aan de impact. Twee verschillende voorbeelden van deze aanpak zijn Theory of Change en small wins. Bij een Theory of Change breng je systematisch in kaart wat de samenhang is tussen activiteiten, uitkomsten en context van een vooraf opgesteld langetermijndoel, waardoor je een

logische reeks van interventies (de *outcome pathway*) kunt kiezen die leidt tot het langetermijndoel. Die outcomes monitor je nauwkeurig en je reflecteert op wat de uitkomst betekent voor het gehele bouwwerk, in feite dus een meer formatief perspectief. Bij small wins (zie ook 'Lineland') gaat het om kleinschalige radicale vernieuwingen die je moet spotten (of initiëren) en die vervolgens aandacht moeten krijgen om ze te verspreiden, verbreden en verdiepen. Het is een groeitraject van klein lokaal naar groot nationaal om aan transities te werken.

De tweede aanpak is voorwaardenscheppend, in die zin dat er in het onderzoeksproject alles aan wordt gedaan om te zorgen dat de impact kan ontstaan. Een goed voorbeeld hiervan is de methode van *contribution mapping*, in deze methode ligt de nadruk op zogeheten *alignment efforts*: het oplijnen van het onderzoek met onderzoeksagenda's, actoren binnen en buiten het onderzoeksproject, betrekken van potentiële gebruikers bij het onderzoeksproces, et cetera (zie verder onder 'Aantoonbaar'). Een ander voorbeeld is de methode *linkage and exchange framework*, die ook zoekt naar hoe je onderzoeksresultaten kunt verbinden en uitwisselen met maatschappelijke actoren en instituten (zie Pedersen, Grønvad & Hvidtfeldt, 2020). Deze voorwaardenscheppende aanpak is bij NWO het uitgangspunt: 'NWO beoordeelt onderzoek daarom niet op het behalen van maatschappelijke impact, maar wel op de inspanningen die zijn verricht en successen die zijn behaald in het bevorderen van het kennisbenuttingsproces' (2019, p. 4). Waarin dan weer de stellingname in het rapport *Maatschappelijke impact in Kaart* van de KNAW doorklinkt: 'richt ex ante evaluatie niet op het in kaart brengen van de maatschappelijke impact zelf maar op de factoren en processen die de kans op maatschappelijke impact vergroten' (Van de Sanden et al., 2018, p. 20).

De derde aanpak verschuift de focus van impact van 'ver weg' naar het 'hier-en-nu'. De methode SIAMPI is hier het meest expliciet in: 'We propose a shift from a focus on social impact towards a focus on "productive interactions". This way, a much more realistic approach is offered, that is closer to the actual process that the researcher is able to influence, that is closer to the actual practice of the researcher doing research and interacting with stakeholders. De facto social impact in the sense of an objective and measurable effect is not neglected, but since those effects only develop over a long time, it is not our prime focus' (Spaapen & Van Drooge, 2011b, p. 216). SIAMPI stelt dus dynamische interacties en gezamenlijke inspanningen centraal. Het uitgangspunt is dan ook dat men niet zozeer de impact van onderzoek moet bestuderen, als wel de processen die functioneren als proxy voor impact. Sociale effecten van onderzoek zijn het gevolg van ingewikkelde processen van uitwisselingen, van onderwijs en leren, tussen betrokken partijen. Deze effecten zijn direct

zichtbaar en traceerbaar in een concreet onderzoeksproject. Hierbij wordt ook sterk de nadruk gelegd op de context van het onderzoek en het leidend zijn van het leren. Eenzelfde standpunt vinden we bij (Sivertsen & Meijer, 2020): 'We think this burden [bewijs aandragen voor sociale impact] could be relieved by replacing the term "societal impact" with "societal interaction" and a focus on the real and normal organizational level interaction according to the aims and purposes on both sides [science/society]' (p. 68).

De keuze voor een van deze methoden zul je niet alleen vanuit het attributieprobleem moeten bekijken maar ook vanuit de andere onderliggende vragen zoals die hier gesteld worden. Het tweede standpunt van voorwaardenschepend geeft wel een mooie voorzet om op een andere manier te formuleren wat eerder al is betoogd, namelijk de opmerking dat 'actief betrekken van de potentiële gebruikers bij het opstellen van de onderzoeksvraag en de uitvoering van een onderzoeksproject' de kans vergroot op impact (Van de Sanden et al., 2018, p. 20): nee het vergroot niet de kans van maatschappelijke impact, het is de impact, of om preciezer te zijn: de doorwerking.

Aantoonbaar

'Aantoonbaar' gaat over de vraag wat en hoe aan te tonen dat doorwerking plaatsvindt of heeft plaatsgevonden.

Ik prefereer de term 'aantonen' boven de term 'meten', omdat hij neutraler is. Bij 'meten' is de connotatie kwantitatief. 'Aantonen' laat dit open, en kun je zowel kwantitatief als kwalitatief invullen. Voorts laat 'aantonen' ook open welke vorm gekozen wordt om over dit aantonen te rapporteren: een rapport, infographic, tabel met cijfers, videoclip, dashboard of narratief. Deze laatste, narratief, wordt hier niet voor niets genoemd. In veel discussies wordt een oneigenlijke tegenstelling gemaakt tussen indicatoren en een narratief, uit onvrede over (kwantitatieve) indicatoren besluiten mensen dan maar een narratief te schrijven. Een indicator vertelt je waar en hoe je naar iets wilt kijken, dat is een vraag die ook een narratief moet beantwoorden. Wordt die vraag niet beantwoord dan verzandt een narratief vaak in een droge procesbeschrijving (we kregen geld, partijen gevonden voor samenwerking, gingen aan het werk met elkaar in een lab, we hebben verteld over de uitkomsten en toen was het voorbij). De lezer verwonderend achterlatend: 'Wat wil je hier nou mee zeggen?' En dat is exact dezelfde vraag die je aan een indicator kunt stellen.

In de voorafgaande discussie hebben zich een aantal standpunten naar voren gedrongen die al verder richting geven aan wat ik versta onder aantonen. Ten eerste gaat aantonen niet alleen over de grote buitengewone opbrengsten, maar juist ook over de 'kleine' opbrengsten in interacties, initiatieven, innovaties, zeg maar de small wins. Ten tweede is aantonen niet iets wat zich alleen moet richten op eindopbrengsten als een soort *end of pipe quality control*. Als een essentieel kenmerk van doorwerking is dat het al voor en tijdens een onderzoeksprogramma speelt, dan kan het aantonen ook daar beginnen. Ten derde zijn stakeholders een noodzakelijk onderdeel van onderzoek en daarmee van het aantonen, zowel als 'object' (hun bijdrage) en als duider van de waarde van de opbrengsten. Ten vierde is de context benadrukt als belangrijk in de bepaling van de betekenis en waarde van wat er gebeurt, dat zal dus ook moeten doorklinken in het aantonen. Dat kan door 'passende' indicatoren en methodieken te kiezen (zie ook figuur 6 in 'Intermezzo').

In dit 'passende' zit een ander standpunt verborgen, namelijk ten aanzien van de eerder aangehaalde discussie over verscheidenheid en vergelijkbaarheid (zie 'Intermezzo'). Dat de context belangrijk is, kan leiden tot de conclusie dat daarmee een vergelijking niet mogelijk is, immers de ene context is de andere niet. Dit houdt niet automatisch in dat ze daarmee totaal onvergelijkbaar zijn, al is het maar omdat het altijd gaat over de positie van een bepaalde stakeholder (rol hogescholen), een zekere gemeenschappelijkheid in aanpak (praktijkgericht onderzoek) en binnen een bepaald beleidskader (missiegedreven innovatiebeleid). Hoewel ik niet zo ver wil gaan als Stam & Van de Ven (2018) met hun ecosysteemindex, dat die vergelijkbaarheid is uit te drukken in een index. Ik vind dan ook dat het aspect van contextualiteit niet leidt tot een fundamenteel probleem van onvergelijkbaarheid maar tot een technisch probleem van vergelijkbaarheid. De karakterisering als een technisch probleem brengt de pragmatiek naar boven: hoe dan?

De aanpak die vooralsnog is gekozen is het werken aan een toolbox. Een toolbox met een verscheidenheid aan instrumenten voor het realiseren en aantonen van doorwerking. Bij 'Aggregatie' kom ik hierop kort terug voor wat betreft de indicatoren. Voor wat betreft de methodieken is een activiteit in gang gezet om methodieken c.q. tools voor het realiseren en aantonen van doorwerking te inventariseren en te ontsluiten op een eenvoudige manier via de website www.doorwerking-hbo-onderzoek.nl. Eerder zijn al de twee methodieken ter sprake gekomen, namelijk die van small wins (zie 'Lineland') en de doorwerkingsmatrix (zie 'Flatland'). Ik voeg hier twee andere voorbeelden aan toe van methoden die een ander perspectief innemen dan gebruikelijk, door zich meer te richten op het onderzoeksproces zelf in plaats van op impact (contribution mapping),

en het onderzoeksproces ook gebruiken om kwalitatief aan te tonen wat de verandering is die wordt nagestreefd (*most significant change*).

Contribution mapping (Kok & Schuit, 2002) is een aanpak om onderzoeksprojecten te monitoren en te evalueren aan de hand van interacties van betrokken actoren. Deze aanpak presenteert zich als een alternatief voor bestaande methoden die gericht zijn op het vaststellen van impact. De aanpak richt zich namelijk op het onderzoeksproces zelf en op de daarin betrokken actoren en hun handelen. Dit doet de methode door systematisch in kaart te brengen welke inspanningen er worden verricht om tot meer bijdragen te komen vanuit het onderzoek aan een specifiek (maatschappelijk) probleem. De nadruk ligt op het in kaart brengen van wie, wanneer en met wat heeft bijgedragen aan een onderzoeksproject. Dit inzicht is te gebruiken om bijdragen te optimaliseren en ook de impact vast te stellen: hebben alle noodzakelijke partijen bijgedragen? Is er alles aan gedaan om de bijdragen te optimaliseren? Die bewuste inspanningen noemen we alignment efforts. Door deze alignment efforts in kaart te brengen en erop te reflecteren, ontstaat er ook lerend vermogen bij de actoren doordat zij gaan anticiperen op, en veranderingen gaan aanbrengen in, hun eigen handelen. Dit doen ze om tot betere afstemming te komen met de uitvoering van het onderzoek en om de onderzoeksopbrengsten te kunnen gaan gebruiken. De vooronderstelling is dat je door zo nadrukkelijk op het netwerk te sturen in en rond een onderzoeksproject, de impact kunt vergroten. Vooralsnog is contribution mapping als methode (kort) na afloop van een project toegepast, maar het is een methode die ook tijdens de uitvoering zijn nut kan hebben door bijvoorbeeld actief te sturen op de alignment efforts.

Most Significant Change (MSC) (Davies & Dart, 2005) is een kwalitatieve methode om de betekenisvolste veranderingen die een project teweegbrengt te vangen in verhalen. In een gezamenlijke dialoog wordt systematisch gezocht naar betekenisvolle veranderingen die zich voordoen en de gedeelde criteria waarom die verandering als betekenisvol wordt gezien. Dit kan helpen om tijdens het project sturing te geven aan het project en/of om achteraf tot een assessment van het succes van het project te komen. De methode wordt soms ook wel aangeduid als *monitoring without indicators*, een manier om het niet-telbare te laten meetellen. MSC leunt sterk op de participatie van deelnemers en stakeholders in een project, zij bepalen gezamenlijk welke verhalen het betekenisvolst zijn om de meerwaarde van het project aan te tonen. Verhalen verzamel je op gebieden waarbij je impact wilt bereiken door projectdeelnemers en doelgroepen de open vraag te stellen 'wat was in jouw ogen de belangrijkste verandering die zich de afgelopen drie maanden heeft voorgedaan als gevolg van het project?' De verzamelde verhalen bespreek je onderling en je kiest het

beste verhaal, waarbij je nadrukkelijk de criteria waarom het verhaal het beste is expliceert. De discussie hierover is al waardevol voor het vormen van een gezamenlijk beeld over wat het project moet bereiken en over de criteria voor wat een betekenisvolle verandering is. Dit proces wordt ook wel *dynamic values inquiry* genoemd. Door MSC te gebruiken tijdens een project, ontstaat er een gezamenlijke focus op wat belangrijk wordt gevonden, waarom het belangrijk wordt gevonden en welke verhalen het project wil opleveren. In die zin is MSC meer gericht op tussentijdse resultaten dan op de traditionele output van een project, en ook meer op leren tijdens dan verantwoordt achteraf.

Aggregatie

Aggregatie gaat over de vraag hoe opbrengsten van individuele onderzoekers of onderzoeksprojecten kunnen worden 'opgeteld' om tot een uitspraak te komen die een meer omvattend geheel representeert.

Het onderzoek van een onderzoeker is veelal ingebed in een groter geheel, het is bijvoorbeeld onderdeel van een werkpakket dat onderdeel is van een project dat weer onderdeel is van een programma, dat weer onderdeel is van een landelijke agenda. Parallel daaraan, of er dwars doorheen, lopen dan nog eens onderzoeksgroepen (lectoraten), organisatie-eenheden (kenniscentra), publiek-private samenwerkingen (Centres of Expertise) en instituten (hogescholen) die ook verantwoording moeten/willen afleggen over de opbrengsten van het onderzoek. In deze wirwar van verantwoordelijkheden, verplichtingen, verwachtingen, (strategische) keuzes, activiteiten en allerhande kwaliteitscycli, indicatorenlijstjes en registratiesystemen is het onvermijdelijk dat de handen weleens in de lucht worden geworpen van moedeloosheid hoe dit allemaal aan elkaar te praten, bijvoorbeeld bij een bezoek van een visitatiecommissie. Op de vraag of dit aan elkaar te praten is, is vanuit het perspectief van doorwerking het antwoord: ja en nee.

Het ja-antwoord is tweeledig. Ten eerste gaat het 'optellen' van verschillende resultaten makkelijker als dit kwantitatieve gegevens zijn, zoals omvang van het onderzoek (middelen, mensen). Dit kun je vanaf het niveau van een lectoraat relatief eenvoudig optellen tot landelijke cijfers, zoals ook gebeurt. Het BKO draagt op dit soort gegevens aan te leveren voor de 'landelijke monitoring', en menig rapport over de ontwikkeling van het praktijkgericht onderzoek doet dit aan de hand van dit soort gegevens (De Jonge, 2016; Cloosterman et al., 2019). Dit kan wel tot een tunnelvisie leiden doordat een zoektocht ontstaat

naar dé alles overkoepelende index (zie eerdere discussie over een index voor ecosystemen in ‘Spaceland’), of hooguit drie want het moet overzichtelijk en bestuurbaar zijn. Het kan ook leiden tot steeds verder ‘decontextualiseren’ van de cijfers (Coombs, 2019) waardoor er steeds meer afstand ontstaat tot de werkvloer (‘Wat zeggen die cijfers nou?’). Ten tweede kan aggregeren omdat het eerder gedaan is. Er zijn diverse voorbeelden waarin grote programma’s met complexe samenwerkingen integraal zijn geëvalueerd, alleen blijkt dat dat erg arbeidsintensief is: ‘The most robust and sophisticated approaches are labour-intensive and not always feasible or affordable’ (Greenhalgh et al., 2016, p. 1). Een goed voorbeeld van hoe arbeidsintensief het kan zijn, is ASIRPA (Joly et al., 2015). ASIRPA staat voor *Socio-Economic Analysis of the Impacts of Public Agricultural Research*, en is een project geweest van het Franse landbouwonderzoeksinstituut. Een ingewikkeld traject waarvan ik 1 aspect wil uitlichten, namelijk dat het project een aanpak heeft gehanteerd om 33 casestudies te aggregeren tot uitspraken over de impact van een onderzoeksinstituut als geheel. Dit is gedaan door standaardisatie op twee manieren: door identificatie en systematische analyse van gedeelde kenmerken in alle cases én de keuze van enkele analytische tools. De eerste was mogelijk doordat alle cases gingen over hetzelfde domein (landbouw) en bestond bijvoorbeeld uit een systematische analyse via een chronologie in ontwikkelingen, waarbij het niet ging om attributie maar om contributie van de casestudies. Ook impact pathways werden gebruikt, maar wel met een specifieke interpretatie (non-lineair, formatief, invloed van ecosysteemelementen, et cetera). De analyse leverde uiteindelijk zo’n honderd variabelen op voor de casestudies die in een database werden gecodeerd, waarop analyses werden uitgevoerd, ook kwantitatieve.

Het nee-antwoord is ook tweeledig. Ten eerste geeft de optelling van kwantitatieve gegevens op zijn minst een onvolledig beeld van wat bereikt is met het onderzoek. Alleen de ware ‘inputgelovige’ (‘We pompen er zoveel geld in, dan moet het wel impact hebben.’) kan daarmee weggkomen. Ook kwantificering aan de achterkant – bij de output bijvoorbeeld het tellen van publicaties – geeft een beeld, maar een onvolledig beeld; zeker in relatie tot wat er gezegd is over doorwerking. Eerder is al een lans gebroken voor contextualisering van onderzoek als het gaat over doorwerking. En nog preciezer: welke activatie, relevantie, gebruik en verandering het onderzoek als effect heeft. Dit grijpt terug op de conclusie van onder andere Pedersen, Grønvad & Hvidtfeldt (2020): ‘A key finding of the literature review therefore is that societal impact does not occur in a social or cultural vacuum, but is realized in a network of interacting actors, interests, values, and institutions’ (p. 12). Die netwerken en (productieve) interacties verschillen per onderzoekssituatie, per domein, per regio, et cetera. Wat

als waardevol wordt gezien, vindt in die communicatie- en handelingsruimte plaats. Het magische woord 'opschalen' legt dan ook vaker de onmacht bloot om individuele successen over te planten naar andere contexten. Hiermee is niet gezegd dat kwalitatieve eigenschappen niet op een bepaalde manier geaggregeerd kunnen worden, dat dit fundamenteel onmogelijk zou zijn. Het methodische probleem is alleen ingewikkelder dan bij kwantitatieve gegevens, of in ieder geval arbeidsintensiever.

Ten tweede hebben de verschillende 'lagen', daar waar het gaat om aggregatie, vaak ook andere doelstellingen, die misschien in elkaars verlengde liggen maar wel anders geframed worden en tot een ander narratief leiden. Een instelling die de omvang van het onderzoek (aantal fte's) en binnengehaalde financiering vanuit de tweede geldstroom als indicatoren kiest, kan hiermee het beleid van focus en massa onderbouwen. Maar deze indicatoren ook gebruiken voor aansturing en evaluatie van de lectoraten is een armzalige aanpak. Voor de lector zijn dit soort indicatoren randvoorwaarden om zijn of haar bijdrage aan maatschappelijke uitdagingen te leveren. Geen onbelangrijke randvoorwaarden maar wel randvoorwaarden. Met verschillende doelstellingen als profilering (instituut), verbinding met het onderwijs (kenniscentra) en met onderzoek de beroepspraktijk innoveren (lector) komen ook andere vormen van bewijslast en hoe deze te verzamelen en te interpreteren. Het is ook de praktijk 'that methods for tracking and assessing research impact vary significantly at the project-, programme-, or system-level.' (Pedersen, Grønvd & Hvidtfeldt, 2020, p. 2). Met andere woorden: je moet het niet altijd willen kunnen optellen.

Het grijze gebied tussen ja en nee betekent dat er ook mogelijkheden zijn om dat wat strakker naar elkaar toe te trekken. Dan moeten we wel voorbij naïeve uitspraken komen zoals 'Innovaties zijn zo verschillend dat je ze niet kunt optellen. Vormen één nieuwe smaak yoghurt plus één hernieuwbare bron van energie samen twee innovaties?' (Rosenthal & Bovens, 2018, p. 9). Dit is typisch een uitspraak vanuit een raamwerk *end of pipeline*: als de output zo verschillend is dan zijn ze ook totaal niet vergelijkbaar, laat staan op te tellen. Doorwerking richt zich ook op het proces, en het kan zomaar zijn dat twee totaal verschillende eindproducten eenzelfde, of in ieder geval een vergelijkbaar, proces hebben doorlopen dat ons veel leert over wat werkt en wat niet.¹⁰

10 In feite is dit de idee van De Valorisatiefabriek. Een idee waar ik zo'n tien jaar geleden mee heb geflirt: een fabriek die gespecialiseerd is in een aanpak om allerlei beloftevolle ideeën naar de markt te brengen, of dit nu yoghurt is of hernieuwbare energie. Die specialisatie bestaat uit kennis, methoden en netwerken om dit te doen, als proces, ongeacht het product. De url bezit ik nog steeds...

Er zijn mogelijkheden het kwantitatieve/kwalitatieve en de verschillende lagen met hun eigen doelstellingen dichterbij elkaar te brengen. In feite zijn daar al een aantal aanzetten toe gegeven met: 1) de criteria voor praktijkgericht onderzoek die in kaart zijn gebracht (tabel 2), waarover lectoraten, beleidsmedewerkers, bestuurders in gesprek kunnen gaan over wat ze belangrijk vinden en om *common ground* te vinden en activiteiten op te lijnen; 2) de ecosysteemelementen (tabel 6) als mogelijke knoppen waaraan gedraaid kan worden, wat ook tot een discussie kan leiden wie waarop stuurt en waarom, binnen een hogeschool of groter samenwerkingsverband; 3) de doorwerkingsmatrix (figuur 4) die als conceptueel model geen specifieke onderzoeksmethodologie of specifiek onderzoeksproces veronderstelt en daarmee onderzoekers vanuit verschillende achtergronden en disciplines in gesprek kan laten gaan over welk effect ze beogen en zo een gezamenlijker narratief en gedeelde indicatoren (categorieën) kunnen kiezen. Tot slot kan ook nog worden genoemd de uitwerking van de 39 subcategorieën van indicatoren (*exploratief onderzoek #3*) in 8 hoofdcategorieën die op geaggregeerd niveau ook de discussie over indicatoren kan stimuleren (Van Vliet, 2021).

Het zijn allemaal onderdelen van een taalspel dat op gang gebracht moet worden in 'den brede' binnen een hogeschool, tussen hogescholen en ook met subsidieorganen. Het zijn (deel)oplossingen die in ieder geval perspectief geven, en ook laten zien dat gedeelde methodieken voor doorwerking ook meer aggregatiemogelijkheden creëren. Het is zaak deze deeloplossingen meer te verbinden met de vijf hoofdvragen waarop we hier zijn uitgekomen. Dat is nog een taak maar te doen. Dat baseer ik op onder andere het voorbeeld van ASIRPA. Zonder nu ASIRPA als 'lichtend voorbeeld' op een voetstuk te plaatsen, zie ik daarin, los van alle technische details, dat die aanpak werkt omdat hij antwoorden geeft op de vragen die hier gesteld worden. Die aanpak neemt immers duidelijk stelling in hoe hij naar kennisproductie en -benutting kijkt (analysekader), neemt een formatief standpunt in zonder het summatieve van tafel te vegen (attitude), richt zich op contributie niet op attributie (attributie), gebruikt een mix van kwalitatieve en kwantitatieve methoden van rapporten (aantoonbaar), en komt zo tot een 'optelsom' van het effect van een onderzoeksinstituut op basis van individuele casestudies (aggregatie).

Epiloog

Het aantonen van de bijdrage van onderzoek is ingewikkeld. Dat is geen conclusie maar een vertrekpunt geweest voor deze (voor)beschouwing. Maar waar zit die ingewikkeldheid dan in? Die ingewikkeldheid zit in een samenhangend geheel van veronderstellingen, aannames en vraagstukken: van kennistheoretische vragen hoe kennis ontstaat tot praktische vraagstukken hoe dingen 'optellen'. Dit is samengebracht in de vijf A's van: analysekader, attitude, attributie, aantonen en aggregaties, al voorsortierend op een eenvoudige tentamenvraag... Het dominante antwoord dat zich in vele praktijken heeft genesteld slaat het ingewikkelde plat in een lijn. Als houvast. Het woord doorwerking is gebruikt als koevoet om dit open te breken, niet in de laatste plaats om meer recht te doen aan de praktijk van het onderzoek aan hogescholen. Een aantal vergezichten, dimensies, zijn geopend: van doorwerking 'in de breedte' en indicatoren voor de effecten van directe interacties, tot meer handelingsperspectieven dan alleen nieuwe kennis om veranderingen te bewerkstelligen. Hoe dit allemaal in elkaar 'klikt' is een opdracht die zich ook in de praktijk zal moeten bewijzen. Met altijd de mogelijkheid dat iets over het hoofd is gezien, een perspectief, een dimensie, die door vooringenomenheid slechts als een vage dissonant wordt gezien. Zoals een vierdimensionaal object dat slechts zijn schaduw werpt in een driedimensionale ruimte. Een schaduw waar we naar staren zonder het object echt te kennen, dat we voorzichtig beroeren met een vingertop, en ons afvragen wat de precieze vorm en inhoud is. Ik weet het niet.

Bronnen

- Abbott, E. A. (1884). *Flatland. A Romance of Many Dimensions* (2nd Edition). Oxford University Press.
- Aldersey-Williams, H. (2020). *Een eeuw van licht. Het leven van Christiaan Huygens*. Thomas Rap.
- Andriessen, D., & Franken, A. (2015). Zie kennis als een proces. Hoe meet en vergroot je het effect van praktijkgericht onderzoek? *Th&ma*, (5), 42-47.
- Arkesteijn, M., Van Mierlo, B., & Leeuwis, C. (2015). The need for reflexive evaluation approaches in development cooperation. *Evaluation*, 21(1), 99-115.
- Balconi, M., Brusoni, S., & Orsenigo, L. (2010). In defence of the linear model: An essay. *Research policy*, 39(1), 1-13.
- Berends, F., & Van Delft, D. (2019). *Lorentz. Gevierd fysicus, geboren verzoener*. Prometheus.
- Biggar Economics (2014). *Economic impact of the Amsterdam Universities and University Medical Centres*. Biggar Economics
- Bijker, W. E. (1997). *Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change*. MIT press.
- Borgdorff, H., Van Staa, A. L., & Van der Vos, J. (2007). Kennis in context. Onderzoek aan hogescholen. *Th&ma*, 15(5), 10-17.
- Bornmann, L. (2013). What is Societal Impact of Research and How Can It Be Assessed? A Literature Survey. *Journal of American Society for Information Science and Technology*, 64(2), 217-233.
- Brouns, M. (2016). Van Olympus naar Agora. Een frisse blik op praktijkgericht onderzoek. *Th&ma*, (4), 69-74.
- Cipolla, C. M. (2016). *De wetten van menselijke stupiditeit*. Amsterdam University Press.
- Cloosterman, E., Gielen, M., Van den Toren, J. P., & Van der Starre, B. (2019). *De stand van praktijkgericht onderzoek in Nederland*. Birch.
- Coombs, S. K. (2019). *Towards Evaluating the Research Impact of Dutch Universities of Applied Sciences: How do we begin?* Saxion Hogescholen. (Ongepubliceerd)
- Coombs, S. K., & Meijer, I. (2021). Towards evaluating the research impact made by University of Applied Sciences. *Science and Public Policy*, 1-9. <https://doi.org/10.1093/scipol/scab009>
- Davies, R., & Dart, J. (2005). The 'Most Significant Change' (MSC) Technique.
- De Jong, S. P., Van Arensbergen, P., Daemen, F., Van der Meulen, B., & Van den Besselaar, P. (2011). Evaluation of research in context: an approach and two cases. *Research Evaluation*, 20(1), 61-72.

- De Jonge, J. (2016). Praktijkgericht onderzoek bij lectoraten van hogescholen. *Feiten en Cijfers*, 19. Rathenau Instituut.
- De Poot, H. (2019). *PPS Impactmeting Katapult Netwerk*. Nobis.
- De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, 25(2), 161-169.
- Deuten, J., Koens, L., Faasse, P., & Van der Meulen, B. (2015). *Protocol voor de monitoring en evaluatie van de Toegepast Onderzoeksorganisaties in Nederland*. Rathenau Instituut.
- ERiC. (2011). *Handreiking. Evaluatie van maatschappelijke relevantie van wetenschappelijk onderzoek*. ERiC.
- Finne, H., Day, A., Piccaluga, A., Spithoven, A., Walter, P., & Wellen, D. (2011). *A composite indicator for knowledge transfer. Report from the European Commission's expert group on knowledge transfer indicators*. EU.
- Franken, A., Andriessen, D., Van der Zwan, F., Kloosterman, E., & Van Ankeren, M. (2018). *Meer waarde met hbo. Doorwerking praktijkgericht onderzoek van het hoger beroepsonderwijs*. Vereniging Hogescholen.
- Friedman, M. (2005). *Trying hard is not good enough*. Fiscal Policy Studies Institute.
- Godin, B. (2006). The linear model of innovation: The historical construction of an analytical framework. *Science, Technology & Human Values*, 31(6), 639-667.
- Godin, B. (2010). *National innovation system: A note on the origins of a concept. Project on the Intellectual History of Innovation*. (Manuscript)
- Godin, B., & Doré, C. (2005). *Measuring the impacts of science: Beyond the economic dimension*. INRS Urbanisation, Culture et Société. HIST Lecture, Helsinki Institute for Science and Technology Studies, Helsinki, Finland.
- Graeber, D., & Wengrow, D. (2021). *The dawn of everything. A new history of humanity*. Allen Lane.
- Greenhalgh, T., Ouseiko, P. V., Fahy, N., Shaw, S., Kerr, P., Rushforth, A. D., Channon, K. M., & Kiparoglou, V. (2017). Maximising value from a United Kingdom biomedical research centre: study protocol. *Health Research Policy and Systems*, 15(1), 1-17.
- Greenhalgh, T., Raftery, J., Hanney, S., & Glover, M. (2016). Research impact: a narrative review. *BMC Medicine*, 14(1), 1-16.
- Greven, K., & Andriessen, D. (2019). *Practice-based Research Impact Model for Evaluation: PRIME*. EAIR 41st Annual Forum, Leiden, The Netherlands.
- Holi, M. T., Wickramasinghe, R., & Van Leeuwen, M. (2008). *Metrics for the evaluation of knowledge transfer activities at universities*. Library House.
- James, C. (2011). *Theory of change review. A report commissioned by Comic Relief*. Comic Relief.
- Joly, P.-B., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., & Matt, M. (2015). ASIRPA: A comprehensive theory-based approach to assessing the societal impacts of a research organization. *Research Evaluation*, (24), 440-453.

- Kabinet-Rutte 3. (2019). *Versterken van onderzoeks- en innovatie-ecosystemen*. Den Haag.
- Keustermans, L., Wells, G., Maes, K., Ruiter, E., Alexander, D., Meads, C., & Noble, A. (2018). *Impact and the next Framework Programme for Research and Innovation (FP9)*. LERU.
- Kline, S. J., & Rosenberg, N. (1986). An overview of innovation. In R. Landau & N. Rosenberg, *The positive sum strategy*. National Academy Press. (275-305)
- KNAW, VSNU & NWO. (2016). *Standard Evaluation Protocol 2015 – 2021. Protocol for research assessments in the Netherlands*. KNAW, VSNU & NWO.
- Kok, M. O., & Schuit, A. J. (2012). Contribution mapping: a method for mapping the contribution of research to enhance its impact. *Health Research Policy and Systems*, 10(21).
- Lau, A. M. S. (2016). 'Formative good, summative bad?' – A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, 40(4), 509-525.
- McCawley, P. F. (2001). *The Logic Model for program planning and evaluation*. University of Idaho.
- Ministerie van OCW, Nationaal Regieorgaan Praktijkgericht Onderzoek SIA & Vereniging Hogescholen. (2019). *Verkenning praktijkgericht onderzoek op hogescholen*. Den Haag.
- Moerman, P., Vos, J., & Boode, F. (2016). *Samenwerking in cijfers. Tussen beroepsopleiding en bedrijfsleven*. Katapult.
- Nederland Ondernemend Innovatieland Innovatieplatform. (2009). *Van voornemens naar voorsprong: Voorstel voor een Nederlandse valorisatieagenda*. Interdepartementale programmadirectie Kennis en Innovatie.
- NWO. (2019). *Maatschappelijke impact via kennisbenutting*. NWO.
- NWO. (2020). *Knowledge and Innovation Covenant. Funding instruments & impact strategy*. NWO.
- Pedersen, D. B., Grønvad, J. F., & Hvidtfeldt, R. (2020). Methods for mapping the impact of social sciences and humanities – A literature review. *Research Evaluation*, 0(0), 1-18. <https://doi.org/10.1093/reseval/rvz033>
- Pijlman, H., Andriessen, D., Goumans, M., Jacobs, G., Majoor, D., Cornelissen, A., de Jong, H. et al. (2017). *Advies werkgroep Kwaliteit van Praktijkgericht Onderzoek en het Lectoraat*. Vereniging Hogescholen.
- QRiH. (2017). *Handleiding evaluatie van geesteswetenschappelijk onderzoek volgens het SEP*. <https://www.qrih.nl/nl/over-qrih/de-handleiding>
- Raftery, J., Hanney, S., Greenhalgh, T., Glover, M., & Blatch-Jones, A. (2016). Models and applications for measuring the impact of health research: update of a systematic review for the Health Technology Assessment programme. *Health Technology Assessment*, 20(76).
- Reeler, D. (2007). *A three-fold theory of social change*. The Community Development Resource Association.
- Reiner, C., Bekke, H., Hooghiemstra, E., Van Mil, T., de Ruiter, H., & Rullens, L. (2019). *Centres of Expertise: groeibriljant voor excellente samenwerking in het hbo. In allianties werken aan maatschappelijke impact voor de toekomst*. Vereniging Hogescholen.

- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29-48.
- Rosenthal, U., & Bovens, J. J. G. (2018). *Verspreiding. De onderbelichte kant van innovatie*. AWTi.
- Rosenthal, U., Heimovaara, S., Van der Giessen, A.M., & Knops, H. P. A. (2019). Het stelsel op scherp gezet. *Th&ma*, (3), 43-49.
- Scholtes, A., Van Vught, F., De Haas, M., & et al. (2011). *The EDUPROF project: developing indicators of applied research*. European Union, Education and Culture DG.
- Schot, J., & Steinmueller, W. E. (2016). *Framing innovation policy for transformative change: Innovation policy 3.0*. SPRU Science Policy Research Unit, University of Sussex.
- Sivertsen, G., & Meijer, I. (2020). Normal versus extraordinary societal impact: how to understand, evaluate, and improve research activities in their relations to society? *Research Evaluation*, 29(1), 66-70.
- Smeulders, A. (2019). *De HvA van waarde*. Hogeschool van Amsterdam.
- Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation: a review of societal impact assessment methods. *Research Evaluation*, 30(3), 323-335.
- Spaapen, J., & Van Drooge, L. (2011a). *SIAMPI final report: Social Impact Assessment Methods for research and funding instruments through the study of productive interactions between science and society*.
- Spaapen, J., & Van Drooge, L. (2011b). Introducing 'productive interactions' in social impact assessment. *Research Evaluation*, 20(3), 211-218.
- Termeer, C., & Dewulf, A. (2019). A small wins framework to overcome the evaluation paradox of governing wicked problems. *Policy and Society*, 38(2), 298-314.
- Termeer, K. (2019). *Het bewerkstelligen van een transitie naar kringlooplandbouw*. Wageningen University & Research.
- Van Boxsel, M. (2016). *De draagbare encyclopedie van de domheid*. Querido.
- Van Bree, T., Speijer, F., Suurs, R., & Willems, M. (2020). *Regionale innovatie-ecosystemen. Onderzoek naar optimale vormgeving van en dynamiek in regionale ecosystemen*. TNO.
- Van de Sanden, R., Arends, I., Van Erp, J., Franken, A., Van Laarhoven, H., Meijers, A., Mol, A., & Soete, L. (2018). *Maatschappelijke impact in kaart*. KNAW.
- Van den Akker, W., & Spaapen, J. (2017). *Productive interactions: societal impact of academic research in the knowledge society. LERU position paper*. LERU.
- Van den Bosch, S., & Rotmans, J. (2008). *Deepening, Broadening and Scaling up: a Framework for Steering Transition Experiments*. TNO.
- Van den Toren, J. P., & Van Vliet, H. (2021). *Naar een nieuw impactkader voor Regieorgaan SIA*. Regieorgaan SIA.
- Van der Zwan, F., & Van der Vos, J. (2015). *Notitie valorisatie-indicatoren voor bcoz*. Vereniging Hogescholen.

- Van Drooge, L., & Spaapen, J. (2017). Evaluation and monitoring of transdisciplinary collaborations. *The Journal of Technology Transfer*. <https://doi.org/10.1007/s10961-017-9607-7>
- Van Drooge, L., Vandenberg, R., Zuijdam, F., Mostert, B., Van der Meulen, B., & Bruins, E. (2011). *Waardevol. Indicatoren voor Valorisatie*. STW, Rathenau Instituut en Technopolis.
- Van Mierlo, B., Regeer, B., Van Amstel, M., Arkesteijn, M., Beekman, V., Bunders, J., De Cock Buning, T., Elzen, B., Hoes, A., & Leeuwis, C. (2010). *Reflexieve monitoring in actie. Handvatten voor de monitoring van systeeminnovatieprojecten*. Boxpress.
- Van Staa, A., & Mobach, M. P. (2020). Randvoorwaarden voor succesvol praktijkgericht onderzoek. *Thema*, 27(1), 61-67.
- Van Vliet, H. (2019). *Doorwerking. Deconstructie & Reconstructie*. Hogeschool van Amsterdam. (white paper).
- Van Vliet, H. (2020). *Kritische reflectie. Een oneerlijke maar leerzame kritische reflectie op de Kritische Zelfreflecties Kenniscentra 2018*. Hogeschool van Amsterdam.
- Van Vliet, H. (2021). *Kritische reflecties: nwo-model van impact*. Regieorgaan SIA.
- Van Vliet, H. (2022). *Kritische reflecties: analyse van de SPRONG-regeling en -aanvragen*. Regieorgaan SIA.
- Van Vliet, H., & Slotman, R. (2006). De regio als basis voor innovatie in het mkb. *Th&ma*, 6(1), 30-36.
- Van Vliet, H., Wakkee, I., Fukkink, R., Teepe, R., & Van Outersterp, D. (2020). *Rapporteren over doorwerking van Praktijkgericht Onderzoek*. Hogeschool van Amsterdam.
- Von Hippel, E. (1994). 'Sticky information' and the locus of problem solving: implications for innovation. *Management Science*, 40(4), 429-439.
- Vankan, A., Den Hertog, P., Jansen, M., De Boer, P. J., & Smeitink, A. (2020). *Onderzoeks- en innovatie-ecosystemen in Nederland*. Dialogic.
- Vereniging Hogescholen. (2015a). *#hbo2025 Wendbaar en Weerbaar*. Vereniging Hogescholen.
- Vereniging Hogescholen. (2015b). *Brancheprotocol Kwaliteitszorg Onderzoek 2016-2022. Kwaliteitszorgstelsel Praktijkgericht Onderzoek Hogescholen*. Vereniging Hogescholen.
- VSNU. (2013a). *Een Raamwerk Valorisatie-indicatoren*. VSNU.
- VSNU. (2013b). *Keuzemenu valorisatie indicatoren*. VSNU.
- VSNU. (2017). *VSNU-handreiking tabellen Standaard Evaluatie Protocol (SEP) 2015-2021*. VSNU.
- Wakkee, I., Van der Sijde, P., & Nuijens, N. (2013). *Valorisatie in Nederland. Exploratieve verkenning van het landschap van Valorisatieprogramma's*. Vrije Universiteit Amsterdam.
- Wetenschappelijke Raad voor het Regeringsbeleid (WRR). (2013). *Naar een lerende economie. Investeren in het verdienvermogen van Nederland*. Amsterdam University Press.
- W.K. Kellogg Foundation (2004). *Logic Model Development Guide*. W.K. Kellogg Foundation.

Het aantonen van de bijdrage van onderzoek is ingewikkeld, dat is geen conclusie maar een vertrekpunt geweest voor de voorbeschouwing in deze lectorale rede. Sturende vraag daarbij is: waar zit die ingewikkeldheid dan in? Het dominante antwoord dat zich in vele praktijken heeft genesteld slaat het ingewikkelde plat in een lijn. Als houvast. Het woord doorwerking wordt gebruikt als koevoet om dit open te breken,

niet in de laatste plaats om meer recht te doen aan de praktijk van het onderzoek aan hogescholen. Er opent zich zo een andere manier van kijken naar de bijdrage van praktijkgericht onderzoek aan veranderingen: van doorwerking 'in de breedte' en indicatoren voor de effecten van directe interacties, tot de meerdere handelingsperspectieven dan alleen nieuwe kennis om veranderingen te bewerkstelligen.

