

De evaluatie van universitair onderzoek Methodiek voor het incorporeren van de maatschappelijke waarde in onderzoek

Drs. Frank J.M. Wamelink
Dr. Jack B. Spaapen

sci_Quest, Maarssen-Amsterdam

Mei 1999

Colofon

De evaluatie van universitair onderzoek
Methodiek voor het incorporeren van de maatschappelijke waarde van onderzoek

Deze publikatie is een bewerking van een studie die medio 1998 is uitgevoerd in opdracht van de Landbouwuniversiteit Wageningen, de Nationale Raad voor Landbouwkundig Onderzoek en de Commissie van Overleg Sectorraden door:

Drs. F.J.M. Wamelink
Dr. J.B. Spaapen
sci_Quest, Maarssen-Amsterdam

Personen die worden bedankt voor hun medewerking

De begeleidingscommissie:

Mevrouw Ir. C.T. Kandou, Landbouwuniversiteit Wageningen
de heer Dr. B.J. Blaauboer, VSNU
de heer Ir. N.A. Dijkveld Stol, NRLO
de heer Drs. P. Morin, Commissie van Overleg Sectorraden

Voor het aanleveren van informatie:

Landbouwuniversiteit Wageningen, afdeling: Onderzoeks- en Onderwijsbeleid
Met name: Mevrouw Ir. C.T. Kandou en mevrouw Ir. I. Hijmans

De betrokken onderzoekers die tijd hebben vrijgemaakt om de gevraagde informatie te verzamelen en aan te leveren.

Voor commentaar:

De leerstoelgroep i.o. Wetenschaps- en Technogiedynamica van de Universiteit van Amsterdam.
Het Bureau van de Nationale Raad voor Landbouwkundig Onderzoek
De visitatiecommissie Landbouwwetenschappen

Voor hun bijdrage aan het gebruikersonderzoek:

Huub Dijstelbloem en Drs. Jet Bout (Conscience)

Ontwerp en productie Ottevanger en van de Kuinder
Lithografie en druk Ronaveld

**Nationale Raad voor
Landbouwkundig Onderzoek**
Postbus 20401
2500 EK Den Haag
tel.: 070 378 56 53
internet: <http://www.agro.nl/nrlo/>

**Commissie van Overleg
Sectorraden**
Postbus 25000
2700 LZ Zoetermeer
tel.: 079 323 35 40

ISBN: 90 - 5059 - 099 - 3

Overname van tekstdelen is toegestaan, mits met bronvermelding.
NRLO-rapport nr. 99/12, Den Haag, mei 1999

Deze publikatie is een bewerking van een studie die medio 1998 is uitgevoerd in opdracht van de LUW, de NRLO en de COS. Die studie betrof een zogenaamde additionele beoordeling van landbouwkundig onderzoek in het kader van de reguliere VSNU-onderzoeksvisitatie. Het ging om twaalf onderzoeksprogramma's die najaar 1998 werden beoordeeld door een internationale visitatiecommissie. Ten behoeve van die commissie heeft bureau sci_Quest een onderzoek uitgevoerd, dat tot doel had de commissie te voorzien van systematische informatie omtrent de 'inbedding' en *performance* van onderzoeksprogramma's in de voor hen relevante maatschappelijke omgeving. Tevens is ten behoeve van de commissie een survey gehouden onder gebruikers van het onderzoek van de twaalf groepen teneinde de aansluiting van onderzoeksaanbod op de maatschappelijke vraag te analyseren. Het onderzoek als geheel beoogde de commissie te ondersteunen bij het beoordelen van verschillen in onderzoeksoriëntatie tussen onderzoeksprogramma's, en van de prestaties van programma's in relatie tot hun maatschappelijke missie of doelstellingen.

De relatie tussen onderzoeksprogramma's en de hen omringende context staat centraal in de aanpak. Enerzijds is de interactie tussen onderzoek en omgeving in kaart gebracht aan de hand van indicatoren, welke samengebond zijn in een radar-achtig figuur. Anderzijds is die interactie via een survey onder gebruikers meer kwalitatief in beeld gebracht.

Het resultaat van het onderzoek is verschenen in een rapport dat aan de visitatiecommissie ter hand is gesteld en haar heeft geholpen de programma's te beoordelen. Belangrijk daarbij is dat de beoordelingscommissie en de beoordeelde de profielen en de gebruikersanalyse aanvaardden als een overzichtelijke en overtuigende representatie van de activiteiten en de *performance* van de groep, en dat deze als uitgangspunt kan worden genomen voor de beoordeling.

Teneinde de discussie over evaluatie van wetenschappelijk onderzoek in de context van maatschappelijke vraagstellingen te bevorderen hebben NRLO en COS aan sci_Quest gevraagd het rapport om te werken naar een publikatie van meer algemene strekking. De nadruk ligt hierbij niet meer op een oordeel over de twaalf onderzoeksprogramma's, maar op een uiteenzetting van de aanpak. Wel zijn twee bestudeerde programma's bij wijze van voorbeeld in een bijlage bij dit rapport besproken.

Wij hopen dat deze publikatie zal bijdragen aan het inzicht in de problematiek van het evalueren van wetenschappelijk onderzoek in de context van maatschappelijke vraagstellingen, en aan de discussie die hierover in brede kring plaatsvindt in het kader van wetenschaps- en innovatiebeleid.

Discussie over de haalbaarheid van de methode kan niet stoppen bij aspecten van dataverzameling en methodiek. Minstens even belangrijk is dat alle betrokken groeperingen overeenstemming bereiken over het uitgangspunt dat de relevante context van onderzoeksprogramma's onderling sterk kunnen verschillen en dat zulks gevolgen heeft voor de beoordelingsprocedure.

Bij het tot stand komen van dit rapport hebben de opstellers veel steun gehad aan het commentaar van de leden van de begeleidingscommissie. Ook de commentaren van de VSNU-beoordelingscommissie en enkele sleutelfiguren uit de landbouwwereld waren van grote waarde, zowel waar het de inhoud als de vormgeving betreft. Desalniettemin blijven de auteurs volledig verantwoordelijk voor de inhoud van het rapport.

Prof.Dr.Ir. A. Rörsch
voorzitter NRLO

Prof.Dr. H.O. Voorma
voorzitter COS

Inhoudsopgave

	Ten Geleide	3
	Inhoudsopgave	5
0	Samenvatting	7
0.1	Maatschappelijke betekenis en onderzoeksbeoordeling	7
0.2	Opzet van het onderzoek	7
0.2.1	Het 'Research Embedment and Performance Profile' (REPP)	8
0.2.2	Het gebruikersonderzoek	8
0.2.3	Terugkoppeling op de missie van het programma	9
0.3	Bevindingen in het licht van ontwikkeling naar een reguliere beoordelingspraktijk	9
1	Ruimte voor de maatschappelijke betekenis van onderzoek in de beoordeling	11
1.1	Inleiding	11
1.2	De maatschappelijke betekenis van landbouwwetenschappelijk onderzoek	11
1.3	De onderzoeksopdracht	11
1.4	Het theoretisch kader	12
1.5	Indeling van deze nota	12
2	Het 'Research Embedment and Performance Profile' (REPP)	13
2.1	Profiel van onderzoek in verschillende sociale domeinen	13
2.2	Domeinen en indicatoren voor 'inbedding' en <i>performance</i>	13
2.2.1	Wetenschap en 'gecertificeerde' kennis	14
2.2.2	Opleiding en training	14
2.2.3	Innovatie en professionals	15
2.2.4	Beleid en maatschappelijke vraagstukken (<i>Public policy</i>)	15
2.2.5	Interne en externe samenwerking en zichtbaarheid (<i>Collaboration and visibility</i>)	15
2.3	Het radar-profiel (REPP)	16
2.4	Methodische overwegingen	16
2.4.1	Uniformiteit van de data	16
2.4.2	Verifiëren van 'inbedding' en ' <i>performance</i> '	16
2.4.3	Omvang van de groepen	17
2.4.4	Vergelijkbare schalen	17
2.4.5	Afgrenzing van de programma's	18
2.4.6	Statische representatie en dynamiek	18
2.4.7	Zelfbeeld van de groep	18
2.5	Typologie van groepen op grond van de interactie-patronen	18
3	Gebruikersonderzoek	21
3.1	Het 'gebruik' van wetenschappelijk onderzoek	21
3.2	Gebruikersonderzoek in het kader van de LUW-beoordeling	21
3.3	Deel 1: Kartering. Structuur van de gebruikersomgeving	23
3.4	Deel 2: Survey. Kwalitatieve analyse van de interactie	24
4	Terugkoppeling	27
4.1	Inleiding	27
4.2	'Embedment', ' <i>performance</i> ' en leeromgeving combineren	27

5	Bevindingen in het licht van ontwikkeling naar een reguliere beoordelingspraktijk	29
5.1	Inleiding	29
5.2	De aanpak algemeen	29
5.3	REPP	30
5.4	Gebruikersonderzoek	31
5.5	Terugkoppeling	32
5.6	Implementatie in een reguliere beoordelingspraktijk: technische en culturele haalbaarheid	32
	Afkortingenlijst	34
	Literatuur	35
Bijlage 1:	Twee voorbeelden van een REPP	39
Bijlage 2:	Gebruikersonderzoek: twee voorbeelden; vragenlijst	47
Bijlage 3:	Gebruikte data en bronnen	53
Bijlage 4:	Theoretische achtergrond bij de methodiek	57
Bijlage 5:	Mogelijke transformaties van de REPPS	59
	Noten	60

0 Samenvatting

0.1 Maatschappelijke betekenis en onderzoeksbeoordeling

In dit rapport wordt een methode beschreven om de maatschappelijke betekenis van wetenschappelijke onderzoeksprogramma's systematisch en in overeenstemming met de oriëntatie van het programma te incorporeren in onderzoeksbeoordelingen. Deze methode is ontwikkeld in het kader van de VSNU onderzoeksbeoordeling Landbouwwetenschappen 1998^①. De daar toegepaste methodiek wordt in deze nota gerapporteerd met het oog op ontwikkeling naar een reguliere beoordelingspraktijk.

Grote delen van het wetenschappelijke onderzoek zijn niet alleen op het wetenschappelijk domein gericht, maar ontwikkelen zich ook in samenhang met maatschappelijke terreinen. Daarom wordt gezocht naar methoden van onderzoeksbeoordeling waarin de maatschappelijke betekenis van onderzoek systematisch en evenwichtig wordt geïncorporeerd^②. Evenwichtig impliceert rekening houden met de oriëntatie van een onderzoeksprogramma, die kan variëren op de schaal van fundamenteel naar toegepast.

De methode die hier wordt gepresenteerd wijkt op veel punten af van een methodiek die 'objectieve' indicatoren voor kwaliteit construeert (bibliometrie), of een methodiek die alleen aansluit bij de criteria van de collega wetenschappers (*peer review*). In de methode wordt geprobeerd om de beoordelingscommissie de data zo aan te leveren dat een systematisch beeld ontstaat van de gehele context waarin de groep functioneert, wetenschappelijk en overige gebruikers. Zo wordt de groep betrokkenen die kan worden meegewogen in de beoordeling sterk verbreed.

0.2 Opzet van het onderzoek

Het onderzoek is uitgewerkt in drie deelonderzoeken. In de eerste plaats is een specifiek 'profiel' van de onderzoeksprogramma's geconstrueerd. In de tweede plaats is een gebruikersanalyse

uitgevoerd waarin belanghebbende gebruikers^③ (*stakeholders*) is gevraagd naar de betekenis van het programma voor hun activiteiten. In de derde plaats vindt een terugkoppeling plaats naar de missie van het onderzoeksprogramma op grond van de resultaten van de eerste twee deelonderzoeken. De eerste twee deelonderzoeken betreffen de interactie tussen onderzoekers en gebruikers, maar het REPP kiest het perspectief van de onderzoekers, terwijl het gebruikersonderzoek kijkt vanuit het standpunt van de overige actoren in het innovatieproces. Het REPP is gebaseerd op de verzameling van veel, meest kwantitatieve data, het gebruikersonderzoek is het resultaat van een relatief eenvoudig *survey*. Het eerste deelonderzoek leidt tot een grafische representatie van de onderzoekseenheid in relatie tot de omringende context, het tweede onderdeel leidt tot een kwalitatieve identificatie van de innovatie-context in termen van leeromgevingen. Het derde deelonderzoek, de terugkoppeling, is bedoeld om de resultaten van beide andere onderdelen met elkaar in verband te brengen en zo een discussie tussen belanghebbenden mogelijk te maken.

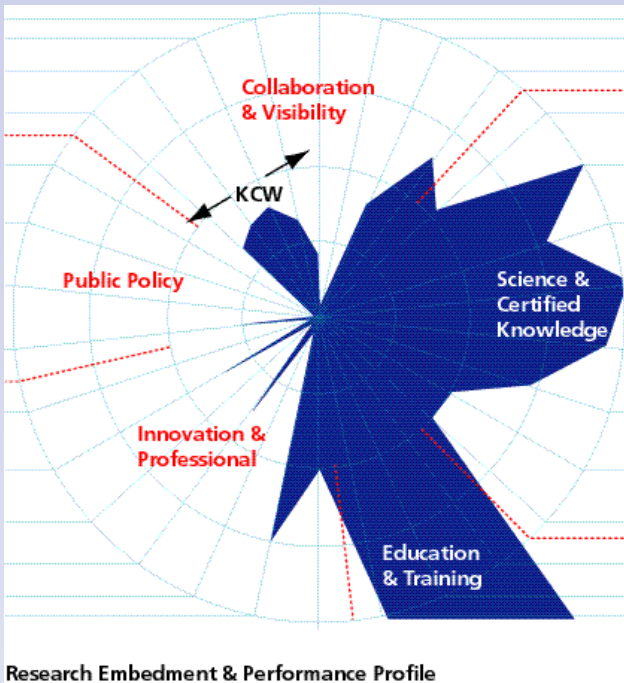
Het uitgangspunt is dat onderzoeksprogramma's zich ontwikkelen in wederzijdse interacties met de brede omgeving waarin zij zijn ingebed. Het succes van een onderzoeksprogramma hangt af van de wijze waarop men in het onderzoek aansluit bij thema's die in de omgeving aan de orde zijn en van de wijze waarop deze omgeving de kennisproducten accepteert en consolideert (zie bijlage 4 voor theoretische achtergrond).

0.2.1 Het 'Research Embedment and Performance Profile' (REPP)

Het profiel beoogt een representatie te zijn van de soort en mate van activiteit (*performance*) van een programma. In het REPP worden vijf domeinen onderscheiden met eigen criteria voor de beoordeling van kennisproducten. De domeinen zijn verbonden met de verschillende sectoren van maatschappelijke actoren waar een onderzoeksprogramma relaties mee kan hebben. In deze domeinen bestaan verschillende verwachtingspatronen ten aanzien van het onderzoek. Deze

verschillen beïnvloeden de evaluatiecriteria die in de domeinen van belang worden geacht. De volgende domeinen worden onderscheiden:

Figuur 0.1



- Wetenschap en gecertificeerde kennis.
- Onderwijs en training.
- Innovatie en professionals.
- Beleid en maatschappelijke vraagstukken.
- Interne (binnen de instelling) en externe samenwerking en zichtbaarheid⁴.

Voor ieder domein wordt een reeks indicatoren bepaald. Daarmee ontstaat een systematisch beeld van productie en impact in elk van de domeinen, en tevens van de *performance* van het programma in zijn specifieke omgeving.

Het profiel geeft zo een beeld van de domeinen waarin het programma zich meer of minder sterk ontwikkelt. Het programma in figuur 0.1 is bijvoorbeeld zeer sterk gericht op wetenschap en op onderwijs en training; het is daar in hoge mate ontwikkeld over vrijwel alle indicatoren; het heeft een kleine uitschieter in het domein van 'innovatie en professionals' en vertoont een matige tot geringe interactie met het KCW⁵. In deze voorbeeld-grafiek zijn de namen van de indicatoren weggelaten om de systematiek van het profiel over te brengen.

0.2.2 Het gebruikersonderzoek

Wat het profiel niet aangeeft is in hoeverre de interacties waarin het programma betrokken is ook leiden tot min of meer stabiele vormen van samenwerking waarin de belangen en gezichtspunten van gebruikers en andere belanghebbenden *stakeholders* aan bod komen. Hierdoor wordt niet (of onvoldoende) zichtbaar wat de rol van de gebruikers is in de ontwikkeling van (kennis)producten. Daarbij spelen vragen als: hebben gebruikers invloed op de onderzoeksagenda, is er voorzien in *feedback* vanuit de gebruikers op de innovatie, enzovoorts. Daarom is een onderzoek onder gebruikers en andere belanghebbenden een noodzakelijke aanvulling op het REPP.

De gebruikersanalyse kent twee onderdelen:

1) het in kaart brengen van de gebruikersomgeving en 2) het bevragen van een aantal geselecteerde gebruikers. In het eerste onderdeel worden gebruikers geïdentificeerd en ingedeeld in categorieën aan de hand van de projectlijsten. Zo ontstaat een beeld van wie het onderzoek gebruikt en op welke wijze. In het tweede onderdeel worden gebruikers benaderd met een vragenlijst waarin een aantal aspecten van de interactie met de onderzoeksgroep aan de orde komt, bijvoorbeeld: de wijzen van interactie; het doel van de interactie; de aard van het gebruik en het nut van het onderzoek voor de gebruiker. Het uiteindelijke doel van dit onderdeel is het karakteriseren van de interactie tussen onderzoekers en omgeving in termen van leerprocessen. Uitgangspunt is dat het onderzoek alleen de weg naar de praktijk vindt wanneer de verschillende actoren die betrokken zijn bij het innovatieproces (zowel onderzoekers als gebruikers) van elkaar leren. Deze leerprocessen zullen verschillen vertonen in verschillende fasen van het onderzoeksproces. We onderscheiden drie vormen van leerprocessen:

- 1 Een fase waarin behoeften worden gearticuleerd.
- 2 Een fase waarin nadere afstemming plaats van belangen van actoren.
- 3 Een experimentele fase.

0.2.3 Terugkoppeling op de missie van het programma

Tenslotte worden de resultaten van het REPP en het gebruikersonderzoek met elkaar in verband gebracht. Deze vergelijking is gebaseerd op de verwachting dat groepen met een bepaald profiel een leeromgeving hebben die bij dat profiel 'past'. Dat wil bijvoorbeeld zeggen dat van groepen die in hun profiel *industry driven* blijken, wordt verwacht dat zij zich in een leeromgeving bevinden waarin experimenteel testen dominant is. Geanalyseerd dient dus te worden in hoeverre de gevonden REPP-profielen en typen leeromgeving met elkaar overeenkomen. Het doel van de vergelijking is discussie met de onderzoeksgroepen op gang te brengen over hun activiteiten en producten in het licht van hun missie.

0.3 Bevindingen in het licht van ontwikkeling naar een reguliere beoordelingspraktijk

In het onderzoek dat is uitgevoerd voor de Landbouwuniversiteit bleek het mogelijk om de interacties met de brede omgeving van onderzoekprogramma's op betrouwbare wijze in kaart te brengen. De VSNU-beoordelingscommissie was van mening dat de geconstrueerde profielen (REPP) een duidelijke meerwaarde vertegenwoordigden ten opzichte van de zelfevaluatie van de groepen. Het REPP heeft een dialoog bevordert tussen de commissie en het onderzoeksprogramma waarin de missie, de claims over de resultaten en de verificatie daarvan tegelijk aan de orde zijn en in verband worden gebracht met toekomstige mogelijkheden van de groep.

De VSNU-beoordelingscommissie onderkende de aanvullende waarden van de uitgevoerde gebruikersanalyse. De grafieken van gebruikers (institutionele achtergrond en rollen) gaven een overzicht dat niet (of met veel inspanning) kan worden verkregen op grond van de zelfevaluatie. De gebruikersrapporten - waarin de resultaten van de interviews per onderzoeksprogramma werden samengevat - bevatte volgens de commissie veel bruikbare aanvullende informatie.

De clustering van onderzoeksgroepen in klassen bleek een weinig uitgesproken typologie van interactiepatronen op te leveren. Programma's zijn complexer en dynamischer dan kan worden

gerepresenteerd in deze classificaties.

De classificatie in termen van leerprocessen bleek wel mogelijk, maar de betrouwbaarheid hiervan werd beperkt door het relatief gering aantal geïnterviewde gebruikers.

Het routinematig toepassen van deze methodiek stuit op de moeilijkheid dat data voor belangrijke indicatoren op dit moment nog niet stelselmatig worden opgenomen in verslaglegging en administraties. Het verzamelen van de benodigde data heeft in dit project veel inspanning gekost. Deels is dit veroorzaakt doordat de methodiek nog moest worden ontwikkeld en als gevolg daarvan meer data en informatie zijn verzameld dan strikt genomen noodzakelijk was. Door deze bredere exploratie van de data kan nu wel worden aangegeven op welke punten aanpassingen denkbaar zijn in de verzameling van data om in evaluaties eenvoudiger het profiel van groepen te construeren en de gebruikersomgeving te analyseren.

Een discussie over de haalbaarheid van deze methode kan niet stoppen bij aspecten van dataverzameling en methodiek. Belangrijker is nog dat alle betrokken groepen overeenstemming bereiken over het uitgangspunt dat de relevante contexten van onderzoekprogramma's onderling sterk kunnen verschillen en dat dit gevolgen heeft voor de beoordelingsprocedure.

1 Ruimte voor de maatschappelijke betekenis van onderzoek in de beoordeling

1.1 Inleiding

In dit rapport wordt een methode besproken om de maatschappelijke betekenis van wetenschappelijke onderzoeksprogramma's systematisch te incorporeren in onderzoeksbeoordelingen. Deze methode is ontwikkeld in het kader van de VSNU onderzoeksbeoordeling Landbouwwetenschappen 1998[®]. De toegepaste methodiek wordt hier beschreven met het oog op ontwikkeling naar een reguliere beoordelingspraktijk.

In dit hoofdstuk wordt eerst kort het karakter van landbouwwetenschappelijk onderzoek beschreven, voornamelijk in relatie tot de maatschappelijke relevantie van het onderzoek. Vervolgens wordt de onderzoeksopdracht besproken en volgt een toelichting op de indeling van de notitie.

1.2 De maatschappelijke betekenis van landbouwwetenschappelijk onderzoek

Landbouwwetenschappen hebben een duidelijke maatschappelijke missie. Onderzoeksvragen worden mede ontwikkeld in interactie met verschillende belanghebbenden uit de samenleving. Landbouwwetenschappen zijn bijvoorbeeld verbonden met de landbouwsector die onder grote druk staat om zich aan te passen aan de veranderende maatschappelijke eisen (duurzaamheid, gezondheid, veiligheid, grotere rol van de vraagkant, meer marktwerking, enz.). Mede daardoor zijn het milieu, het landschap en natuurbehoud nadrukkelijk op de onderzoeksagenda gekomen. Landbouwwetenschappen worden nauw betrokken bij de herformulering van de problemen en een herdefinitie van de oplossingen. Dat komt ook tot uitdrukking in de missie van de Landbouwuniversiteit Wageningen:

De Landbouwuniversiteit wil de wetenschappelijke kennis ontwikkelen en uitdragen, die de samenleving nodig heeft om op duurzame wijze te voorzien in haar behoefte aan voldoende en gezond voedsel en een goed leefmilieu voor mens, plant en dier[®].

Dit heeft consequenties voor de organisatie en de aard van het onderzoek. Het onderzoek aan de Landbouwuniversiteit is multidisciplinair georganiseerd en wordt in veel gevallen uitgevoerd in een toegepaste context. De aandacht van het onderzoek in deze sector is in de laatste 10 tot 15 jaar verbreed naar landgebruik en de voedselketen en omvat nu ook milieu- en gezondheidsvraagstukken en sociaal-economisch gebruik van de 'groene-ruimte'.

1.3 De onderzoeksopdracht

De Landbouwuniversiteit wil de VSNU onderzoeksbeoordeling[®] dan ook afstemmen op de missie en de organisatie van het eigen landbouwwetenschappelijk onderzoek. sci_Quest is gevraagd om een aanvullende analyse uit te voeren aan de hand waarvan de visitatie-commissie het maatschappelijk belang van het programma systematisch zou kunnen meewegen in de evaluatie[®].

De opdracht aan sci_Quest was drieledig[®]:

- 1 De beoordelingscommissie voorzien van systematische informatie waardoor het maatschappelijk belang van onderzoek geïncorporeerd zou kunnen worden in de totale beoordeling van de programma's, waaronder dus ook de wetenschappelijke kwaliteit.
- 2 Deze data zo te presenteren dat de commissie in staat is om de legitieme verschillen in de oriëntatie van groepen op maatschappelijke dan wel wetenschappelijke vragen in de beoordeling te betrekken.

- 3 Informatie te verzamelen die relevant is voor de beoordeling van de *performance* van het programma in relatie tot de missie.

1.4 Het theoretisch kader

Als theoretisch kader is gekozen voor het uitgangspunt dat onderzoeksprogramma's zich ontwikkelen in wederzijdse interacties met de brede omgeving waarin zij zijn ingebed. Het succes van een onderzoeksprogramma hangt af van de wijze waarop men in het onderzoek aansluit bij thema's die in die omgeving aan de orde zijn en van de wijze waarop deze omgeving de kennisproducten accepteert en consolideert (zie bijlage 4 voor een nadere uitwerking van de theoretische achtergrond).

1.5 Indeling van deze nota

Deze vraagstelling is in drie deelonderzoeken uitgewerkt. In de eerste plaats is het 'profiel' van de onderzoeksprogramma's in zijn context geconstrueerd. Daarin ligt de nadruk op de activiteiten van een programma in verschillende domeinen. In de tweede plaats is een gebruikers-evaluatie uitgevoerd waarin belanghebbende gebruikers is gevraagd naar de betekenis van het programma voor hun activiteiten. Het '*Research Embedment and Performance Profile*' (REPP) wordt in hoofdstuk 2 besproken, de gebruikersstudie komt in hoofdstuk 3 aan de orde. In hoofdstuk 4 wordt vervolgens kort ingegaan op de wijze waarop per groep de resultaten van het REPP en de gebruikersanalyse met elkaar in verband zijn gebracht (deelonderzoek 3, terugkoppeling). In het slothoofdstuk wordt de bruikbaarheid van de methode in onderzoeksbeoordeling en onderzoeksbeleid besproken. Dat gebeurt mede aan de hand van een meta-evaluatie die is uitgevoerd naar de rol van de gepresenteerde methode in de beoordeling van de landbouwwetenschappelijke programma's.

In de hoofdtekst ligt het accent op de hoofdlijnen van het onderzoek. Om een indruk te geven van de resultaten van het Landbouwwetenschappelijk onderzoek zijn die van enkele programma's opgenomen in de bijlagen 1 en 2.

2 Het 'Research Embedment and Performance Profile' (REPP)

2.1 Profiel van onderzoek in verschillende sociale domeinen

Kennisproductie in onderzoeksprogramma's ontwikkelt zich in interactie met een brede omgeving waarin verschillende sociale domeinen kunnen worden onderscheiden. Bij de beoordeling van onderzoek zou dan ook de mate waarin de kennisproductie zich oriënteert op elk van die domeinen moeten worden betrokken. Om zo'n brede beoordeling mogelijk te maken is het 'Research Embedment and Performance Profile' (REPP) ontwikkeld.

Het REPP is op hoofdlijnen gebaseerd op de 'compass card' methodiek van Callon *et al.* (1994), die de diversiteit in kennisproductie bij een aantal onderzoeksgroepen op een systematische wijze in kaart brachten. Via de 'compass card of research' wordt de volle breedte van mogelijke sociale contexten voor kennisproductie en beoordeling van onderzoek opgenomen: van de honorering van kennisclaims in de internationale wetenschappelijke gemeenschap via de bijdrage aan concurrentievoordeel in de professionele en commerciële context naar de articulatie en oplossing van maatschappelijke of beleidsproblemen. Zo kan van elke mogelijke constellatie van kennisproductie in relatie tot een relevante sociale context een profiel worden geconstrueerd.

De 'compass card of research' onderscheidt in de context van een onderzoeksprogramma een vijftal sociale domeinen, waarin eigen criteria voor de beoordeling van kennisproducten zijn ontwikkeld. Deze zijn enigszins aangepast terug te vinden in de vijf domeinen die in voor het REPP worden gebruikt:

Sociale domeinen van kennisproductie:

- 1 **Science and certified knowledge:** productie van kennisclaims ter validatie door de wetenschappelijke gemeenschap (*peers*)
- 2 **Education and training:** productie van geschoolde onderzoekers en vaardigheden (in het Engels vaak *embodied of tacit knowledge* genoemd)
- 3 **Innovation and professional:** productie van kennis met het oog op concurrentie-voordeel
- 4 **Public policy:** productie van kennis en vaardigheden ten behoeve van beleids- of maatschappelijke doelstellingen
- 5 **Collaboration and visibility:** hier wordt de 'interne' oriëntatie en *performance* in de bijdrage aan de doelstellingen van de Landbouwuniversiteit Wageningen vergeleken met de oriëntatie op andere Nederlandse instellingen en buitenlandse instellingen.

De studie van Callon *et al.* beperkt zich tot domein-indicatoren waarvoor de data beschikbaar zijn in databanken. In het geval van de Landbouwuniversiteit is een survey onder de onderzoeksprogramma's uitgezet om extra data te verzamelen. Dat was nodig omdat niet alle door ons gewenste data administratief worden bijgehouden. Bovendien was het zo mogelijk om te reflecteren op de meest adequate data en indicatoren.

2.2 Domeinen en indicatoren voor inbedding en performance

Indicatoren voor de verschillende domeinen zijn ontwikkeld langs vier 'kanalen' van interactie die door Callon *et al.* worden onderscheiden: teksten, mensen, artefacten en financiering. Voor elk van deze kanalen wordt per domein bezien welke indicatoren mogelijk en zinvol zijn. In onze benadering verwijzen we naar het patroon van interactie- en communicatieprocessen van een onderzoeksprogramma als de 'inbedding' van het

programma. Naar het gebruik en de 'impact' verwijzen we als de *performance* van het programma. 'Inbedding' en *performance* zijn niet geheel onafhankelijk van elkaar. Sommige kanalen van interactie vertegenwoordigen deze beide kanten. De financiering van contractonderzoek bijvoorbeeld is zowel een 'inbedding' van het onderzoeksprogramma in thema's van de specifieke context als een indicatie voor de (verwachte) bruikbaarheid van het onderzoek bij de opdrachtgevers.

Er is een rijke set beschikbaar aan indicatoren en indicaties voor interactie en impact. Voorbeelden zijn: co-publicaties, gedeelde onderzoeksstaf, samenwerking met de professionele sector of met het bedrijfsleven, onderzoek in opdracht, professionele publicaties, wetenschappelijke artikelen, mobiliteit van de staf, adviesrelaties en lidmaatschappen van beleidsremia, deelname aan speciale programma's, publicaties in gerefereerde tijdschriften en patenten.

Veel van deze indicaties of indicatoren worden al gebruikt in beoordelingsprocedures voor wetenschappelijke programma's. De indicatoren verschillen wel in de mate van robuustheid. Het belangrijkste verschil tussen onze aanpak en andere op dit punt is dat indicatoren niet worden gebruikt om 'kwaliteit' of een van de beleidsmatig verdunde vormen van kwaliteit te representeren. Wij construeren een specifieke constellatie 'profiel' van kennisproductie in relatie met de context, waarin de strategische keuzen van het programma herkenbaar zijn. Bij ieder van de onderscheiden domeinen is een relevante set indicatoren gekozen die inzicht kunnen geven in de mate waarin de 'inbedding' en de *performance* in ieder van die domeinen is geëvolueerd. Hieronder worden per domein de belangrijkste indicatoren aangegeven. Deze hoeven niet altijd allemaal te worden geanalyseerd. Per geval moet een voor de specifieke vraagstelling relevante keuze worden gemaakt. Verderop geven wij de precieze keuze die wij in het geval van de LUW hebben gemaakt.

2.2.1 Wetenschap en 'gecertificeerde' kennis

Toetsing en kwaliteitscontrole zijn binnen de wetenschappelijke gemeenschap geïnstitutionaliseerd in verschillende vormen. Een van de

bekendste is het *referee-process* voor de acceptatie van artikelen. Daarnaast zijn er gevestigde processen van middelentoedeling op grond van criteria voor wetenschappelijke reputatie of excellentie. De bibliometrie heeft een groot aantal indicatoren ontwikkeld die inzage geven in de verspreiding van kennisclaims en de productie voor de internationale wetenschappelijke gemeenschap. Bijvoorbeeld de aantallen referenties die een indicatie geven van de omvang van de gemeenschap waarin de kennisclaim *impact* heeft. Het aantal publicaties in internationale wetenschappelijke tijdschriften karakteriseert het wetenschappelijke engagement van het programma. Het hebben van een positie als lid van de redactie van een wetenschappelijk tijdschrift geeft wetenschappers de rol van 'poortwachters' van de wetenschappelijke maatstaf. Deze lidmaatschappen bevestigen de oriëntatie van de groep op de wetenschappelijke gemeenschap, evenals de uitstroom van staf naar onderzoeksinstituten. Daarnaast is ook de samenwerking van groepen met andere universitaire groepen een indicatie voor de academische richting van het onderzoek. De wetenschappelijke reputatie van het programma kan worden gehonoreerd uit financieringsbronnen met een meer fundamenteel wetenschappelijke signatuur.

2.2.2 Opleiding en training

In het domein van opleiding en training staat de 'productie' van gekwalificeerde onderzoekers centraal. Het aantal promovendi en onderzoekers in opleiding geeft een goede indicatie voor de omvang van deze activiteit en de middelen die voor deze activiteit ter beschikking worden gesteld. Daarnaast zijn de dissertaties als proeve van bekwaamheid een indicatie voor het rendement van die opleiding. Voorts kan de externe affiliatie (herkomst) van studenten en/of onderzoekers op verschillende manieren worden geanalyseerd. De mate van externe financiering van de opleiding, en de mate van bijdragen aan opleidingen elders kunnen eveneens als indicatie gelden.

2.2.3 Innovatie en professionals

De interactie met het domein van de professionele sector is zeer gevarieerd. Mobiliteit van onderzoekers naar posities in bedrijven of professionele (intermediaire) organisaties indiceert een stroom van vaardigheden en kennis naar dit domein, net als

lidmaatschappen van wetenschappelijke adviesraden. Samenwerking in projecten met het bedrijfsleven is een indicatie voor gedeelde thema's in het onderzoek. Professionele publicaties geven aan dat kennisclaims worden vertaald naar relevantie voor de commerciële sector. Financiering van dat contract-onderzoek bevestigt dat deze commerciële partijen die oriëntatie van het onderzoek willen honoreren. Patenten en royalty contracten zijn de meest directe indicatie voor de commerciële waarde van de kennis en de technologie die door het programma worden ontwikkeld.

2.2.4 Beleid en maatschappelijke vraagstukken *Public policy*

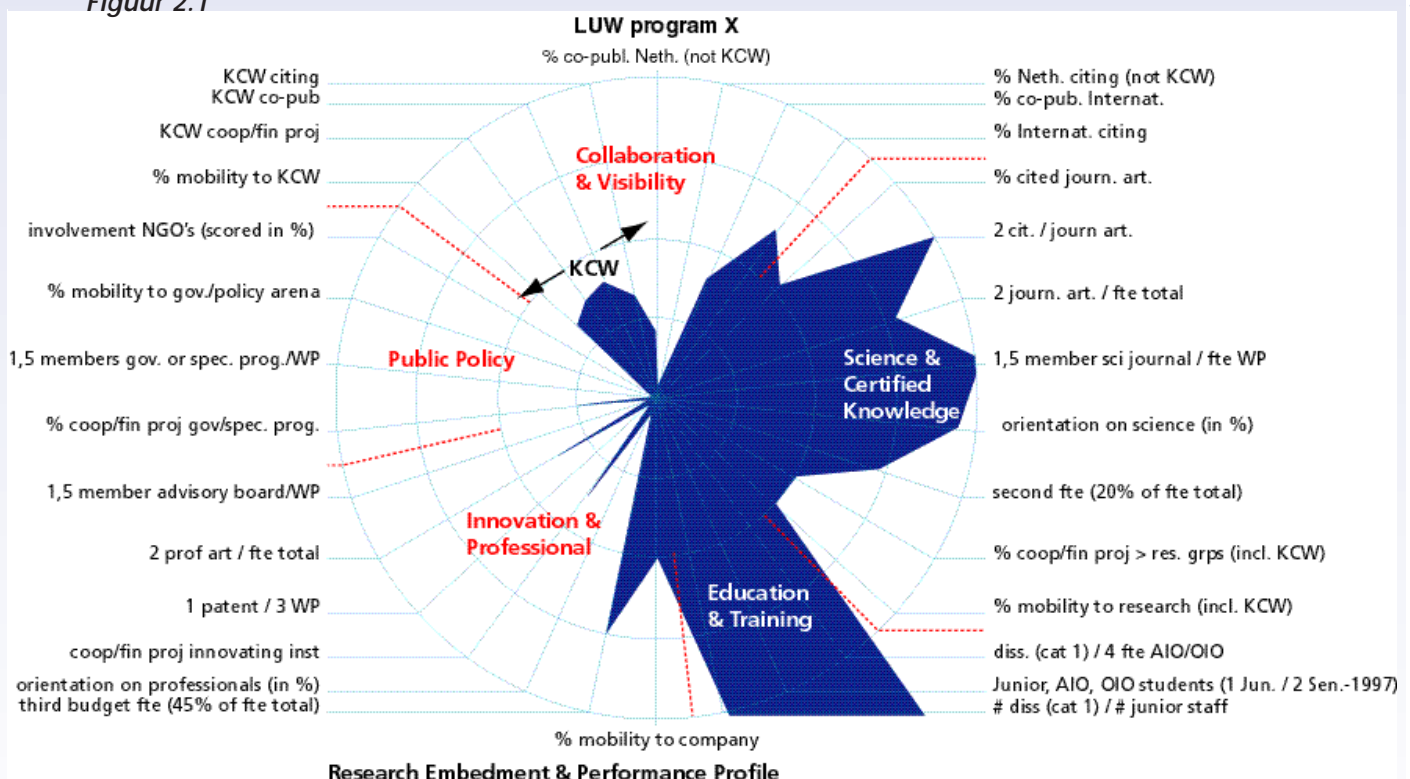
Het domein van de bijdragen aan beleidsvraagstukken is nog het meest onderontwikkeld wat betreft de robuustheid van de beschikbare indicatoren. Er is bijvoorbeeld geen equivalent van het patent of octrooi, gevestigde indicatoren voor commercieel succes. Ook vormen beleidsrapporten geen categorie die systematisch wordt onderscheiden van de overige publicaties. Wel kunnen lidmaatschappen van overheidsorganen en non-gouvernementele organisaties worden onderscheiden, net als mobiliteit van onderzoekers naar dit domein. Een concrete indicator voor de bijdrage aan het beleidsdomein is de gerichte opdrachtfinanciering. Het is opvallend dat een belangrijke actor in het vestigen van de evaluatieve

praktijken zo weinig is vertegenwoordigd in de administratieve classificatie die wordt gebruikt om over het onderzoek te rapporteren. Alhoewel veel onderzoeksgroepen een groot deel van de onderzoeksinspanning richten op de thema's van het beleid en op de maatschappelijke vragen blijft in de rapportage relatief onzichtbaar langs welke kanalen wordt bijgedragen aan dit domein.

2.2.5 Interne en externe samenwerking en zichtbaarheid (*Collaboration and visibility*)

Dit domein is toegevoegd aan de bestaande domeinen om inzicht te krijgen in welke mate de groepen zijn 'ingebed' in de instelling waar zij deel vanuit maken. In dit geval gaat het daarbij om Wageningen University Research Centre (Kenniscentrum Wageningen, KCW), waartoe eveneens de DLO instellingen behoren. De programma's van deze instelling realiseren in samenhang een overkoepelende missie in het onderzoek. Een vergelijkingsmaatstaf voor deze interne 'inbedding' kan worden gevonden in de interacties met andere groepen in het nationale kennissysteem en de internationale omgeving. Het gaat hier om de interacties met deze drie niveaus die in de omgeving van het programma kunnen worden onderscheiden. Indicatoren zijn mobiliteit van onderzoekers, samenwerking en het aanhalen van publicaties uit het betreffende programma.

Figuur 2.1



2.3 Het radar-profiel (REPP)

De radar-profielen (REPPs) zijn geconstrueerd om de complexe en multivariate activiteiten van onderzoeksgroepen op een relatief eenvoudige, inzichtelijke, redelijk volledige en vergelijkbare wijze te presenteren. In een radar-diagram worden data uitgezet als een functie van een statistische 'gemiddelde groep'. Dit maakt het radar-diagram geschikt om de overeenkomsten en verschillen in datasets te tonen. De oppervlakte geeft aan in welke domeinen een groep het sterkst tot ontwikkeling is gekomen. Deze oppervlakte is tevens een resultante van de onderlinge bevestiging (*corroboration*, zie verderop) van de scores op meerdere indicatoren: alleen bij die bevestiging van de indicatoren komt een oppervlakte tot stand^{①②}. Naast de constructie van het specifieke profiel van een programma kunnen verschillende profielen ook onderling worden vergeleken.

De constructie van het REPP gebeurt met de grootste zorgvuldigheid, maar het dient wel bekeken te worden met in het achterhoofd de gedachte dat een REPP-figuur het resultaat is van een bepaald kader dat wij als uitgangspunt hebben genomen en de keuzen die op grond daarvan door ons zijn gemaakt. Hieronder geven we aan de hand van een aantal methodische overwegingen inzicht in dat afwegingsproces. In figuur 2.1 wordt een voorbeeld gegeven van een volledig REPP.

2.4 Methodische overwegingen

2.4.1 Uniformiteit van de data

Een belangrijk criterium bij de keuze van de indicatoren is dat deze op vergelijkbare wijze voor ieder van de programma's kunnen worden vastgesteld. Een vereiste is dus dat de meetinstrumenten voor dataverzameling tussen groepen op stabiele wijze kunnen worden toegepast. Daarnaast moeten de benodigde data op relatief eenvoudige wijze kunnen worden verzameld. Dat impliceert dat soms moet worden aangesloten bij gevestigde categorieën in de administratie en verslaglegging van onderzoek door de programma's, de instelling, het landelijk wetenschapsbeleid of externe databestanden of bibliografische conventies. Afhankelijk van de

omvang van de beoordeling is het niet altijd mogelijk om zelf data te verzamelen. In het geval van Wageningen is dat overigens via een enquête onder groepen wel gedaan^{①②}. Informatie over de mobiliteit van onderzoekers was bijvoorbeeld niet centraal beschikbaar evenals informatie over adviesrelaties in de vorm van redacteurschappen of lidmaatschap van wetenschappelijke adviesraden. Daarnaast zijn samenwerkingsverbanden en de financiering van projecten nader in kaart gebracht aan de hand van deze enquête. Niet alle gewenste informatie kon worden geleverd. Lidmaatschappen konden bijvoorbeeld alleen betrouwbaar over het laatste jaar voor alle groepen worden verzameld.

2.4.2 Verifiëren van inbedding en 'performance'

In de domeinen wordt telkens met een aantal indicatoren gewerkt. Aan de hand van de vier interactiekanalen van Callon *et al.* wordt een selectie gemaakt van een beperkt aantal indicatoren. Deze indicatoren worden zo gekozen dat ze de mate van 'inbedding' en *performance* onderling kunnen bekrachtigen *corroboration*. In het domein van wetenschap en gecertificeerde kennis wordt bijvoorbeeld een hoog gemiddelde aan citaties bevestigd door een hoge productie van wetenschappelijke artikelen, door relatief veel redacteurschappen, door een aanzienlijk aandeel financiering uit meer wetenschappelijk georiënteerde fondsen, door veel samenwerking met andere universitaire groepen, door een relatief hoge uitstroom van personeel naar onderzoeksinstellingen, enzovoorts. Voor ieder van de domeinen is gezocht naar interactiekanalen die binnen dat domein relevant zijn en die tevens voldoen aan het criterium van uniforme data. In bijlage 3 is een tabel (Pagina 54 en 55) met de gebruikte indicatoren opgenomen. In kolom twee wordt daar verwezen naar de gebruikte data, in de laatste kolom wordt aangegeven in welk domein deze indicatoren zijn gebruikt.

2.4.3 Omvang van de groepen

Absolute aantallen publicaties of samenwerkingsrelaties geven geen indruk van de mate van interacties met een domein in de context van de groep. Daarvoor moet rekening gehouden worden met de omvang van het programma. In veel gevallen is hierbij gebruik gemaakt van de onderzoekscapaciteit die in de groep beschikbaar is

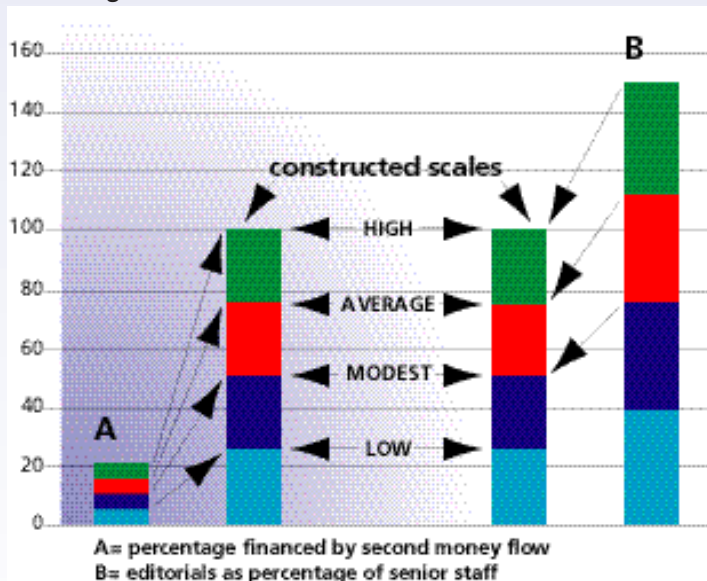
(in fte's). In andere gevallen is gefractioneerd door gebruik te maken van relevante delen van de capaciteit die wordt ingezet voor de betreffende interactie (bijvoorbeeld alleen de senior-staf in het geval van redacteurschappen). In andere gevallen wordt een indicator verdeeld over de domeinen. Dan wordt de fractie van 100% gebruikt als indicatie voor de 'inbedding' in het betreffende domein. In de derde kolom van de tabel in bijlage 3 zijn de berekeningen van het schaal-effect opgenomen.

2.4.4 Vergelijkbare schalen

'Inbedding' en *performance* worden in ieder van de domeinen geconstrueerd aan de hand van beschikbare, verzamelde en relevante data over de interacties met dat domein. Onderlinge vergelijkbaarheid vereist dat voor een programma *de mate van interactie of performance* wordt vastgesteld. Bij een aantal indicatoren is dat nogal ongecompliceerd, daarbij kan eenvoudig worden vastgesteld wat het percentage interacties is dat kan worden toegedeeld aan een bepaald domein. Bijvoorbeeld het percentage citaties door internationale groepen ten opzichte van de Nederlandse groepen. Het 'einde van de schaal' is hier altijd 100% (zie bijlage 3: kolom *end of scale*). In andere gevallen is het gecompliceerder om de mate van interactie conceptueel betekenisvol te construeren. Zo kan bijvoorbeeld worden vastgesteld wat het percentage is van lidmaatschappen van redactiesraden van wetenschappelijke tijdschriften ten opzichte van de omvang van de wetenschappelijke staf. Een andere mogelijkheid is het vaststellen van het percentage van de financiering afkomstig van fondsen met een wetenschappelijke signatuur

(hier geoperationaliseerd als de tweede geldstroom). Beide indicatoren zouden de mate van 'inbedding' in het wetenschappelijk domein operationaliseren. Het conceptueel probleem is dat het niet duidelijk is wat de percentages betekenen. Wanneer het aantal redacteurschappen 150% (1,5 redacteurschap per fte senior staf) is van de omvang van de vaste staf, is dat dan hoog of laag? Wanneer 20% wordt gefinancierd uit fondsen met een wetenschappelijke signatuur hoe verhoudt zich dat tot de 150% redacteurschappen? Om de onderlinge vergelijkbaarheid te bevorderen hebben we de indicatoren getransformeerd naar een geconstrueerde schaal die de waarden van 'NUL' tot en met 'HOOG' vertegenwoordigt voor een gemiddeld onderzoeksprogramma. Wegens tijdgebrek hebben we dit moeten uitvoeren zonder systematisch onderzoek en is de transformatie gebaseerd op geïnformeerde schattingen¹³. In dit geval is de 150% redacteurschappen en de 20% financiering uit de tweede geldstroom getransformeerd naar de score 'HOOG' voor een 'statistisch' gemiddeld programma in de sector landbouwwetenschappen. Cijfermatig worden de percentages 150% en 20% getransformeerd naar 100 als getalmatige representatie van een hoge score ten opzichte van die 'gemiddelde groep'. Alle tussenliggende scores worden op gelijke wijze getransformeerd: 10% financiering uit de tweede geldstroom wordt 50 op de schaal van de gemiddelde groep en representeert een bescheiden niveau van wetenschappelijke financiering. Vijftien procent financiering wordt 75 op de geconstrueerde schaal en representeert een gemiddelde niveau. Zie figuur 2.2 voor de uitwerking van de transformatie van deze beide schalen. De getallen op de geconstrueerde schaal zijn betekenisloos; alleen de positie op de schaal en de corresponderende kwalitatieve aanduiding zijn nog van conceptueel belang. De transformatie is dus gebaseerd op het vaststellen van een criterium gewicht voor een gemiddelde groep in deze sector. In kolom vier van de tabel in bijlage 3 wordt daarnaar verwezen als *end of scale*. De indicatoren die op een directe berekening van het percentage zijn gebaseerd corresponderen zonder transformatie met de geconstrueerde schalen. We hebben hier de percentages 25%, 50%, 75% en 100% dan ook corresponderend gelabeld, respectievelijk: 'laag', 'bescheiden', 'gemiddeld' en 'hoog'¹⁴.

Figuur 2.2 Transformation of scales



2.4.5 Afgrenzing van de programma's

Onderzoeksprogramma's kunnen op zeer verschillende manieren zijn 'verankerd' in de institutionele omgeving. Het ene programma moet gekarakteriseerd worden als een netwerk van samenwerkende groepen (bijvoorbeeld, interdisciplinaire stimuleringsprogramma's), andere programmatische verbanden zijn thematische georganiseerd en overstijgen de indeling in afdelingen, vakgroepen en leerstoelgroepen (bijvoorbeeld de organisatie van VF programma's binnen de LUW). In het geval van de LUW gaat het om onderzoeksprogramma's die samenvallen met functionele eenheden van onderzoek op dagelijkse basis. Institutioneel zijn deze programma's nu gevestigd als 'leerstoelgroepen'. Deze eenheid van organisatie is echter pas geïntroduceerd. De dataverzameling is *'with hindsight'* aangepast aan deze nieuwe institutionele eenheden. Door alle betrokkenen wordt echter erkend dat daarmee de programma's die samenwerken aan de realisatie van één onderzoeksprogramma (en de afgelopen vijf jaar hebben samengewerkt) worden gerepresenteerd in de data. Dit betekent dat een programma op een betekenisvolle wijze 'productie-eenheid' is geweest, door een eigen managementstructuur op een dagelijkse basis en werkend aan gezamenlijke doelstellingen voor de laatste vijf jaar¹⁶.

2.4.6 Statische representatie en dynamiek

Om een statistisch betrouwbare representatie van het interactiepatroon mogelijk te maken zijn de data voor vijf jaren samengenomen. Daarop is een uitzondering gemaakt voor die indicatoren waarvoor de data niet betrouwbaar over de gehele periode konden worden verzameld. In die gevallen is het laatste jaar (1997) als meetjaar genomen. Het interactieprofiel is daarmee een statische representatie van het programma. Gezien de looptijd van onderzoek -de meeste projecten lopen vier jaar of meer- is het niet zinvol om representaties te maken van kortere perioden. Die reconstructies zouden teveel te lijden hebben van fluctuaties op indicatoren die te maken hebben met korte termijn effecten. Voorbeelden zijn: 'generaties' promovendi die elkaar opvolgen; het organiseren van conferenties of het hebben van themanummers waardoor de output van een jaar zwaar in één van de indicatoren terechtkomt. Dergelijke korte termijneffecten moeten worden 'uitgemiddeld' om

een goed beeld van het profiel van de groep te krijgen. Daarnaast blijft het noodzakelijk om iets te weten van de ontwikkeling in de groep. Heeft zich in de onderzochte periode een duidelijke cesuur voorgedaan in de trend, bijvoorbeeld door het vertrek van de hoogleraar? Hoe heeft de personeelsopbouw zich ontwikkeld? Fluctueert de output sterk, is er groei of krimp zichtbaar? Om deze vragen te beantwoorden zijn rechte tellingen van enkele input en output indicatoren voor elk van de onderzochte jaren in een grafiek gezet. Uit deze grafieken kunnen de trends eenvoudig worden afgelezen.

2.4.7 Zelfbeeld van de groep

Aan de groepen is ook gevraagd om een 'zelfbeeld' te geven van de 'inbedding' van de groep in de omgeving. Deze zijn uitgezet in een 'ster-diagram'. Dit 'zelfbeeld' had het oogmerk naast het REPP betrokken te worden in een discussie over de oriëntatie van de groep op zijn omgeving. Het beeld dat men van zichzelf bij anderen wil overbrengen is een belangrijke factor in de keuze van een bepaald gedragsalternatief. Het is daarom relevant om dat in de evaluatie aan de orde te stellen. Deze opzet is door de beoordelingscommissie herkend maar beide diagrammen zijn dermate onvergelijkbaar dat het niet mogelijk was om de representaties op elkaar te betrekken. Een meer vergelijkbare uitwerking zou een leerzame ervaring kunnen opleveren.

2.5 Typologie van groepen op grond van de interactiepatronen

De gevonden REPP-profielen geven een karakteristiek beeld van een door een groep onderzoekers ontwikkelde strategie. Het lag in de bedoeling om uit deze groepseigen profielen enkele meer algemene classificaties te destilleren die kenmerkend zijn voor een bepaald interactiepatroon met de omgeving. Daarbij stonden ons uit de literatuur afkomstige voorbeelden voor ogen als: *science-driven*; *industry-driven* of *policy-driven*. In deze studie zouden deze classificaties door clustering van de empirisch vastgestelde profielen worden vastgesteld. Daaraan zou dan vervolgens een daarop toegesneden maatstaf voor de beoordeling kunnen worden verbonden.

Tabel 2.1 Typologie van interactiepatronen

Naam van de groep	Prominent domein	Secondair domein
Crop and Grassland Science	Wetenschappelijk & onderwijs	Innovatie/ professional
Plant Production Systems	Wetenschappelijk & onderwijs	Externe financiering, beleidsgeoriënteerd
Greenhouse Horticulture	Wetenschappelijk & onderwijs	-
Vegetative Propagation	Wetenschappelijk direct gerelateerd aan innovatie & onderwijs	-
Theoretical Production Ecology	Wetenschappelijk & onderwijs	Externe financiering (beleid – innovatie)
Soil Tillage	Innovatie	Internationale projecten, beleid
Systems & Control	Integratie: wetenschappelijk & innovatie (& beleid)	Interne (KCW) samenwerkingsverbanden
Farm Technology	Innovatie (& wetenschappelijk)	Beleid en onderwijs
Applied Physics	Wetenschappelijk	Externe financiering Interne (KCW) en internationale samenwerkingsverbanden (beleid)
Erosion and Soil & Water Conservation	Onderwijs (opbouw groep)	Proefstations -> beleid
Irrigation and Water Engineering Group	Beleid (& onderwijs)	Onderwijs (pas prominent na de recente verjonging van de groep)
Technology and Agrarian Development	Beleid & onderwijs	(innovatie -> bedrijven betrokken in beleidsrelevant onderzoek)

De constructie van de klassen is uitgevoerd door een eenvoudige clustering van groepen aan de hand van primaire en secundaire domeinen in het interactieprofiel (zie: tabel 2.1). Dat resulteerde in de volgende vijf klassen voor onderzoeksprogramma's:

- 1 Producenten van wetenschappelijke kennis van belang voor innovatieve, beleidsmatige en professionele actoren. Het onderzoeksprogramma is geïnspireerd door de algemene problemen van deze actoren. Die vraagstelling wordt echter altijd vertaald naar wetenschappelijke vragen en methoden.
- 2 Producenten gericht op het ontwikkelen van zowel wetenschap als innovatie om problemen die direct uit de professionele praktijk of uit het bedrijfsleven afkomstig zijn op te lossen.
- 3 Programma's waar de nadruk ligt op scholing en training van competente onderzoekers.
- 4 Groepen die gericht zijn op het ophelderen van vraagstukken in de 'techno-economische beleidsarena'. Zij combineren sociaal-weten-

- 5 schappelijke en technische kennis in de analyse van technologieontwikkeling.
- 5 Programma's gericht op ontwikkelingswerk en technische artefacten. Het belangrijkste doel is het ontwikkelen van een nieuwe praktijk of een nieuw artefact.

Alle onderzochte programma's zijn bij een van deze klassen ingedeeld. Als gevolg van de kleine set programma's die is bestudeerd en de enigszins intuïtieve manier waarop de clustering is uitgevoerd moet geconcludeerd worden dat hier nader onderzoek noodzakelijk is om tot een betrouwbare indeling te komen. In het kader van dit onderzoek heeft het een nuttige functie gehad in de demonstratie van de belangrijke verschillen in de oriëntatie van de groepen op hun omgeving. Opvallend is dat de clustering laat zien dat het kader voor de probleemdefinitie van het onderzoek een belangrijk onderscheidend kenmerk is. Wordt het probleem vertaald naar wetenschappelijke methoden en vragen, komen ze direct uit de

professionele praktijk of wordt gewerkt aan artefacten en procedures? Daarmee worden de verschillen tussen drie klassen al verklaard.

3 Gebruikersonderzoek

3.1 Het 'gebruik' van wetenschappelijk onderzoek

In het REPP wordt de structuur van de kennisproductie in relatie tot een relevante omgeving in kaart gebracht. Dat gebeurt aan de hand van specifieke indicatoren die in een vijftal sociale domeinen worden geselecteerd. Het REPP beperkt zich tot een kwantificering van interacties. Wat het niet aangeeft is in hoeverre die interacties ook leiden tot min of meer stabiele vormen van samenwerking waarin de belangen en gezichtspunten van gebruikers en andere belanghebbenden *stakeholders* aan bod komen. Hierdoor wordt niet (of onvoldoende) zichtbaar wat de rol van de gebruikers is in de ontwikkeling van (kennis)producten. Daarbij gaat het om vragen als: hebben gebruikers een substantiële invloed op de onderzoeksagenda, en hoe is er voorzien in feedback vanuit de gebruikers op onderzoek en innovatie. Om dergelijke vragen te beantwoorden, is een onderzoek onder gebruikers en andere belanghebbenden een noodzakelijke aanvulling op het REPP.

Bij 'gebruik' en 'gebruiker' zou men gemakkelijk alléén kunnen denken aan eindgebruik(er). In het geval van universitair onderzoek is die vorm van gebruik in veel gevallen nog ver weg. Dat kan zelfs het geval zijn bij onderzoek dat expliciet op innovatie is gericht. Bovendien gaat de beperking tot eindgebruiker voorbij aan het gegeven dat veel innovatie het resultaat is van wederzijdse beïnvloeding van verschillende deelnemers aan het innovatieproces, onderzoekers en andere actoren. Er zijn dus bij het onderzoek betrokkenen die als belanghebbende van het onderzoek kunnen worden aangemerkt zonder dat zij daarmee onmiddellijk 'gebruiker' in een meer materiële zin zijn. Er is dan geen sprake van gebruik van kennisproducten, maar er wordt, bijvoorbeeld vanuit een specifieke expertise of ervaring, bijgedragen aan het tot stand komen van nieuwe inzichten die het onderzoek en/of sociaal-economische veranderingsprocessen ten goede komen. Het begrip 'gebruiker' wordt daarom in deze gebruikersanalyse verbreed naar het gehele veld van de belanghebbenden, ook

wel *stakeholders* genoemd. Daaronder rekenen wij bijvoorbeeld zowel collega-onderzoekers als organisaties, al dan niet financierend, met een algemene maatschappelijke missie (bijvoorbeeld het bevorderen van wetenschappelijk onderzoek op een bepaald terrein).

Onderzoek wordt hier gezien als een onderdeel van het innovatie-proces dat voortschrijdt via interactie tussen een veelheid van actoren, wetenschappers, technici, praktijkmensen, beleidsactoren, het publiek.¹⁶ Een gebruikersonderzoek is dus een breed onderzoek. Het betreft in beginsel alle actoren die aan een innovatie werken. Deze brede benadering impliceert uiteraard dat rekening moet worden gehouden met de verschillen tussen gebruikers voor wat betreft aard en doel van hun betrokkenheid bij de innovatie.

3.2 Gebruikersonderzoek in het kader van de LUW-beoordeling

Het gebruikersonderzoek in het kader van de LUW-beoordeling is op twee manieren als aanvulling op het REPP gehanteerd. Ten eerste omdat de focus is gericht op de gebruikers (en niet op de onderzoeksgroep). In de tweede plaats omdat in dit deelonderzoek een kwalitatief oordeel wordt gezocht over de interactie.

In het kader van een analyse van de interactie tussen een onderzoeks-programma en haar omgeving kunnen verschillende dimensies worden onderscheiden. In een TNO-STB onderzoek uit 1996 worden de volgende zeven dimensies onderscheiden:

- 1 Deelnemende actoren en hun connecties.
- 2 Noviteit van de innovatie.
- 3 Perspectief op en timing van de gebruikersbetrokkenheid.
- 4 Leeromgeving en aard van het leren.
- 5 Interactie-mechanismen.
- 6 Technologie-transfer.
- 7 Beleidsinterventies.

De vraagstelling bij het LUW-onderzoek is grotendeels beperkt tot de dimensies 1, 4 en 5. De dimensies 2 en 6 zijn vooral van belang in een vergevorderde fase van de technologie-ontwikkeling, wanneer de producten of diensten naar de markt zijn of worden gebracht. In het geval van universitair onderzoek is dat meestal nog niet aan de orde. De dimensies 3 en 7 komen wel aan de orde in onze benadering, maar niet als aparte dimensie. In het licht van onze vraagstelling zijn de dimensies enigszins anders georganiseerd en benoemd. Daarbij komen we tot de volgende vijf dimensies:

- 1 Structuur van de gebruikersomgeving: de bij het onderzoek betrokken actoren en de rollen die gebruikers kunnen aannemen (gebruikerskarakteristiek).
- 2 Interactiekanaal van gebruikers en onderzoekers, dwz de wijzen waarop communicatie plaats vindt tussen gebruikers en onderzoekers.
- 3 Methoden en mechanismen om gebruik te bevorderen, c.q. gebruikers te betrekken.
- 4 Aard van het gebruik, d.w.z. van welke output van onderzoek maakt men gebruik en met welk doel.
- 5 Oordeel van gebruikers over de relatie met onderzoekers.

Elk voor zich geven deze dimensies informatie over aspecten van de interactie tussen onderzoekers en gebruikers. De eerste dimensie is onderzocht via een *analyse van de projectlijsten van onderzoeksprogramma's*. In die lijsten worden inhoudelijke en administratieve gegevens met betrekking tot alle onderzoeksprojecten van een groep geregistreerd. De structuur is op twee manieren in kaart gebracht: langs de lijn van sponsoring (welke actoren steken geld in het onderzoek), en langs de lijn van rollen die gebruikers kunnen spelen in het innovatieproces. Er zijn drie rollen onderscheiden: collega-onderzoeker (niet uit hetzelfde programma); intermediair (niet noodzakelijkerwijze financieel); eindgebruiker (niet noodzakelijkerwijze betalend).

Om de overige dimensies te onderzoeken is een *survey onder gebruikers* uitgevoerd. Een centraal doel van de survey is het identificeren van een karakteristieke 'leeromgeving' voor een bepaald onderzoeksprogramma. Deze leeromgevingen zijn - theoretisch- kenmerkend voor een bepaalde fase van

innovatie. Het hanteren van het concept 'leeromgeving' -en de daarmee verbonden karakterisering van innovatie als leerproces- heeft tot doel verwachtingen te toetsen die men kan hebben ten aanzien van de wijze waarop de interactie is georganiseerd in verschillende fasen van het innovatieproces. Bijvoorbeeld: Van een onderzoeksprogramma dat sterk gericht is op experimentele ontwikkeling, kan men verwachten dat er een veel nauwere samenwerking bestaat met een duidelijk aanwijsbare groep (potentiële) eindgebruikers dan van een programma dat zojuist een nieuw terrein heeft betreden. In dat laatste geval valt eerder te verwachten dat er lossere contacten zijn met een brede omgeving (overheid, intermediairen, gebruikers) en dat de contacten zijn gericht op de ontwikkeling van ideeën omtrent een gewenste ontwikkeling. In de verschillende leerfasen zullen uiteraard de verwachtingen van gebruikers ten aanzien van onderzoekers verschillen. Men moet zich wel bedenken dat dit een analytisch instrument is: een onderzoeksprogramma kan tegelijkertijd in verschillende leeromgevingen functioneren, de bevindingen wijzen er zelfs op dat dit meestal het geval is.

De verschillen tussen leeromgevingen hangen dus samen met de verschillende fasen in een innovatieproces. In de beginfase zal waarschijnlijk vooral het identificeren van allerlei mogelijke maatschappelijke en technologische obstakels en mogelijkheden centraal staan. Naarmate de innovatie meer vorm krijgt zal het leerproces meer zijn gericht op praktische toepassingen en het testen van apparatuur, modellen e.d.

In navolging van de TNO/STB studie worden drie leeromgevingen onderscheiden (1996: 183-4).

Elk van deze leeromgevingen wordt gekarakteriseerd door een bepaalde, kenmerkende, vorm van interactie. De leeromgevingen zijn:

- A **Articulatie:** Omgeving waarin verwachtingen ten aanzien van nieuwe technologische ontwikkelingen kunnen worden geuit en gearticuleerd.
- B **Afstemming:** Omgeving waarin een technologische *nexus* wordt gecreëerd: een intermediair coördinatie-mechanisme waarin de verschillende belangen op elkaar worden afgestemd.
- C **Testfase:** Omgeving waarin een *niche* wordt gevormd en waarin praktische experimenten buiten het lab worden gedaan om nieuwe variaties uit te testen.

De kenmerkende interactievormen zijn respectievelijk voor A. discussieplatform, voor B. (semi)permanente organen, en voor C. test-sites buiten het laboratorium (het onderzoeksprogramma). De leeromgevingen maken het mogelijk verbindingen te leggen tussen bepaald wetenschappelijk onderzoek, de praktijk en onderzoeks- en innovatiebeleid. Ze worden in de terugkoppeling, het derde deelonderzoek van onze benadering (zie hoofdstuk 4) vergeleken met de REPP-profielen en kunnen aldus het kwantitatieve beeld dat daarin wordt gegeven van de onderzoeksorganisatie aanvullen.

Samengevat: Er worden vijf dimensies onderscheiden in de interactie tussen onderzoeksprogramma en omgeving. De eerste (structuur) is aan de hand van de projectlijsten separaat onderzocht van de andere vier, die onderwerp zijn van een survey onder gebruikers. Het gebruikersonderzoek als geheel bestaat derhalve uit de volgende twee delen:

Deel 1: **Kartering**

Het in kaart brengen van de structuur van de omgeving. Gebruikers worden op twee manieren geïdentificeerd: naar type sponsor en naar hun rol in het innovatieproces.

- Dataverzameling met behulp van projectlijsten
- Vaststellen van de institutionele achtergrond van de sponsors
- Identificatie van de rollen van gebruikers (wij onderscheiden er 3: collega-onderzoeker; intermediair, eindgebruiker)
- Grafische representatie van gebruikersomgeving

Deel 2: **Survey**

Een interviewronde onder een selectie van gebruikers. Vragen betreffen vier onderscheiden dimensies: interactiemechanismen; methoden om gebruikers te betrekken; aard van gebruik; oordeel van gebruiker. Belangrijk doel: identificeren leeromgeving

- Selectie respondenten survey
- Opstellen vragenlijst langs vier dimensies
- Uitvoeren survey (telefonisch of via e-mail)
- Dataverwerking, identificeren leeromgevingen
- Rapportage

3.3 Deel 1: **Kartering** Structuur van de gebruikersomgeving

In dit deel wordt een beeld geconstrueerd van de maatschappelijke omgeving van onderzoeksprogramma's. In de eerste plaats willen we weten *wie* het onderzoek gebruikt. Hiertoe hebben we in dit geval er voor gekozen de institutionele achtergrond van sponsors vast te stellen (zie verderop). In de tweede plaats willen we weten wat de *rol* is die gebruikers in het innovatieproces spelen. Dat is van belang om inzicht te verwerven in het traject van onderzoeksvraagstelling tot praktische toepassing.

De gebruikers worden dus op twee manieren in kaart gebracht. In de eerste plaats wordt de institutionele achtergrond van sponsors geïdentificeerd. Financiële ondersteuning is een beperkt criterium, hierdoor worden bepaalde betrokkenen uitgesloten. Maar vooralsnog is het het meest betrouwbare dat op grond van de projectlijsten is te verkrijgen. Om ook niet-sponsors betrouwbaar te classificeren (en dit zou inderdaad een vollediger beeld geven) is het noodzakelijk nader onderzoek te doen onder onderzoekers en gebruikers die worden vermeld als belanghebbend (maar niet betalend). Dat is in dit geval niet gedaan. Wel zijn de gebruikers nog op een andere wijze geïdentificeerd aan de hand van de rol die ze spelen in het innovatieproces. Daartoe zijn drie 'hoofdrollen' onderscheiden: collega-onderzoeker intermediair en eindgebruiker. Op deze plaats gaan we niet in op de specifieke resultaten van de studie maar vermelden slechts als algemeen resultaat dat bij de onderzochte groepen de categorie 'collega-onderzoekers' bij driekwart van de onderzoeksgroepen 60% of meer scoorde. Dat is niet onverwacht voor universitair onderzoek. Maar ook de rol van 'eindgebruiker' scoorde redelijk hoog bij veel programma's, vaak tussen de 20 en 40%. En dat is minder gebruikelijk voor universitair onderzoek, en bevestigt het gebruikers-georiënteerde karakter aan van onderzoek in de landbouwwetenschappen.

Kartering van gebruik aan de hand van sponsoring. Het ging ons in dit geval om een indicatie van betrokkenheid, niet van omvang. Dat wil zeggen dat niet bedragen zijn geteld (trouwens toch niet zonder meer betrouwbaar uit de projectlijsten te

halen), maar actoren zijn geïdentificeerd die meebetalen aan (delen van) onderzoeksprogramma's. Zoals gezegd kiezen we met financiers voor beperkt criterium, maar wel een dat redelijk betrouwbaar is. Andere vormen van samenwerking en kennisuitwisseling blijven hier dus buiten beschouwing (maar komen wel aan bod bij de rollen).

Na het identificeren van de financiers worden ze ingedeeld naar institutionele achtergrond, en daarna in een beperkt aantal klassen. Ten behoeve van deze studie is gekozen voor een indeling in vijf categorieën actoren (sponsors): het KennisCentrum Wageningen (KCW); overige Nederlandse publieke sponsors (ODP); internationale publieke sponsors (PINT), bedrijfsleven (IND) en non-gouvernementele organisaties (NGO). In het staafdiagram is de omvang van de betrokkenheid van financiers in beeld gebracht aan de hand van het aantal projecten dat wordt ondersteund. Het aantal sponsors overtreft het aantal projecten ruimschoots. Daaruit is af te lezen dat veel projecten gemengd worden gefinancierd.

Uit de resultaten van de kartering in deel 1 ontstaat het beeld dat de twaalf groepen een zeer gevarieerde omgeving hebben, waarin grote aantallen, zeer verschillende, gebruikers zijn te onderscheiden. Voor een deel wordt dit veroorzaakt door onze ruime definitie van 'gebruiker'. Maar voor een (belangrijker) deel ligt het in de brede belangstelling voor landbouw-

kundig onderzoek: Het bedrijfsleven (boeren en industrie), nationale en internationale instanties en programma's, ngo's, veel buitenlandse kennisinstellingen en individuen.

Daarbinnen kunnen enkele dominante groepen gebruikers worden onderscheiden (bijvoorbeeld het KCW), maar dat neemt niet weg dat de gebruikersomgeving van landbouwkundig onderzoek bestaat uit zeer uiteenlopende actoren. In figuur 3.1 is een overzicht gegeven van sponsorcategorieën van de 12 onderzochte LUW programma's.

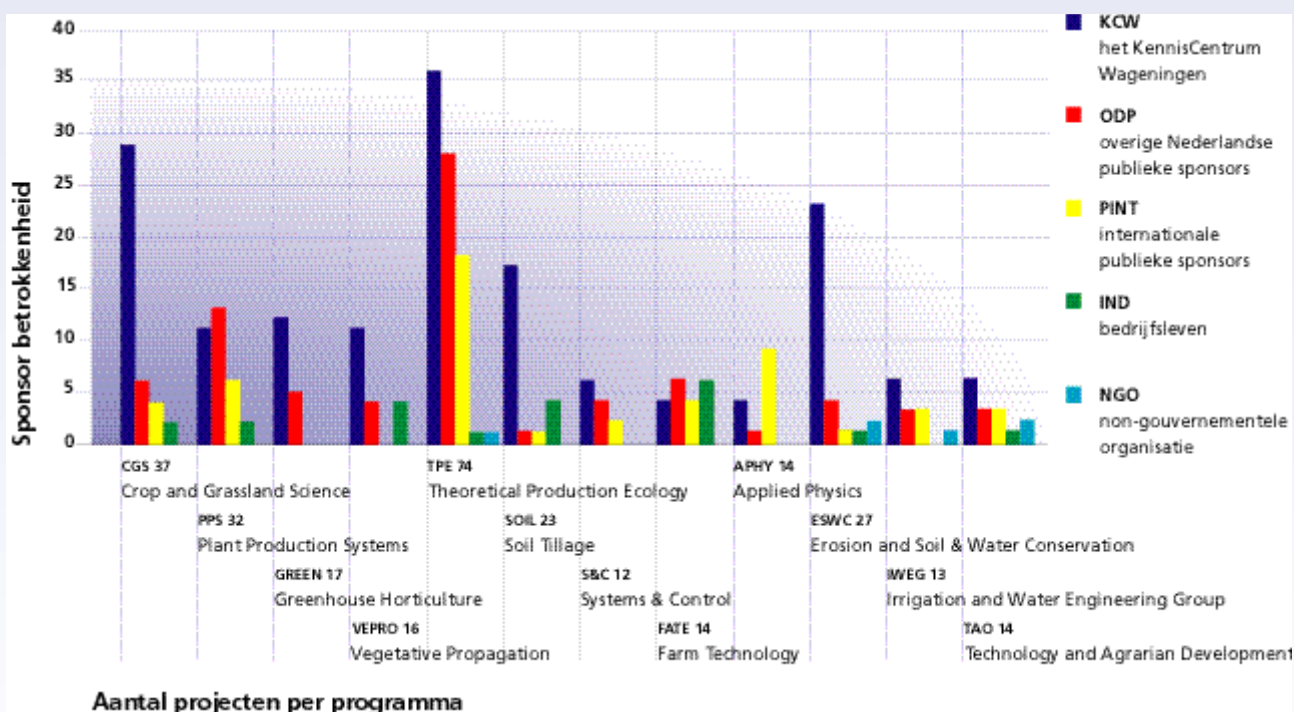
3.4 Deel 2: Survey

Kwalitatieve analyse van de interactie

In het totaal werden ongeveer 200 gebruikers geïdentificeerd en geclassificeerd ten behoeve van de gebruikerskartering. Om een beeld te krijgen van de inhoudelijke betrokkenheid van deze actoren bij het onderzoek zijn per programma 3 gebruikers geïnterviewd (telefonisch of per e-mail). Daarbij is voor elk programma geprobeerd om de geïnterviewde gebruikers zo te kiezen dat elk van de volgende drie categorieën is vertegenwoordigd: publieke (onderzoeks)sector, private sector, en internationale programma's en organisaties. In het totaal werden 36 gebruikers geïnterviewd.

De vragenlijst had een semi-gestructureerd karakter om de uitvoerbaarheid van een telefonische/ e-mail

Figuur 3.1 Sponsor categorieën



enquête onder een groot aantal gebruikers mogelijk te maken. De vragenlijst (zie bijlage 2) die we in dit geval hebben ontwikkeld is gebaseerd op de laatste vier dimensies van interactie zoals die hierboven al zijn besproken (punten 2 t/m 5). We herhalen deze vijf punten hier kort:

- 1 Structuur van de gebruikersomgeving, zowel voor wat betreft de soorten actoren als de soorten rollen (deel 1 van het gebruikersonderzoek).
- 2 Interactiekanaalen van gebruikers en onderzoekers, d.w.z. de wijzen waarop communicatie plaats vindt tussen gebruikers en onderzoekers.
- 3 Methoden en mechanismen om gebruik te bevorderen, c.q. gebruikers te betrekken.
- 4 Aard van het gebruik, dwz van welke output van onderzoek maakt men gebruik en met welk doel.
- 5 Oordeel van gebruikers over de relatie met onderzoekers.

Tezamen geven de vier laatste aspecten een beeld van de leeromgeving van het betreffende onderzoek. Hieronder lichten we de samenhang tussen deze aspecten en de 'leeromgeving' kort toe.

Ad 2

Interactiekanaalen zijn als het ware de 'dragere' van de interactie. Afhankelijk van de fase van innovatie kan het hierbij gaan om heel algemene, voorlichtende bijeenkomsten, of om testen in de praktijk. Naast het identificeren van soorten gaat het hier ook om de frequentie waarmee men een kanaal gebruikt, en om het doel waarom men het gebruikt.

Ad 3

Hierbij gaat het om twee onderscheiden maar verwante zaken. Aan de ene kant wil men weten of en hoe gebruikers bij het onderzoek zijn betrokken (de aard van de betrokkenheid). Aan de andere kant is er een scala van (onderzoeks)methodieken die er op gericht zijn de betrokkenheid van gebruikers bij onderzoek te vergroten.

Ad 4

Hier vragen we naar wat het nu eigenlijk is dat wordt gebruikt, m.a.w. welke vorm van output van de onderzoeksgroep is relevant voor een gebruiker. Daarnaast vragen we hoe van de output gebruik

wordt gemaakt, en wie de eventuele eindgebruikers zijn.

Ad 5

Het gebruikersoordeel over de relatie met de onderzoeksgroep betreft de sterke en zwakke punten in de samenwerking, in de groep, en in de producten. Voorts wordt gevraagd naar de mogelijkheden om de onderzoeksagenda te beïnvloeden, en naar de wens om in de toekomst te blijven samenwerken.

De resultaten van de interviews geven in de eerste plaats een gedetailleerd beeld van de waarde van het onderzoeksprogramma voor de omgeving, *vanuit* die omgeving. Deze resultaten zijn per programma samengevat in kleine rapportjes en in deze vorm aangeboden aan de commissie. De beoordelingscommissie had op deze manier de beschikking over informatie over de gebruikersomgeving van het onderzoeksprogramma en van de waardering van het programma door de gebruikers.

Op grond van deze resultaten worden groepen gekarakteriseerd in termen van een leeromgeving. Vaak zal het hierbij gaan om toewijzing op grond van een dominant kenmerk. Veel groepen hebben immers activiteiten in verschillende fasen van het innovatieproces. Via de indeling in leeromgevingen kan een terugkoppeling met het REPP plaats vinden (zie volgende hoofdstuk).

De survey dient aldus twee doelen: het geeft een kwalitatief beeld; en het geeft de mogelijkheid om dat beeld te vergelijken met het beeld uit het eerste deelonderzoek.

In bijlage 2 worden bij wijze van voorbeeld de *survey*-resultaten voor twee onderzoeksprogramma's samengevat.

De algemene conclusies en discussie naar aanleiding van het gebruikersonderzoek worden besproken in hoofdstuk 5. Het gebruikersonderzoek leverde voor de beoordelingscommissie zinvolle aanvullende informatie op, met name voor wat betreft de mening van de gebruikers over hun betrokkenheid bij het onderzoek en de bijdrage die het onderzoek levert aan de doelstellingen van de gebruikers. Deze informatie is uiteraard ook een empirische toets van de opmerkingen die door de programma's in de zelfevaluaties zijn gemaakt over

de maatschappelijke relevantie van hun onderzoek. De vijf dimensies van interactie zijn een belangrijk analytisch onderscheid om deze gebruikersomgeving aan de hand van algemene karakteristieken te kunnen beschrijven.

4 Terugkoppeling

4.1 Inleiding

Ondanks de verschillen tussen de twee hiervoor behandelde deelonderzoeken zijn het beide representaties van de twaalf onderzoeksprogramma's. Het ligt dan ook voor de hand om de twee representaties van ieder onderzoeksprogramma met elkaar in verband te brengen. Deze vergelijking wordt hier aangeduid met 'terugkoppeling'.

De terugkoppeling dient twee doelen. In de eerste plaats een vergelijking tussen REPP en gebruikersonderzoek. In de tweede plaats om de deelonderzoeken zodanig te integreren dat een eenvoudige vergelijking tussen deze empirische resultaten en de missie van de onderzoekseenheid mogelijk wordt.

4.2 'Embedment', 'performance' en leeromgeving combineren

In overeenstemming met het werk van andere wetenschapsonderzoekers, bijvoorbeeld Callon et al. (1994) en Joly en Mangematin (1996)^{①②}, zijn aan de hand van de REPPs enkele typen interactie profielen vastgesteld. Binnen elk van deze klassen worden relaties op een specifieke manier aangegaan, en wordt kennis op een specifieke manier uitgewisseld en ontwikkeld.

Aanvankelijk bestond bij ons de verwachting dat het mogelijk zou zijn om op grond van deze clustering te komen tot een heldere indeling van de groepen. Daarbij stond een indeling vergelijkbaar met de driedeling *science-driven*, *policy-driven* en *industry-driven* voor ogen. Deze indeling zou dan in verband gebracht worden met de driedeling in leeromgevingen uit het gebruikersonderzoek: 'articulatie', 'afstemming' en 'experiment'. Die twee driedelingen kunnen dan met elkaar worden vergeleken. We meenden hier ten minste een heuristische indicatie te kunnen construeren bijvoorbeeld voor de fase waarin het innovatieproces zich bevindt. Zo zal bijvoorbeeld een onderzoekseenheid die aan het begin van een innovatief traject staat haar onderzoeksverwachtingen *articuleren* naar een brede groep actoren,

en omgekeerd zullen deze actoren hun verwachtingen uiten ten aanzien van de technologische vernieuwing en/of implementatie. Later in de ontwikkeling kan een politiek/maatschappelijke *afstemming* plaatsvinden, dikwijls via intermediaire organisaties.

De clustering van de gevonden profielen bleek echter niet eenvoudig. Er waren vijf klassen nodig om de profielen adequaat te representeren. Daarbij bleken deze klassen minder duidelijk te differentiëren dan wij hadden gehoopt en ook minder sterk samen te hangen met mogelijke fasen in het innovatietraject.

Ook de driedeling in leeromgeving bleek niet zomaar toe te passen. In de eerste plaats omdat voor elk programma slechts drie gebruikers konden worden geïnterviewd. De classificatie hangt daardoor teveel af van deze beperkte survey. Daarnaast bleken een aantal onderzoeksprogramma's gelijktijdig betrokken te zijn in verschillende leeromgevingen. In die gevallen werd het programma geklasseerd onder de dominante leeromgeving.

Deze hele operatie was erop gericht de beoordelingscommissie te voorzien van een eenvoudige samenvatting van de bevindingen die kan worden teruggekoppeld op de missie van het programma. In dat proces ging echter erg veel informatie verloren. De uitkomsten van deze vergelijkingen kunnen dan ook alleen dienen als indicatieve basis voor een terugkoppeling van de missie van het onderzoekprogramma op de empirisch vastgestelde resultaten. In een dergelijke discussie kan de toekomstige richting van het onderzoeksprogramma centraal staan in relatie tot de missie en de bestaande 'inbedding' in de omgeving.

5 Bevindingen in het licht van ontwikkeling naar een reguliere beoordelingspraktijk

5.1 Inleiding

In opdracht van de Landbouwwetenschappelijke Landbouwwetenschappen van de VSNU. In deze notitie wordt de methodiek die is gebruikt besproken met het oog op mogelijke toepassing in andere sectoren van wetenschappelijk onderzoek. Voordat de implicaties van een bredere toepassing van de methodiek worden besproken, is het in de eerste plaats noodzakelijk om vast te stellen welke bijdrage het rapport heeft geleverd aan de beoordelingsprocedure en hoe de representaties zijn ontvangen door de betreffende onderzoeksgroepen. Om deze vragen te beantwoorden is een beperkte *meta-evaluatie* uitgevoerd^{1⑧}. De voornaamste punten die hieruit naar voren kwamen worden hieronder besproken in het licht van een algemenere toepassing van de gevolgde aanpak. Afsluitend worden enkele opmerkingen gemaakt over de haalbaarheid van de voorgestelde 'contextuele' beoordeling van wetenschappelijk onderzoek in relatie tot de huidige beoordelingspraktijk in het wetenschapsbeleid.

5.2 De aanpak algemeen

Er bestaat overeenstemming bij betrokkenen over het feit dat REPP en gebruikersonderzoek een antwoord kunnen zijn op een beleidsprobleem dat in het wetenschapsbeleid en de evaluatie van onderzoek op de agenda staat. Men constateert daar dat in de gebruikte methoden voor onderzoeksbeoordeling het beleidsmatig en

maatschappelijk belang van onderzoek onderbelicht blijft. Onderzoek is niet alleen op het wetenschappelijk domein gericht maar ontwikkelt zich ook in andere domeinen. Juist de extrinsieke of contextuele waarde van het programma dat in die domeinen aan de orde is blijft in de huidige procedure te zeer buiten beeld.

De radar-diagrammen (REPPs) reconstrueren de 'inbedding' en *performance* van onderzoeksprogramma's in de verschillende domeinen die in de omgeving van een programma kunnen worden onderscheiden. Zij geven aldus een zo volledig mogelijk beeld van de wetenschappelijke en maatschappelijke oriëntatie van een onderzoeksprogramma. Het gebruikersonderzoek vult de gevonden resultaten in het REPP aan, in het bijzonder waar het de verwachtingen van de omgeving betreft ten aanzien van het onderzoek. Zo kan ten behoeve van de beoordeling van een onderzoeksprogramma de evaluatieve maatstaf worden verbreed en worden toegesneden op de voor het betreffende programma specifieke mix van domeinen. Ook kan aan de hand van de resultaten van beide deelonderzoeken de realisatie van de missie van het programma in meer detail worden besproken.

De VSNU-beoordelingscommissie was van mening dat de geconstrueerde profielen (REPP) hier een duidelijke meerwaarde vertegenwoordigden ten opzichte van de zelfevaluatie door de groepen. De commissie vond dat men daar de werkzaamheden op nogal eenzijdige wijze presenteert, namelijk primair gericht op de wetenschappelijke productie van programma's. Die eenzijdigheid was echter mede het gevolg van het bij de wetenschappers bestaande beeld van de wijze waarop rekenschap dient te worden gegeven van onderzoeksactiviteiten. Men heeft het idee dat het accent ligt op indicatoren voor wetenschappelijke reputatie en productie. Hierdoor raakt de maatschappelijke rol van onderzoek op de

achtergrond en wordt daarover slechts beperkt gerapporteerd. De commissie merkte hierover op dat een cultuurverandering noodzakelijk is waarin de profielen een rol kunnen spelen.

Dit illustreert dat het REPP een dialoog kan bevorderen tussen de commissie en het onderzoeksprogramma waarin de missie, de claims over de resultaten en de verificatie daarvan tegelijk aan de orde zijn en in verband worden gebracht met toekomstige mogelijkheden van de groep.

Belangrijk daarbij is dat de beoordelingscommissie en de beoordeelde onderzoeksprogramma's de profielen en de gebruikersanalyse aanvaarden als een overzichtelijke en overtuigende representatie van de activiteiten en de *performance* van de groep, en dat deze als uitgangspunt kan worden genomen voor de beoordeling. Dat was zeker het geval in de studie uitgevoerd voor de LUW-beoordeling. Maar deze gold dan ook als een experiment. Wel gaven de profielen in een aantal gevallen aanleiding tot commentaar vanuit de groepen. Maar meestal ging het er daarbij om bepaalde aspecten nog eens toe te lichten of te onderstrepen en daarmee het profiel in feite aan te scherpen.

Naast de statische profielen bestond er bovendien behoefte aan enkele representaties van de ontwikkeling in het programma. Hieraan is voldaan door toevoeging van eenvoudige trend-grafieken (zie bijlage 1, voorbeeld-cases).

5.3 REPP

Een zorgvuldige *selectie en afweging van de indicatoren* zijn noodzakelijk. Daarbij moet bij voorkeur gebruik gemaakt kunnen worden van bestaande administratieve en of bibliografische categorieën om te voorkomen dat buitensporig grote inspanningen moeten worden geleverd voor de dataverzameling. Wel kan worden overwogen waar aanpassingen in de administratieve categorieën of het gedetailleerder rapporteren over onderzoek kan bijdragen aan meer betekenisvolle informatie¹⁹. Naast de beschikbaarheid van betrouwbare, goed geüniformeerde publicatiebestanden bleken van groot belang: een goede project-administratie (samenwerkingsverbanden en uitwisselingen); een personeels-administratie (mobiliteit van onderzoekers); het bijhouden van lidmaatschappen en nevenfuncties van personeels-

leden en een gedetailleerde financiële administratie waardoor uitsplitsing van de financiële bronnen mogelijk wordt²⁰.

Het is ook wenselijk dat een meer evenwichtige verdeling van het aantal indicatoren over de domeinen tot stand wordt gebracht.

Complexiteit van het profiel. De REPP-profielen geven een beeld van de verschillende constellaties die bestaan in de interacties van een programma met zijn omgeving. Door de veelheid van indicatoren is het echter niet altijd eenvoudig om aan de hand van de profielen een typering of interpretatie van die een constellatie te geven. Hier stuit men op de complexiteit van de onderzoeksprogramma's en hun nogal unieke karakter die moeilijk in algemene classificaties is onder te brengen. De vergelijking van de groepen en het principe van vergelijking van 'like-with-like' in evaluaties komt hiermee in gevaar. Wel is opgemerkt dat vergelijking in deze evaluaties vaak een geconstrueerde gelijkheid is.

Het '*radar-oppervlakte-probleem*' is naar voren gebracht. Conceptueel lijkt direct duidelijk wat de oppervlakte in de radar-grafiek betekent. Er is echter een complexe samenhang van de oppervlakte met: a) de mate van activiteit of *performance* en b) de onderlinge bevestiging van indicatoren (corroboratie is de technische term) door volledigheid in die activiteit. Er ontstaat namelijk pas een oppervlakte bij scores op naast elkaar liggende indicatoren. Alhoewel de volgorde van de indicatoren met het oog op *corroboration* is gekozen kan die volgorde juist voor een specifieke groep niet geschikt zijn. Ook is het mogelijk om op grond van andere argumenten tot een andere volgorde te komen. Dit probleem kan worden opgelost door het radar-diagram meer als histogram (en minder als polygoon) te construeren. In bijlage 5 is een mogelijke aanpassing in die richting weergegeven (figuur, bijlage 5.1.a-c).

De transformatie naar de geconstrueerde schalen in het radar-diagram (figuur 2.1 in hoofdstuk 2) leek tot enige verschillen in de sensitiviteit van de indicatoren (assen) te leiden. De 'sensitiviteit' van de indicatoren moet daarom systematischer worden 'gekalibreerd'. Voor elk van de indicatoren zou op grond van onderzoek, of op grond van een brede consultatie moeten worden vastgesteld bij welke waarde sprake is van een 'gemiddelde' of 'hoge'

score voor een gemiddelde groep. De programma's kunnen dan worden uitgezet tegen een helder bepaalde referentiewaarde.

Een *confrontatie tussen het zelfbeeld* van de groep en de *empirische reconstructie* van de interactiepatronen is pas mogelijk wanneer deze meer systematisch op elkaar worden afgestemd. Een mogelijke ontwikkeling in die richting is uitgewerkt in figuur, bijlage 5.1.c. De empirisch vastgestelde mate van interactie is binnen de domeinen gemiddeld wat resulteert in de zwarte contour. Daarnaast is in de stippellijn het zelfbeeld van de groep weergegeven. Een dergelijk zelfbeeld zou geconstrueerd kunnen worden aan de hand van een daarvoor ontwikkelde enquêtevraag. Bijvoorbeeld zoals in de huidige studie is gevraagd naar de wetenschappelijke en naar de professionele oriëntatie.

Er is naar voren gebracht dat de *statistiek van de kleine getallen* de robuustheid zou kunnen aantasten. Inderdaad waren enkele van de onderzochte groepen erg klein. Alhoewel het hier gaat om eenvoudige beschrijvende statistiek is het goed om beter na te gaan wanneer de onderlinge vergelijking van de profielen in gevaar komt door verschillen in schaal. Wij hebben dat op grond van onze bevindingen nog niet kunnen vaststellen.

De profielen suggereren wellicht dat er sprake is van een *coherente*, bewust gekozen, en door het gehele programma gedragen *strategie* in beeld wordt gebracht. Maar het kan natuurlijk evenzeer een optelling zijn van ongecoördineerde handelingen van (eigenzinnige) individuele leden van een programma. We benadrukken dat wij geen enkele veronderstelling over een programmatische coherentie gebruiken. Het is bijvoorbeeld best mogelijk dat flexibele 'entrepreneurs' zich sneller aanpassen aan de veranderende omgevingseisen (milieu, natuurbehoud en landschap) dan meer planmatig werkende onderzoekers. Hier is met de representatie dus geen enkel waardeoordeel bedoeld maar slechts een representatie van de optelling van empirische resultaten van strategisch handelen. Overigens blijven we wel van mening dat deze resultaten moeten worden vergeleken met de doelstellingen van de groep.

5.4 Gebruikersonderzoek

Het gebruikersonderzoek is opgezet als aanvulling op het REPP. Het is in het kader van het LUW-onderzoek beperkt opgezet. Dat zit hem dan vooral in het aantal gebruikers dat in de beperkte tijd kon worden geïnterviewd (30-35). Op een totaal van ongeveer 200 gebruikers is dat overigens nog niet zo erg slecht. Maar per groep betekende dit een beperking tot maar drie gebruikers. In het bijzonder voor de grotere groepen blijft dan een deel van de activiteiten buiten beeld. In dit geval was dat overigens niet heel bezwaarlijk, omdat het doel primair was het bieden van inzicht in kwalitatieve aspecten van de interactie met de gebruikers. Overigens is de volledige gebruikersomgeving ook kwantitatief in kaart gebracht en gepresenteerd in grafieken.

Naar aanleiding van de vragenlijst zijn kritische opmerkingen gemaakt over het feitelijke (eind)gebruik. Sommigen merkten op dat de vragenlijst ophoudt waar het spannend wordt, nl. wat doet men met het onderzoek, en hoe helpt het onderzoek gebruikers bij hun eigen praktijk, wat hebben ze er precies aan. Hoewel vragen met die strekking wel zijn opgenomen in de vragenlijst (zie bijlage 2 de vragenlijst) is het duidelijk dat ze niet voldoende scherp waren gesteld. De redenen die hiervoor kunnen worden aangevoerd zijn dat enerzijds bij de opstelling van de vragen teveel is gericht op aspecten van interactie (en niet van het nut of iets dergelijks), en anderzijds dat de verwachtingen ten aanzien van het 'gebruik' van universitair onderzoek te weinig uitgingen van gebruik in de zin van (praktische) toepassing en teveel van gebruik in de zin van verdere ontwikkeling. Dat neemt niet weg dat in een volgend geval meer aandacht moet zijn voor wat er feitelijk gebeurt met het resultaat van onderzoek, c.q. de producten van de onderzoekseenheid. Een (telefonische/elektronische) survey blijkt een geschikt middel om redelijk snel een redelijk adequaat beeld te krijgen van een aantal aspecten betreffende de interactie tussen onderzoekers en omgeving, zeker wanneer het om een groot aantal verschillende gebruikers gaat. Feitelijk gebruik is veel moeilijker te onderzoeken. Daarvoor zouden case-studies nodig zijn. Gezien de tijdsinvestering die daarvoor nodig is moet dat bij een onderzoeksevaluatie beperkt blijven tot enkele specifieke gevallen.

De VSNU-beoordelingscommissie onderkende de aanvullende waarde van de uitgevoerde gebruikersanalyse. De grafieken van gebruikers (institutionele achtergrond en rollen) gaven een overzicht dat niet (of met veel inspanning) kan worden verkregen op grond van de zelfevaluatie. De gebruikersrapporten -waarin de resultaten van de interviews per onderzoeksprogramma werden samengevat- bevatte veel bruikbare aanvullende informatie.

5.5 Terugkoppeling

De combinatie van de resultaten van beide deelonderzoeken in een terugkoppeling op elkaar was bedoeld om 'inbedding', *performance* en leeromgeving te evalueren in het licht van de missie van het programma. In dit onderdeel van de studie ontstonden echter grote spanningen tussen het streven naar reductie van complexiteit voor overzichtelijkheid voor de beoordelingscommissie en de prijs die daarvoor aan nuance en detail moet worden betaald. Van de klassen die uit de gevonden REPPs zijn gedestilleerd vond men in het algemeen dat teveel van de specifieke kenmerken van het programma verloren gingen. De gevonden klassen hebben wel als *eye-opener* gewerkt voor de eigenaard van de groepen in de sector van de landbouwwetenschappen. Er bleek ook een spanning te ontstaan tussen de ideaaltypische klassen *industry-driven*, *policy-driven*, etc. die ons voor ogen stonden als ijkpunt voor een evaluatieve maatstaf en de mogelijkheden die de empirisch aangetroffen profielen boden tot classificatie. Daarnaast ligt een aan het profiel corresponderende weging van activiteiten en output ligt niet zomaar voor de hand.

De analytische classificatie om de gebruikersomgeving te karakteriseren werden wel geaccepteerd, maar hier deed zich het probleem voor dat de empirische fundering te mager werden gevonden. Het is de vraag of interviews met enkele gebruikers in alle gevallen een veralgemenisering rechtvaardigen. Classificaties behoeven een steviger theoretisch en empirisch fundament. Wel kunnen de klassen een oriëntatiepunt geven aan een discussie tussen beoordelaars en onderzoekers over de empirische realisatie van de missie van het programma. Voor de empirische verificatie van die realisatie moet dan echter worden teruggegrepen op het oorspronkelijke REPP en gebruikersrapport.

5.6 Implementatie in een reguliere beoordelingspraktijk: technische en culturele haalbaarheid

De besproken beoordelingsprocedure is zo ingericht dat de maatschappelijke betekenis van het onderzoek op evenwichtige wijze kan worden meegewogen in de evaluatie. Een discussie over de haalbaarheid van de methodiek kan niet stoppen bij aspecten van dataverzameling en methodiek (die hierboven al zijn behandeld). Minstens even belangrijk is het dat alle betrokken groeperingen overeenstemming bereiken over het uitgangspunt dat de relevante contexten van onderzoeksprogramma's onderling sterk kunnen verschillen, en dat zulks gevolgen heeft voor de beoordelingsprocedure. De ene groep is meer op toepassingen en het bedrijfsleven gericht, de andere groep is sterk op de internationale wetenschappelijke gemeenschap georiënteerd. We willen overigens niet treden in discussies over de mate waarin een 'wetenschappelijk gehalte' in iedere vorm van onderzoek wordt gehandhaafd, maar slechts vaststellen dat er verschillen bestaan.

Die verschillen komen ook tot uitdrukking bij groeperingen die eigen verwachtingen en vragen hebben naar het programma en alleen bepaalde opbrengsten honoreren. In interacties met sommige groeperingen staat maatschappelijke relevantie centraal, voor andere zijn de fundamentele wetenschappelijke vragen (d.w.z. onafhankelijk van maatschappelijk nut, waarde enz.) belangrijker. Bij de beoordeling van het programma zouden de behoeften en inzichten van alle relevante groeperingen *stakeholders* moeten kunnen meewegen. In de methode die hier wordt gepresenteerd wordt geprobeerd om de beoordelingscommissie de data zo aan te leveren dat een systematisch beeld ontstaat van de gehele context waarin de groep functioneert, wetenschappelijk en overige gebruikers. Zo wordt de groep betrokkenen die kan worden meegewogen in de beoordeling sterk verbreed.

Deze benadering wijkt op veel punten af van een methodiek die 'objectieve' indicatoren voor kwaliteit construeert (bibliometrie), of een methodiek die alleen aansluit bij de criteria van de

collega wetenschappers (*peer review*).

De methodiek die hier wordt gepresenteerd streeft juist naar een maatstaf die is toegesneden op de specifieke context van het onderzoeksprogramma. Het gaat ons er echter niet om oppositie te voeren tegen bestaande methoden, maar om een verbreding en aanvulling. Voorwaarde is wel dat wetenschap en overheid erkennen dat ook andere groeperingen in de samenleving vanuit hun specifieke ervaring een oordeel kunnen hebben over de kwaliteit en de relevantie van wetenschappelijke onderzoek, en deel nemen in de ontwikkeling van onderzoek naar toepassing (innovatie).

In de landbouwwetenschappen –met een lange traditie van werken aan maatschappelijke vragen– was het bij deze studie niet moeilijk om alle gezaghebbende instituties te overtuigen van het nut en de noodzaak van een dergelijke verbreding van de beoordelingen. In andere velden zal dat naar verwachting op meer culturele weerstanden stuiten.

Concluderend: In algemene zin is de haalbaarheid van deze methode afhankelijk van twee voorwaarden: technische en culturele.

De technische betreffen enerzijds verbeteringen in de methodische aanpak zoals hierboven besproken, anderzijds dient de administratie van onderzoek te worden afgestemd op de vereisten van data-verzameling voor deze methode. Cultureel is een omslag nodig bij de onderzoekers en gezaghebbende instellingen om de beoordelingen zodanig te verbreden dat ook de rol en het oordeel van andere betrokkenen dan wetenschappers in de evaluatie worden betrokken.

Afkortingenlijst

AIO:	Assistent in Opleiding
APHY:	Applied Physics
CGS:	Crop and Grassland Science
COS:	Commissie van Overleg Sectorraden
DLO:	Dienst Landbouwkundig Onderzoek
ESWC:	Erosion and Soil & Water Conservation
EU:	Europese Unie
FATE:	Farm Technology
Fte:	Full time equivalent
GREEN:	Greenhouse Horticulture
IMAG-DLO	DLO Instituut voor Milieu- en Agritechniek
IWEG:	Irrigation and Water Engineering Group
KCW:	KennisCentrum Wageningen (nu: Wageningen UR)
LNV:	Ministerie voor Landbouw, Natuurbeheer en Visserij
LUW:	Landbouwniversiteit Wageningen
NOVEM:	Nederlandse Onderneming voor Energie en Milieu
NRLO:	Nationale Raad voor Landbouwkundig Onderzoek
OIO:	Onderzoeker in opleiding
PAGV:	Proefstation voor de Akkerbouw en de Groenteteelt
PPS:	Plant Production Systems
REPP:	Research Embedment and Performance Profile
S&C:	Systems & Control
SCI:	Science Citation Index
SOIL:	Soil Tillage
SSCI:	Social Science Citation Index
TAO:	Technology and Agrarian Development
TPE:	Theoretical Production Ecology
VEPRO:	Vegetative Propagation
VSNU:	Vereniging van Samenwerkende Nederlandse Universiteiten
Wageningen UR:	Wageningen University Research Centre
WP:	Wetenschappelijk personeel

Burrell, G. and G. Morgan (1985),

Sociological Paradigms and Organisational Analysis, Elements of the sociology of corporate life, Gower 1985 (oorspronkelijk gepubliceerd bij Heinemann 1979)

Callon, M. et al. (1991),

'Tools for the evaluation of technological programmes: an account of work done at the Centre for the Sociology of Innovation', *Technology Analysis & Strategic Management*, vol. 3, no. 1, 3-41

Callon, M., P. Laredo and V. Rabeharisoa (1991),

'The management and evaluation of technological programmes and the dynamics of techno-economic networks: the case of AFME'. Internal research paper, Parijs

Callon, M. (1992),

'The Dynamics of Techno-Economic Networks', in: R Coombs, P. Saiotti, V. Walsh (eds.), *Technological change and company strategies*, London, Harcourt Brace.

Callon, M. (1994),

'Is Science a Public Good?', *Science, Technology and Human Values* 19 (4), 395-424

Callon, M., P. Larédo and P. Mustar (1994)

'Panorama de la science Française', *La Recherche*, 264, 378-384. (In het kader van deze studie is gebruik gemaakt van een Engelstalig conferentie paper over de compass card method, zie hieronder bij Larédo)

Chatelin, Y. en R. Arvanitis (1992)

'Representing scientific activity by structural indicators: the case of Côte d'Ivoire 1884-1968', *Scientometrics*, vol. 23, no. 1, 235-247

Crow, M and B. Bozeman (1987)

'R&D laboratory classification and public policy, the effects of environmental context on laboratory behavior', *Research Policy*, 16, 229-258.

Enzing, C.M., I. Limpens, and H.L. Schaffers (1997)

Methode voor kwaliteitsbeoordeling van de agrokennisinfrastructuur. Een haalbaarheidsstudie op basis van 3 case studies, NRLO report 97/23, Den Haag

Freeman, C. (1974)

The economics of industrial innovation, Penguin, Harmondsworth, UK.

Freeman, C. (1991)

'Networks of innovators: a synthesis of research issues', *Research Policy*, 20, 499-514

Gibbons, M. et al. (1994)

The new production of knowledge. The dynamics of science and research in contemporary societies, SAGE, Londen

Grin, J. en H. van der Graaf (1996)

'Technology assessment as learning', *Science, Technology and Human Values*, 21 1, blz. 72-99

Hertog, P. den et al. (1996)

User involvement in RTD concepts, practices and policy lessons, TNO/STB Apeldoorn

Inkster, I (1991)

Science and technology in history: an approach to industrial development. MacMillan, London.

- Joly, P.B. en V. Mangematin (1996)**,
'Profile of public laboratories, industrial partnerships and organisation of R&D: the dynamics of industrial relationships in a large research organisation', *Research Policy*, 25, 901-922
- Klein, J.T. (1990)**
Interdisciplinarity. History, theory and practice. Wayne State University Press, Detroit
- Kleisen, C.M. (1996)**
Impressies uit het veld. Interviews for the national foresight study of Agricultural Sciences, NRLO report 96/19, Den Haag/Amsterdam
- Knorr-Cetina, K (1990)**
'The couch, the cathedral and the lab: on the relationship between experiment and laboratory in science'. In A. Pickering ed. *Science as practice and culture*, Chicago University Press, Chicago
- Knorr Cetina, K (1982)**
'Scientific communities or transepistemic arenas of research? A critic of quasi economic models of science', *Social Studies of Science*, 12, p101-130
- Koningsveld, H. (1996)**
Landbouwwetenschap aan de universiteit van 2010. Essay for the national foresight study of Agricultural Sciences, NRLO report 96/15, Den Haag/Amsterdam
- Landbouwniversiteit Wageningen (1996)**
Strategisch Plan 'Richting 2000'
- Landbouwniversiteit Wageningen (1996)**
Instellingsplan 1997-2000
- Laredo, P. et al. (undated)**
'Defining the strategic profile of research labs: the research compass card', Paris
- Latour, Bruno & Steve Woolgar, (1979)**
Laboratory Life, Princeton University Press, Princeton
- Latour, B. (1996)**
Aramis or the love of technology, Harvard University Press, Cambridge, MA
- Le Pair, C. (1988)**
'The Citation Gap of Applicable Science', in: *Handbook of Quantitative Studies of Science and Technology*, Raan, A.F.J. van (ed.), Elsevier, Amsterdam.
- LNV-kennisbeleid ter discussie**
Aanzet tot maatschappelijk debat, *Spil*, 137-138/139-140, 1996, 1-2, themanummer
- Looze, M-A. de (1996)**
'Scientometrics as a tool for analysis of industrial relationships of two departments in a major French applied research institute, 1988-1992', *Scientometrics*, vol. 37, no. 2, 349-360
- Meulen, Barend van der, (1992)**
Beoordelingsprocessen in wetenschap; de constructie van kwaliteit door wetenschap, overheid en industrie, proefschrift, Universiteit van Twente, Enschede
- Meulen, B.J.R. van der (1996)**
Kennisproductie als wetenschap en praktijk: aard en verandering van de landbouwwetenschappen, division study for the national foresight study of Agricultural Sciences, NRLO report 96/18, Den Haag/Amsterdam
- Meulen, B.J.R. van der en A. Rip (1997)**
Maatschappelijke kwaliteit van onderzoek tussen verantwoording en management. Een inventarisatie van beoordelingspraktijken, Universiteit Twente, Enschede
- Meulen, B.J.R. van der, and A. Rip (undated)**
Indicatoren en indicaties voor de beoordeling van de maatschappelijke kwaliteit van onderzoek, Universiteit Twente, Enschede
- Ministerie van Landbouw (1995)**
Natuurbeheer en Visserij, *Nota Dynamiek en Vernieuwing*, Den Haag

Ministerie van Landbouw, Natuurbeheer en Visserij
LNV-Kennisbeleid tot 1999

Tweede Kamer der Staten-Generaal, vergaderjaar
1995-1996, 24 430, nrs 1 en 2, Den Haag

**Ministeries van Onderwijs, Cultuur en
Wetenschappen, Economische Zaken en Landbouw
en Natuurbeheer en Visserij (1995)**

*Kennis in Beweging. Over kennis en kunde in de
Nederlandse economie,*
Den Haag

**Ministerie van Onderwijs, Cultuur en
Wetenschappen,**

Wetenschapsbudget 1997
Tweede Kamer, vergaderjaar 1996-1997,
25 008, nrs. 1-2

**Ministerie van Landbouw, Natuurbeheer en
Visserij, Directie Wetenschap en Kennisoverdracht
(1995)**

*Agrokennisnetwerkvorming in de regio en de rol
van de betrokkenen daarbij. Een verkenning,*
Den Haag

**Ministerie van Landbouw, Natuurbeheer en
Visserij, Van institutioneel naar functioneel denken
(1995)**

Eindrapport van de projectgroep *kennistransfor-
matie en transfer*, aktie 9 en 10
landbouwkennisbeleid,
Den Haag

Morgan, G. (1989)

Images of Organization, Sage Publications,
London

Morgan, G. (1989)

Creative Organization Theory, Sage Publications,
Newbury Park

Nelson, R.R. and S.G. Winter (1977)

'In search of a useful theory of innovation',
Research Policy, 6, 36-76

Nelson, R.R. (1987)

*Understanding technical change as an evolutionary
process*, North Holland, Amsterdam

Nelson, R.R. (1993)

National systems of innovation,
Oxford University Press, New York

NRLO

*Landbouwkundig onderzoek op weg naar de
21^e eeuw. Meerjarenvisie Landbouwkundig
onderzoek 1995-1998*,
's Gravenhage (ongedateerd)

OCV (1996)

Wageningen in profiel. Landbouwwetenschappen
in 2010; de positie van de LUW, eindrapport
Overlegcommissie Verkenningen, Amsterdam

OCV/NRLO (1996)

De positie van de LUW in de toekomst.
Tussenrapportage in het kader van de verkenning
'Landbouwwetenschappen in 2010, en de positie
van de LUW',
Amsterdam

Peper, A. (1996)

Duurzame kennis, duurzame landbouw, Een advies
aan de minister van LNV over de kennisinfrastruc-
tuur van de landbouw in 2010 (onder meer advies
over fusie LUW/DLO tot Kenniscentrum
Wageningen)

Peters, I., P. Groenewegen en N. Fiebelkorn (1998)

'A comparison of networks between industry and
public sector research in materials technology and
biotechnology',
Research Policy, 27, 255-271

Rabbinge, R. (1996)

*Vernieuwing van de kennis-innovatiepyramide en
de positie van de LUW*, essay voor de OCV,
NRLO rapport 96/15, Den Haag/Amsterdam

Shinn, T. (1999)

'Change or mutation? Reflections on the
foundations of contemporary science', in *Social
Science Information*
38 1 blz. 149-176

Spaapen, J. en F. Wamelink (1998)

*Assessment of the embedment and performance of
12 agricultural research programs in the societal
and scientific environment*,
Amsterdam, sci_Quest

SPRU (1996)

The relationship between publicly funded basic research and economic performance, report prepared for HM the Treasury

Verkaik, A.P. (1997)

Uitdagingen en concepten voor toekomstig landbouwkennisbeleid, NRLO report 97/17, The Hague

VSNU, (1994)

Quality Assessment of Research. Onderzoekbeoordeling. Protocol 1994, Utrecht 1994

Noten

- ① Het volledige verslag is getiteld 'Assessment of the embedment and performance of 12 agricultural research programs in the societal and scientific environment' (Spaapen en Wamelink 1998).
- ② Zie bijvoorbeeld het Wetenschapsbudget 1997 p.42.
- ③ Daarbij is gebruik van onderzoek breed opgevat en niet alleen beperkt tot eindgebruiker.
- ④ De methodiek die hier wordt ontwikkeld is voor een deel afkomstig uit Callon et. al. 1994.
- ⑤ KCW staat voor 'Kenniscentrum Wageningen', tegenwoordig 'Wageningen University Research Centre'.
- ⑥ Het volledige verslag is getiteld 'Assessment of the embedment and performance of 12 agricultural research programs in the societal and scientific environment' (Spaapen en Wamelink 1998).
- ⑦ Zie Instellingsplan 1997-2000, p. 6).
- ⑧ Binnen het huidige VSNU protocol wordt de mogelijkheid geboden aan instellingen om de maatschappelijke relevantie van hun onderzoek nader te onderbouwen (VSNU protocol 1997).
- ⑨ Het disciplineprotocol is in deze richting bijgesteld. De beoordelingscommissie werd daarin gevraagd de gegevens over de maatschappelijke relevantie in de beoordeling te betrekken, i.c. de resultaten van deze studie.
- ⑩ De opdracht werd niet alleen door de LUW gegeven, maar ook door de NRLO en de COS.
- ⑪ Dit is een bewuste keuze geweest, maar het heeft wel commentaar opgeroepen. Het is de vraag of de representatie van het oppervlak zo sterk mag samenhangen met deze onderlinge bevestiging van interacties door indicatoren die enigszins toevallig naast elkaar staan. Dit punt komt later nog ter sprake. (zie ook bijlage 5).
- ⑫ Zie bijlage 3 voor een systematisch overzicht van de data en de bronnen van deze informatie.
- ⑬ Dat heeft ook het commentaar opgeroepen dat de sensitiviteit van de indicatoren onderling verschillend is. Hier is nader onderzoek voor de ontwikkeling van de methode gewenst.
- ⑭ Getalsmatig lijkt dit een wat vreemde labeling ('gemiddeld' hoort eerder bij 50 dan bij 75%), maar gezien de data die door de indicatoren worden gerepresenteerd en rekening houdend met een 'statistisch' gemiddelde groep leek dit de beste labeling. Maar er zijn nog wel wat problemen op te lossen in de transformatie en de labeling van de schaal.
- ⑮ Overigens wordt met de studie naar de interactie patronen geen organisatiekundige karakterisering van de groep gegeven. De uitkomsten van de analyse zeggen niets over de coherentie van het programma, interne samenwerking, taakverdeling, enzovoorts. De relatie tussen de managementstijl van het programma en de mate van 'inbedding' kan ook niet eenvoudig worden gelegd. In de organisatiesociologie wordt er bijvoorbeeld op gewezen dat kennis-intensieve organisaties en toegepast onderzoek wordt geconfronteerd met de eis om zich aan te passen aan een snel veranderende omgeving waarbij iedere keer nieuwe 'opportunities' moeten worden gezocht voor de bestaande vaardigheden en competenties. Innovatieve onderzoeksprogramma's in de meer toegepaste domeinen zullen in dat proces in constante flux en transformatie zijn. In de managementstijl zou een dergelijk programma meer als een capaciteitsgroep kunnen verschijnen dan als een coherente groep (zie ook 'mode 2' bij Gibbons et al. 1994). Juist ook om die reden wordt in deze studie de nadruk gelegd op de interacties met de omgeving als een maatstaf voor de mate waarin de 'missie' van het programma wordt gerealiseerd.
- ⑯ Het belang van de interactie tussen onderzoekers en gebruikers voor succesvolle innovatie is reeds lang aangetoond. (zie onder meer TNO-STB 1996: 3) Onderzoek naar die interactie tussen actoren in het kader van evaluatie is relatief nieuw.
- ⑰ Joly en Mangematin (1996) komen in hun studie van de INRA-instituten in 'plant breeding' en biochemie¹⁸ op grond van een aantal variabelen tot een indeling van drie typen publieke onderzoekslaboratoria:
 - (1) gericht op de professionele sector,
 - (2) gericht op fundamenteel onderzoek, en
 - (3) gericht op algemene vraagstukken van de industriële sector.
- ⑱ In deze meta-evaluatie is commentaar op de studie verzameld van: de VSNU beoordelingscommissie Landbouw-wetenschappen; de begeleidingscommissie; de betrokken onderzoeks-programma's; het bureau van de NRLO; en enkele sleutelfiguren uit de landbouwwereld; enkele beleids-medewerkers van de LUW en leden van de vakgroep Wetenschaps- en Technolgie-dynamica van de UvA.
- ⑲ Hier wordt geen enkel oordeel gegeven over de kwaliteit van de administratie (de publicatielijsten bleken bijvoorbeeld zeer volledig en uniform), het gaat hier om de mogelijkheden die de administratie biedt om te aggregeren, te desaggregeren en te classificeren.
- ⑳ Toch moet hier ook worden opgemerkt dat wij ons ervan bewust zijn dat de administratieve inspanning in verhouding moeten staan met wat het oplevert aan informatie.

Noten (vervolg)

- ②① Dit was geen uitzondering, veel van de programma's lieten hoge score's zien in dit domein. Deels is dit toe te schrijven aan de speciale PhD programma's, deels is het criterium gewicht aan de lage kant gekozen.
- ②② We gebruiken hier de impactfactoren zoals berekend door ISI voor de 1996 Journal Citation Records (let op de impactfactoren kunnen over de jaren schommelingen vertonen). Ook de indeling van tijdschriften naar velden van onderzoek door ISI wordt hier overgenomen.
- ②③ In de survey werd aan de programma's gevraagd om 100% activiteit te verdelen over wetenschappelijk gericht onderzoek en professioneel gericht onderzoek, daarnaast kon een eigen categorie worden toegevoegd.
- ②④ We zijn ons ervan bewust dat de periode van investeren en de periode van productie hier niet overeenkomen (zoals wel wordt verondersteld in deze berekening). De effecten van groei of terugval moeten daarom in de beschouwing worden betrokken.
- ②⑤ Alleen voor dit jaar werd daarover betrouwbaar gerapporteerd in de zelfevaluatie aan de VSNU.
- ②⑥ In hoeverre mode 2 werkelijk een 'nieuwe' vorm van kennisproductie is, waarin wetenschap als een aparte institutionele en cognitieve categorie vervaagt, wordt overigens door sommigen ernstig in twijfel getrokken. Shinn (1999) betoogt dat de veranderingen die zich nu in de kennisproductie lijken voor te doen historische precedents hebben.